Forensic Science International: Genetics 23 (2016) 55–63

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

Research paper

Developmental validation studies of epigenetic DNA methylation markers for the detection of blood, semen and saliva samples

Deborah S.B.S. Silva[a,b], Joana Antunes[a], Kuppareddi Balamurugan[c], George Duncan[d], Clarice S. Alho[b], Bruce McCord[a,*]

[a] Department of Chemistry and Biochemistry, Florida International University, Miami, FL, USA
[b] Faculty of Biosciences, Laboratory of Human and Molecular Genetics, PUCRS, Porto Alegre, Brazil
[c] School of Criminal Justice, University of Southern Mississippi, Hattiesburg, MS, USA
[d] DNA Laboratory, Broward County Sheriff's Office, Fort Lauderdale, FL, USA

ARTICLE INFO

Article history:
Received 8 June 2015
Received in revised form 23 January 2016
Accepted 26 January 2016
Available online 1 February 2016

Keywords:
DNA methylation
Body fluid
Pyrosequencing
Developmental validation

ABSTRACT

Determining the type and origin of body fluids in a forensic investigation can provide important assistance in reconstructing crime scenes. A set of epigenetic markers, ZC3H12D, BCAS4 and cg06379435, have been developed to produce unique and specific patterns of DNA methylation that can be used to identify semen, saliva, and blood, respectively. To ensure the efficacy of these markers, developmental validation studies were performed to determine the conditions and limitations of this new tool for forensic analysis. DNA was extracted from human samples and bisulfite modified using commercial bisulfite modification kits. Specific primers were used to amplify the region of interest and the methylation profile of the CpG sites were determined by pyrosequencing. The percent methylation values at each CpG site were determined in multiple samples and averaged for each tissue type. The versatility of these new markers is presented by showing the results of validation studies on sensitivity, human specificity, stability and mixture resolution. When testing the markers using different organisms, we did obtain positive results for certain non-human primate samples, however, all other tested species were negative. The lowest concentration consistently detected varied from 0.1 to 10 ng, depending on the locus, indicating the importance of primer design and sequence in the assay. The method also proved to be effective when inhibitors were present in the samples or when samples were degraded by heat. Simulated case- samples were also tested. In the case of mixtures of different cell types, the overall methylation values varied in a consistent and predictable manner when multiple cell types were present in the same sample. Overall, the validation studies demonstrate the robustness and effectiveness of this new tool for body fluid identification.

© 2016 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

Body fluids recovered from crime scenes are considered among the most important types of evidence in forensic cases. They contain DNA evidence that may identify a suspect or victim and exonerate an innocent person. Moreover, determining the type and origin of a biological material can help reconstruct crime scenes [1–3]. Identifying body fluids can be a critical factor in criminal investigations as the presence of skin cells may indicate innocent transfers of DNA while blood or other body fluids can indicate that a criminal act has taken place. However, many times it can be difficult to identify a forensic stain due to its similarities with other fluids or substances. Additionally, when trace levels of body fluids are present, visualization may be impossible and serological tests can fail to produce a result [1].

Recent efforts in researching new assays to identify body fluids have focused on epigenetic DNA based markers. While RNA and protein markers may be used for body fluid identification, DNA methylation presents the ideal methodology since it provides quantitative results, is less prone to variations in expression, and has greater long term stability [4,5]. In addition, as the extracted DNA target is already present in the laboratory, only minor sample processing is needed. This is important, since in many cases the forensic sample available for analysis is limited.

DNA methylation is an epigenetic modification that is involved in transcriptional regulation. A typical mammalian genome has approximately 3 billion base pairs, with a G/C content of about 40%.

* Corresponding author at: Department of Chemistry and Biochemistry, Florida International University, 11200 SW 8th St., Miami, FL, USA. Fax: +1 3053483772.
E-mail address: mccordb@fiu.edu (B. McCord).

http://dx.doi.org/10.1016/j.fsigen.2016.01.017
1872-4973/© 2016 Elsevier Ireland Ltd. All rights reserved.

Methylated cytosines typically occur as CpG dinucleotides and the presence of a methyl group at the C-5 position of the molecule tends to reduce gene expression. The CpG dinucleotide distribution is uneven, with several short DNA elements having a much higher density of CpG dinucleotides than other regions of the genome, forming so-called CpG islands. Most of these islands are located near transcription start sites. The methylated form, 5-methylcytosine (5-mC) correlates with cytosine in the same way that thymine correlates with uracil, with no effect on the base pairing [6–9]. Although much is still to be understood about the mechanisms by which methylation affects gene expression, it is known that methylation is important in cell differentiation, and genomic loci are differentially methylated between tissues. Because of this, differentiation of methylation patterns can provide the basis of an assay for body fluid identification.

The most common body fluids found at crime scenes are blood, semen and saliva. A set of epigenetic markers have been developed, ZC3H12D, BCAS4, and cg06379435, which produce unique and specific patterns of DNA methylation that can be used to identify semen, saliva and blood, respectively [10,11]. The process involves bisulfite modification, followed by amplification by the polymerase chain reaction (PCR), and pyrosequencing. The results permit quantitative determination of methylation patterns at each locus for each body fluid type. The data demonstrate the applicability of epigenetic markers as an assay for trace body fluid identification. However, to ensure the efficiency of these epigenetic markers, developmental validation studies need to be performed to determine the conditions and limitations of this new tool for forensic analysis.

2. Materials and methods

2.1. Sample collection

Biological samples (semen, blood and saliva) were obtained from volunteer donors. Buccal cells were collected using cotton swabs. Blood was obtained by finger pricking and blood cells were collected using cotton swabs. Male volunteers were given a specimen jar and requested to donate a semen sample. All participants signed informed consent statements prior to sample collection. All samples were collected after obtaining appropriate Institutional Review Board approvals from Florida International University (IRB-13-0555), The University of Southern Mississippi (protocol #12010303) and Pontifícia Universidade Católica do Rio Grande do Sul (CONEP #723.619/CEP #845.747).

2.2. DNA extraction and modification

DNA extraction was performed using the BioRobot® EZ1 automated purification workstation (Qiagen, CA) and the EZ® DNA Investigator kit according to manufacturer's protocol. DNA was recovered in a final volume of 50 μl and quantified using PicoGreen® fluorescence (Life Technologies, CA) [10].

