Athina Vidaki*
Federica Giangasparo
Denise Syndercombe Court

Department of Pharmacy and Forensic Science, King's College London, Franklin-Wilkins Building, London, UK

Received June 3, 2016
Revised August 21, 2016
Accepted August 22, 2016

Research Article

Discovery of potential DNA methylation markers for forensic tissue identification using bisulphite pyrosequencing

The presence of specific body fluids at crime scenes could be linked with particular types of crime, therefore attributing a DNA profile to a specific tissue could increase the evidential significance of a match with a suspect. Current methodologies such as tissue-specific mRNA profiling are useful but drawbacks include low tissue specificity and applicability to degraded samples. In this study, the potential of 11 tissue-specific differentially methylated regions, initially identified following large-scale methylation analysis of whole blood, buccal cells and sperm, was explored in order to identify markers for blood, saliva and semen. Bisulphite pyrosequencing analysis supported previous findings, but tissue-specific differentially methylated regions for blood and buccal cells did not show enough specificity to be proposed as markers for blood and saliva, respectively. For some CpGs, a large inter-individual variation in methylation levels was also observed. Two of the semen markers (cg04382920 and cg11768416) were used for further validation on a large set of stains. These two semen-specific assays showed high sensitivity (as low as 50 pg) and stability. Future experiments will shed light on the usefulness of these markers in forensic casework.

Keywords:
Bisulphite pyrosequencing / DNA methylation / Forensic / Tissue identification
DOI 10.1002/elps.201600261

Additional supporting information may be found in the online version of this article at the publisher's web-site

1 Introduction

The weight of forensic scientific evidence could be enhanced in court if the cellular origin of a body fluid stain was identified. Alongside DNA typing, information regarding the cellular origin of a recovered biological stain would be very beneficial if attempting to reconstruct events that have taken place at a crime scene. The presence of specific body fluids could be linked with particular types of crime, for example the presence of semen might indicate sexual assault. However, sometimes it is very difficult, if not impossible, to attribute a DNA profile recovered from a stain to a specific body fluid type, limiting the evidential significance of a match. New solutions are needed to eliminate doubts as to whether the presence of an obtained DNA profile is a truly meaningful event in the context of the case, or simply the result of an unrelated, 'innocent' contact.

While forensic mRNA profiling using tissue-specific mRNA markers is a useful confirmatory test due to its great sensitivity and potential applicability [1, 2], it does not come without drawbacks. Issues have been raised regarding tissue-to-tissue specificity and its applicability in extensively degraded samples where RNA is more difficult to obtain. Such constraints are particularly important when re-examining 'cold cases', where only DNA has been retained and current methods cannot be used. A method that would exploit tissue-specific variations in the stability of DNA molecule would be helpful overcoming the limitations of existing methods and

Correspondence: Dr. Athina Vidaki, Department of Pharmacy and Forensic Science, King's College London, Franklin-Wilkins Building, 150 Stamford Street, London, UK
E-mail: athina.a.vidaki@kcl.ac.uk

Abbreviations: **BLM**, blood methylated marker; **BLU**, blood unmethylated marker; **BUM**, buccal methylated marker; **BUU**, buccal unmethylated marker; **NGS**, next-generation sequencing; **SEM**, semen methylated marker; **SEU**, semen unmethylated marker; **tDMR**, tissue-specific differentially methylated region

*Current address: Dr. Athina Vidaki, Department of Genetic Identification, Erasmus MC University Medical Center Rotterdam, P.O. Box 2040, 3000 CA, Rotterdam, the Netherlands
E-mail: a.vidaki@erasmusmc.nl

providing a direct link between the recovered DNA and its source.

It is known that epigenetics, and specifically DNA methylation, is one of the main mechanisms responsible for cell differentiation and differential gene expression [3, 4]. Researchers have investigated how DNA methylation regulates gene expression, which is mostly by silencing (or in some cases activating) gene transcription [5]; furthermore, there are various published studies investigating large-scale methylation patterns across various tissues that have either revealed tissue-specific [6, 7] or age-associated CpG sites [8, 9]. Therefore, differential DNA methylation patterns could offer an important alternative to mRNA profiling when differentiating forensically relevant tissues. Indeed, Frumkin et al. were the first to explore the possibility of DNA methylation-based forensic tissue identification [10]. Following initial screening of >200 CpG islands, a total of 38 genomic loci demonstrated differential amplification signals, 15 of which were subsequently used in a proposed tissue identification assay. Even though variability in methylation ratios, due to inter-individual variation or stochastic PCR effects, was observed, authors suggested that each tissue type had a distinct methylation profile. Further analysis also revealed that 100% identification could be achieved by using only seven of the 15 loci. The proposed method would also be very promising as a forensic application since it utilises the same platform used in standard STR profiling; however, incomplete digestion by the methylation-sensitive restriction enzyme could also lead to erroneous results. Furthermore, methylated loci not amplified because of degradation in aged samples could be mistakenly identified as unmethylated.

In addition, Wasserstrom et al. [11] developed a DNA methylation-based semen test (Nucleix DSI-Semen kit), which tested whether five genomic loci could be used to successfully distinguish between semen and non-semen samples with high accuracy. While a comprehensive validation study on the kit's performance illustrated that the required starting DNA material can be as low as 62 pg [12], a test that could simultaneously identify all body fluids would be preferable. In another study, Lee et al. tested previously reported tissue-specific differentially methylated regions (tDMRs) and proposed a different methodological approach that included bisulphite sequencing [13]. As the authors suggested, two previously reported testis-specific DMRs (*DACT1* and *USP49*) could be applied for semen identification, the presence of an unmethylated *HOXA4* tDMR could possibly be used to exclude the presence of blood, while *PFN3* tDMR could potentially be useful for the identification of vaginal secretions. Although the results were promising and were validated by two more sensitive multiplex assay systems [14], sex differences and inter-individual variations were once again observed.

It is evident that, with the exception of semen, determining highly tissue-specific tDMRs utilising CpG sites for the rest of the biologically more complex body fluids, such as saliva, still remains a challenge. While recent efforts have identified new potential markers for forensic use, mainly from genome-wide DNA methylation data [15–17], the identification of more suitable CpG sites would be beneficial. The ability to test multiple tissue-specific CpG sites per body fluid/tissue is needed in addition to deal with problems associated with inter-individual variation and mixtures.

The aim of this study was to identify novel tissue-specific differentially methylated CpG sites by analysing published methylation data obtained from blood, semen and buccal cells and subsequently validating their respective methylation levels through bisulphite pyrosequencing. The specificity and sensitivity of promising markers was evaluated by testing a large set of samples from a wide range of tissues and body fluids as well as aged stains.

2 Materials and methods

2.1 Sample collection

Biological samples included in this study were collected following full ethical approval by the appropriate Research Ethics Subcommittee at King's College London (BDM RESC 13/14-30). Full informed consent was obtained from all donors prior to sample collection. In total, 100 volunteers of both sexes and various ethnic backgrounds with ages ranging from 16 to 70 years participated in this study. Individuals had the choice to donate one or more body fluids/tissues including whole blood, saliva, buccal cells, seminal fluid, vaginal fluid, menstrual secretion, skin, and urine. Up to 20 mL of whole blood were collected by a trained phlebotomist in a clinical setting. Buccal cell samples were collected by rubbing a buccal swab in the inner cheeks for 20 s (mainly containing buccal cells), while saliva samples were collected by depositing mouth liquid (~1 mL) in a suitable receptacle (that could contain other non-epithelial cells or compounds, such as white blood cells). All other body fluid samples were collected either using a cotton swab or a suitable receptacle by the participants in the privacy of their homes. Information such as subject's gender, ethnicity and age were also recorded if possible. For validation purposes, a set of aged samples previously collected were also used, including semen samples stored at –20°C for 16 years and dried bloodstains stored at ambient temperature for 20–21 years.

