



Characterization of DNA methylation-based markers for human body fluid identification in forensics: a critical review

Farzeen Kader¹ · Meenu Ghai¹ · Ademola O. Olaniran²

Received: 13 March 2019 / Accepted: 15 October 2019 / Published online: 12 November 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract

Body fluid identification in crime scene investigations aids in reconstruction of crime scenes. Several studies have identified and reported differentially methylated sites (DMSs) and regions (DMRs) which differ between forensically relevant tissues (tDMRs) and body fluids. Diverse factors affect methylation patterns such as the environment, diets, lifestyle, disease, ethnicity, genetic variation, amongst others. Thus, it is important to analyse the stability of markers employed for forensic identification. Furthermore, even though epigenetic modifications are described as stable and heritable, epigenetic inheritance of potential markers for body fluid identification needs to be assessed in the long term. Here, we discuss the current status of reported DNA methylation-based markers and their verification studies. Such thorough investigation is crucial to develop a stable panel of DNA methylation-based markers for accurate body fluid identification.

Keywords DNA methylation · Forensic science · Body fluid identification · Mutations · Genetic variation · Heritability

Introduction

DNA methylation

Epigenetics is the study of reversible, heritable changes that influence gene regulation without altering the underlying DNA sequence [1]. Perhaps, the most well characterised is DNA methylation, an epigenetic modification whereby a methyl group is covalently added to cytosine in DNA [2]. Most DNA methylation occurs in CpG dinucleotides [3]; however, methylation elsewhere (CpT, CpA and CpC) has

been documented [4, 5]. The human genome harbours millions of CpGs which may exist in a methylated, semi-methylated or unmethylated state depending on the chromosomal location, alleles, cell type or developmental phase [6–8]. Introns, 3' untranslated regions (UTRs) and intergenic sequences are severely depleted in CpGs, whereas exons have a somewhat higher density of CpGs [9]. It is estimated that between 60 and 90% of CpGs are methylated. Unmethylated CpGs were originally thought to be grouped in regions known as CpG islands (CGIs); however, CGIs have also been reported to exist in methylated states. CGIs are often located in promoter regions and control gene expression via DNA-differentiated methylation levels [10]. CGIs are 300–3000 bp long, with a CpG content of above 50%. DNA methylation is essential in mammalian development, partaking in numerous processes including genome stability, X-chromosome inactivation, genomic imprinting, aging and carcinogenesis [2, 3]. Aberrant DNA methylation patterns are said to be implicated in numerous diseases such as cancers [11], autoimmune diseases [12], diabetes [13], neurodegenerative [14] and psychological disorders [15].

DNA methylation patterns are the collective result of methylation establishment and maintenance [extensively described in 6, 16–18]. Such established methylation patterns may be stored as well as stably inherited and may affect an organism phenotype [19]. Even though DNA methylation is heritable

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00414-019-02181-3>) contains supplementary material, which is available to authorized users.

✉ Meenu Ghai
ghai@ukzn.ac.za

¹ Discipline of Genetics, School of Life Sciences, College of Agriculture, Engineering and Science, University of KwaZulu-Natal (Westville Campus), Private Bag X54001, Durban, Republic of South Africa

² Discipline of Microbiology, School of Life Sciences, College of Agriculture, Engineering and Science, University of KwaZulu-Natal (Westville Campus), Private Bag X54001, Durban, Republic of South Africa

over cell divisions, the methylome undergoes changes during development. During early embryogenesis, the oocyte and sperm initially display high methylation levels. Remodelling of methylation occurs in the first cell divisions of the zygote when both maternal and paternal genomes become demethylated. The maternal genome undergoes passive loss of methylation, while the paternal genome is actively demethylated during several DNA replications. Methylation levels are re-established during the blastocyst stage, during which cells also lose totipotency. While these dramatic fluctuations in methylation are not repeated, local changes in DNA methylation do occur in each cell lineage, thereby resulting in specific methylomes which may enable prediction of cell type [18, 20].

Furthermore, DNA methylation patterns are dynamic and may vary due to several factors. These comprise genetic variation including single-nucleotide polymorphisms (SNPs) [21], diet, nutrition and alcohol [22, 23], lifestyles, stress [8, 24], smoking [25, 26], drugs and substance abuse [27] amongst others. Several studies have also highlighted age-related changes in global DNA methylation [17, 28, 29]. Age and DNA methylation are generally thought to exhibit an indirect relation, i.e. the aging of cells is coupled with the loss of DNA methylation [30]. This was evident in a study by Florath et al. [31] who found a negative correlation between age and DNA methylation at 43 CpG sites in six genes. One study by Day et al. [32] presented age-related DNA methylation changes to be tissue-specific. With progressing age, CpG sites outside of CGIs are said to lose methylation, while the opposite was true for the CGIs [30, 33].

DNA methylation has been shown to affect chromatin availability to transcription factors and binding of regulatory proteins, thereby playing a role in the regulation of gene expression. This control of gene transcription is critical for normal human development and cellular differentiation [33]. Unmethylated sequences are generally expressed, while methylated promoters are generally hindered and thus methylated sequences undergo transcriptional repression [34–36]. Conversely, gene body methylation has been reported to show a positive correlation with gene expression. DNA methylation is more predominant within gene bodies than at promoters [37]. While most CGIs display low methylation, a small percentage does acquire methylation during development. Additionally, a significant number of CGIs are differentially methylated between healthy tissues and cell types [38, 39]. DNA methylation may also be present in CGI shores, which are regions 0–2 kb from CGIs, CGI shelves which are regions 2–4 kb from CGIs and other open sea regions which are CpG sites harboured in the genome randomly. Flanking regions of a CGI have been found to exhibit stable methylation which may also play a role in gene regulation [40–42].

Epigenome-wide association studies (EWAS) have found numerous individual CpG sites and genomic regions that

show different methylation patterns between human tissue and body fluids. These are known as differentially methylated sites (DMSs) and tissue-specific differentially methylated regions (tDMRs), respectively [43, 44]. tDMRs are mainly found at the margins of CpG islands and both CpG and G/C content is lower than that of surrounding regions. tDMRs are thought to afford cells with an epigenetic memory by generating cell-type specific hypo- and hypermethylation patterns [35, 45]. While the distinct active regulatory role of tDMRs is not fully established, these differentially methylated regions are thought to function by either attracting or preventing the binding of specific factors in a methyl-dependent manner. Major links between gene silencing and tDMRs have been identified [46, 47]. Based on differential methylation profiles, DMSs and tDMRs enable distinction between tissues and fluids [44, 48, 49]. It is this factor, along with potential to estimate age, that is now exploited in forensic science.

DNA methylation-based markers in forensic sciences

The recovery and subsequent analysis of human body tissues and fluids from crime scenes are perhaps one of the most crucial preliminary steps of a forensic investigation [50, 51]. The transfer of skin, blood, sweat, urine and saliva from human to object may indicate and help decipher events of a physical confrontation. Similarly, semen, vaginal fluid and menstrual blood assist in the reconstruction of sexual assault events. Body fluids can aid in the reconstruction of crime scenes and ease the process of identifying individuals who were present and involved in the crime [35, 52–54]. Identification of body fluids, along with linking donors to said body fluids, is not mediocre tasks; they may have localized on material or surfaces, be mixed with other fluids or contaminants, present in minimal quantities and in sub-optimal quality [41, 55], all of which present restraints to exact identification.

Studies show that tDMR-based markers present salient features for the identification of forensically relevant body fluids and in estimating individual age [56, 57]. However, the focus of the present review is on the use of DNA methylation-based markers for body fluid identification.

Forensic samples are not always present in high quantity and quality, and it is imperative to restrict consumption and degradation of valuable evidence. DNA methylation-based assays are compatible with existing short tandem repeat (STR) typing protocols and allow multiplexing which permits identification of several body fluids at the same time [58–61]. Mixture analysis has also been reported by several studies [62, 63]. However, even though DNA itself is stable, some methylation assays require bisulphite conversion which may degrade DNA. In addition, sometimes low-input protocols might be challenging to work with [64, 65].

Frequently employed methods for the identification and quantification of DNA methylation levels in forensics are

discussed in [Supplementary Information](#). The most recent studies on current DNA methylation-based markers, verification of marker stability and results of mixture analysis are critically discussed in the succeeding pages.

The current status of DNA methylation-based markers for forensic application

Frumkin et al. [58] were the first forensic-based study to report differentially methylated genomic loci between venous blood, saliva, semen, skin epidermis, vaginal fluid, menstrual blood and urine using methylation-sensitive restriction enzyme-PCR (MSRE-PCR). Fifteen loci which contained the restriction enzyme for *HhaI* (GCGC) were selected by an in-house developed software program for the tissue identification assay. However, only four proved promising. Varying ratios of methylation were observed in the target tissues and body fluids. The ratio of methylation levels between *L91762/L68346* was lower in semen than all other fluids, whereas high methylation ratios of the same loci, *L91762/L68346* and *L76138/L26688* were characteristic of skin epidermis samples. The results showed distinction between semen and other fluids and skin epidermis from other fluids, respectively. However, a study by Gomes et al. [66] failed to reproduce these findings, since their results showed that high methylation ratios of *L91762/L68346* were characteristic of not only skin, but also saliva samples, rendering the marker non-specific.

Madi et al. [67] employed bisulphite modification and pyrosequencing to research differentially methylated regions that were previously reported by Eckhardt and colleagues [44]. The group found several CpG sites in the *ZC3H12D*, *FGF7*, *C20orf117* and *BCAS4* genes which displayed differential methylation profiles of saliva, blood, semen and skin tissue. Five CpG sites in *ZC3H12D* exhibited semen specificity. *FGF7* also enabled identification of semen with hypermethylation relative to blood, saliva and skin. Blood was positively identified by hypermethylation of a locus within *C20orf117*. Accurate differentiation from skin was inconclusive in the experiment. Therefore, this marker requires more vigorous testing for identification of blood or skin. Eckhardt et al. [44] also showed high levels of methylation at *C20orf117* in CD4+ and CD8+ lymphocytes relative to skin and sperm. Eckhardt et al. [44] found highest methylation of a region 5' upstream of *BCAS4* in semen; however, the same tDMR showed highest methylation at five CpG sites in saliva when compared to other fluids tested by Madi et al. [67]. Despite the confusion between Eckhardt et al. [44] and Madi et al. [67], in a follow-up study by Antunes et al. [68], the same markers *ZC3H12D* and *FGF7*, *C20orf117* and *BCAS4* were proven useful for semen, blood and saliva identification, respectively.

Testing aged samples by Antunes et al. [68] *ZC3H12D*, *FGF7*, *C20orf117* and *BCAS4* markers were tested to determine mean percent methylation of 9-year-old blood samples, 20-year-old blood and semen samples. Interestingly, methylation patterns were observed to be unwavering over such long periods of time; percent methylation was the same as samples that were recently collected.