Because standard PCR amplification does not preserve DNA methylation sites, genomic DNA extracts were treated with sodium bisulfite in order to convert unmethylated cytosines to uracil using the EpiTect® Fast DNA Bisulfite Kit (Qiagen Inc, CA). For all tests conducted in this research, except for the sensitivity test, the standard manufacturer's protocol for the conversion of around 400 ng DNA in a maximum volume of 40 μl was used. In this method the converted uracils are then replaced by thymine during the PCR amplification process. The bisulfite modified DNA was amplified by site specific PCR primers designed to amplify the bisulfite modified target regions [12].

2.3. Markers

As previously described, a set of epigenetic markers, ZC3H12D, BCAS4, and cg06379435 have been developed to produce unique and specific patterns of DNA methylation that can be used to identify semen, saliva, and blood, respectively.

Markers ZC3H12D and BCAS4 were described by Madi et al. [10] and both targeted 5CpG sites in the sequence to be analyzed. Marker cg06379435 was described by Park et al. [11] and targeted only one CpG site. We have investigated this locus and added 4 additional CpG sites around the cg06379435 site.

BCAS4 and cg06379435 PyroMark® CpG assays were custom designed in house using the Pyromark® assay design software (Qiagen). A predesigned ZC3H12D PyroMark® CpG assay was available from the Qiagen GeneGlobe Web Portal (Table 1). The reverse primers were biotin labeled allowing for the production of biotinylated PCR products necessary for the pyrosequencing reaction [10].

2.4. PCR and pyrosequencing reactions

Singleplex PCR reactions were performed using the PyroMark® PCR Kit (Qiagen) in a GeneAmp® PCR System 9700 (Applied Biosystems, Foster City, CA). For all samples, 1.2 μl of modified DNA template was added to the PCR mixture in a total reaction volume

Table 1
Panel of markers used in this study.

Marker	PCR and sequencing primers	CpG sites to be analyzed in the nucleotide dispensation order (underlined)
BCAS4	Forward primer- AGTGGGTGAGGTTGTGAAATGT Reverse primer- CCCATCCTACTAAAACATCTAATT Seq. primer- AGTTTTTTGGTGAAGTTTAT	TAT**CG**TATCGAGAGAGATCGAGAT**CG**TATAT**CG**TATAGAT**CG**T**CG**AT
cg06379435[a]	Forward primer- AGTAGAGGTGGGGGTTAATAATT Reverse primer- ACCACACAACAAAACAACTATCTC Seq. primer- GTTAGGAAAGAAAAATGTAATTTA	T**CG**GGATAAT**CG**GTGGAATTTTAGG**CG**TGGGA**CG**GTTGT**CG**GA
ZC3H12D	Proprietary[b]	T**CG**T**CG**AGTAT**CG**T**CG**T**CG**

[a] TargetID is a unique ID in the Illumina Human Methylation 450 K bead array; 4th CpG site in the sequence provided is the target in Park et al. [11].
[b] Primer sequences are proprietary (Qiagen Inc.) and not available.

Table 2
Methylation profiles of the three loci investigated in this study.

Marker	Body fluid	CpG (mean % methylation ± SD)				
		CpG 1	CpG 2	CpG 3	CpG 4	CpG 5
ZC3H12D	Semen	5.4 ± 4.0	5.3 ± 3.9	6.7 ± 4.3	6.3 ± 4.0	5.1 ± 3.7
	[a]Threshold	14	13	15	14	13
Specific for semen	Blood	94.0 ± 1.6	94.0 ± 2.7	100.0 ± 0.1	97.0 ± 7.9	86.0 ± 3.5
	Saliva	81.0 ± 4.1	78.0 ± 4.5	99.0 ± 2.1	79.0 ± 4.9	82.0 ± 3.4
		CpG 1	CpG 2	CpG 3	CpG 4	CpG 5
cg06379435	Blood	24.0 ± 7.8	22.0 ± 6.7	33.0 ± 7.4	30.0 ± 8.2	49 ± 12
	[a]Threshold	8.2	8.4	18	14	25
Specific for blood	Semen	3.4 ± 1.7	2.4 ± 1.8	2.7 ± 1.1	1.8 ± 1.3	3.1 ± 1.2
	Saliva	8.7 ± 7.0	2.6 ± 1.4	6.0 ± 3.9	3.5 ± 2.6	7.7 ± 4.7
		CpG1	CpG4	CpG5	CpG6	CpG7
BCAS4	Saliva	64.0 ± 7.1	27.0 ± 5.6	16.0 ± 4.5	45.0 ± 6.7	11.0 ± 2.7
	[a]Threshold	49	16	7.2	31	6.1
Specific for saliva	Blood	6.1 ± 1.4	3.2 ± 2.8	2.7 ± 1.3	6.4 ± 1.9	2.4 ± 1.6
	Semen	3.9 ± 1.6	2.5 ± 1.1	3.0 ± 1.1	5.6 ± 1.6	2.3 ± 0.9

[a] Each CpG site from each marker is given a specific threshold value due to the fact that each CpG has its unique level of methylation. A threshold value is assigned as the average methylation level ± 2 SD depending on whether hypo- or hyper-methylation is present [16].

of 15 μl. The complete PCR protocol outlined by the manufacturer was followed (Qiagen).

Pyrosequencing was performed following completion of the PCR reactions using a Pyromark® Q24 pyrosequencer (Qiagen) following the protocols recommended by the manufacturer. Once the pyrosequencing was complete, the percent methylation was calculated automatically by the Pyromark® Q24 software and was displayed as a pyrogram [10].

2.5. Validation studies

We tested 15 samples of each cell type (blood, saliva and semen) to be used in these studies. The pyrosequencing data generated for each marker was analyzed and compared to other samples in the validation studies performed to ensure they do not differ statistically.

2.5.1. Body fluid specificity

The specificity of the markers was previously tested against blood, semen and saliva [10,11]. In this study we tested the performance of the markers against 3 samples of each: menstrual blood vaginal epithelia and nasal epithelia/secretion.