2.2 DNA sample preparation

Depending on the type of sample and size/quantity, available genomic DNA was extracted using either the QIAamp DNA Investigator kit (Qiagen) or the BioRobot EZ1® DNA Investigator kit (Qiagen) according to the manufacturer's recommendations. To assess the quantity of DNA in the resulting solution, samples were quantified using the Quantifiler® Human DNA Quantification kit (ThermoFisher Scientific) according to the protocol. In order to convert differences in DNA methylation to differences in DNA sequence, DNA

samples are treated with sodium bisulphite, which converts unmethylated cytosines into uracil, while the methylated ones remain unchanged. In this study, the MethylEdge Bisulfite Conversion kit (Promega) was applied. The kit specifies that the method can be used for 100 pg–2 μg starting DNA material (optimal range 200–500 ng), providing DNA recovery of around 80%. For the purpose of this study, depending on the sample type 1–100 ng of each DNA sample was converted using the 'modified conversion protocol' as suggested by the manufacturer. The amount of DNA to be treated depended on the experiment and availability (especially in the cases of aged stains); for example, to test the specificity of the selected markers, a total of 1 ng of each DNA sample was used for bisulphite conversion, while to test the sensitivity of the assays decreasing amounts (as low as 10 pg) were used. Bisulphite-treated DNA samples were eluted in 20 μL of elution buffer and stored at –20°C for up to 1 month. Together with the samples, pre-defined DNA methylation controls (0–100%) (EpigenDx) were also used in order to assess both bisulphite conversion efficiency and the linearity of methylation quantification of selected assays.

2.3 Selection of candidate CpG markers

Combining available data from two studies analysed by different large-scale methodologies for three forensically relevant tissues (whole blood, buccal cells and sperm), the methylation status of a total of 3305 CpGs was obtained. Data were normalised using appropriate methods in order to account for batch effects or study-to-study technical variations as previously described [6, 8]. Methylation data for seven blood and two sperm samples were recovered from the study by Rakyan et al. [6], while methylation values for ten buccal swabs were gathered from Rakyan et al. [8]. No data were available for body fluids like vaginal fluid and menstrual blood, therefore it was decided that these tissues would be included in the validation phase only.

2.4 Bisulphite Pyrosequencing® assay design

In this study, primers were designed to amplify bisulphite-treated DNA; design parameters needed to be adjusted to account for the generally low efficiency of bisulphite PCR, common mis-priming events as well as the non-specific amplification due to the T-richness of the bisulphite-treated DNA sequences. There is various primer design software designed for methylation analysis; however, the online tool BiSearch was considered as most suitable and selected for assay design. Its search algorithm has the unique ability of analysing the proposed primer pairs for potential mis-priming sites and therefore resulting non-specific amplification products on both bisulphite-treated DNA strands. [18]. A total of 12 assays were designed to investigate the CpG sites identified through the analysis of genome-wide methylation data; each bisulphite Pyrosequencing® assay includes a 10× PCR primer set (forward and reverse) as well as a 10× sequencing primer. Although we were interested only in these 12 CpG sites, assay design allowed for the co-analysis of a few adjacent CpG sites per assay resulting in the investigation of a total of 48 analysed CpGs. To simplify data analysis, the assays were named according to the tissue they were specific for (BL for blood, SE for semen and BU for buccal cells) and if they were expected to be unmethylated (-U-) or methylated (-M-) in that specific tissue. For example, the blood methylated marker 1 (BLM1) (cg13763232) assay indicates that the investigated CpG site is expected to be methylated in blood, while being unmethylated in semen and buccal cell samples. Sequential numbers were used to indicate different CpG sites being examined. Information regarding all designed assays is presented in Table 1 and Supporting Information Table 1.

2.5 Amplification of bisulphite-converted DNA

In this study, ZymoTaq premix (Zymo Research) was utilised as it contains a heat-activated, "hot start" DNA polymerase that reduces the occurrence of non-specific product or primer–dimer formation when amplifying bisulphite-converted DNA. Starting with the standardised PCR conditions suggested by the manufacturer, all bisulphite PCR assays were optimised before analysis in order to avoid mis-priming, primer self-annealing or the formation of non-specific PCR products that could affect the accuracy of methylation detection. Each assay was optimised using an annealing temperature gradient, various concentrations of $MgCl_2$ and primer, as well as different PCR cycling conditions. The optimisation of the assay designed for the BLM2 marker was very challenging, with the amplification efficiency being very low at all tested PCR conditions and it was decided that this marker is excluded from further analysis. Failure to optimise this particular assay is believed to be due to its complex, CpG-rich DNA sequence. All other 11 assays performed well resulting in the detection of only the desired PCR product.

Briefly, each PCR reaction consisted of 12.5 μL of ZymoTaq PreMix, 1 μL of 25 mM $MgCl_2$ for a final concentration of 2.75 mM (since the ZymoTaq™ Premix also contains 1.75 mM $MgCl_2$), 1 μL of each PCR primer (for a final concentration of 0.4 μM), 1 μL of bisulphite DNA template and 8.5 μL of nuclease-free water, for a total reaction volume of 25 μL. The thermocycling program used was as follows: 95°C for 10 min, followed by 45 cycles of 94°C for 30 s, T_m for 30 s (SEU1—48°C (where stands for semen unmethylated marker), BLU1—50°C (where BLU stands for blood unmethylated marker), BLM1, BLU2, SEM1 (where SEM stands for semen methylated marker), BUM1 (where BUM stands for buccal methylated marker), BUM2, BUU1 (where BUU stands for buccal unmethylated marker), BUU2—55°C, SEM2—57°C and SEU2—61°C), 72°C for 35 s and a final extension step of 72°C for 2 min. Specifically for BLU1, BUM1,

Table 1. Designed bisulphite PCR assays

Assay	CpGs		Primer sequence (5′→3′)	Length (bp)	GC (%)	Converted Cs	PCR product (bp)
BLM1	4	F	TAGTTGATATTGGTTTGGTA	20	30	5	159
		R	CAAATAACTCAATTTCTCTAC	21	28.6	3	
BLM2	2	F	AAGTGTTGGGATTTTAGGAGT	21	38.1	2	180
		R	CCTCTTAATTTTCTTTTAAAAAC	23	21.7	1	
BLU1	3	F	GGTTTATTGTTTTGTATTAT	20	20	6	127
		R	AAATTCTCCAACACCACC	18	44.4	2	
BLU2	5	F	GAGTTATTTTTTTTGGTGTTGGAT	24	29.2	8	188
		R	ACATCCCCTTAAATTACTTT	20	30	4	
SEM1	3	F	ATGATTTAGTGGTTGGTAGGAA	22	36.4	3	147
		R	AACACCCCTAAAAAAAAC	18	33.3	7	
SEM2	3	F	AGTAAGTAGGAAGTGAATTGA	21	33.3	3	89
		R	ATATCTCAAAACAACCCAAA	20	30	6	
SEU1	10	F	TTTTATTAGAAAGTTTAGG	19	21.1	7	280
		R	ACAACAATAACTAAAAATAAATAC	24	16.7	6	
SEU2	5	F	GGAGGTTGTTTTTTTTTTGGTTT	23	30.4	6	134
		R	CTACCAACACCTTCCTCC	18	55.6	1	
BUM1	6	F	GTAGAGTTTTATTTTTTGTT	20	20	7	357
		R	CTCCTCCACCATAACCTA	18	50	3	
BUM2	4	F	TAGAGATAGATGGGTTTG	19	36.8	3	112
		R	CTAAATTCCTACAATATTCC	20	30	4	
BUU1	1	F	GAAAGGTGAGTTATAGAATAGTT	23	30.4	3	198
		R	CAAAATAAATCTCTCCCTT	19	31.6	2	
BUU2	2	F	TTGAGATGTTATAAGAGTATTGG	23	30.4	5	196
		R	ACTACTCCCTAAAAAAAC	18	33.3	7	

BUM2, BUU1 and BUU2, the annealing and extension step of each cycle was 40 s. Following amplification, the quality of PCR products was assessed on a 2% agarose gel.