Lee et al. [69] selected five previously reported tDMRs to differentiate between saliva, blood, semen, vaginal fluid and menstrual blood. tDMRs in the *USP49* and *DACT1* genes (ubiquitin-specific peptidase 49 and Dapper 1 isoform II, respectively) were selected as semen-specific markers as they were identified to be testes-specific by Kitamura et al. [70] and tDMRs for *PFN3* (profilin III), *PRMT2* (protein arginine N-methyltransferase II) and *HOXA4* (homeobox A4) genes were chosen as blood-specific markers as different methylation patterns in blood, spleen and brain tissues were observed by Illingworth et al. [71]. Bisulphite treatment, PCR and sequencing showed that *DACT1* and *USP49* were unmethylated in over 90% of clones from semen and hypermethylated in almost all blood, saliva, vaginal fluid and menstrual blood clones. The *PFN3* marker displayed moderate methylation (65%) in vaginal fluid whereas high methylation (80%) was observed in other tissues and fluids. *PRMT2* was hypermethylated in vaginal fluid and menstrual blood and demonstrated great differences between semen/vaginal fluid and semen/menstrual blood. Similar results were reported by An et al. [55] and Choi et al. [56]. The *HOXA4* tDMR displayed high degrees of methylation in blood and female saliva and was hypomethylated in vaginal fluid and menstrual blood. The latter two markers did not show specificity for a single body fluid; however, the authors suggested that low methylation of *HOXA4* and high methylation of *PRMT2*, *USP49* and *DACT1* could be used to confirm the presence of vaginal fluid and menstrual blood.

Sensitivity and forensic simulation by Choi et al. [56] Using a multiplexed MSRE-PCR and a complementary bacterial DNA-based assay, DNA methylation patterns could be generated for saliva and semen with just 500 pg (0.5 ng) or more of starting DNA and 250 pg (0.25 ng) of starting DNA from vaginal fluid was sufficient, showing that minimal amounts of DNA could be used for analysis. Mixtures of saliva and semen and semen and vaginal fluid were clearly distinguished by amplification of *L81528* (semen-specific) in this study. A single post-coital penile sample and three post-coital vaginal samples were tested for an artificial sexual assault case and results showed a mixed sample profile as low peaks were observed for *PFN3* (vaginal fluid-specific) and the semen-specific *L81528* marker.

Park and colleagues [72] used the HumanMethylation 450 K bead array to identify eight CpG sites which were differentially methylated between saliva, blood, semen and

vaginal fluid. All chromosomal regions and details of the markers are shown in Table 1. Overall, each marker displayed hypermethylation for its target body fluid with significantly reduced methylation in the others. Both saliva and blood markers showed above 50% methylation levels in respective target body fluids while less than 10% methylation were observed in other fluids. Semen and vaginal fluid markers showed above 90% and 65% methylation, respectively, in target body fluids, whereas below 16% methylation were observed in other fluids [72]. Sensitivity tests by Park et al. [72] showed that at least 10 ng of starting DNA was sufficient for precise distinction of fluids.

Lee et al. [59] identified 64 potential differentially methylated CpGs between saliva, blood, semen, vaginal fluid and menstrual blood using the HumanMethylation 450 K bead array. The 64 sites were selected as they exhibited more than 30% differences in methylation values between saliva, blood and vaginal fluid, whereas a 50% threshold was used for potential semen-specific candidate CpGs. The group examined methylation of the 64 candidate CpGs in 151 more body fluids using bisulphite sequencing and Methylation SNaPshot (Methylation Sensitive Single Nucleotide Primer Extension, MS-SNuPE) and identified a subset of CpG sites that showed most body fluid specificity. Only one CpG site (cg17621389) showed semen-specific hypomethylation, whereas seven others showed target body fluid-specific hypermethylation. These included one saliva-specific marker that was 2 bp downstream of cg09652652, i.e. cg09652652-2d, two blood-specific markers (cg06379435 and cg01543184), two semen-specific markers (cg26763284 and cg17610929) and two vaginal fluid-specific markers (cg09765089-231d and cg26079753-7d). Two markers of this panel, cg06379435 (blood-specific) and cg17610929 (semen-specific) were also included in the study by Park and colleagues [72]. Some semen samples showed almost complete methylation not only at semen-specific sites (cg26763284 and cg17610929), but also at one blood-specific marker (cg01543814) with hypomethylation at other markers. Four semen samples showed low methylation signals at one blood-specific marker (cg06379435) and vaginal fluid-specific markers (cg09765089-231d and cg26079753-7d). Such sporadic methylation profiles exhibited by semen could have been due to a small number of white blood cells that may have been present. Additionally, the authors [59] admittedly had trouble with menstrual blood, as 2 out of 11 samples exhibited similar methylation profiles as vaginal fluid, and 9 out of 11 menstrual blood samples also showed methylation signals at the blood-specific markers (cg06379435 and cg01543184). This was attributed to disregarding the time of collecting menstrual blood during the menstrual cycle, and small sample size (11 menstrual blood samples out of a total 151 samples). The group also exposed samples to the environment by placing samples in shaded areas for 75 days to test the efficacy of

the multiplex SNaPshot (MS-SNuPE) assay. Similar to their previous research [55], all fluids except saliva were successfully analysed. When testing sensitivity of the method, successful DNA methylation profiling results were achieved with approximately 0.5 ng or more bisulphite-converted DNA [59].

Since the markers studied by Lee et al. [59] required extra interrogation for positive identification of body fluids, the research was repeated by the same group [60] by profiling DNA methylation levels in a total of 70 samples using the HumanMethylation 450 K bead array. The authors found two sites in the *SLC26A10* gene (cg09696411 and cg18069290) which showed potential for menstrual blood specificity. The specificity of the CpGs was again tested in 125 vaginal fluid and 201 menstrual blood samples, and both showed menstrual blood-specific hypermethylation. The authors found that methylation levels were highest when sample collection was performed on the second day of menstrual bleeding, whereas specificity of the markers was reduced on the fourth and fifth days. The multiplex SNaPshot (MS-SNuPE) reaction from their preceding research [59] was also modified. In the new assay performed on 229 body fluids, the semen hypomethylation marker cg17621389 was removed, the blood marker cg01543184 (which also showed semen-specificity) was replaced with cg08792630 (which was also studied by Park et al. [72]), and the two new menstrual blood-specific markers were added. This resulted in all markers successfully exhibiting body fluid-specific hypermethylation.

Another study based on tDMRs identified by Eckhardt et al. [44] examined five regions in the *BIK*, *CYTH4*, *GAS2L1*, *MDFI* and *OSM* genes [74]. The group used bisulphite pyrosequencing and found that all CpG sites showed differential methylation patterns between saliva, blood, semen, vaginal fluid and menstrual blood. However, the fourth CpG site in *CYTH4* showed blood-specific hypomethylation, three CpGs in *GAS2L1* showed blood-specific hypermethylation and four sites in *MDFI* showed menstrual blood-specific hypomethylation. Regions examined in the *BIK* and *OSM* genes did not show specificity towards a single body fluid [74].

Antunes et al. [77] developed a high-resolution melt analysis (HRMA) assay to confirm the use of a 91-bp region located in an intron of *ZC3H12D* as a semen-specific marker. In previous research using pyrosequencing [67], an intron in the *ZC3H12D* gene showed hypermethylation in saliva and blood, but hypomethylation in semen. Indeed, the same results were observed by Antunes et al. [77]. During sensitivity analysis, 1 ng of starting genomic DNA was found to be the minimum amount necessary for amplification using the developed HRM assay. Generally, in forensic casework, degraded and contaminated samples are recovered from crime scenes. Such contaminants may contain inhibitors, which, when co-isolated with DNA could decrease amplification efficiency. Inhibitors may act by either binding DNA or by diminishing

Table 1 Summary of previously reported body fluid-specific DMSs, chromosomal locations, sample sizes and methods of analysis

Marker/Illumina ID of CpG site	Chromosome location/ region ^a	Gene/locus	Nearest CGI according to UCSC ^b	Specific genomic locations according to Ensembl (GRCh37)	Saliva (n)	Blood (n)	Semen (n)	Vaginal fluid (n)	Menstrual blood (n)
cg09652652-2d	Chr3:194408845-47	<i>FAM43A</i>	CpG 237	Located within CGI which overlaps exon 1/1 of gene	40	40	49	11	11
cg26107890	Chr3:124860686	<i>SLC12A8</i>	CpG 48	Located within CGI, Intron 5/13 of <i>SLC12A8</i>	20	20	20	20	149
cg20691722	Chr3:181421427	<i>SOX2OT</i>	-	Intron 1-2 of <i>SOX2-OT</i> -255, transcript ENST00000630150.2	20	20	20	20	-
cg21597595-4d	Chr2:5506229-33	<i>SOX11</i>	-	5' upstream sequence of gene	100	100	91	55	96
cg15227982-137d	Chr10:104535855-992	<i>WBPI1/OPAIL</i>	-	3' site, Intron	100	100	91	55	96
cg09107912	Chr6:159590060	<i>FNDC1</i>	CpG 162	5' upstream sequence, 369 bp before exon 1 ENSE00001204099	15	15	10	10	15
cg16732616	Chr1:50886783	<i>DMRTA2</i>	CpG 480	Exon ENSE00001557667	15	15	10	10	15
-	Chr22:37705392	<i>CYTH4</i>	-	Intron 9-10 before exon ENSE00003519453	20	20	10	10	7
cg08792630	Chr6:108883909	<i>FOXO3</i>	-	Intron 1-2 before exon ENSE00003656576	20	20	20	20	149
cg06379435	Chr19:3344273	<i>NFIC</i>	-	5' upstream sequence, 15,288 bp before exon ENSE00002967604	20	20	20	20	-
cg06379435, -7d,	Chr19:3344273-80, 87				40	40	49	11	11
cg26285698-14d,	Chr16:29757335-49, 74	<i>C16orf54</i>	-	5' upstream sequence, 8 bp before exon ENSE00001299431	100	100	91	55	96
cg03363565-59d	Chr16:474529-88	<i>RAB11FIP3</i>	-	5' upstream sequence, before exon ENSE0000222336	100	100	91	55	96
cg24124443	Chr12:107711649	<i>BTBD11</i>	CpG 216	5' upstream sequence, 541 bp before exon ENSE00001790674	15	15	10	10	15
cg01607849	Chr12:6976130	<i>TP11</i>	-	5' upstream sequence, 153 bp before exon ENSE00001164420	15	15	10	10	15
-	Chr12:73985697/Chr3:12175317 (hg18)	L91762(<i>CAPS2</i>)/L68346 (<i>TTIMP4</i>)	<i>TTIMP4</i> -CpG 27	<i>CAPS2</i> (Intron 8-9 before exon ENSE00001243541)/ <i>TTIMP4</i> (Exon ENSE00001030692)	14	14	11	-	-
cg23521140	Chr2:37617601	-	-	Intergenic	20	20	20	20	-
cg17610929	Chr2:220379044	<i>ACCN4/ASJC4</i>	-	5' upstream sequence, before exon ENSE00001181728	20	20	20	20	-
cg26763284	Chr8:145018185	<i>PLEC</i>	CpG 54	Exon ENSE00001369893	40	40	49	11	11
cg17621389	Chr6:147728290	Pseudogene <i>YAP1P1</i>	CpG 23	Exon ENSE00001560366	40	40	49	11	11
cg05261336	Chr7:1032905	-	-	Located after gene <i>C7orf50</i>	15	15	10	10	15
cg22407458-138d	Chr8:145018185/323	<i>PLEC</i>	CpG 54	Spans exon ENSE00001369893	20	20	20	20	149
cg22407458-288d	Chr6:35109122-410	<i>TCPI1</i>	CpG 75	Exon ENSE00002723532	100	100	91	55	96
cg05656364	Chr2:85804732	<i>VAMP8</i>	-	3' UTR, Exon ENSE00001076027	100	100	91	55	96
cg04382920	Chr4:128651527	<i>SLC25A31</i>	CpG 54	5' upstream sequence, 6 bp away from exon ENSE00000999807	11	9	9	10	10
cg11768416	Chr12:91348697	<i>CCER1</i>	-	Intron 1-2 before exon ENSE00002378999	20	15	13	10	15