2.5.2. Species specificity

Species specificity studies were performed with non-human DNA samples of dog, cat, mouse, chicken, bovine, equine, pig, mouse, chimpanzee, orangutan, gorilla and a microbial pool (*Escherichia coli, Staphylococcus aureus, Enterococcus faecali* and *Pseudomonas aeruginosa*). A human blood sample was used as positive control.

Table 3
Methylation profiles of the three investigated loci when testing menstrual blood, vaginal epithelia and nasal epithelia/secretion samples. Comparison with methylation levels of the three main fluids investigated in this study.

Marker	Body fluid	CpG (mean % methylation ± SD)				
		CpG 1	CpG 2	CpG 3	CpG 4	CpG 5
ZC3H12D	Semen	5.4 ± 4.0	5.3 ± 3.9	6.7 ± 4.3	6.3 ± 4.0	5.1 ± 3.7
	Blood	94.0 ± 1.6	94.0 ± 2.7	100 ± 0.0	97.0 ± 7.9	86.0 ± 3.5
Specific for sêmen	Saliva	81.0 ± 4.1	78.0 ± 4.5	99.0 ± 2.1	79.0 ± 4.9	82.0 ± 3.4
	Menstrual blood	80.0 ± 4.7	79.0 ± 6.2	98.0 ± 2.9	77 ± 12	74.3 ± 3.5
	Nasal epithelia	83.0 ± 6.4	83.0 ± 6.8	99.0 ± 1.2	78 ± 11	77.0 ± 6.9
	Vaginal epithelia	77.0 ± 9.7	77 ± 10	97.0 ± 2.5	73 ± 10	80.0 ± 4.0
		CpG 1	CpG 2	CpG 3	CpG 4	CpG 5
cg06379435	Semen	3.4 ± 1.7	2.4 ± 1.8	2.7 ± 1.1	1.8 ± 1.3	3.1 ± 1.2
	Blood	24.0 ± 7.8	22.0 ± 6.7	33.0 ± 7.4	30.0 ± 8.2	49 ± 12
Specific for blood	Saliva	8.7 ± 7.0	2.6 ± 1.4	6.0 ± 3.9	3.5 ± 2.6	7.7 ± 4.7
	Menstrual blood	10 ± 12	12 ± 19[a]	12 ± 18	8.0 ± 12	23 ± 30
	Nasal epithelia	5.7 ± 5.5	5.0 ± 6.9	9.7 ± 11	7.7 ± 11	18 ± 22
	Vaginal epithelia	2.3 ± 0.6	1.3 ± 0.6	3.0 ± 1.0	2.0 ± 1.0	8.3 ± 7.8
		CpG1	CpG4	CpG5	CpG6	CpG7
BCAS4	Semen	3.9 ± 1.6	2.5 ± 1.1	3.0 ± 1.1	5.6 ± 1.6	2.3 ± 0.9
	Blood	6.1 ± 1.4	3.2 ± 2.8	2.7 ± 1.3	6.4 ± 1.9	2.4 ± 1.6
Specific for saliva	Saliva	64.0 ± 7.1	27.0 ± 5.6	16.0 ± 4.5	44.0 ± 6.7	11.0 ± 2.7
	Menstrual blood	35.0 ± 8.5	15.0 ± 4.6	11.0 ± 3.1	27.0 ± 5.5	8.0 ± 1.0
	Nasal epithelia	39 ± 21	12.0 ± 8.0	9.0 ± 3.6	33 ± 20	6.7 ± 3.5
	Vaginal epithelia	36 ± 17	13.0 ± 7.5	11.0 ± 4.9[b]	30 ± 11	6.7 ± 3.2

[a] The methylation level of CpG2 did not significantly differ when menstrual blood samples were compared to blood samples ($p > 0.05$).
[b] The methylation level of CpG5 did not significantly differ when vaginal epithelia samples were compared to saliva samples ($p > 0.05$).

Table 4
Methylation profiles of the three investigated loci when using low amounts of pre-modified DNA.

Marker	Samples	CpG (Mean % Methylation ± SD)				
ZC3H12D		CpG 1	CpG 2	CpG 3	CpG 4	CpG 5
Marker for semen						
	Semen control samples[a]	5.4 ± 4.0	5.3 ± 3.9	6.7 ± 4.3	6.3 ± 4.0	5.1 ± 3.7
	Semen samples (DNA input = 0.1 ng)	4.7 ± 4.0	4.7 ± 3.8	4.3 ± 3.5	3.7 ± 2.1	3.0 ± 4.4
cg06379435		CpG 1	CpG 2	CpG 3	CpG 4	CpG 5
Marker for blood	Blood control samples[a]	24.0 ± 7.8	22.0 ± 6.7	33.0 ± 7.4	30.0 ± 8.2	49 ± 12
	Blood samples (DNA input = 10 ng)	23.0 ± 8.7	22 ± 10	26.0 ± 8.3	23.0 ± 9.5	36.0 ± 7.5
BCAS4		CpG1	CpG4	CpG5	CpG6	CpG7
Marker for saliva	Saliva control samples[a]	64.0 ± 7.1	28.0 ± 5.6	16.0 ± 4.5	44.0 ± 6.7	12.0 ± 2.7
	Saliva samples (DNA input = 10 ng)	66.0 ± 6.0	29.0 ± 4.9	16.0 ± 1.0	44.0 ± 9.5	10.0 ± 6.5

[a] DNA input of control samples ranged from 100 to 500 ng.

2.5.3. Sensitivity studies

According to the SWGDAM guidelines [13] it is important to evaluate the limits of DNA quantities to be used in the reaction in order to obtain reliable results. Madi et al. [10] processed a saliva sample five separate times and, for each replicate, the amount of DNA subjected to bisulfite modification was varied (385, 100, 50, 10 and 5 ng). The samples were amplified with the ZC3H12D PyroMark® CpG Assay. All replicates showed no significant differences in the methylation levels observed at each CpG site. Park et al. [11] tested marker cg06379435 and were able to obtain successful pyrosequencing results when using 500–10 ng of premodified DNA. However, when working with forensic samples, it is important to test the assay for its sensitivity limits and to examine the lowest levels of DNA possible while maintaining reliability in the results. For this validation study we therefore decided to test the following DNA quantities: 10, 5 and 1 ng. Three distinct samples for each body fluid were processed for each DNA amount tested.