2.6 Template preparation for Pyrosequencing® reactions

PCR products were converted to single-strands through biotin-streptavidin selective binding: 10 μL of PCR products were mixed with 3 μL of Streptavidin Sepharose High Performance Beads (GE Healthcare) and 37 μL of PyroMark Binding Buffer (Qiagen) as well as 30 μL of distilled water for a total volume of 80 μL. The solutions were then vortexed at 1000 rpm for 30 min using a clear non-skirted 96-well plate (Starlab) to allow for efficient binding of the PCR products to the beads. Afterwards, the beads were isolated and captured utilising a Vacuum Prep Workstation (Qiagen) according to the manufacturer's recommendations. Sequencing reactions were performed by mixing the single-stranded templates with 11.5 μL of PyroMark Annealing buffer (Qiagen) and 0.5 μL of the appropriate 10 μM sequencing primer. To ensure complete denaturation of DNA templates, these were heated at 80°C for 3 min. Reactions were left at room temperature for 5 min in order for the sequencing primer to bind before loading them on a PyroMark MD Pyrosequencer (Qiagen). All four nucleotides together with the enzymes and substrates (PyroMark Gold Q96 reagents; Qiagen) were placed into the instrument before analysis following the manufacturer's instructions.

Technical issues that needed to be resolved in this step included wide peaks, peaks with 'ski-slope' effect or low signal detected for the internal 'dead dispensations', all of which were eliminated following optimisation of the pyrosequencing reactions. Supporting Information Fig. 1 illustrates example pyrograms™ for all designed assays as obtained from the analysis of one blood sample. Generally, and as expected, the developed assays demonstrated different PCR efficiencies as indicated by the peak heights, with SEU1 showing on average the lowest peak heights taking into account all samples.

2.7 Data analysis

Only pyrograms™ passing the instrument's quality control were used for analysis; this included the detection of desired peaks (no signals in dead injections) as well as detected bisulphite conversion rates of >95%. Bisulphite conversion rates were calculated using the peak heights of thymine and cytosine of each bisulphite conversion control (non-CpG site cytosines) and were on average >98%. If bisulphite conversion rates were lower than 95%, the treatment with sodium bisulphite was repeated. For each CpG site, the degree of methylation was measured as a frequency of C/T signals in the form of peak heights. Pyrosequencing® dispenses both thymine (T) and cytosine (C) for each CpG site; the peak height of thymine corresponds to the unmethylated sequence (C), whereas the peak height of cytosine corresponds to the methylated sequence (mC). The principle of

methylation detection can be described by the ratio of mC:C at each CpG site using the following formula:

$$\%\text{Methylation} = \left[(\text{Peak height } C) \,/\, (\text{Peak height } C + \text{Peak height } T)\right] \times 100$$

For this analysis, the dedicated CpG methylation software PyroMark CpG SW 1.0 (Qiagen) was employed. Detected methylation data distribution was visualised by employing box-and-whisker plots. Lastly, for the most promising markers and as mentioned above, DNA standards of known methylation levels (0–100%) were used to assess the linearity of methylation quantification. The observed methylation ratio was plotted against the expected and the best-fitted regression line (linear) was chosen. For data analysis, the IBM SPSS v.22 software was employed.

3 Results

3.1 Selection of genetic loci

In order to identify potential CpG sites of interest, the average methylation level of each tissue for every CpG site was used. Initially, the criterion of choosing suitable CpG sites was set as a minimum of 70% methylation difference between the tissue in question and the remaining two, demonstrating an 'on/off' methylation pattern. For example, a 'good' blood-specific marker would be one that showed to be methylated in blood (>80%) and unmethylated in sperm and buccal cells (<10%) or vice versa. As a result, 14 blood-specific, 20 saliva-specific and 365 semen-specific CpG sites were identified. The number of semen-specific CpG sites was relatively high, which could reflect the different and unique functions of sperm DNA; thus, for semen only, the minimum difference of methylation levels was increased to 85% and the number of potential CpG sites chosen decreased to 22. Both tissue-specific methylation and de-methylation were observed, although the latter was observed comparatively more often. Supporting Information Table 2 summarises the observed methylation values for the top candidate markers for these three tissues.

For the purpose of this study, four CpG sites demonstrating the highest methylation difference between the above-mentioned tissues were selected for each body fluid (two unmethylated and two methylated markers) providing a total of 12 potentially tissue-specific CpG sites for the Pyrosequencing® assay design (Supporting Information Table 3). The online Ensembl genome browser, and in particular the human GRCh37/hg19 genome, was used to obtained the required genetic information for assay design. The exact chromosomal locations were confirmed and the surrounding DNA sequences were identified. Information on the genes and their function, as well as known SNPs included in the regions of interest, was also obtained. Most of the selected CpGs belong to protein-coding genes and are usually located within their 5′ end or within the main body of the gene.

3.2 Verification of methylation patterns using pyrosequencing

To verify the DNA methylation patterns of the selected CpG sites reported by Rakyan et al. [6, 8] and confirm that they are specific for blood, semen and buccal cells (Supporting Information Table 3), a set of samples were analysed for each assay. It was important to confirm the methylation levels of the proposed CpG sites in these three types of tissues first, before any analysis is performed using other forensically relevant tissues such as menstrual blood or vaginal fluid. However, since the reason for choosing markers differentially methylated in buccal cells was their potential use in the identification of saliva, a set of saliva samples were co-analysed. The aim of this experiment was not only to verify the previously reported methylation but also perhaps potentially select the best CpG sites of each assay for further validation. Most assays also investigate the methylation status of adjacent CpG sites and their potential in identifying the tissue of origin was also tested. For this analysis, 1 ng of each DNA sample was bisulphite treated and eluted in 20 μL, followed by the amplification of all 11 assays in singleplex reactions. Depending on the assay and the number of different tissues tested, the number of samples analysed per assay ranged from 50 to 110 aiming for at least ten samples per tissue.

3.2.1 Potential saliva-specific CpG sites

The markers cg15731815 (BUM1) and cg08258650 (BUM2) have been reported to be methylated in buccal cells (0.84 and 0.77 proportionately), while demonstrating low methylation levels in blood and semen (0.05–0.08) (Supporting Information Table 3). Although in the current study the expected methylation status in blood and semen was confirmed, the methylation ratio in buccal cells in this study was lower (average of 0.47 and 0.48 accordingly) (Fig. 1A and B). Therefore, we can report that both these markers can act as buccal-specific markers as there is still a sufficient difference compared to blood and semen. However, the analysis of saliva samples resulted in methylation levels ranging from 0 to 0.4 (mean of 0.17 and 0.12, respectively); although higher than blood and semen, it is believed that this methylation difference is too narrow for conclusive results. Finally, given that the identification of buccal cells in a forensic scenario is not frequently applicable, these markers were excluded from further analysis. Additionally, the marker cg05761971 (BUU1) had been previously reported to be unmethylated in buccal cells (0.11), while being highly methylated in blood and semen (0.87 and 0.88, respectively). Similar to above, the methylation difference in this tissue was smaller than expected as we found an average methylation ratio of 0.37 in buccal cells (Fig. 1C). Furthermore, BUU1 also demonstrated a very large inter-individual methylation variation when saliva was analysed ($n = 15$), hence it was concluded not suitable for saliva identification. Moreover, although cg16779976 (BUU2) was reported

Assay	CpG site	Number of samples			
		Whole blood	Semen	Buccal cells	Saliva
BUM1	cg15731815	14	12	8	4
BUM2	cg08258650	15	11	7	12
BUU1	cg05761971	14	10	8	15
BUU2	cg16779976	11	9	9	10