Table 1 (continued)

Marker/Illumina ID of CpG site	Chromosome	Coordinate	Gene	Region	Urine (n)	Skin (n)	Buccal cells (n)	Target body fluid	Methylation status	Method of analysis	Reference		
cg01774894	Chr17	46674395	<i>LOC404266</i> (<i>HOXB-A53</i>)	Intron 2-3 before exon	-	-	-	Saliva	High	20	20	20	-
cg14991487	Chr2	176987405	<i>HOXD9</i>	Exon	CpG 172	-	-	Saliva	High	20	20	20	-
cg14991487-85d	Chr2	176987405-90								100	100	91	55
cg03874199-212d	Chr2	176964457-669	<i>HOXD12</i>	5' UTR, Exon	CpG 118	-	-	Saliva	High	100	100	91	55
cg26079753-7d	Chr12	54355528	<i>HOTAIR</i>	3' UTR	-	-	-	Saliva	High	40	40	49	11
cg09765089	Chr7	27291346	-	Intergenic	CpG 89	-	-	Saliva	High	20	20	20	149
cg09765089-231d	Chr7	27291346-577								40	40	49	11
cg25416153	Chr18	43355414	-	Intergenic	CpG 37	-	-	Saliva	High	15	15	10	10
cg18069290	Chr12	58013539	<i>SLC26A10</i>	5' UTR before exon	CpG 41	-	-	Saliva	High	20	20	20	149
cg09696411	Chr12	58013517	<i>SLC26A10</i>	5' UTR before exon	CpG 41	-	-	Saliva	High	20	20	20	149
cg09696411-25d, 40d	Chr12	58013517-42, 57	<i>SLC26A10</i>	5' UTR before exon	CpG 41	-	-	Saliva	High	100	100	91	55
cg09652652-2d	-	-	-	Saliva	-	-	-	Saliva	High	B, C			Lee et al. [59]
cg26107890	-	-	-	Saliva	-	-	-	Saliva	High	B, C			Lee et al. [60]
cg20691722	-	-	-	Saliva	-	-	-	Saliva	High	A			Park et al. [72]
cg21597595-4d	-	-	-	Saliva	-	-	-	Saliva	High	A			Park et al. [72]
cg15227982-137d	-	-	-	Saliva	-	-	-	Saliva	High	B, C			Forat et al. [73]
cg09107912	-	-	-	Saliva	-	-	-	Saliva	High	B, C			Forat et al. [73]
cg16732616	-	-	-	Saliva	-	-	-	Saliva	High	C, D			Lin et al. [62]
-	-	-	-	Blood	-	-	-	Blood	Low	C, D			Lin et al. [62]
cg08792630	-	-	-	Blood	-	-	-	Blood	High	A			Fu et al. [74]
cg06379435	-	-	-	Blood	-	-	-	Blood	High	A			Park et al. [72]
cg06379435	-	-	-	Blood	-	-	-	Blood	High	B, C			Lee et al. [60]
cg06379435, -7d, 14d	-	-	-	Blood	-	-	-	Blood	High	A			Park et al. [72]
cg26285698-14d, 39d	-	-	-	Blood	-	-	-	Blood	High	B, C			Lee et al. [59]
cg033363565-59d	-	-	-	Blood	-	-	-	Blood	High	B, C			Lee et al. [60]
cg24124443	-	-	-	Blood	-	-	-	Blood	High	B, E			Watanabe et al. [75]
cg01607849	-	-	-	Blood	-	-	-	Blood	Low	B, C			Forat et al. [73]
-	-	-	-	Blood	-	-	-	Blood	High	B, C			Forat et al. [73]
cg23521140	-	-	-	Semen	-	11	-	Blood	High	C, D			Lin et al. [63]
-	-	-	-	Semen	-	-	-	Blood	High	C, D			Lin et al. [63]
-	-	-	-	Semen	-	-	-	Semen	Low	F			Frumkin et al. [58]
-	-	-	-	Semen	-	-	-	Semen	High	A			Park et al. [72]

Table 1 (continued)

cg17610929	-	-	-	Semen	High	A	Park et al. [72]
cg26763284	-	-	-	Semen	High	B, C	Lee et al. [59]
cg17621389	-	-	-	Semen	High	B, C	Lee et al. [60]
cg05261336	-	-	-	Semen	High	C, D	Lin et al. [62]
cg26763282-138d	-	-	-	Semen	High	B, C	Lee et al. [59]
cg22407458-288d	-	-	-	Semen	Low	B, C	Lee et al. [59]
cg05656364	-	-	-	Semen	High	C, D	Lin et al. [62]
cg04382920	12	2	5	Semen	High	B, C	Forat et al. [73]
cg11768416	15	10	10	Semen	Low	A	Vidaki et al. [76]
cg01774894	-	-	-	Vaginal fluid	High	A	Vidaki et al. [76]
cg14991487	-	-	-	Vaginal fluid	High	A	Park et al. [72]
cg14991487-85d	-	-	-	Vaginal fluid	High	B, C	Forat et al. [73]
cg03874199-212d	-	-	-	Vaginal fluid	High	B, C	Forat et al. [73]
cg26079753-7d	-	-	-	Vaginal fluid	High	B, C	Lee et al. [59]
cg09765089	-	-	-	Vaginal fluid	High	B, C	Lee et al. [60]
cg09765089-231d	-	-	-	Vaginal fluid	High	B, C	Lee et al. [59]
cg25416153	-	-	-	Vaginal fluid	High	C, D	Lin et al. [62]
cg18069290	-	-	-	Vaginal fluid	High	B, C	Lee et al. [60]
cg09696411	-	-	-	Vaginal fluid	High	C, D	Lin et al. [62]
cg09696411-25d, 40d	-	-	-	Menstrual blood	High	B, C	Lee et al. [60]
	-	-	-	Menstrual blood	High	B, C	Lee et al. [60]
	-	-	-	Menstrual blood	High	B, C	Forat et al. [73]

Note: Illumina ID is a unique ID in the Illumina Human Methylation 450K bead array

ng not given

^ahg19 unless otherwise stated; method of analysis: (a) bisulphite pyrosequencing, (b) methylation SNaPshot, (c) bisulphite sequencing, (d) methylation-specific PCR, (e) methylation-sensitive restriction enzyme PCR, (f) high-resolution melt analysis

the activity of the Taq polymerase enzyme [77]. A simple clean-up step might not be adequate enough to remove inhibitors which bind to DNA [77]. One substance which is known to bind DNA and hence negatively affect PCR is humic acid [78, 79]. Thus, to determine if the clean-up steps in bisulphite conversion were sufficient to remove humic acid, it was added before bisulphite conversion. Addition of humic acid did not inhibit amplification and proved that clean-up steps after bisulphite conversion were sufficient to remove any humic acid that may be co-extracted with genomic DNA. To determine if the presence of humic acid affects PCR, humic acid was also added after conversion. While the PCR efficiency declined, it did not affect the melting curve and regardless of the humic acid, as semen still showed lower melting temperature and methylation levels when compared to saliva and blood. The study was performed to portray the development, sensitivity, accuracy and efficiency of the assay itself; however, the authors brilliantly demonstrated the use of the assay by validating a previously reported semen-specific marker [77].

Two markers, one specific to saliva (BCAS4) and one for semen (ZC3H12D) from Madi et al. [67], and one marker specific to blood (cg06379435 and four adjacent CpG sites) from Park et al. [72] were analysed by Silva et al. [63]. In this extensive report, bisulphite pyrosequencing was used to perform several developmental validations to determine the efficacy of the markers. BCAS4 showed saliva-specific hypermethylation, cg06379435 showed blood-specific hypermethylation and ZC3H12D showed semen-specific hypomethylation. Similar results were obtained even when vaginal fluid, menstrual blood and nasal secretion were included in the reaction; however, a CpG site in the BCAS4 locus (saliva-specific) showed similar methylation levels when compared to vaginal fluid, and one CpG site in the cg06379435 locus (blood-specific) showed similar methylation values in menstrual blood. Similar to their previous work [77], the group tested the effects of inhibitors such as hematin and humic acid which could affect PCR amplification [63, 78]. When added before bisulphite modification followed by amplification, these inhibitors had no discernible effects on methylation levels. This was possibly due to the clean-up step prior to amplification, because when both inhibitors were added after bisulphite modification, the amplification process failed, and results were distorted. There were no significant differences in methylation levels for all three markers even when samples were exposed to heat for 10–25 minutes.

Mixture analysis, forensic simulation and human specificity of markers by Silva et al. [63] In a mixture analysis, methylation levels showed intermediate percentages when compared to pure samples, and hence, ZC3H12D (semen-specific), BCAS4 (saliva-specific) and cg06379435 (blood-specific) could not accurately distinguish the target body fluids in mixtures. Analysis was also performed for a saliva swab obtained

from a coffee drink and blood and semen that were retrieved from cotton fabrics. Methylation levels of the saliva and blood markers were still higher in target body fluids when compared to others, and the semen marker still showed low methylation in semen with high methylation in other fluids. Two novel features of this study included analysis of the same markers and body fluids in another laboratory wherein all results were reproduced, as well as a species-specificity test. Methylation of the markers was profiled in non-human samples such as mouse, dogs, cats, chicken, bovine, equine, pig, chimpanzee, orangutan, gorilla, *Escherichia coli*, *Staphylococcus aureus*, *Enterococcus faecalis* and *Pseudomonas aeruginosa*. Only the non-human primates showed results which was expected given the close evolutionary relation to humans [80]; however, cg0679435 (blood-specific) was not recognized in all primates, and ZC3H12D (semen-specific) was not recognized in orangutan samples. No results were observed for all other non-human species; thus, the authors declared the markers suitable for forensic casework [64]. However, further mixture analysis is necessary, as it was not successful for differentiation of body fluids in the study.

Antunes et al. [81] examined a sub-region of the PFN3 (vaginal fluid-specific) marker found by Lee et al. [69]. The marker was previously shown to display vaginal fluid-specific intermediate hypomethylation. In this study, 10 CpGs out of the previously reported 42 CpGs located in the 5' upstream sequence were examined by pyrosequencing. Similar to the previous reports, the marker did not show complete hypomethylation, but rather intermediate methylation (25–55%) when compared to saliva, blood which showed > 60% methylation levels, while semen showed lower than 10% methylation. The study found that 5 ng of starting genomic DNA was sufficient to obtain the above results.

Mixture analysis and human specificity of markers by Antunes et al. [81] During the mixture analysis by Antunes et al. [81], saliva, blood and semen were mixed in varying concentrations. The results showed that each fluid displayed intermediate methylation levels and accurate discrimination was not possible; hence, it was suggested that upon deconvoluting mixtures, more than one marker specific to each tested body fluid should be analysed. Similar to their previous work [63], a unique aspect of the study included a test of the influence of non-human DNA on the specificity of the marker. Only chimpanzee DNA showed similar results to human DNA. The cats, cows, orangutan and gorilla DNA yielded non-specific results which the pyrosequencing software declared as unknown as it did not match the human reference sequences, while none of the other non-human DNA yielded any results.