2.5.4. Stability studies and case-type samples

Forensic samples are often exposed to environmental and chemical insults, and may contain impurities which can act as PCR inhibitors. In this study, we simulated casework samples to evaluate the robustness and stability of the markers.

To assess the effect of the inhibitors on PCR, we tested 2 different solutions: hematin (100 mM in 0.1 N sodium hydroxide) and humic acid (1 mg/ml in water). From these stock solutions, we prepared subsequent dilutions in water in a final concentration of 0.08 mM (hematin) and 0.24 mg/ml (humic acid) [14]. Aliquots of 2 μl of the inhibitors were added before or after bisulfite conversion. For this study, we tested three distinct samples for each body fluid.

To test the stability of the markers under degraded conditions, samples with genomic DNA extracts were heated in deionized water at 95 °C for 10, 15, 20, and 25 min to simulate natural DNA fragmentation [14]. For this study, we tested three distinct samples for each body fluid.

Also, DNA from the following simulated case-type samples was extracted and evaluated for all tested markers:

i) 200 μl of blood in 100% cotton fabric;
ii) 200 μl of semen in 100% cotton fabric;
iii) Saliva swab of lid of a coffee drink.

2.5.5. Mixture studies

DNA mixtures were prepared in order to determine if there are differences in methylation values when multiple cell types were present in the same sample. DNA from blood, saliva and semen were tested in different ratios of mixtures:

i) Blood + saliva: ratios of 75% blood: 25% saliva—50% blood: 50% saliva—25% blood: 75% saliva
ii) Blood + semen: ratios of 75% blood: 25% semen—50% blood: 50% semen—25% blood: 75% semen
iii) Saliva + semen: ratios of 75% saliva: 25% semen—50% saliva: 50% semen—25% saliva: 75% semen
iv) Blood + saliva + semen: ratios of 75% blood: 12.5% saliva: 12.5% semen—50% blood: 25% saliva: 25% semen—75% saliva: 12.5% blood: 12.5% semen—50% saliva: 25% blood: 25% semen—75% semen: 12.5% blood: 12.5% saliva—50% semen: 25% blood: 25% saliva

In addition to these mixtures, each sample was analyzed individually as a positive control.

2.5.6. Reproducibility

This study was performed by testing 24 samples (saliva $n = 8$, blood $n = 8$, semen $n = 8$). The pyrosequencing data generated by our laboratory (Florida International University) and an independent laboratory (University of Southern Mississippi) was analyzed and compared to assess the reliability and the production of concordant results between different laboratories [15].

2.6. Data analysis

The CpG sites were examined and, for each marker, the methylation values were averaged and listed for each cell type along with standard deviations. Using a one-way ANOVA, the methylation values were compared across CpG sites of each marker between controls and other samples to determine if they differed significantly. A t-test was used to compare results obtained by the two different laboratories in the reproducibility study. A p-value of <0.05 was assumed as significant for all tests conducted.

3. Results and discussion

3.1. Control samples

A set of 15 samples of each cell type (blood, saliva and semen) were analyzed and methylation profiles of each locus were obtained to be used as controls. Table 2 presents the pyrosequencing data generated for each marker.

3.2. Body fluid specificity

Besides previously testing the performance of the markers against blood, semen and saliva, we also tested the markers against menstrual blood, vaginal epithelia and nasal epithelia/secretion (Table 3). Most of the CpG sites presented significant differences in the methylation levels when compared with the different types of