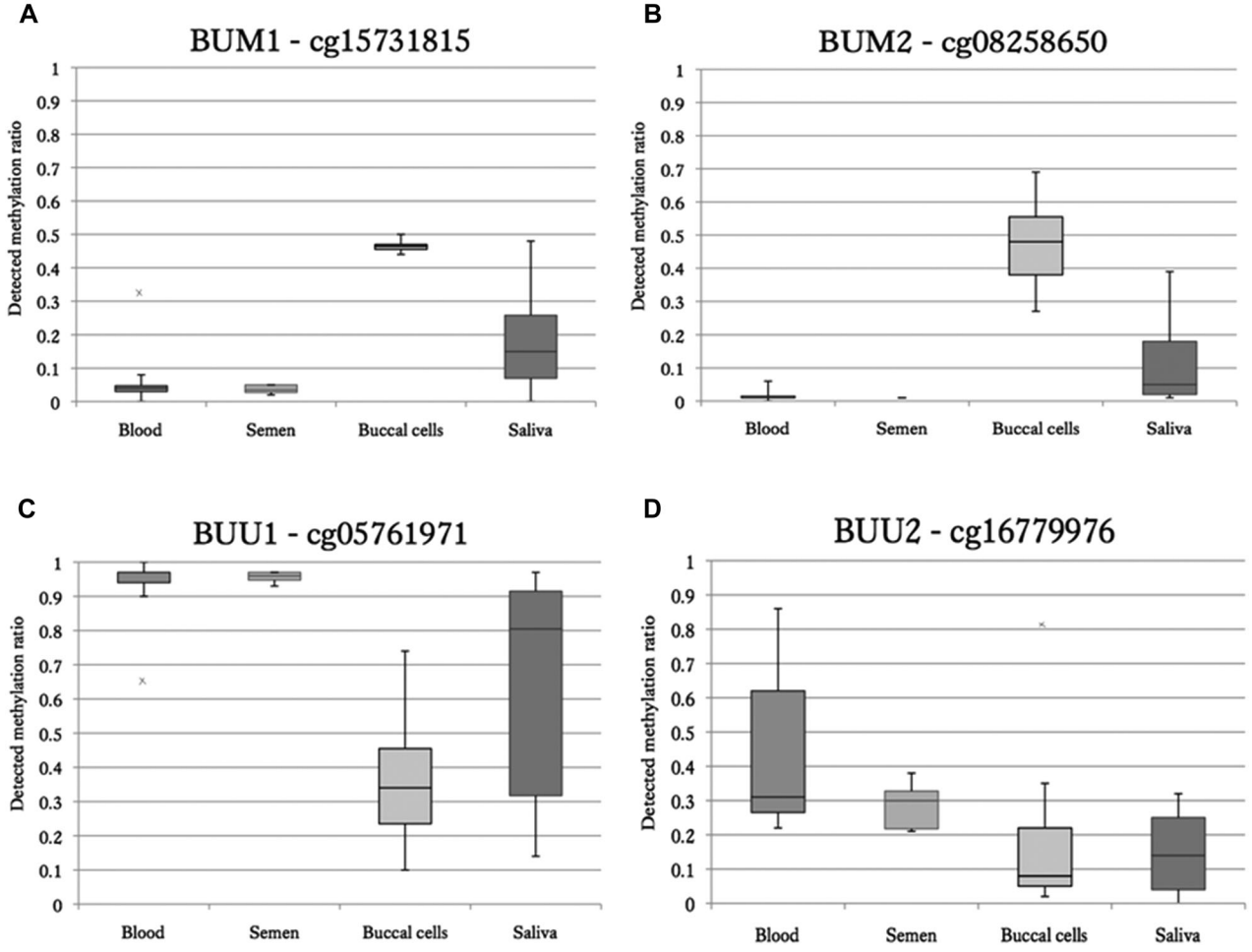

Figure 1. Methylation levels of proposed buccal cell specific markers. Observed methylation ratio of each CpG site (A-BUM1, B-BUM2, C-BUU1 and D-BUU2). Data are presented in the form of box-and-whisker plots showing the first and third quartiles (boxes), the median (horizontal line) and minimum and maximum (error bars) methylation values detected. Outliers (methylation value ≥3SD) are shown separately as '*x*' dots.

to be methylated in both blood and semen by genome-wide analysis, this observation was not confirmed in this study. The methylation levels of buccal cells and saliva were similar and slightly lower than the other two body fluids but not discriminatory enough to include the marker in further analysis (Fig. 1D).

In conclusion, although the methylation patterns of three out of four selected CpG sites were successfully verified in blood and semen, the methylation difference compared to buccal cells was smaller than previously reported. Additionally, it was clear that saliva did not share the same methylation profile with buccal cells but, especially for BUM1 and BUM2, gave very similar methylation values with blood. This leads to the conclusion that possible presence of leucocytes in saliva (e.g. due to gum bleeding) could be the reason for such observation. It has previously been shown that up to 74% of the recovered DNA in saliva can come from white blood cells [19]. The amount of 'contaminating' blood cells or

other cell types rather than buccal cells in saliva could vary between individuals and this could potentially explain the large inter-individual methylation ratio observed in some cases.

3.2.2 Potential blood-specific markers

Although there were a total of four potentially blood-specific CpG sites initially selected, only three of them could be verified since the BLM2 PCR assay failed the optimisation step. CpGs cg17518965 (BLU1) and cg26285698 (BLU2) were reported to be non-methylated in blood (0.02 and 0.09, respectively) while being highly methylated in semen and buccal cells (0.86–0.96). In this study, the methylation profile of blood was confirmed, however the methylation levels in semen and buccal cells demonstrated large inter-individual variation (Fig. 2A and B). Furthermore, saliva samples once again demonstrated very low methylation, even though it was originally thought that they would be methylated at these two chosen CpG sites. Even though, it was thought that BLU1 could act as a semen-specific marker, it was decided it was not desirable to have similar methylation profiles in semen and buccal cell samples from a forensic perspective, and especially in cases of complex or mixed stains. Tests of a small number of vaginal and skin samples resulted in similar methylation values as buccal cells, which could be explained by the shared epithelial tissue origin (data not shown).

Interestingly, there were two semen samples that showed very low methylation (outliers, Fig. 2A), either because of natural variation in methylation levels (16.7% of semen samples tested) and/or possible presence of blood in semen. Previous research has reported the presence of white blood cells in semen samples of infertile men or men with bacterial infections [20]. This could be partially supported when comparing the DNA yield following DNA isolation using the same staring material from all semen samples. These two samples resulted in a tenfold decrease in DNA yield compared to the average yield, potentially indicating a smaller number of spermatozoa present. Lower sperm count is often associated with infertility problems and it has been previously shown that alternations in sperm DNA methylation at particular loci are common with low sperm motility and different types of male infertility [21–23].

On the other hand, the findings regarding the third marker cg13763232 (BLM1) were more promising. Initial results analysing the four tissues revealed a distinct blood methylation profile as all blood samples resulted in >0.85 methylation ratio while semen and buccal cells verified the expected low methylation (0.12 and 0.21, respectively) (Fig. 2C). Even though the methylation levels of saliva samples ranged between 0.16 and 0.87 (mean = 0.60), it was decided that the investigation of other forensically relevant tissues was vital before final conclusions regarding this marker's specificity are made. Thus, an additional set of 34 samples (nine vaginal fluid, 14 menstrual blood, five skin and six urine samples) were analysed, the detected methylation levels of which are shown in Fig. 2C.

As results revealed, BLM1 was shown to be highly methylated only in blood, while the rest of the tested tissues demonstrated various profiles being either non-methylated or partially methylated. Similar methylation patterns were also obtained for the three adjacent CpG sites, indicating a shared profile within the locus. However, and for example in the case of skin samples, a wide range of methylation ratio (0.13–0.94) was observed. From a forensic perspective and considering that skin/'touch DNA' contamination often occurs in surface swabs, employing the BLM1 assay would not be sufficient to confirm the possible presence of a minute bloodstain with confidence.