Watanabe et al. [75] developed a quantitative real-time PCR-based assay to examine the specificity of the potential blood-specific CpG site cg06379435 as well as CpGs located 7 bp and 14 bp downstream. These 3 CpG sites were analysed

in saliva, blood, semen and vaginal fluid samples. The marker was indeed shown to be specific for blood as methylation levels were highest in blood compared to other fluids. Since mixtures of body fluids are usually found at crime scenes, the group resorted to analysing rs7359943, which is the SNP adjacent to cg06379435 to identify semen and blood in a mixture. For this, a DNA sample mixed with 50% of semen DNA with the rs7359943 AA genotype and 50% of blood DNA with the rs7359943 GG genotype was used. The methylation levels of all three CpG sites were lower in AA allele clones (semen) and higher in the GG allele clones (blood). While the method was successful and depicts the advantages of using genetic information along with epigenetic markers, it is only applicable when the identity of the mixed body fluids is known, and alleles from one individual can be distinguished from the genotype of other individuals. Furthermore, while the study was informative, menstrual blood was not included in this research. Blood is a component of menstrual blood, and confusion between the body fluids is common. It is essential to consider menstrual blood when interrogating a blood-specific marker, especially since many researchers do have difficulty in differentiating between these two body fluids [59, 62].

Age and sensitivity testing by Watanabe et al. [75] The cg06379435 marker also proved to be effective and showed low methylation levels when tested against saliva and semen samples which were stored at room temperature from between 4 months and 29 years, whereas high methylation levels were observed for 29-year-old blood stains. Upon testing sensitivity, various concentrations (0.1–10 ng) of pooled blood DNA were evaluated. The group found that just 1 ng of starting DNA was sufficient for blood identification. The accuracy of the marker was tested by mixing various concentrations of blood with semen; the results showed that blood was identified only when blood/semen was 100:80% and 80:20%.

Two previously reported markers (cg17610929 for semen and cg09765089 for vaginal fluid) along with six novel body fluid-specific markers were studied by Lin et al. [62] for forensic body fluid identification. The authors selected the markers from public datasets of the HumanMethylation 450 K, namely the GPL13534 platform in NCBI Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/geo>). The markers were selected with a criteria of a high methylation value in target fluids and nearly zero values in other fluids (Table 1). The group developed a 10-plex Methylation-Specific PCR assay which included cg09107912 and cg16732616 for saliva, cg24124443 and cg01607849 for blood, cg17610929 and cg05261336 for semen, cg09765089 and cg25416153 for vaginal fluid, along with a bisulphite conversion control (beta-actin) and a gender-specific marker (Amelogenin). The markers were tested against 65 samples of saliva, blood, semen, vaginal fluid and menstrual blood. A MS-SNuPE

reaction was used to analyse the multiplex products, and each marker showed high methylation in target body fluids while low methylation levels were found in other fluids. However, menstrual blood samples also showed methylation peaks at both blood and vaginal fluid markers, indicating that menstrual blood was composed of those fluids.

Forensic simulation and mixture analysis by Lin et al. [62] A novel aspect of this study was analysis of actual forensic case-work samples (fabrics cut from underwear, swabs, tissues) and thereafter comparing the results obtained from conventional catalytic, immunological and microscopic tests. While the conventional tests gave ambiguous results, the 10-plex DNA methylation-based assay enabled accurate identification of all samples present on fabric from undergarments, swabs and tissue papers. Upon mixture analysis, when mixed in equal quantities, each marker enabled accurate identification of respective target fluids. In fact, semen was even detected when present at only 1.89% concentration in mixtures, reinforcing the reliability of the technique.

Forat and colleagues [73] performed two extensive genome-wide methylation profiling using the HumanMethylation 27 K and 450 K arrays. The group reported differentially methylated sites specific to saliva, blood, semen, vaginal fluid and menstrual blood. The description of each marker and their methylation status is described in Table 1. These researchers used both bisulphite sequencing and MS-SNuPE to analyse methylation levels of the markers. cg21597595-4d and cg15227982-137d were both hypermethylated in saliva; cg26285698-14d and cg26285698-39d were hypomethylated in blood, whereas cg03363565-59d was hypermethylated in blood; cg22407458-288d and cg05656364-283d were hypomethylated and hypermethylated in semen, respectively; cg14991487-85d and cg03874199-212d were both hypermethylated in vaginal fluid; and cg09696411-25d and cg09696411-40d were both hypermethylated in menstrual blood. Blood and menstrual blood exhibited similar methylation profiles for the cg03363565-59d marker, and methylation differences were rather small between the menstrual blood markers and vaginal fluid markers for the target fluids. Evidently, venous blood, menstrual blood and vaginal fluid are not always well discriminated [59, 62]. Similar to Lee et al. [59], the authors credited this to incorrect sampling on tampons and time of collection; since samples were collected on days 1–5, whereas most women regard day 2 as strongest bleeding day and day 5 is usually disregarded. A step-wise analysis of body fluids in question was recommended, beginning with cg09696411 for menstrual blood and then using other markers to confirm the presence of menstrual blood. Methylation levels found by the MS-SNuPE assay were compared to standard bisulphite sequencing, Roche 454 sequencing and Illumina MiSeq-sequencing [73]. The group found

that all the next -generation sequencing (NGS) platforms yielded the same results as the MS-SNuPE assay.

Forensic simulation and disease analysis by Forat et al. [73]

During a forensic simulation, a total of 75 body fluids were left in dry (left at room temperature), humid (in an exicator) and outdoors (on ground) conditions. The methylation levels of the body fluid-specific CpGs were determined several times over a 6-month period. All methylation levels were found to be unaffected in dry conditions whereas humid conditions in the exicator seemed to affect methylation levels the most. Still, humidity-related changes did not affect the discriminatory power of the markers in target body fluids. The markers also facilitated detection of target body fluids when the target body fluids comprised a minimum of 20% or more of mixtures.

A previous study had shown that methylation levels of a tumour cell is decreased to 40–60% of the healthy cell [82]; thus, the potential effect of some common tumours on methylation levels of the loci was tested [73]. Some of the markers indeed showed variation; in cervix carcinoma samples, both vaginal fluid markers showed decreased methylation values in vaginal fluid, when compared to healthy controls. This decreased methylation resulted in overlapping of methylation levels with the menstrual blood marker (cg09696411). Aberrant methylation due to disease must be considered when using this marker as false negatives may greatly compromise the test. Additionally, methylation of the semen markers also fluctuated in blood obtained from patients with chronic lymphocytic leukaemia (CLL), chronic myelocytic leukaemia and myeloproliferative syndrome, and one CLL sample showed a slightly increased methylation signal with the menstrual blood marker (cg09696411). Zhang et al. [83] previously reported that methylation levels of loci may be controlled or altered by genetic variants; thus, a NGS experiment was performed to determine if genetic factors would affect the methylation levels of the candidate CpGs [73]. Briefly, library preparation was completed using the NEBNext™ Ultra DNA Library from Illumina. Five PCRs were performed using equal amounts of pooled bisulphite converted DNA from 10 participants. Following PCR, the amplicons were purified and used as input for the library preparation. NGS analysis was performed on the MiSeq instrument (Illumina) using 2 × 250-bp paired end sequencing [73]. A mutation del-91 (Chr 12:57619697, hg38) that was found in only 2.6% of menstrual blood reads was shown to decrease the methylation level of the menstrual blood marker (cg09696411). Additionally, a sequence variant G>A 167 (Chr 10:102776279, hg38) found in 14.7% of reads decreased methylation of cg15227982-137d (saliva-specific). Such reduction in methylation levels due to the genetic variants would greatly reduce the sensitivity of the test. Finally, a sequence variant T>C 38 (Chr 2:5366146, hg38) which occurred in 4% of reads increased the methylation level of cg21597595-4d (saliva-specific); however, since

this was a hypermethylation marker anyway, the results would not be distorted. This comprehensive research study [73] highlighted the importance of considering the effects of external factors, disease, genetic factors and different methods when researching differential methylation for body fluid identification.

Vidaki and colleagues [76] obtained methylation data for buccal cells, blood and semen from two previous studies [48, 84] and selected 11 potential body fluid-specific differentially methylated CpG sites. These sites were analysed against saliva, blood, semen, vaginal fluid, menstrual blood, buccal cells, skin and urine. The study included four markers for buccal cells, three markers for blood and four markers for semen (Table 1). Markers for buccal cells were inconclusive; however, upon testing for blood specificity, the cg13763232 site showed potential as a hypermethylation marker. The site was methylated only in blood, while in other fluids, a partial methylation or unmethylation was observed. In fact, three adjacent CpGs showed blood-specific hypermethylation. One caveat was that skin samples did also show a wide range of methylation (0.13–0.94); thus, the locus cannot be used confidently. More rigorous testing of the cg13763232 marker with a larger sample size is necessary. For identification of semen, only hypomethylation markers were found to be highly semen-specific; in fact, nine sites around cg04382920 and four sites around cg11768416 showed semen-specific hypomethylation.

Sensitivity and analysis of aged samples by Vidaki et al. [76]

Two markers (cg04382920 and cg11768416) underwent extensive validation using a sensitivity analysis and testing against aged samples. The markers were shown to provide stable methylation results even when 0.05 ng starting DNA was used in a mixture with blood. Additionally, no false negatives, outliers or significant differences in methylation were detected when the markers were tested on semen samples that were stored at – 20 °C for almost 16 years, as well as blood stains on fabric that was kept at ambient temperature for more than 20 years [76].

Holtkotter and colleagues [61] employed a unique approach—the group first performed a literature search to identify most promising methylation-based markers for saliva, blood, semen and menstrual blood. A total of 13 markers were chosen from Lee et al. [69], An et al. [55], Park et al. [72] and Lee et al. [59]. Only four markers, cg09652652-2d (saliva-specific), cg06379435 (blood-specific), cg26763284-138d (semen-specific) and cg26079753-7d (menstrual blood-specific), were subsequently used in a multiplex SNaPshot (MS-SNuPE) assay wherein each marker successfully distinguished the respective target body fluid. However, for the blood marker (cg06379435), semen also showed a partial methylation profile. It was therefore suggested to use the blood marker in combination with the semen marker

(cg26763284-138d), since semen would display much higher methylation levels.

Mixture analysis and forensic simulation by Holtkotter et al. [61] During mixture analysis, each marker was able to distinguish the target fluid. The saliva-specific marker cg09652652-2d showed similar methylation profiles in saliva and buccal cells; thus, differentiation between those fluids was not possible. The multiplex assay was also applied to mock crime scene samples, which were created by mixing various ratios of fluids and then applying to cellulose swabs. Once again, the markers successfully discriminated each respective body fluid [61]. It is indeed advantageous that methylation of several CpG sites in each region was analysed in the bisulphite PCR reaction as this makes the markers more reliable than relying on a single CpG site. However, the SNaPshot primers are designed to terminate one base pair upstream of the target cytosine which means that the method facilitates analysis of only one target CpG site [85]. It is definitely not ideal to rely solely on one CpG site. Furthermore, while these four markers were shown to be robust and reliable in a multiplex assay, the location of the markers was not given much consideration. Both saliva and semen-specific markers were located in exons and these regions are said to be more prone to mutations [86]. The blood and menstrual blood-specific markers were located in 5' and 3' regions, respectively. Untranslated regions play an essential role in health and disease. These regions house upstream open reading frames, internal ribosome entry sites, are GC rich and influence the rate of translation. Mutations in these regions are fairly common, and pivotal in the onset of diseases such as Alzheimer's disease, breast cancer, bipolar affective disorder, congenital heart disease, among others [87, 88]. Whether sequence alterations in the UTRs of the blood and menstrual blood-specific markers occur frequently and/or result in aberrant protein expression requires additional studies; however, these factors must be considered prior to validating them.

The location of these tDMRs in the human genome, related CGIs and genes as well as the number of samples used have been summarised in Table 2.