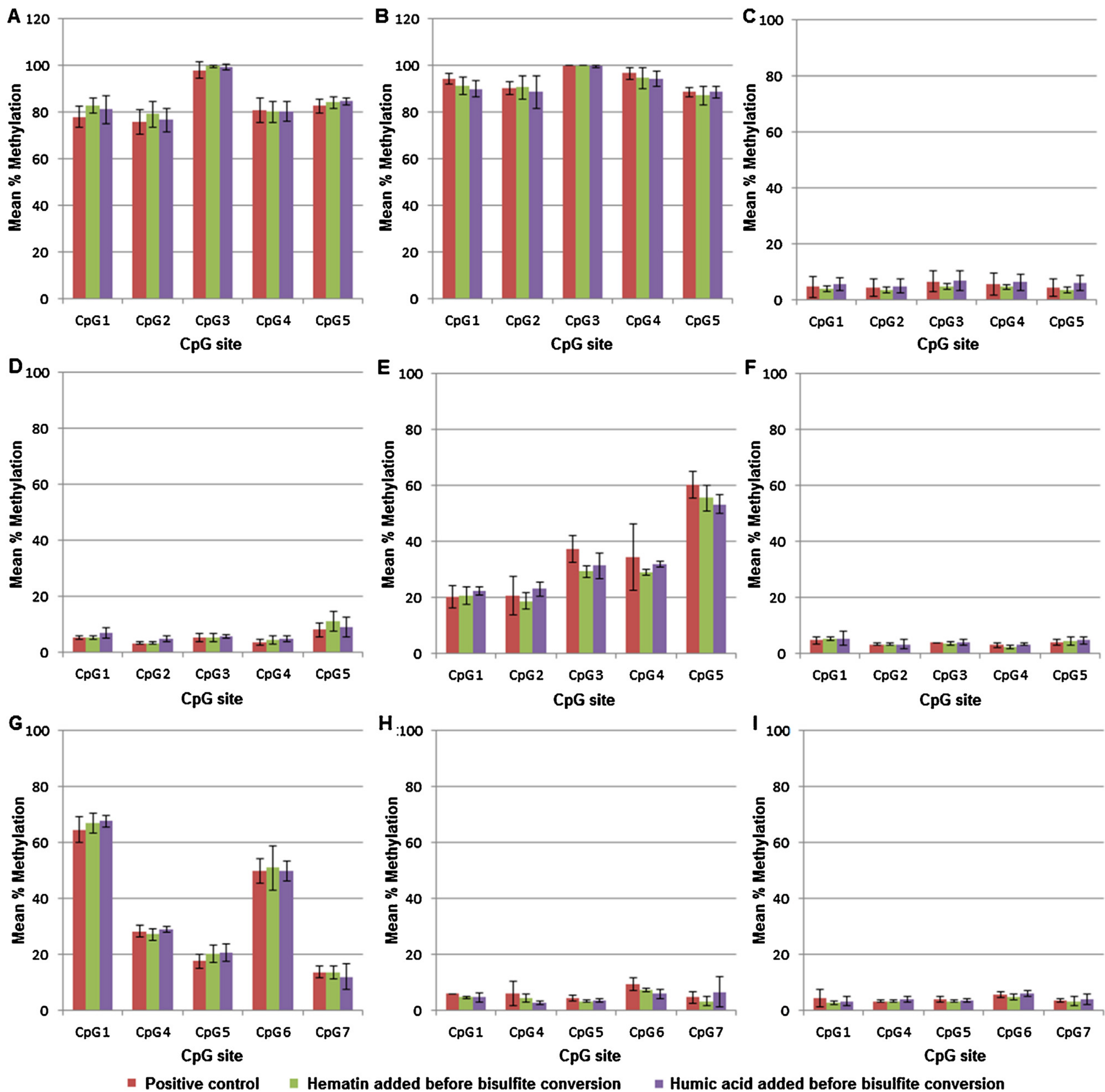

Fig. 1. Mean methylation levels of samples in which PCR inhibitors were added before bisulfite modification. There were no significant differences between the control and the tested samples. (A) Marker ZC3H12D: saliva samples. (B) Marker ZC3H12D: blood samples. (C) Marker ZC3H12D: semen samples. (D) Marker cg06379435: saliva samples. (E) Marker cg06379435: blood samples. (F) Marker cg06379435: semen samples. (G) Marker BCAS4: saliva samples. (H) Marker BCAS4: blood samples. (I) Marker BCAS4: semen samples.

fluids. For marker BCAS4, the methylation level of CpG5 did not significantly differ when saliva samples were compared to vaginal epithelia samples ($p > 0.05$). For marker cg06379435, the methylation level of CpG2 did not significantly differ when blood samples were compared to menstrual blood samples ($p > 0.05$). However, all CpGs within each marker must be evaluated to obtain an indication of the body fluid present. Our results show that the majority of CpGs within each marker presents different mean methylation levels between the target body fluid and all other body fluids tested (Table 3).

In this study, we calculated the threshold as ± 2 SD from the average methylation percent for each CpG, according to what has been published in terms of DNA methylation and threshold calculation [16]. However, in our study multiple CpG sites are available at each locus for each body fluid; therefore we suggest that the user can adjust the threshold levels to ± 1 SD, making the test more discriminatory. At present, in order to have a conclusive result for the identification of a body fluid, all CpG sites of a marker should reach the threshold value; otherwise the test should be indicated as inconclusive.

This study represents the first validation for the use of DNA methylation patterns in specific genome locations for body fluid discrimination. In order to achieve a higher discriminatory power for a wider range of body fluids, additional genome locations need