3.2.3 Potential semen-specific markers

For the identification of semen, two (cg04382920-SEU1 and cg11768416-SEU2) reported to have no methylation in semen (<0.02) and two (cg01318557-SEM1 and cg05656364-SEM2) with more than 0.9 methylation were tested. The exact opposite methylation status was shown in blood and buccal cells. These patterns were successfully verified indicating the potential of these CpGs in the identification of semen. However, to establish their specificity a total of 110 samples including various tissue types were analysed (Fig. 3). When testing for SEM1 and SEM2, all non-semen tissues including whole blood, buccal cells, saliva, vaginal fluid, menstrual blood, skin and urine demonstrated very low levels of methylation, while semen resulted in a mean of 0.40 and 0.59, respectively. However, as shown in Fig. 3A and B, there was a large methylation range among semen samples.

Occasionally, in SEM1 there were a few outliers obtained for blood, saliva and urine, which is believed to be due to natural methylation variability; for urine, the presence of contaminating semen cannot also be excluded since the particular sample belongs to a male volunteer. Although semen clearly demonstrated a different distribution in methylation compared to other tissues, it was believed that the observed outliers could introduce uncertainties in confirming semen using this CpG site. Similarly, for SEM2 there were two saliva samples resulting in 0.32 and 0.35 methylation ratio, respectively, while there were two semen samples that were found completely unmethylated. Interestingly, these two semen samples were the same with the ones showing an 'opposite' methylation profile in the BLU1 assay and were the ones with low sperm count. Since these samples represent 15% of the total analysed semen samples, it was thought that this marker will not be included in further validation in this study, but increasing the sample size in future studies is vital to assess how significant these outliers are.

On the other hand, as illustrated in Fig. 3C and D, both SEU1 and SEU2 seemed to be highly specific markers for semen. The results of the genome-wide methylation analysis were confirmed for both markers, although the obtained methylation for SEU2 in non-semen tissues was slightly lower than originally reported (0.75). There was one saliva sample (5%) that showed low methylation levels, but this is believed

Assay	CpG site	Number of samples							
		Whole blood	Semen	Buccal cells	Saliva	Vaginal fluid	Menstrual blood	Skin	Urine
BLU1	cg17518965	15	12	9	19	N/A	N/A	N/A	N/A
BLU2	cg26285698	15	12	10	20	N/A	N/A	N/A	N/A
BLM1	cg13763232	15	12	10	20	9	14	5	6

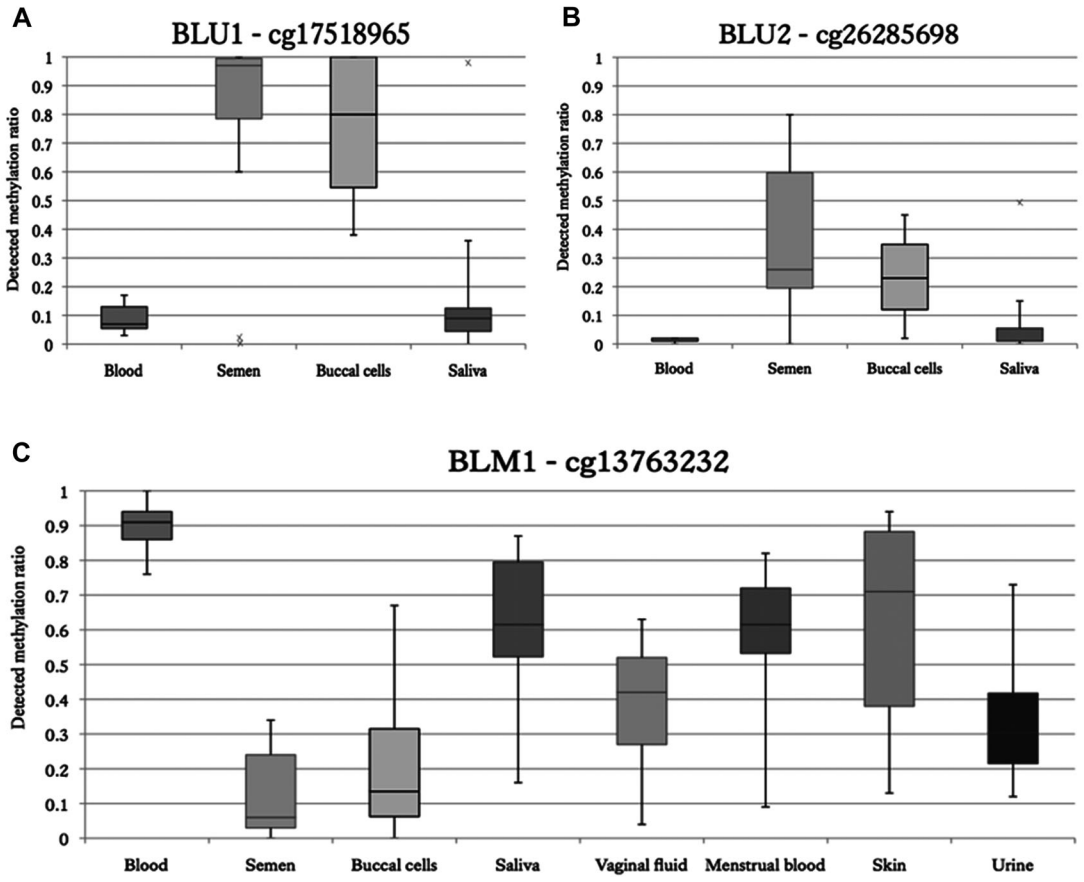

Figure 2. Methylation levels of proposed blood-specific markers. Observed methylation ratio of each CpG site (A-BLU1, B-BLU2 and C-BLM2). Data are presented in the form of box-and-whisker plots showing the first and third quartiles (boxes), the median (horizontal line) and minimum and maximum (error bars) methylation values detected. Outliers (methylation value ≥3SD) are shown separately as '*x*' dots.

to be due to natural methylation variability. Thus, we can conclude that using these two CpG sites (SEU1 and SEU2) no false negative results were obtained and only one out of the total 154 non-semen samples resulted in a false-positive identification of semen, which is particularly important in a forensic scenario. Interestingly, most of the co-analysed adjacent CpG sites (nine CpGs in SEU1 assay and four CpGs in SEU2 assay) also demonstrated the observed semen-specific methylation pattern; therefore, they can be used all together as semen-specific loci. In this study, we aimed for the highest methylation difference possible (on/off methylation), which for these two assays was successfully obtained. However, the threshold of the methylation differences between the semen and non-semen samples was CpG specific, with some demonstrating higher average differences (>0.7).

3.3 Validation of semen-specific assays (SEU1 and SEU2)

As previously shown, CpGs cg04382920 (SEU1) and cg11768416 (SEU2) demonstrated a semen-specific DNA methylation profile; however, in order to implement such markers in forensic casework, extensive validation of the associated methylation assays is required. Initial validation of these assays included testing their accuracy through linearity

A

Assay	CpG site	Number of samples							
		Whole blood	Semen	Buccal cells	Saliva	Vaginal fluid	Menstrual blood	Skin	Urine
SEM1	cg01318557	10	12	15	12	10	15	9	7
SEM2	cg05656364	15	14	15	20	10	15	10	10
SEU1	cg04382920	9	9	12	11	10	10	2	5
SEU2	cg11768416	15	13	15	20	10	15	10	10