Considerations for development of DNA methylation-based markers for application in forensic science

Assessing the stability of DNA methylation markers

For a marker to be deemed reliable and confidently applied in routine forensic casework, rigorous testing is crucial. While it is true that some researchers do attempt to validate the markers using simulated forensic conditions (discussed below), it remains to be seen if any marker will fulfil such high requirements.

The chromosomal location of a potential marker is a first and foremost factor. Ideally, a body fluid-specific CpG site or tDMR should not be located in a region with a high mutation frequency, aptly coined 'a mutation hotspot' [89]. Mutation rates at any given region are a reflection of the stability and sensitivity of nucleotides to mutagenic agents, and the fidelity and efficiency of DNA replication and repair machinery [90, 91]. Studies show that mutation/substitution rates at A/T base pairs are between 25 and 85% lower than G/C base pairs (excluding CpG sites) [92, 93]. Specifically of interest would be that although primarily in single-stranded DNA, a cytosine followed by guanine is approximately 10 times more mutable than a cytosine followed by adenine or thymine. This hypermutability is due to spontaneous deamination of methylcytosines into thymine [91, 94]. Thus, the rate of CpG mutation is a function of the rate of DNA melting, which in turn is affected by the local base composition wherein G/C base pairs are stronger than A/T pairs [95, 96]. There have been numerous surveys of the human genome to determine which regions undergo mutation and thereby facilitate evolutionary divergence. These studies show that synonymous sites in exons and protein coding genes produce considerably higher estimates of evolutionary divergence when compared to introns and pseudogenes [86, 97–100]. Additionally, with the exception of regulatory sites, intergenic DNA also shows lower evolutionary divergences than those seen in coding sequences [86, 101–103]. Currently, most body fluid-specific markers for forensic use are located near or within exons (Tables 1 and 2). Exons are known to harbour numerous polymorphisms [104]; in fact, Ng et al. [105] examined a single human exome (exons in the genome) and found approximately 12,500 coding variants that can affect gene expression and protein function.

Considering the above, it would be beneficial that a DNA methylation-based body fluid-specific marker be located in a region containing more A/T pairs than C/G pairs, in an intron, pseudogene or intergenic region of DNA, but preferably not in an exon. With cytosine itself being more mutable than any other base [86, 91], it is compulsory to investigate whether the regions within and neighbouring the tDMR/marker are prone to mutation.

Preferably, a marker should facilitate absolute unambiguous identification of the target body fluid based on the differential methylation levels between the fluids. Since body fluids are generally recovered in mixtures, it is ideal that a marker shows only a complete on/off methylation status. This means that the marker should ideally exhibit complete methylation (> 90%) in the target fluid and unmethylation (< 10%) in non-target fluids, or vice versa [62, 76]. This should hold true even if the DNA recovered from body fluids is of a degraded/contaminated nature or is present in varying concentrations and mixtures. The ability of the marker to identify the body fluid should not falter using various methods to analyse the

methylation levels, or with factors known to affect DNA methylation, such as genetic variation [106], diets, lifestyle, environment [24], age [21, 107], disease [16, 108, 109], ethnicity [6, 107, 110, 111] and smoking [25, 112], to name a few. However, to the best of our knowledge, no DNA methylation-based marker for forensic use has undergone such extensive research which considers all the above stipulations.

Reproducibility of methylation patterns using various DNA methylation-based assays

Exposure of samples to the environment for a certain number of days to determine effects of time and age, mixture analysis as well as sensitivity tests to determine minimum concentration of DNA required are rather standard and common aspects in the above studies [55, 56, 59, 72, 75–77]. However, another essential aspect is to test whether methylation results are reproducible using different DNA methylation-based assays. Reed et al. [113] compared methylation levels generated by bisulphite PCR and sequencing (BSP) with methylation levels obtained from pyrosequencing. In this report, higher variation of methylation levels was obtained from BSP possibly due to the bacterial cloning step, whereas this step is needless for pyrosequencing. Each method will present its own advantages and limitations, and it is to be expected that the use of numerous methods to measure DNA methylation will not present the exact same results [114]. As an example, the vaginal fluid-specific marker cg09765089 was examined by Lee et al. [59] on 11 samples using bisulphite sequencing and multiplex MS-SNuPE. This study found approximately 35% methylation in vaginal fluid and 37% methylation in menstrual blood. Lin et al. [62] studied the same marker on 10 samples using methylation-specific PCR combined with multiplex MS-SNuPE and found approximately 39.6% methylation in vaginal fluids. However, this study did not test menstrual blood samples.

Silva et al. [63] performed analysis of body fluid-specific markers in another laboratory to assess the reliability and production of similar results between different laboratories. Their results were indeed reproduced in another facility for the cg06379435 (blood-specific), BCAS4 (saliva-specific) and ZC3H12D (semen-specific) markers. Depending on the extent of the variation in methylation levels using different methods, equal sample types and numbers, data normalization, base alignments, corrections and quality controls might alleviate these errors to a certain degree [57, 115].

Additionally, while some markers have been repeatedly studied, they are investigated by the same groups of researchers using the same methods. These include cg09652652-2d (saliva-specific), cg06379435 (blood-specific) and cg17610929 (semen-specific) which were studied by Lee et al. [59] and Lee et al. [60] using targeted bisulphite sequencing and methylation SNaPshot as well as

ZC3H12D and FGF7 (semen-specific), C20orf117 (blood-specific) and BCAS4 (saliva-specific) markers which have been repeatedly studied by Madi et al. [67], Silva et al. [63] and Antunes et al. [68] by bisulphite pyrosequencing. It is imperative that more than one method be used in a single study to prove that regardless of the method of analyses, the methylation results of a particular marker in a specific body fluid will be reproducible; thus, the marker will not waver in accuracy to identify the target fluid. From the above studies, only Forat et al. [73] reproduced results of their MS-SNuPE assay using bisulphite sequencing, pyrosequencing and the MiSeq platform by Illumina.

Selecting regions of DNA that house multiple CpG sites

Numerous studies using the Illumina HumanMethylation arrays find potential CpG sites at which methylation may be specific to a body fluid (Table 1). However, ideally, a marker comprising methylation patterns at more than one CpG site that are found to exhibit body fluid-specific methylation patterns would be sought after. In fact, upon experimentation, Lee et al. [59], Lee et al. [60] and Forat et al. [73] who found that the body fluid-specific CpG sites were not those identified by the Illumina array, but rather a few bases downstream (for example; cg09652652-2d which is saliva-specific; and cg09765089-231d and cg26079753-7d which are vaginal fluid-specific). Instead of only the single CpG sites, perhaps the entire region (a few hundred base pairs) may facilitate body fluid identification. Additionally, mixture analysis by Silva et al. [63] and Watanabe et al. [75] was not successful using single CpG sites. Placing confidence on a region/locus as opposed to a single CpG site would surely be more desirable in forensic casework. Only few studies have considered this; Park et al. [72] assessed between two and four CpGs around the eight markers that were analysed and found them to also be hypermethylated in target fluids. Vidaki et al. [76] also identified four and nine neighbouring CpGs in two semen-specific hypomethylated markers (cg11768416 and cg04382920) and proved that the regions around the sites were also specific. However, despite Antunes et al. [81] examining regions containing five CpG sites, this group also struggled with deconvoluting mixtures.

Assessing environmental influences and disease-related changes of methylation patterns in tDMRs

Forat and colleagues [73] examined environmental effects such as dried, humid and wet conditions on DNA methylation levels of markers; these researchers showed that humidity did indeed influence methylation levels of their menstrual blood marker (cg09696411), vaginal fluid marker (cg14991487) and

Table 2 Summary of previously reported body fluid-specific tDMRs, chromosomal locations, sample sizes and methods of analysis

Chromosome location/region ^a	Gene/locus	Nearest CGI according to UCSC ^b	Specific genomic locations according to Ensembl (GRCh37)	Saliva (n)	Blood (n)	Semen (n)	Vaginal fluid (n)	Menstrual blood (n)	Buccal cells (n)
Chr20:48844209-48844366 (hg18)	<i>BCAS4</i>	Upstream of CGI 92	CGI upstream of gene <i>BCAS4</i>	11	10	10	-	-	-
Chr20:35491003-35491032 (hg18)	<i>C20orf117 / SOGAI</i>	CpG 209	Exon ENSE00000800452	11	10	10	-	-	-
Chr22:29708324, Chr22:29708344, Chr22:29708365	<i>GAS2L1</i>	-	Exon ENSE00001551291	15	15	15	3	3	-
Chr15:49714275-49714404	<i>FGF7</i>	-	5' upstream sequence before exon ENSE00002572210	11	10	9	-	-	-
Chr6:149778103-149778127	<i>ZC3H12D</i>	CpG 41	Intron 3-4 before exon ENSE00000813607	11	10	10	-	-	-
Chr21:46905841-46906149 (hg18)	<i>PRMT2</i>	CpG 55	Primarily located in intron 9/10 but covers boundary of exon ENSE00001051356	-	10	7	-	-	9
Chr14:58182690-58182995 (hg18)	<i>DACT1</i>	CpG 50	Within Exon 4 ENSE00001352768	15	15	15	3	3	-
Chr6:41881884-41882111 (hg18)	<i>USP49</i>	CpG 46	Exon 4 ENSE00001086408	ng	ng	ng	ng	-	-
Chr19:47395603-47395731	<i>L81528/DEDD2</i>	CpG 27	Exon 5 ENSE00001358691	ng	ng	ng	ng	-	-
Chr5:176759049-176760021 (hg18)	<i>PFN3</i>	CpG 82	Located 5'upstream, overlaps exon ENSE00001410937	46	46	35	6	6	-
Chr6:41604585, Chr6:41604602, Chr6:41604604, Chr6:41604614	<i>MDF1</i>	CpG 168	5' UTR	20	8	12	10	10	7

Chromosome location/region ^a	Urine (n)	Target body fluid	Methylation status	Method of analysis	Reference
Chr20:48844209-48844366 (hg18)	11	Saliva	High	A	Madi et al. [67]
	9	Saliva	High	A	Antunes et al. [68]
	-	Saliva	High	A	Silva et al. [63]
Chr20:35491003-35491032 (hg18)	1	Blood	High	A	Madi et al. [67]
Chr22:29708324, Chr22:29708344, Chr22:29708365	-	Blood	High	A	Fu et al. [74]
Chr15:49714275-49714404	8	Semen	High	A	Madi et al. [67]
	8	Semen	High	A	Antunes et al. [68]
Chr6:149778103-149778127	8	Semen	Low	A	Madi et al. [67]
	8	Semen	Low	A	Antunes et al. [68]
	8	Semen	Low	A	Antunes et al. [77]
	-	Semen	Low	G	Antunes et al. [63]
	-	Semen	Low	A	Silva et al. [63]
Chr21:46905841-46906149 (hg18)	-	Semen	Low	B, D	Lee et al. [69]
Chr14:58182690-58182995 (hg18)	-	Semen	Low	B, D	Lee et al. [69]

Table 2 (continued)

Chr6: 41881884–41882111 (hg18)	-	-	Semen	Low	C, F	An et al. [55]
	-	-	Semen	Low	F	Choi et al. [56]
	-	-	Semen	Low	B, D	Lee et al. [69]
	-	-	Semen	Low	C, F	An et al. [55]
	-	-	Semen	Low	F	Choi et al. [56]
Chr19: 47395603–47395731	-	-	Semen	High	F	Choi et al. [56]
Chr5: 176759649–176760021 (hg18)	-	-	Vaginal fluid	Low	B, D	Lee et al. [69]
	-	-	Vaginal fluid	Low	B, F	An et al. [55]
	-	-	Vaginal fluid	Low	F	Choi et al. [56]
Chr6: 41604585, Chr6: 41604602, Chr6: 41604604, Chr6: 41604614	-	-	Vaginal fluid	Low	A	Antunes et al. [81]
	-	-	Menstrual blood	Low	A	Fu et al. [74]

Note: Illumina ID is a unique ID in the Illumina Human Methylation 450K bead array

ng not given

^a hg19 unless otherwise stated; method of analysis: (a) bisulphite pyrosequencing, (b) methylation SNaPshot, (c) bisulphite sequencing, (d) methylation-specific PCR, (e) methylation-sensitive restriction enzyme PCR, (f) high-resolution melt analysis

saliva marker (cg21597595). The concepts of testing effects of inhibitors such as humic acid which can be found freely in soil and influences of non-human DNA were only tested in two studies by the same group [63, 81]. Disease is known to cause aberrant methylation patterns in humans [15, 108, 109]; yet, only Forat and colleagues [73] have compared methylation levels between healthy and cancer-inflicted individuals, wherein cervix carcinoma was shown to affect methylation levels in vaginal fluid. No other study in forensics has considered that individuals with disease will exhibit different methylation levels when compared to healthy individuals, and this would drastically reduce the reliability of body fluid-specific markers.