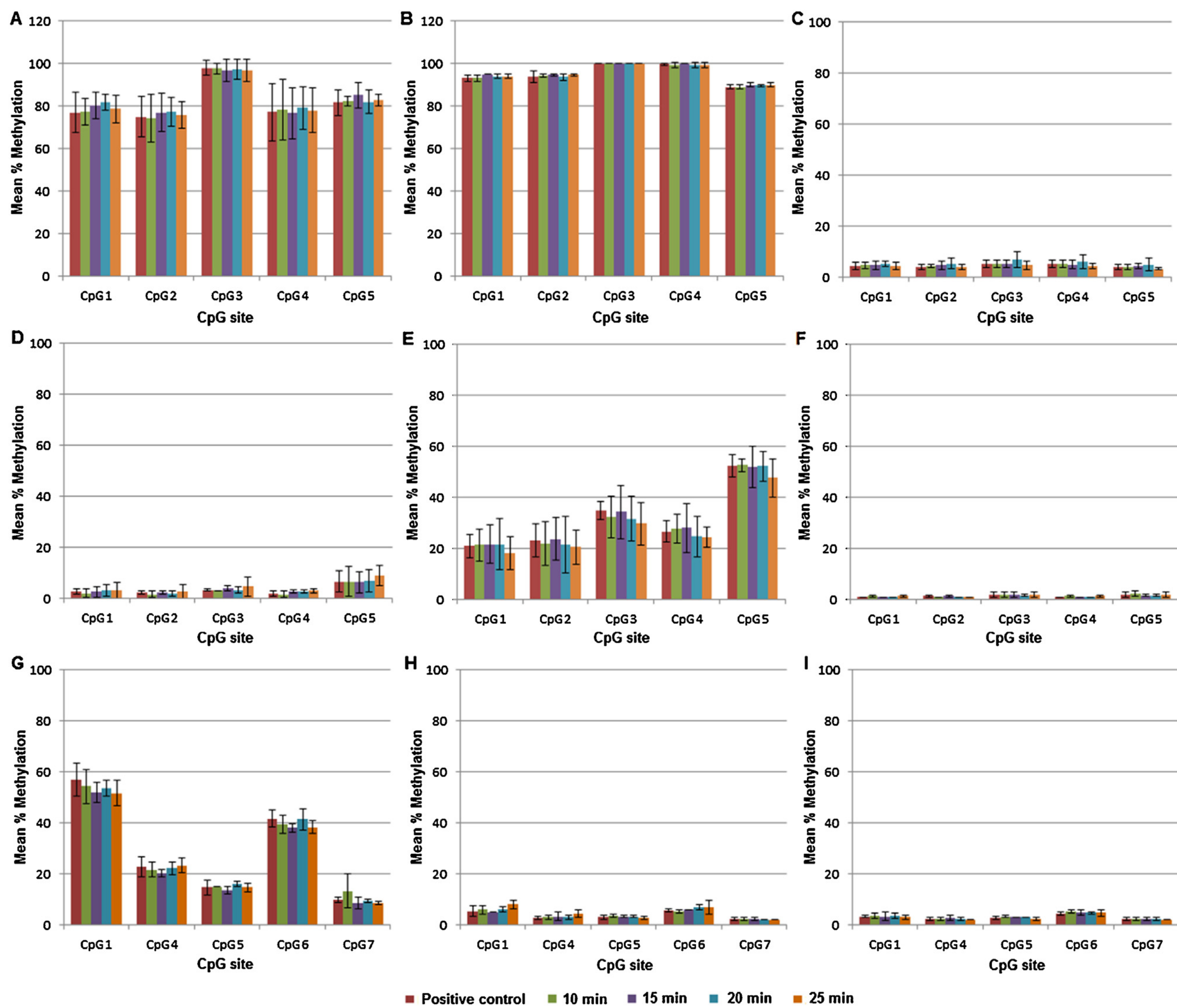

Fig. 2. Mean methylation levels of samples degraded by heat with different time points. There were no significant differences between the control and the degraded samples. (A) Marker ZC3H12D: saliva samples. (B) Marker ZC3H12D: blood samples. (C) Marker ZC3H12D: semen samples. (D) Marker cg06379435: saliva samples. (E) Marker cg06379435: blood samples. (F) Marker cg06379435: semen samples. (G) Marker BCAS4: saliva samples. (H) Marker BCAS4: blood samples. (I) Marker BCAS4: semen samples.

to be determined using array technology and NGS and validated using pyrosequencing. With the advancements in next-generation sequencing we estimate that the number of research studies addressing this issue for forensic purposes will increase.

3.3. Species specificity

Several non-human DNA samples were tested to assess the specificity of the markers. Non-human primate samples amplified the target regions but not all showed pyrosequencing results. For

Table 5
Methylation levels of simulated case-type samples.

Marker	Samples	CpG (% methylation)				
ZC3H12D		CpG 1	CpG 2	CpG 3	CpG 4	CpG 5
Marker for semen	200 μl of semen in 100% cotton fabric	15	10	16	14	13
	200 ul of blood in 100% cotton fabric	93	96	100	99	85
	Saliva swab of lid of a coffee drink	72	73	100	79	85
cg06379435		CpG 1	CpG 2	CpG 3	CpG 4	CpG 5
Marker for blood	200 μl of semen in 100% cotton fabric	10	7	9	5	15
	200 μl of blood in 100% cotton fabric	23	20	36	33	49
	Saliva swab of lid of a coffee drink	5	6	10	7	17
BCAS4		CpG1	CpG4	CpG5	CpG6	CpG7
Marker for saliva	200 μl of semen in 100% cotton fabric	5	5	5	11	3
	200 μl of blood in 100% cotton fabric	6	7	6	7	6
	Saliva swab of lid of a coffee drink	50	19	17	70	24

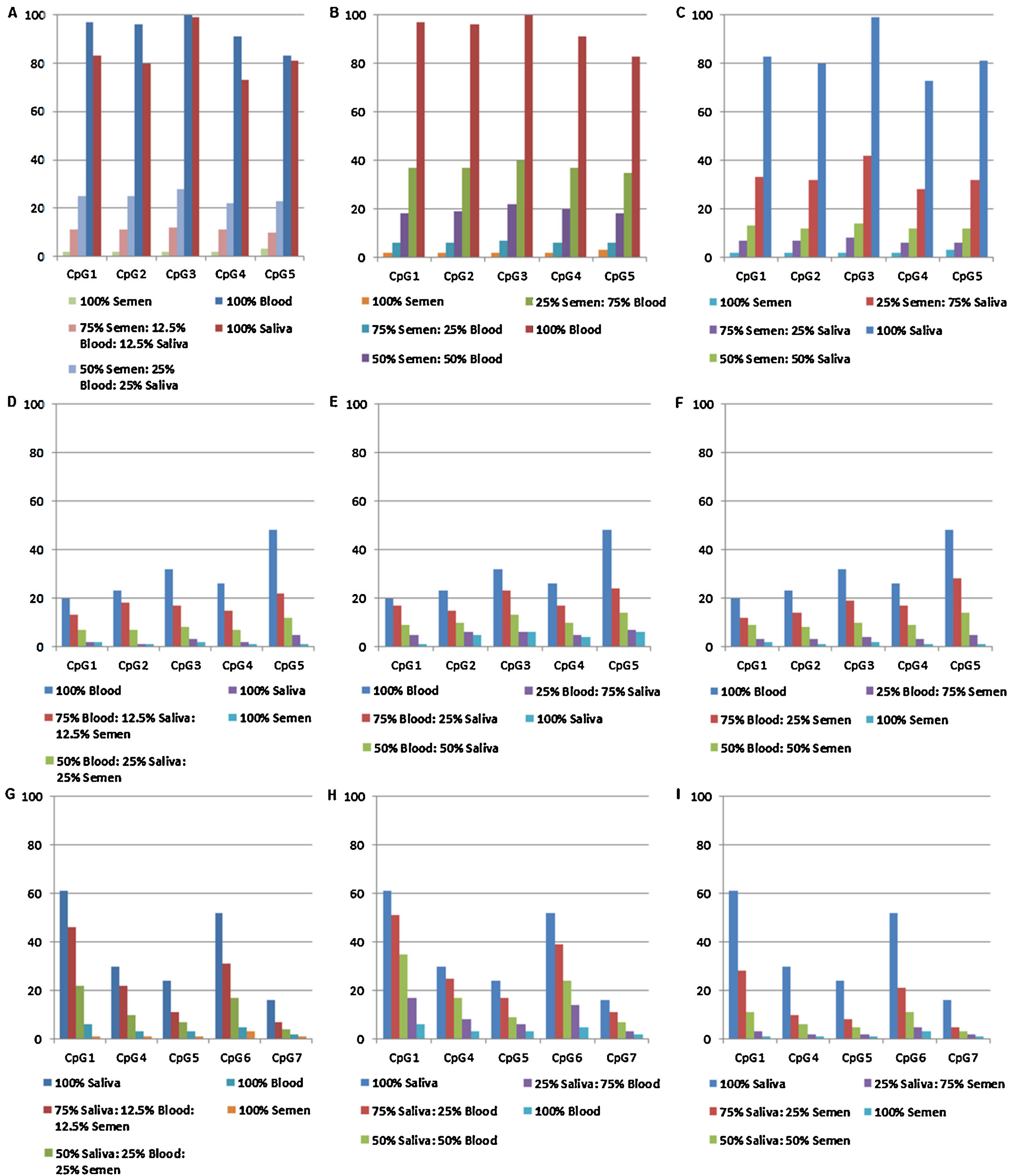

Fig. 3. Mean methylation levels of different ratios of DNA mixtures. (A) Marker ZC3H12D: semen + blood + saliva. (B) Marker ZC3H12D: semen + blood. (C) Marker ZC3H12D: semen + saliva. (D) Marker cg06379435: blood + saliva+ semen. (E) Marker cg06379435: blood + saliva. (F) Marker cg06379435: blood + semen. (G) Marker BCAS4: saliva + blood+ semen. (H) Marker BCAS4: saliva + blood. (I) Marker BCAS4: saliva + semen.