B

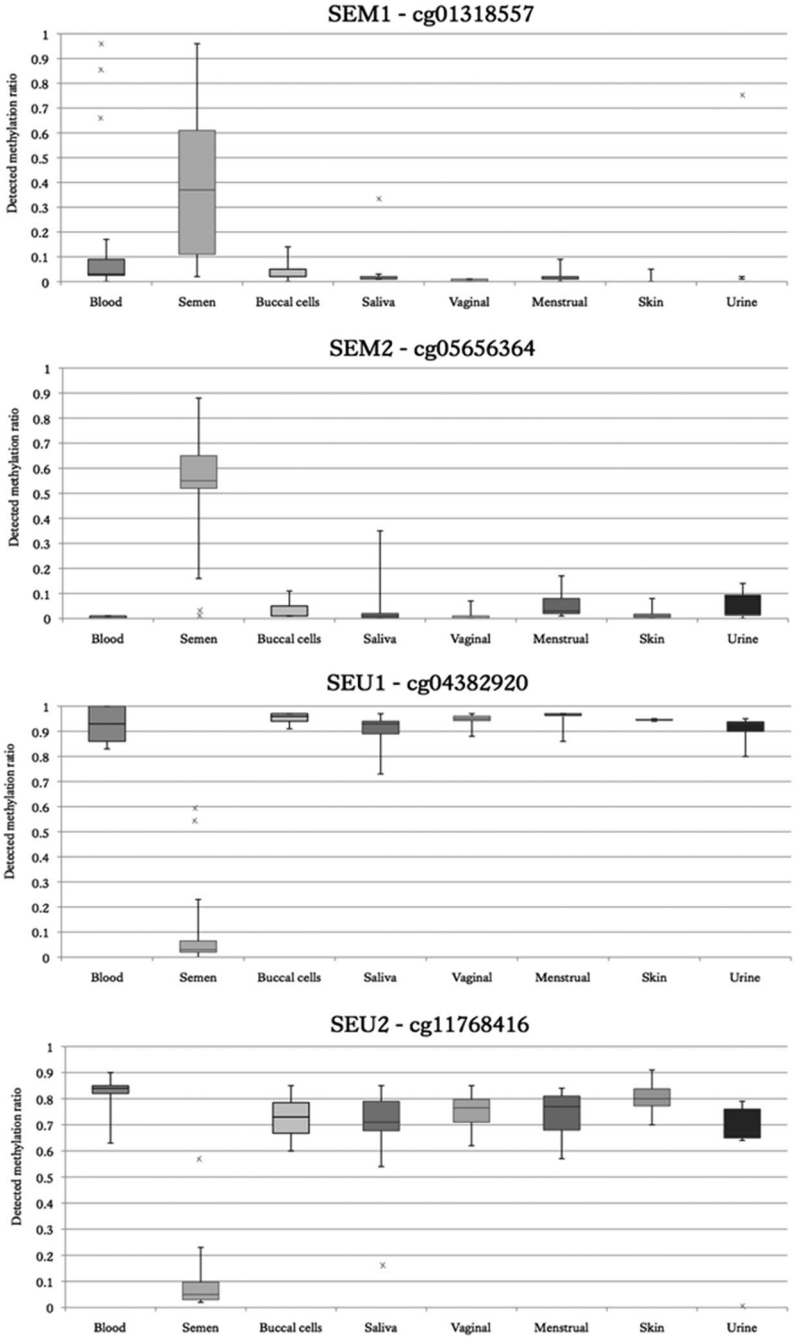

Figure 3. Methylation levels of the proposed semen-specific markers. (A) Number of samples used and (B) observed methylation ratio of each CpG site (SEM1, SEM2, SEU1 and SEU2) in various body fluids/tissues. Data are presented in the form of box-and-whisker plots showing the first and third quartiles (boxes), the median (horizontal line) and minimum and maximum (error bars) methylation values detected. Outliers (methylation value ≥3SD) are shown separately as '*x*' dots.

analysis, sensitivity as well as testing aged semen to evaluate potential applicability in cold cases.

3.3.1 Linearity of methylation quantification

To assess the accuracy of SEU1 and SEU2 assays, 100 ng of each DNA methylation control (0–100%) was analysed in duplicate. The mean and standard deviation of each standard was then calculated taking into account all CpG sites included in the sequences (10 CpGs in SEU1 and five CpGs in SEU2) and used to assess the linearity of quantifying methylation ratios. As shown in Supporting Information Fig. 2, both assays resulted in linear quantification (SEU1: $R^2 = 0.97$ and SEU2: $R^2 = 0.99$) and therefore no evidence of amplification bias commonly observed in other bisulphite PCR assays in the literature [24]. Methylation quantification was more challenging and less accurate when analysing partially methylated standards, which could be explained by stochastic variations or pipetting errors during amplification.

3.3.2 Sensitivity analysis

The sensitivity of the proposed semen-specific methylation assays was assessed by analysing decreasing amounts of starting DNA material (10 ng, 1 ng, 500 pg, 100 pg, 50 pg and 10 pg) in duplicate. For this experiment, non-semen DNA (blood) was used to assess if false-positive results would be obtained due to the low amounts of DNA used. As shown in Supporting Information Fig. 3A, successful amplification and the expected blood methylation pattern for SEU2 was obtained with as low as 50 pg of starting DNA, which corresponds to around ten cells. In particular, for the semen-specific CpG site cg11768416, the detected DNA methylation levels were constantly highly methylated for all tested amounts (ranging from 0.87 to 0.97), which fully corresponds to the expected blood methylation of this marker as obtained in the specificity experiment (Fig. 3). This small variation in DNA methylation as resulted when using decreasing starting DNA amount was also obtained for the neighbour CpGs in the SEU2 assay. This is very promising for the analysis of low-quantity or degraded samples, indicating the assay's potential applicability in forensic casework. Bisulphite conversion, as the first step of the proposed method, is essential in determining the overall sensitivity of bisulphite pyrosequencing assays. In our experiments, the detected conversion seemed to be slightly affected by the starting DNA amount, nevertheless an average of 91.5% conversion using all three controls included in the SEU2 assay was obtained for all amounts (Supporting Information Fig. 3B). It should be noted that for the first bisulphite control, conversion rates were higher (average of 96.2%) while the third control resulted in a mean rate of 85.4%. It is believed that this is due to low S/N (low peak heights) or due to signal carry over from previous unincorporated cytosines. Similar results were obtained also for the SEU1 assay, where for all amounts high methylation levels of cg04382920 (>92%) were detected in blood (data not shown).

3.3.3 Stability of CpG methylation and aged samples

To further test the SEU1 and SEU2 markers' DNA methylation stability as well as the assays' potential applicability in forensic casework, a set of aged and potentially degraded samples were tested. DNA from a set of nine semen samples was extracted shortly after collection and following storage at –20°C for 1 year. Although the obtained DNA amount per extracted microlitre of semen decreased over time, following the treatment of 10 ng of DNA the methylation status for all semen samples was found to be <0.2 methylated in both loci. No significant difference was obtained between the methylation values of fresh and stored semen samples ($p > 0.05$). In addition, a set of four semen stains on fabric (cotton) stored at –20°C for almost 16 years were also analysed. No false-negative results were obtained as all four semen stains produced low methylation levels (<0.1). Furthermore, together with the semen stains, another five blood stains on fabric (cotton) stored at ambient temperature for more than 20 years were analysed. Once again, no false positive results were obtained as all five blood stains gave high methylation levels (>0.6), despite their lower peak heights (<70 rfus), which could be explained by the low DNA quantity recovered from these stains. The detected methylation patterns for both semen and blood matched those obtained by freshly collected body fluid samples, highlighting the stability of DNA methylation when samples are stored both at room temperature and –20°C.

4 Discussion

Although forensic researchers have investigated in depth the potential of applying mRNA profiling in tissue identification [2, 25] and research on developing robust assays is still ongoing [26, 27], it is thought that in certain forensic scenarios, such as in cold cases or cases that involved degraded stains, its application would not be feasible. For this purpose, DNA methylation patterns among tissues were evaluated for their potential in differentiating tissues. Over the last few years, both the potential and challenges of such approach have been apparent [10, 13]. As an example, even though Frumkin et al. proposed the use of seven loci for the simultaneous identification of blood, semen, saliva, skin, urine, menstrual blood and vaginal secretion, it was reported by Gomes et al. that they failed to reproduce their results on skin identification [28].