Genetic variation

Genetic variation such as single-nucleotide polymorphisms (SNPs) is known to influence methylation levels to some degree [83, 116, 117], but this impact is seldom investigated in forensic-based research. Watanabe et al. [75] failed to distinguish between body fluids in mixtures, and therefore resorted to searching for nearby SNPs. However, the group did not test the effects of polymorphisms on methylation, but merely used it as complementary method. One study, Forat et al. [73], indeed found that a mutation (del-91, Chr 12:57619697, hg38) reduced methylation levels of a menstrual blood marker, and a sequence variant (T>C 38, Chr 2:5366146, hg38) increased methylation levels of a saliva-specific marker. Such findings prove that SNPs do affect methylation and are certainly a factor to be reckoned with. Even though Forat et al. [73] stated that there is no reported SNP that confirmedly alters methylation of CpG sites found by the Illumina array, if a potential marker is located in a region of a gene that is highly variable or undergoes mutation frequently, the marker cannot be relied on. However, it must be noted that some SNPs can also influence the expression of remote genes located at a distance, instead of the expressions of the genes that actually harbour them [118, 119]. Furthermore, SNPs affect heritability of methylation sites, which is also a factor that affects the stability of DNA methylation-based markers, as elaborated next.

Heritability

Heritability of DNA methylation is the proportion of variance explained by additive genetic factors. In humans, low ‘epigenetic heritability’ was reported by Gervin et al. [120]. These researchers obtained blood from 49 monozygotic (MZ) twin pairs and 40 dizygotic (DZ) twin pairs to investigate inter-individual variation and heritable patterns of DNA methylation. CpG islands (CGIs) and 5′ regions exhibited low methylation, while conserved noncoding regions had intermediate methylation and randomly selected CpG sites within the major

histocompatibility complex (but outside the CGIs, 5' and non-coding regions) showed high methylation levels. Lower within-pair differences in DNA methylation were observed for MZ twins than that of DZ twins, in the conserved noncoding regions and randomly selected CpG sites compared to CGIs and 5' regions. The authors stated that low heritability levels of between 2 and 16% were observed across the four distinct regions [120]. Another twin-based study by Bell and colleagues [121] examined blood samples from 172 female twins (43 DZ pairs, 33 MZ pairs and 20 singletons) at 26,690 promoter CpG sites and found a mean genome-wide heritability of 18%.

For forensic use of a tDMR-based marker, the heritability of methylation at that marker must be thoroughly investigated and confirmed. This was certainly not the case in two studies investigating a CpG site in the promoter of *AXL* receptor tyrosine kinase by Boks et al. [122] and Breton et al. [123]. Boks and colleagues [122] examined whole blood obtained from 23 MZ twin pairs and 23 DZ twin pairs to investigate heritability. Twenty-three percent of sites analysed showed significant heritability, with most significant heritability (0.94) observed at sites in *AXL*, amongst other genes. However, contradictory results were obtained when the same site was assessed by Breton and colleagues [123]. This study, however, examined a different body fluid, buccal cells which were obtained from 16 MZ twins and 20 DZ twins. Such large discrepancies of heritability were attributed to different body fluids assessed, underlying genetic distributions between individuals as well as polymorphic imprinting, wherein in some individuals, *AXL* may be mono-allelically expressed whereas in others, it would be biallelically expressed [122, 123]. Even though these studies were not based on forensic applications of DNA methylation, they portrayed important considerations to avoid discrepancies when investigating heritability of particular CpG sites.

Several studies have shown that *cis*-acting (closely located) genetic variation such as SNPs can account for a large proportion of variation in DNA methylation [83, 116, 124]. The true extent of this influence is not known, but is dependent on the individual, tissue, location and functional genomic context of the CpG site. Quon and colleagues [124] assessed heritability of methylation levels of 21,000 CpG sites in four regions of the brain: the cerebellum, frontal cortex, caudal pons and temporal cortex from 150 unrelated individuals. This research identified 636, 654, 600 and 812 heritable DNA methylation sites/loci in the cerebellum, frontal cortex, caudal pons and temporal cortex, respectively. These heritable sites were found to be enriched in open chromatin regions and known binding sites of transcription factors, which infers functional roles of some of the sites. Heritable methylation sites were found to have a high number of SNPs within a 50-kb window than non-heritable methylation loci, suggesting that the higher the number of SNPs in the region, the more likely it was to find heritable methylation loci. The estimated heritability of all

methylation loci that were thought to be heritable across all brain regions was nearly 30%; whereas heritability across loci including those which were not thought to be heritable was less than 3%. Across all four regions, 181 loci were heritable, whereas 207 loci were heritable across at least three of the brain regions [125].

Rowlatt et al. [126] assessed over 196,000 CpG sites in healthy colorectal tissue samples obtained from Colombian participants to examine phenotypic profiles, genetic effects and regional genomic heritability. This research found that CpG sites located in regions of low CpG content exhibited great variation, higher methylation and were more likely to be heritable than compared to CpG sites located in CpG-rich regions. CpG sites located in intergenic regions displayed higher methylation levels and were more likely to be heritable than those in transcription start sites or intragenic regions. The group also found that genetic variants in genomic risk regions for colorectal cancer can also affect methylation levels in healthy tissues. For example, methylation levels of cg15193198 and cg24112000 were affected by SNPs rs4925386. Both of these methylation sites were found to be heritable; however, heritability decreased considerably due to the presence of the SNP [126].

Thus, upon investigating a DNA methylation-based marker for body fluid identification, thorough investigation must be performed to ensure that methylation of the marker is heritable across generations. Alternatively, it might be wise to select a marker based in a region which is epigenetically stable, even throughout environmental and nutritional fluctuations.

Finally, other vital yet overlooked aspects regarding DNA methylation-based markers are the impact of diets, lifestyle and ethnicity on the stability of the markers. No study based on DNA methylation-based markers for forensic body fluid identification has considered the impact of varying diets, different socio-economic rates, cultures and geographic locations, smoking, drugs, alcohol or ethnic backgrounds on the methylation status of the markers. These factors have repeatedly been shown to alter methylation levels [6, 22, 25, 107, 112].

Future outlook

An all-encompassing study would be one (or preferably a few collaborators in various countries) which selects a panel of potential body fluid-specific markers and performs methylation analysis using several different techniques which will demonstrate the reproducibility of the markers. The tests should be performed on a large number of samples of different fluids, which ought to be obtained from individuals of diverse age groups, ethnicities and geographic locations, healthy individuals as well as those afflicted with various diseases, medicated and non-medicated, smokers and non-smokers,

alcohol-drinkers and non-alcohol drinkers, drug-users and non-drug users. These samples must be corrected for cell type heterogeneity if need be. SNP-mapping as well as determining whether the methylation levels of the particular sites are heritable are also crucial steps to be taken. Heritability of DNA methylation patterns at certain regions requires longitudinal studies, since this is a measure of methylation over several generations. The standard mixture, sensitivity and specificity tests must certainly not be neglected. And finally, once the markers have been completely confirmed to identify body fluids, they should be used in age prediction models to determine efficacy in age estimation, as is currently being widely explored [127–130].

Such large-scale study designs and concomitant rigorous interrogation of the markers would be essential to standardise markers, methods and techniques which is presently done for SNP and STR markers. Once the currently identified DNA methylation-based markers undergo such investigation, it would be conceivable that a dedicated forensic database for DNA-methylation-based markers be developed. Reverting to this database will in turn enable us to place as much confidence on differential DNA methylation-based body fluid identification as is placed currently on databases such as the Y-STR database and Combined DNA Index System (CODIS). Ensuing this intense research, we see much potential for routine application of differential DNA methylation-based body fluid identification in forensic casework.