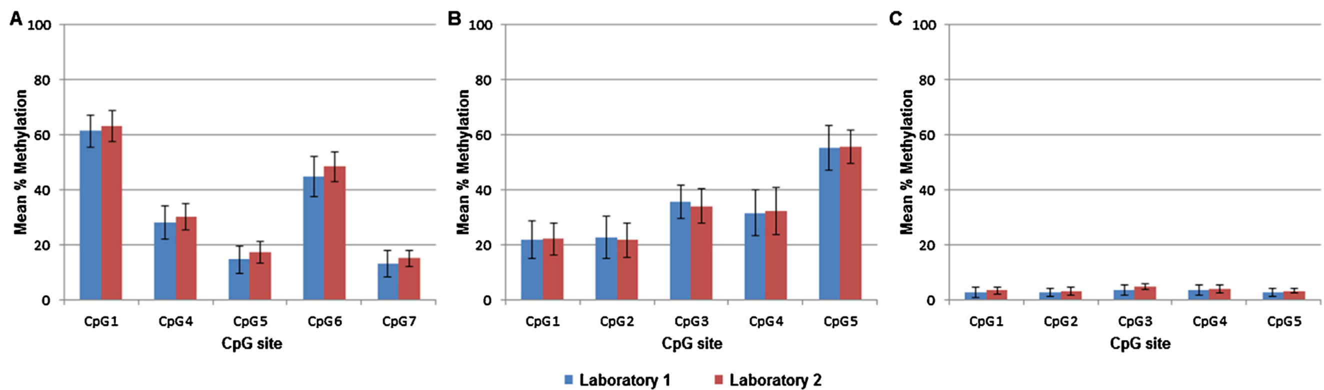

Fig. 4. Comparison between pyrosequencing data generated by two independent laboratories: Laboratory 1 (Florida International University) and Laboratory 2 (University of Southern Mississippi). (A) Methylation data of marker BCAS4 (saliva samples $n = 8$). (B) Methylation data of marker cg06379435 (blood samples $n = 8$). (C) Methylation data of marker ZC3H12D (semen samples $n = 8$).

marker cg0679435, an error occurred as the surrounding reference sequence was not recognized in all primate samples tested. Also, the orangutan sample did not sequence for marker ZC3H12D. All other primate samples presented pyrograms similar to those obtained when testing a human sample. These results are expected since there is a close evolutionary relationship between these animals and humans [15]. In contrast, the other species used in this validation study, which are more evolutionarily distant to humans, failed to show any results. The specificity results obtained here are sufficient to enable the application of these markers in forensic casework.

3.4. Sensitivity studies

The sensitivity was assessed to determine the minimum quantity of DNA with which accurate and reliable results could be obtained. For markers BCAS4 and cg06379435, we were able to obtain successful pyrosequencing results when using 10 ng of pre-bisulfate modified DNA. When the samples were amplified for the ZC3H12D marker, good results were obtained using 1 ng of pre-modified DNA. When the input pre-modified DNA was lowered to 0.1 ng, good quality methylation profiles were obtained (Table 4). All low template samples for the three markers were compared with the control samples and no significant differences were found.

For most of the samples tested, the amplification with the markers BCAS4 and cg06379435 was a success, and we were able to see bands with the correct size on a 2% agarose gel. However, after pyrosequencing, the percent methylation of samples with DNA inputs lower than 10 ng differed from the values presented by the control samples. When working with bisulfite-treated DNA, this so-called PCR-bias is often observed in the amplification step. When using low amounts of DNA, a more efficient amplification can occur with unmethylated alleles when compared to those which are methylated. Alternatively, in some cases, an inverse bias can occur with a deviation toward the methylated alleles [17]. With regard to the samples amplified with the ZC3H12D marker, it was possible to keep the consistency in the results even when using low DNA templates. According to Moskalev et al. [17] the primer design of some markers may lead to a more unbiased amplification of bisulfite-treated DNA. However we cannot prove this assumption for the ZC3H12D PyroMark® CpG assay, since its sequence is proprietary and not provided by the manufacturer of the assay (Qiagen, CA).

3.5. Stability study and case-type samples

When dealing with forensic casework samples, it is common to encounter degraded and contaminated samples. Therefore, we have investigated the effects of PCR inhibitors on the methylation levels. The final inhibitor concentrations tested were: hematin (0.08 mM) and humic acid (0.24 mg/ml). We could observe that when these inhibitors were added prior to bisulfite modification, all samples were amplified and there was no significant difference between the methylation values of the control and the other tested samples (Fig. 1). However, when inhibitors were added after bisulfite modification, the amplification process failed and the pyrosequencing results were similar to those obtained when testing a no template control (water). According to the data provided by the manufacturer (Qiagen, CA), the bisulfite modification kit provides bisulfite conversion as well as cleanup of the modified DNA for subsequent methylation analysis. The cleanup step is likely the reason that PCR inhibitors had no discernable effect on the PCR amplification that followed bisulfite modification.

When testing the stability of the markers under degraded conditions, we noticed that all primers were still able to amplify the target sequence. There were no significant differences in the methylation values for the degraded samples *versus* intact DNA (positive control) (Fig. 2).

Also, all three simulated case-type samples were successfully modified, amplified and pyrosequenced. Table 5 presents the pyrosequencing data obtained for each tested sample.