It is known that DNA methylation regulates gene expression and controls cellular differentiation during development [4, 29]; however, in the medical field differential epigenetic patterns among tissues are mainly assessed in order to evaluate if observed differential patterns in diseased tissues are meaningful events or part of natural tissue-to-tissue

variation. The main challenge faced in this study regarding the identification of suitable CpG sites based on publicly available datasets was the lack of methylation data in forensically relevant body fluids such as vaginal fluid and menstrual blood. Therefore, an extensive validation of the selected markers as well as the identification of novel ones was needed to assess their specificity. However, it was not only the selection of previously reported differentially methylated CpG sites that was necessary but also the development of a method that would allow for the amplification and sequencing of shorter fragments. This ability would increase the proposed method's applicability in forensic casework where samples are often of low quality/quantity. The capability of bisulphite Pyrosequencing® to quantify methylation levels has been previously assessed [30] and its advantages over other methods have also been established [31].

Analysis of genome-wide methylation data gathered from blood, semen and buccal cell samples identified a set of 11 body fluid specific CpG sites. In this case, following analysis with bisulphite Pyrosequencing® only the selected semen markers were confirmed as suitable. This is due to similar methylation profiles being detected among blood and saliva samples, which supports the likely presence of white blood cells in saliva, further supported by the observed difference in methylation between buccal cells and saliva. Nevertheless, cg13763232 (BLM1) demonstrated very high methylation levels (>0.85) only in whole blood and could potentially be used to exclude the presence of menstrual blood, although a larger study is required before conclusions are made. Interestingly, this CpG site belongs to the promoter region of the solute-carrier family 6, member 6 (*SLC6A6*) gene that encodes for a sodium- and chloride-dependent transporter of the neurotransmitters taurine and beta-alanine [32]. Taurine plays an important role in many biological activities including osmoregulation, membrane stabilisation and antioxidation [33]. It has been reported that this gene is regulated by controlling transcription factor binding sites in its promoter; therefore, the involvement of DNA methylation in this control cannot be excluded.

As shown in similar studies [13, 14, 34], identifying semen-specific methylation patterns has shown to be rather simple. This comes as no surprise since sperm cells have a unique composition in terms of histones and proteins. This was highlighted further for DNA methylation, since initial selection of highly differentially methylated sites between semen and blood/buccal cells using the genome-wide methylation data obtained by Rakyan et al. [6, 8] resulted in hundreds of potential markers. All four tested CpG sites demonstrated semen-specific differential methylation; however, two of them (cg01318557-SEM1 and cg05656364-SEM2) did not seem to be sufficiently robust.

Looking at the biological functions of the involved genes for the best two semen-specific CpG sites to help understand and justify the observed differential methylation in semen, some interesting observations can be made. The sequence analysed by the SEU1 assay belongs to a gene encoding for the solute carrier family 25 (mitochondrial carrier; adenine nucleotide translocator), member 31 protein (SLC25A31), which catalyses the exchange of cytoplasmic ADP with mitochondrial ATP across the mitochondrial inner membrane. It is believed that it mediates energy generating and energy consuming processes in the distal flagellum, possibly as a nucleotide shuttle between flagellar glycolysis, protein phosphorylation and mechanisms of motility [35]. It has previously been reported that its SLC25A31 mRNA transcripts are exclusively present in liver, testis and brain [35]; therefore, differential DNA methylation in its promoter (where SEU1 is found) could regulate its expression. A recent study supports this assumption, as SLC25A31 was found to be one of the few differentially methylated genes in semen following a genome-wide analysis of 38 semen samples [36]. It was reported to be involved in spermatogenesis and associated with inflammation and autoimmune processes interfering with fertility. Moreover, cg11768416 (SEU2) belongs to the gene encoding for the coiled coil glutamate rich protein 1 (Ccer1 or C12orf12). Remarkably, this gene has been previously reported to have dense promoter CpG island methylation and gene silencing in normal tissues except testis and sperm [29]; a finding that was confirmed in this study. Authors tested various healthy and diseased tissues and found that hypomethylation of this gene in non-semen tissues is associated with gene activation in cancer.

Furthermore, using the proposed markers, successful detection in aged stains of up to 16 years as well as high sensitivity (using as little as 50 pg of starting DNA material) was obtained making their application in forensic casework possible. Even though both PCR assays resulted in a linear DNA methylation quantification, future validation experiments are necessary in order to determine whether PCR bias occur when using very low DNA input. The sensitivity of DNA methylation assays for tissue identification have also been demonstrated in the study by LaRue et al. [12], where positive results for semen were obtained when using as low as 31 pg using the proposed DSI-Semen™ kit developed by Wasserstrom et al. [11]. Future experiments using semen stains of decreasing size (as measured in microlitres) would fully determine the assays' sensitivity proposed in this study. Novel next-generation sequencing (NGS), including various NGS platforms, has been increasingly used in the forensic field and recent studies have also investigated its applicability in tissue identification using DNA methylation markers [37, 38]. Applying NGS could significantly increase the sensitivity and also allow for analysis of an increased number of tissue-specific markers included in a tissue-ID assay.

It should be noted that caution is needed when applying DNA methylation markers for tissue identification since changes in DNA methylation patterns have been reported in various diseases [39, 40]. With regard to semen detection, it has been shown that alterations in sperm DNA methylation patterns in particular loci are associated with low sperm motility and different types of male infertility [21–23]. In fact, in this study there were two semen samples in particular that showed an 'unexpected' methylation ratio for most semen-specific assays tested. The observations above, including the

implication of some of these markers in cancer via an 'abnormal' methylation profile as well as the possible 'connection' between low sperm counts and altered methylation levels, highlights the need for validation of the proposed markers prior to implementation by analysing diseased samples. Ideally, together with samples from 'healthy' volunteers, samples from patients suffering from various diseases should be co-analysed so that a better representation of the general population is achieved. Because of the ethical restrictions, gathering personal data regarding disease status from crime scene stains is not permitted; therefore, all possible reasons that could lead to altered methylation profiles should be taken into account.

Additionally, it can be concluded that an important drawback of applying DNA methylation profiling into forensic casework is the analysis of mixed stains. Since DNA methylation results are presented in a more quantitative manner compared to the mRNA profiling (percentages rather than the presence/absence of peaks), a result regarding one tissue-specific marker would be insufficient to report the tissue source. Therefore, a serial analysis including at least two or three CpG sites per tissue (similar to proposed mRNA strategies) would be necessary in order to include or exclude the presence of a body fluid and marker-specific ratios employed to investigate possible mixed stains. Future work regarding the tissue-specific DNA methylation profiling could include not only an extensive validation of the blood- and semen-specific markers found in this study but also the potential development of a multiplex PCR assay capable of amplifying all proposed markers in one test, which is more desirable in a forensic setting. Furthermore, future research could consider the investigation of more genomic loci showing potentially better tissue-specific DNA methylation patterns. To account for potential, common inter-individual differences in DNA methylation and, consequently, gene expression, a larger dataset of body fluids and tissues should be collected.

4.1 Conclusion

In conclusion, DNA methylation profiling for use in detecting body fluids and tissues seems promising. Evaluation of the bisulphite Pyrosequencing®-based assays revealed that this method can be highly sensitive; successful DNA methylation profiles were obtained with as low as 50 pg of starting material. In general, it was observed that semen demonstrated clear, differential DNA methylation patterns and was easy to detect with confidence. The two proposed semen-specific markers performed well in mock casework and their DNA methylation profiles were stable for up to 16 years in semen stains.

The authors would like to thank Professor Vardhman Rakyan at the Blizard Institute, Barts and the London School of Medicine and Dentistry for assisting with the access to the initial DNA methylation data as well as Dr. David Ballard, King's College London for his technical support and advice throughout this project.

The authors have declared no conflict of interest.