Funding information This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

- Nustad HE, Almeida M, Canty AJ, LeBlanc M, Page CM, Melton PE (2018) Epigenetics, heritability and longitudinal analysis. *BMC Genet* 19(Suppl 1):77. <https://doi.org/10.1186/s12863-018-0648-1>
- Lalruatfela B (2013) On DNA methylation: An introductory review. *Science Vision* 13:1-7. ISSN (online):2229-6026.
- Ghosh S, Yates AJ, Frühwald MC, Miecznikowski JC, Plass C, Smiraglia D (2010) Tissue specific DNA methylation of CpG islands in normal human adult somatic tissues distinguishes neural from non-neural tissues. *Epigenetics* 5(6):527–538. <https://doi.org/10.4161/epi.5.6.12228>
- Pinney SE (2014) Mammalian Non-CpG Methylation: Stem Cells and Beyond. *Biology (Basel)* 3(4):739–751. <https://doi.org/10.3390/biology3040739>
- Yan J, Zierath JR, Barrès R (2011) Evidence for non-CpG methylation in mammals. *Exp Cell Res* 317(18):2555–2561. <https://doi.org/10.1016/j.yexcr.2011.08.019>
- Kader F, Ghai M (2016) DNA methylation-based variation between human populations. *Mol Genet Genomics* 292(1):5–35. <https://doi.org/10.1007/s00438-016-1264-2>
- Rienius LE, Acevedo N, Joerink M, Pershagen G, Dahlén SE, Greco D et al (2012) Differential DNA methylation in purified human blood cells: implications for cell lineage and studies on disease susceptibility. *PLoS One* 7(7):e41361. <https://doi.org/10.1371/journal.pone.0041361>
- Tammen SA, Friso S, Choi SW (2013) Epigenetics: the link between nature and nurture. *Mol Aspects Med* 34(4):753–764. <https://doi.org/10.1016/j.mam.2012.07.018>
- Bestor TH, Edwards JR, Boulard M (2015) Notes on the role of dynamic DNA methylation in mammalian development. *Proc Natl Acad Sci USA* 112(22):6796–6799. <https://doi.org/10.1073/pnas.1415301111>
- Lövkvist C, Dodd IB, Sneppen K, Haerter JO (2016) DNA methylation in human epigenomes depends on local topology of CpG sites. *Nucleic Acids Res* 44(11):5123–5132. <https://doi.org/10.1093/nar/gkw124>
- Mokarram P, Kumar K, Brim H, Naghibalhossaini F, Saberi-Firoozi M, Nouraie M et al (2009) Distinct high-profile methylated genes in colorectal cancer. *PLoS One* 4:e7012. <https://doi.org/10.1371/journal.pone.0007012>
- Wiley KL, Treadwell E, Manigaba K, Word B, Lyn-Cook BD (2013) Ethnic differences in DNA methyltransferases expression in patients with systemic lupus erythematosus. *J Clin Immunol* 33(2):342–348. <https://doi.org/10.1007/s10875-012-9803-z>
- Chambers JC, Loh M, Lehne B, Drong A, Kriebel J, Motta V, Wahl S, Elliott HR, Rota F, Scott WR, Zhang W, Tan ST, Campanella G, Chadeau-Hyam M, Yengo L, Richmond RC, Adamowicz-Brice M, Afzal U, Bozaoglu K, Mok ZY, Ng HK, Pattou F, Prokisch H, Rozario MA, Tarantini L, Abbott J, Ala-Korpela M, Albeti B, Ammerpohl O, Bertazzi PA, Blancher C, Caiazzo R, Danesh J, Gaunt TR, de Lusignan S, Gieger C, Illig T, Jha S, Jones S, Jowett J, Kangas AJ, Kasturiratne A, Kato N, Kotea N, Kowlessur S, Pitkaniemi J, Punjabi P, Saleheen D, Schafmayer C, Soininen P, Tai ES, Thorand B, Tuomilehto J, Wickremasinghe AR, Kyrtopoulos SA, Aitman TJ, Herder C, Hampe J, Cauchi S, Relton CL, Froguel P, Soong R, Vineis P, Jarvelin MR, Scott J, Grallert H, Bollati V, Elliott P, McCarthy M, Kooner JS (2015) Epigenome-wide association of DNA methylation markers in peripheral blood from Indian Asians and Europeans with incident type 2 diabetes: a nested case-control study. *Lancet Diabetes Endocrinol* 3(7):526–534. [https://doi.org/10.1016/S2213-8587\(15\)00127-8](https://doi.org/10.1016/S2213-8587(15)00127-8)
- Lu H, Liu X, Deng Y, Qing H (2013) DNA methylation, a hand behind neurodegenerative diseases. *Front. Aging Neurosci* 5:85. <https://doi.org/10.3389/fnagi.2013.00085>
- Kader F, Ghai M, Maharaj L (2018) The effects of DNA methylation on human psychology. *Behav Brain Res* 346:47–65. <https://doi.org/10.1016/j.bbr.2017.12.004>
- Morgan HD, Santos F, Green K, Dean W, Reik W (2005) Epigenetic reprogramming in mammals. *Hum Mol Genet* 14(1):R47–R58. <https://doi.org/10.1093/hmg/ddi114>
- Ciccarone F, Tagliatesta S, Caiafa P, Zampieri M (2018) DNA methylation dynamics in aging: how far are we from understanding the mechanisms? *Mech Ageing Dev* 174:3–17. <https://doi.org/10.1016/j.mad.2017.12.002>
- Messerschmidt DM, Knowles BB, Solter D (2014) DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos. *Genes Dev* 28:812–828. <https://doi.org/10.1101/gad.234294.113>
- Schmitz RJ, Schultz MD, Lewsey MG, O'Malley RC, Urich MA, Libiger O, Schork NJ, Ecker JR (2011) Transgenerational epigenetic instability is a source of novel methylation variants. *Science* 334(6054):369–373. <https://doi.org/10.1126/science.1212959>

20. Stricker SH, Götz M (2018) DNA-Methylation: Master or Slave of Neural Fate Decisions? *Front Neuro* 12:5. <https://doi.org/10.3389/fnins.2018.00005>
21. van Dongen J, Nivard MG, Willemsen G, Hottenga JJ, Helmer Q, Dolan CV, Ehli EA, Davies GE, van Iterson M, Breeze CE, Beck S, BIOS Consortium, Suchiman HE, Jansen R, van Meurs J, Heijmans BT, Slagboom PE, Boomsma DI (2016) Genetic and environmental influences interact with age and sex in shaping the human methylome. *Nat Commun* 7:11115. <https://doi.org/10.1038/ncomms11115>
22. King-Batoon A, Leszczynska JM, Klein CB (2008) Modulation of gene methylation by genistein or lycopene in breast cancer cells. *Environ Mol Mutagen* 49:36–45. <https://doi.org/10.1002/em.20363>
23. Perrier F, Viallon V, Ambatipudi S, Ghantous A, Cuenin C, Hernandez-Vargas H, Chajès V, Baglietto L, Matejčić M, Moreno-Macias H, Kühn T, Boeing H, Karakatsani A, Kotanidou A, Trichopoulou A, Sieri S, Panico S, Fasanelli F, Dolle M, Onland-Moret C, Sluijs I, Weiderpass E, Quirós JR, Agudo A, Huerta JM, Ardanaz E, Dorronsoro M, Tong TYN, Tsilidis K, Riboli E, Gunter MJ, Herceg Z, Ferrari P, Romieu I (2019) Association of leukocyte DNA methylation changes with dietary folate and alcohol intake in the EPIC study. *Clin Epigenetics* 11(1):57. <https://doi.org/10.1186/s13148-019-0637-x>
24. Alegria-Torres TAA, Baccarelli A, Bolati V (2011) Epigenetics and Lifestyle. *Epigenomics*. 3:267–277. <https://doi.org/10.2217/epi.11.22>
25. Elliott HR, Tillin T, McArdle WL, Ho K, Duggirala A, Frayling TM, Davey Smith G, Hughes AD, Chaturvedi N, Relton CL (2014) Differences in smoking associated DNA methylation patterns in South Asians and Europeans. *Clin Epigenetics* 6(1):4. <https://doi.org/10.1186/1868-7083-6-4>
26. Tsaprouni LG, Yang TP, Bell J, Dick KJ, Kanoni S, Nisbet J, Viñuela A, Grundberg E, Nelson CP, Meduri E, Buil A, Cambien F, Hengstenberg C, Erdmann J, Schunkert H, Goodall AH, Ouwehand WH, Dermitzakis E, Spector TD, Samani NJ, Deloukas P (2014) Cigarette smoking reduces DNA methylation levels at multiple genomic loci but the effect is partially reversible upon cessation. *Epigenetics*. 9(10):1382–1396. <https://doi.org/10.4161/15592294.2014.969637>
27. Nielsen DA, Yuferov V, Hamon S, Jackson C, Ho A, Ott J, Kreek MJ (2009) Increased OPRM1 DNA methylation in lymphocytes of methadone-maintained former heroin addicts. *Neuropsychopharmacology* 34(4):867–873. <https://doi.org/10.1038/npp.2008.108>
28. Horvath S (2013) DNA methylation age of human tissues and cell types. *Genome Biol* 14:R115. <https://doi.org/10.1186/gb-2013-14-10-r115>
29. Phipson B, Oshlack A (2014) DiffVar: a new method for detecting differential variability with application to methylation in cancer and aging. *Genome Biol* 15(9):465. <https://doi.org/10.1186/s13059-014-0465-4>
30. Thompson RF, Atzmon G, Gheorghie C, Liang HQ, Lowes C, Grealley JM, Barzilai N (2010) Tissue-specific dysregulation of DNA methylation in aging. *Aging Cell* 9(4):506–518. <https://doi.org/10.1111/j.1474-9726.2010.00577.x>
31. Florath I, Butterbach K, Müller H, Bewerunge-Hudler M, Brenner H (2014) Cross-sectional and longitudinal changes in DNA methylation with age: an epigenome-wide analysis revealing over 60 novel age-associated CpG sites. *Hum Mol Genet* 23(5):1186–1201. <https://doi.org/10.1093/hmg/ddt531>
32. Day K, Waite LL, Thalacker-Mercer A, West A, Bamman MM, Brooks JD et al (2013) Differential DNA methylation with age displays both common and dynamic features across human tissues that are influenced by CpG landscape. *Genome Biol* 14(9):R102. <https://doi.org/10.1038/npp.2008.108>
33. Christensen BC, Houseman EA, Marsit CJ, Zheng S, Wrensch MR, Wiemels JL et al (2016) Ageing and environmental exposures alter tissue-specific DNA methylation dependent on CpG island context. *PLoS Genet* 5(8):e1000602. <https://doi.org/10.1371/journal.pgen.1000602>
34. Dmitrijeva M, Ossowski S, Serrano L, Schaefer MH (2018) Tissue-specific DNA methylation loss during ageing and carcinogenesis is linked to chromosome structure, replication timing and cell division rates. *Nucleic Acids Res* 46(14):7022–7039. <https://doi.org/10.1093/nar/gky498>
35. Kader F, Ghai M (2015) DNA Methylation and Application in Forensic Sciences. *Forensic Sci Int* 249:245–265. <https://doi.org/10.1016/j.forsciint.2015.01.037>
36. Prokhortchouk E, Defossez PA (2008) The cell biology of DNA methylation in mammals. *Biochim Biophys Acta* 1783(11):2167–2173. <https://doi.org/10.1016/j.bbamcr.2008.07.015>
37. Jjingo D, Conley AB, Yi SV, Lunyak VV, Jordan IK (2012) On the presence and role of human gene-body DNA methylation. *Oncotarget* 3(4):462–474. <https://doi.org/10.18632/oncotarget.497>
38. Illingworth RS, Bird AP (2009) CpG islands—a rough guide'. *FEBS Lett* 583(11):1713–1720. <https://doi.org/10.1016/j.febslet.2009.04.012>
39. Mendizabal I, Yi SV (2016) Whole-genome bisulfite sequencing maps from multiple human tissues reveal novel CpG islands associated with tissue-specific regulation. *Hum Mol Genet* 25(1):69–82. <https://doi.org/10.1093/hmg/ddv449>
40. Edgar R, Tan PP, Portales-Casamar E, Pavlidis P (2014) Meta-analysis of human methylomes reveals stably methylated sequences surrounding CpG islands associated with high gene expression. *Epigenetics Chromatin* 7(1):28. <https://doi.org/10.1186/1756-8935-7-28>
41. Sijen T (2014) Molecular approaches for forensic cell type identification: On mRNA, miRNA, DNA methylation and microbial markers. *Forensic Sci Int Genet* 18:21–32. <https://doi.org/10.1016/j.fsigen.2014.11.015>
42. Rechache NS, Wang Y, Stevenson HS, Killian JK, Edelman DC, Merino M et al (2012) DNA methylation profiling identifies global methylation differences and markers of adrenocortical tumors. *J Clin Endocrinol Metab* 97(6):E1004–E1013. <https://doi.org/10.1210/jc.2011-3298>
43. Lökk K, Modhukur V, Rajashekar B, Mårtens K, Mägi R, Kolde R, Koltšina M, Nilsson TK, Vilo J, Salumets A, Tõnisson N (2014) DNA methylome profiling of human tissues identifies global and tissue-specific methylation patterns. *Genome Biol* 15(4):r54. <https://doi.org/10.1186/gb-2014-15-4-r54>
44. Eckhardt F, Lewin J, Cortese R, Rakyan VK, Attwood J, Burger M, Burton J, Cox TV, Davies R, Down TA, Haefliger C, Horton R, Howe K, Jackson DK, Kunde J, Koenig C, Liddle J, Niblett D, Otto T, Pettett R, Seemann S, Thompson C, West T, Rogers J, Olek A, Berlin K, Beck S (2006) DNA methylation profiling of human chromosomes 6, 20 and 22. *Nature Genet* 38:1378–1385. <https://doi.org/10.1038/ng1909>
45. Igarashi J, Muroi S, Kawashima H, Wang X, Shinojima Y, Kitamura E, Oinuma T, Nemoto N, Song F, Ghosh S, Held WA, Nagase H (2008) Quantitative analysis of human tissue-specific differences in methylation. *Biochemical and Biophysical Research Communications*. 376:658–664. <https://doi.org/10.1016/j.bbrc.2008.09.044>
46. Cohen NM, Kenigsberg E, Tanay A (2011) Primate CpG Islands are maintained by Heterogeneous Evolutionary Regimes Involving Minimal Selection. *Cell*. 145:773–786. <https://doi.org/10.1016/j.cell.2011.04.024>
47. Ohgane J, Yagi S, Shiota K (2008) Epigenetics: the DNA methylation profile of tissue-dependent and differentially methylated