3.6. Mixtures

Determining the presence of a mixture is considered a difficult task, even in a case of human identification using short tandem repeats. To determine which body fluids are present in a mixture containing multiple tissue sources is even more arduous. In this study, we prepared different ratios of DNA mixtures to analyze if there were differences in the methylation values when multiple cell types are present in the same sample.

For all markers, regardless of whether a mix of two or three body fluids are present (semen, blood and saliva) in one sample, it is possible to recognize that a mixture is present. The methylation levels for mixtures show intermediate percentages when compared to pure samples (Fig. 3). In this study, with the analysis of a single specific marker for a particular body fluid, we were not able to perform complete mixture deconvolution. Additional genome

locations will need to be identified and analyzed in order to fully deconvolute this data, however our results demonstrate that mixtures can be detected for sample discrimination.

3.7. Reproducibility

A concordance study was performed in two independent laboratories by operators with different experience levels. No significant difference was found when we compared the methylation values for each CpG site obtained by both laboratories (*t* test, $p > 0.05$) (Fig. 4).

4. Conclusion

Body fluids recovered from crime scenes can provide crucial information for an investigation, particularly in situations where the presence of a suspects DNA may not in dispute, such as sexual assault and other types of violent crime. Epigenetic markers permit the identification of these biological fluids from trace levels of extracted DNA making them a powerful investigative tool. The validation studies presented here demonstrate robustness and reliability of the tested markers. We present results demonstrating, sensitivity, species specificity, resistance to degradation and the effect of mixtures. Our results demonstrate the capability of these new epigenetic markers in the determination of trace levels of body fluids. They provide the community with specific and sensitive epigenetic methods that can add important information to the trier of fact in forensic analysis.

Conflict of interest

The authors have declared the following potential conflict of interest: They acknowledge the financial support for this project from Qiagen Inc.

Acknowledgments

The authors would like to thank all the volunteers who participated in this study. The authors would like to acknowledge the generous support from Qiagen Inc. with bisulfite modification, PCR, and pyrosequencing and the expertise of all the Qiagen staff, especially Mark Guilliano. Deborah S. B. S. Silva is supported by FAPERGS/CAPES and CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico—modalidade Doutorado Sanduíche), Brazil. This study was supported by funding provided by the National Institute of Justice,Department of Justice, USA, under award number 2012-D1-BX-K018. Points of view in the document are those of the authors and do not necessarily represent the official view of the U.S. Department of Justice.

References

[1] K. Virkler, I.K. Lednev, Analysis of body fluids for forensic purposes: From laboratory testing to non-destructive rapid confirmatory identification at a crime scene, Forensic Sci. Int. 188 (2009) 1–17.

[2] N. Frascione, R. Thorogate, B. Daniel, S. Jickells, Detection and identification of body fluid stains using antibody-nanoparticle conjugates, Analyst 137 (2012) 508–512.

[3] H.Y. Lee, M.J. Park, A. Choi, J.H. An, W.I. Yang, K. Shin, Potential forensic application of DNA methylation profiling to body fluid identification, Int. J. Legal Med. 126 (2012) 55–62.

[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, Forensic Sci. Int. Genet. 3 (2009) 80–88.

[5] D. Frumkin, A. Wasserstrom, B. Budowle, A. Davidson, DNA methylation-based forensic tissue identification, Forensic Sci. Int.: Genet. 5 (2011) 517–524.

[6] D.P. Snustad, M.J. Simmons, Fundamentos de Genética, Guanabara Koogan, Rio de Janeiro, 2001.

[7] B. Alberts, A. Johnson, P. Walter, Biologia Molecular da Célula, Artmed, Porto Alegre, 2004.

[8] T.B. Miranda, P.A. Jones, DNA methylation: the nuts and bolts of repression, J. Cell Physiol. 213 (2007) 384–390.

[9] D. Frumkin, A. Wasserstrom, B. Budowle, A. Davidson, DNA methylation-based forensic tissue identification, Forensic Sci. Int.: Genet. 5 (2011) 517–524.

[10] T. Madi, K. Balamurugan, R. Bombardi, G. Duncan, B. McCord, The determination of tissue-specific DNA methylation patterns in forensic biofluids using bisulfite modification and pyrosequencing, Electrophoresis 33 (2012) 1736–1745.

[11] J. Park, O. Kwon, J.H. Kim, H. Yoo, H. Lee, K. Woo, S. Kim, S. Lee, Y.S. Kim, Identification of body fluid-specific DNA methylation markers for use in forensic science, Forensic Sci. Int.: Genet. 13 (2014) 147–153.

[12] I. Blotta, F. Prestinaci, S. Mirante, A. Cantafora, Quantitative assay of total dsDNA with PicoGreen reagent and real-time fluorescent detection, Ann. Ist. Super Sanità 41 (2005) 119–123.

[13] Revised validation guideline—Scientific Working Group on DNA Analysis Methods (SWGDAM) (2004) Forensic Sci Commun v.6, n.3. Available at http://www.cstl.nist.gov/strbase/validation/SWGDAM_Validation.doc.

[14] J. Wang, B. McCord, The application of magnetic bead hybridization for the recovery and STR amplification of degraded and inhibited forensic DNA, Electrophoresis 32 (2011) 1631–1638.

[15] L. Chaitanya, M. Van Oven, N. Weiler, J. Harteveld, L. Wirken, T. Sijen, P. de Knijff, M. Kayser, Developmental validation of mitochondrial DNA genotyping assays for adept matrilineal inference of biogeographic ancestry at a continental level, Forensic Sci. Int.: Genet. 11 (2014) 39–51.

[16] D. Pellacani, D. Kestoras, A.P. Droop, F.M. Frame, P.A. Berry, M.G. Lawrence, M.J. Stower, M.S. Simms, V.M. Mann, A.T. Collins, G.P. Risbridger, N.J. Maitland, DNA hypermethylation in prostate cancer is a consequence of aberrant epithelial differentiation and hyperproliferation, Cell Death Differ. 21 (2014) 761–773.

[17] E.A. Moskalev, M.G. Zavgorodnij, S.P. Majorova, I.A. Vorobjev, P. Jandaghi, I.V. Bure, J.D. Hoheisel, Correction of PCR-bias in quantitative DNA methylation studies by means of cubic polynomial regression, Nucleic Acids Res. 39 (2011) e77.