5 References

[1] Lindenbergh, A., de Pagter, M., Ramdayal, G., Visser, M., Zubakov, D., Kayser, M., Sijen, T., *Forensic Sci. Int. Genet.* 2012, *6*, 565–577.

[2] Lindenbergh, A., Maaskant, P., Sijen, T., *Forensic Sci. Int. Genet.* 2013, *7*, 159–166.

[3] Plachot, C., Lelievre, S. A., *Exp. Cell Res.* 2004, *298*, 122–132.

[4] Song, F., Mahmood, S., Ghosh, S., Liang, P., Smiraglia, D. J., Nagase, H., Held, W. A., *Genomics* 2009, *93*, 130–139.

[5] Newell-Price, J., Clark, A., King, P., *Trends Endocrinol Metab* 2000, *11*, 142–148.

[6] Rakyan, V. K., Down, T. A., Thorne, N. P., Flicek, P., Kulesha, E., Graf, S., Tomazou, E. M., Backdahl, L., Johnson, N., Herberth, M., Howe, K. L., Jackson, D. K., Miretti, M. M., Fiegler, H., Marioni, J. C., Birney, E., Hubbard, T. J., Carter, N. P., Tavare, S., Beck, S., *Genome Res.* 2008, *18*, 1518–1529.

[7] Thompson, T. M., Sharfi, D., Lee, M., Yrigollen, C. M., Naumova, O. Y., Grigorenko, E. L., *Behav. Genet.* 2013, *43*, 168–176.

[8] Rakyan, V. K., Down, T. A., Maslau, S., Andrew, T., Yang, T. P., Beyan, H., Whittaker, P., McCann, O. T., Finer, S., Valdes, A. M., Leslie, R. D., Deloukas, P., Spector, T. D., *Genome Res.* 2010, *20*, 434–439.

[9] Zykovich, A., Hubbard, A., Flynn, J. M., Tarnopolsky, M., Fraga, M. F., Kerksick, C., Ogborn, D., MacNeil, L., Mooney, S., Melov, S., *Aging Cell* 2014, *13*, 360–366.

[10] Frumkin, D., Wasserstrom, A., Budowle, B., Davidson, A., *Forensic Sci. Int. Genet.* 2011, *5*, 517–524.

[11] Wasserstrom, A., Frumkin, D., Davidson, A., Shpitzen, M., Herman, Y., Gafny, R., *Forensic Sci. Int. Genet.* 2013, *7*, 136–142.

[12] LaRue, B. L., King, J. L., Budowle, B., *Int. J. Legal Med.* 2013, *127*, 299–308.

[13] Lee, H. Y., Park, M. J., Choi, A., An, J. H., Yang, W. I., Shin, K. J., *Int. J. Legal Med.* 2012, *126*, 55–62.

[14] An, J. H., Choi, A., Shin, K. J., Yang, W. I., Lee, H. Y., *Int. J. Legal Med.* 2013, *127*, 35–43.

[15] Park, J. L., Kwon, O. H., Kim, J. H., Yoo, H. S., Lee, H. C., Woo, K. M., Kim, S. Y., Lee, S. H., Kim, Y. S., *Forensic Sci. Int. Genet.* 2014, *13C*, 147–153.

[16] Fu, X. D., Wu, J., Wang, J., Huang, Y., Hou, Y. P., Yan, J., *Forensic Sci. Int. Genet. Suppl. Series* 2015, *5*, e151–e153.

[17] Lee, H. Y., An, J. H., Jung, S. E., Oh, Y. N., Lee, E. Y., Choi, A., Yang, W. I., Shin, K. J., *Forensic Sci. Int. Genet.* 2015, *17*, 17–24.

[18] Tusnady, G. E., Simon, I., Varadi, A., Aranyi, T., *Nucleic Acids Res.* 2005, *33*, e9.

[19] Thiede, C., Prange-Krex, G., Freiberg-Richter, J., Bornhäuser, M., Ehninger, G., *Bone Marrow Transplant.* 2000, *25*, 575–577.

[20] Lackner, J. E., Herwig, R., Schmidbauer, J., Schatzl, G., Kratzik, C., Marberger, M., *Fertil. Steril.* 2006, *86*, 601–605.

[21] Hammoud, S. S., Purwar, J., Pflueger, C., Cairns, B. R., Carrell, D. T., *Fertil. Steril.* 2010, *94*, 1728–1733.

[22] El Hajj, N., Zechner, U., Schneider, E., Tresch, A., Gromoll, J., Hahn, T., Schorsch, M., Haaf, T., *Sex. Dev.* 2011, *5*, 60–69.

[23] Pacheco, S. E., Houseman, E. A., Christensen, B. C., Marsit, C. J., Kelsey, K. T., Sigman, M., Boekelheide, K., *PLoS One* 2011, *6*, 1–10.

[24] Moskalev, E. A., Zavgorodnij, M. G., Majorova, S. P., Vorobjev, I. A., Jandaghi, P., Bure, I. V., Hoheisel, J. D., *Nucleic Acids Res.* 2011, *39*, e77.

[25] Gomes, I., Strohbücker, B., Rothschild, M. A., Schneider, P. M., *Forensic Sci. Int.* 2013, *4*, e180–e181.

[26] Xu, Y., Xie, J., Cao, Y., Zhou, H., Ping, Y., Chen, L., Gu, L., Hu, W., Bi, G., Ge, J., Chen, X., Zhao, Z., *PLoS One* 2014, *9*, e100123.

[27] Jucker, R., Sulzer, A., Voegeli, P., Haas, C., Kratzer, A., *Forensic Sci. Int. Genet. Suppl. Series* 2015, *5*, e381–e383.

[28] Gomes, I., Kohlmeier, F., Schneider, P. M., *Forensic Sci. Int. Genet. Suppl. Series* 2011, *3*, e469–e470.

[29] Shen, L., Kondo, Y., Guo, Y., Zhang, J., Zhang, L., Ahmed, S., Shu, J., Chen, X., Waterland, R. A., Issa, J. P., *PLoS Genet.* 2007, *3*, 2023–2036.

[30] Dejeux, E., El abdalaoui, H., Gut, I. G., Tost, J., in: Tost, J. (Ed.), *DNA methylation: Methods and Protocols (2nd ed)*, Humana Press, USA 2009, pp. 189–205.

[31] Reed, K., Poulin, M. L., Yan, L., Parissenti, A. M., *Anal. Biochem.* 2010, *397*, 96–106.

[32] Han, X., Chesney, R. W., *FEBS Lett.* 2003, *540*, 71–76.

[33] Han, X., Patters, A. B., Jones, D. P., Zelikovic, I., Chesney, R. W., *Acta Physiol.* 2006, *187*, 61–73.

[34] Madi, T., Balamurugan, K., Bombardi, R., Duncan, G., McCord, B., *Electrophoresis* 2012, *33*, 1736–1745.

[35] Dolce, V., Scarcia, P., Iacopetta, D., Palmieri, F., *FEBS Lett.* 2005, *579*, 633–637.

[36] Schutte, B., El Hajj, N., Kuhtz, J., Nanda, I., Gromoll, J., Hahn, T., Dittrich, M., Schorsch, M., Muller, T., Haaf, T., *Andrology* 2013, *1*, 822–829.

[37] Forat, S., Huettel, B., Reinhardt, R., Fimmers, R., Haidl, G., Denschlag, D., Olek, K., *PLoS One* 2016, *11*, e0147973.

[38] Bartling, C. M., Hester, M. E., Bartz, J., Heizer, E., Jr., Faith, S. A., *Electrophoresis* 2014, *35*, 3096–3101.

[39] Tsai, H. C., Baylin, S. B., *Cell Res.* 2011, *21*, 502–517.

[40] Teschendorff, A. E., Jones, A., Fiegl, H., Sargent, A., Zhuang, J. J., Kitchener, H. C., Widschwendter, M., *Genome Med.* 2012, *4*, 1–24.