- regions in cells. *Placenta* 29(Suppl A):S29–S35. <https://doi.org/10.1016/j.placenta.2007.09.011>
48. Rakyan VK, Down TA, Thorne NP, Flicek P, Kulesha E, Graf S et al (2008) An integrated resource for genome-wide identification and analysis of human tissue-specific differentially methylated regions (tDMRs). *Genome Res* 18:1518–1529. <https://doi.org/10.1101/gr.077479.108>
 49. Ziller MJ, Gu H, Müller F, Donaghey J, Tsai LTY, Kohlbacher O, de Jager PL, Rosen ED, Bennett DA, Bernstein BE, Gnirke A, Meissner A (2013) Charting a dynamic DNA methylation landscape of the human genome. *Nature* 500(7463):477–481. <https://doi.org/10.1038/nature12433>
 50. Kayser M, de Knijff P (2011) Improving human forensics through advances in genetics, genomics and molecular biology. *Nat Rev Genet* 12(3):179–192. <https://doi.org/10.1038/nrg2952>
 51. Weyerman C, Ribaux O (2012) Situating forensic traces in time. *Sci Justice* 52:68–75. <https://doi.org/10.1016/j.scijus.2011.09.003>
 52. Virkler K, Lednev IK (2009) Analysis of body fluids for forensic purposes: From laboratory testing to non-destructive rapid confirmatory identification at a crime scene. *Forensic Sci Int* 188:1–17. <https://doi.org/10.1016/j.forsciint.2009.02.013>
 53. An JH, Shin KJ, Yang WI, Lee HY (2012) Body fluid identification in forensics. *BMB Rep* 45:545–553. <https://doi.org/10.5483/BMBRep.2012.45.10.206>
 54. Quinone I, Daniel B (2012) Cell free DNA as a component of forensic evidence recovered from touched surfaces. *Forensic Sci Int Genet* 6(1):26–30. <https://doi.org/10.1016/j.fsigen.2011.01.004>
 55. An JH, Choi A, Shin KJ, Yang WI, Lee HY (2013) DNA methylation-specific multiplex assays for body fluid identification. *Int J Legal Med* 127:35–43. <https://doi.org/10.1007/s00414-012-0719-1>
 56. Choi A, Shin KJ, Yang WI, Lee HY (2014) Body fluid identification by integrated analysis of DNA methylation and body fluid-specific microbial DNA. *Int J Legal Med* 128(1):33–41. <https://doi.org/10.1007/s00414-013-0918-4>
 57. Vidaki A, Kayser M (2018) Recent progress, methods and perspectives in forensic epigenetics. *Forensic Sci Int Genet* 37:180–195. <https://doi.org/10.1016/j.fsigen.2018.08.008>
 58. Frumkin D, Wasserstrom A, Budowle B, Davidson A (2011) DNA methylation-based forensic tissue identification. *Forensic Sci Int Genet* 5:517–524. <https://doi.org/10.1016/j.fsigen.2010.12.001>
 59. Lee HY, An JH, Jung SE, Oh YN, Lee EY, Choi A, Yang WI, Shin KJ (2015) Genome-wide methylation profiling and a multiplex construction for the identification of body fluids using epigenetic markers. *Forensic Sci Int Genet* 17:17–24. <https://doi.org/10.1016/j.fsigen.2015.03.002>
 60. Lee HY, Jung SE, Lee EH, Yang WI, Shin KJ (2016) DNA methylation profiling for a confirmatory test for blood, saliva, semen, vaginal fluid and menstrual blood. *Forensic Sci Int Genet* 24:75–82. <https://doi.org/10.1016/j.fsigen.2016.06.007>
 61. Holtkotter H, Beyer V, Schwender K, Glaub A, Johann KS, Schurenkamp M et al (2017) Independent validation of body fluid-specific CpG markers and construction of a robust multiplex assay. *Forensic Sci Int Genet* 29:261–268. <https://doi.org/10.1016/j.fsigen.2017.05.002>
 62. Lin YC, Tsai LC, Lee JC, Liu KL, Tzen JT, Linacre A et al (2016) Novel identification of biofluids using a multiplex methylation-specific PCR combined with single-base extension system. *Forensic Sci Med Pathol* 12(2):128–138. <https://doi.org/10.1007/s12024-016-9763-3>
 63. Silva DSBS, Antunes J, Balamurugan K, Duncan G, Alho CS, McCord B (2016) Developmental validation studies of epigenetic DNA methylation markers for the detection of blood, semen and saliva samples. *Forensic Sci Int Genet* 23:55–63. <https://doi.org/10.1016/j.fsigen.2016.01.017>
 64. Peat JR, Smallwood SA (2018) Low Input Whole-Genome Bisulfite Sequencing Using a Post-Bisulfite Adapter Tagging Approach. *Methods Mol Bio* 1708:161–169. https://doi.org/10.1007/978-1-4939-7481-8_9
 65. Darst RP, Pardo CE, Ai L, Brown KD, Kladd MP (2010) Bisulfite sequencing of DNA. *Curr Protoc Mol Biol*. Chapter 7 Unit 7.9.1–17. DOI: <https://doi.org/10.1002/0471142727.mb0709s91>.
 66. Gomes I, Kohlmeier F, Schneider PM (2011) Genetic markers for body fluid and tissue identification in forensics. *Forensic Sci Int Genet Suppl Ser* 3:e469–e470. <https://doi.org/10.1016/j.fsigs.2011.09.096>
 67. Madi T, Balamurugan K, Bombardi R, Duncan G, McCord B (2012) The determination of tissue specific DNA methylation patterns in forensic biofluids using bisulphite modification and pyrosequencing. *Electrophoresis* 33:1736–1745. <https://doi.org/10.1002/elps.201100711>
 68. Antunes J, Madi T, Balamurugan K, Bombardi R, Duncan G, McCord B (2013) DNA methylation markers as a powerful technique to discriminate body fluids present in crime scenes. Available at: <http://au.promega.com/~media/files/resources/conference%20proceedings/ishi%2024/oral%20presentations/antunes-manuscript.pdf>. [Downloaded 25 May 2014].
 69. Lee HY, Park MJ, Choi A, An JH, Yang WI, Shin KJ (2012) Potential forensic application of DNA methylation profiling to body fluid identification. *Int J Legal Med* 126:55–62. <https://doi.org/10.1007/s00414-011-0569-2>
 70. Kitamura E, Igarashi J, Morohashi A, Hida N, Oinuma T, Nemoto N et al (2007) Analysis of tissue-specific differentially methylated regions (TDMs) in humans. *Genomics* 89:326–337. <https://doi.org/10.1016/j.ygeno.2006.11.006>
 71. Illingworth RS, Kerr A, DeSousa D, Jorgensen H, Ellis P, Stalker J et al (2008) A novel CpG island set identifies tissue-specific methylation at developmental gene loci. *PLoS Biology* 6:e22. <https://doi.org/10.1371/journal.pbio.0060022>
 72. Park JL, Kwon OH, Kim JH, Yoo HS, Lee HC, Woo KM, Kim SY, Lee SH, Kim YS (2014) Identification of body fluid specific DNA methylation markers for use in forensic science. *Forensic Sci Int Genet* 13:147–153. <https://doi.org/10.1016/j.fsigen.2014.07.011>
 73. Forat S, Huettel B, Reinhardt R, Fimmers R, Haidl G, Denschlag D, Olek K (2016) Methylation markers for the identification of body fluids and Tissues from forensic trace evidence. *PLoS One* 11(2):e0147973. <https://doi.org/10.1371/journal.pone.0147973>
 74. Fu XD, Wu J, Wang J, Huang Y, Hou YP, Yan J (2015) Identification of body fluid using tissue-specific DNA methylation markers. *Forensic Sci Int Genet Suppl Ser* 5:e151–e153. <https://doi.org/10.1016/j.fsigs.2015.09.061>
 75. Watanabe K, Akutsu T, Takamura A, Sakurada K (2016) Evaluation of a blood-specific DNA methylated region and trial for allele-specific blood identification from mixed body fluid DNA. *Legal Med (Tokyo, Japan)* 22:49–53. <https://doi.org/10.1016/j.legalmed.2016.08.004>
 76. Vidaki A, Giangasparo F, Syndercombe Court D (2016) Discovery of potential DNA methylation markers for forensic tissue identification using bisulphite pyrosequencing. *Electrophoresis* 37(21):2767–2779. <https://doi.org/10.1002/elps.201600261>
 77. Antunes J, Silva DS, Balamurugan K, Duncan G, Alho CS, McCord B (2015) High-resolution melt analysis of DNA to discriminate semen in biological stains. *Anal Biochem* 494:40–45. <https://doi.org/10.1016/j.ab.2015.10.002>

78. Matheson CD, Gurney C, Esau N, Lehto R (2010) Assessing PCR Inhibition from Humic Substances. *Open Enzyme Inhib J* 3:38–45. <https://doi.org/10.2174/1874940201003010038>
79. Li LC, Dahiya R (2002) MethPrimer: designing primers for methylation PCRs. *Bioinformatics* 18(11):1427–1431. <https://doi.org/10.1093/bioinformatics/18.11.1427>
80. Chaitanya L, Van Oven M, Weiler N, Harteveld J, Wirken L, Sijen T et al (2014) Developmental validation of mitochondrial DNA genotyping assays for adept matrilineal inference of biogeographic ancestry at a continental level. *Forensic Sci Int Genet* 11:39–51. <https://doi.org/10.1016/j.fsigen.2014.02.010>
81. Antunes J, Silva DS, Balamurugan K, Duncan G, Alho CS, McCord B (2016) Forensic discrimination of vaginal epithelia by DNA methylation analysis through pyrosequencing. *Electrophoresis* 37(21):2751–2758. <https://doi.org/10.1002/elps.201600037>
82. Esteller M, Herman JG (2002) Cancer as an epigenetic disease: DNA methylation and chromatin alterations in human tumours. *J Pathol* 196:1–7. <https://doi.org/10.1002/path.1024>
83. Zhang D, Cheng L, Badner JA, Chen C, Chen Q, Luo W, Craig DW, Redman M, Gershon ES, Liu C (2010) Genetic control of individual differences in gene-specific methylation in human brain. *Am J Hum Genet* 86(3):411–419. <https://doi.org/10.1016/j.ajhg.2010.02.005>
84. Rakyanc VK, Down TA, Maslau S, Andrew T, Yang TP, Beyan H, Whittaker P, McCann O, Finer S, Valdes AM, Leslie RD, Deloukas P, Spector TD (2010) Human aging-associated DNA hypermethylation occurs preferentially at bivalent chromatin domains. *Genome Res* 20(4):434–439. <https://doi.org/10.1101/gr.103101.109>
85. Kaminsky Z, Petronis A (2009) Methylation SNaPshot: a method for the quantification of site-specific DNA methylation levels. *Methods Mol Biol* 507:241–255. https://doi.org/10.1007/978-1-59745-522-0_18
86. Subramanian S, Kumar S (2003) Neutral Substitutions Occur at a Faster Rate in Exons than in Noncoding DNA in Primate Genomes. *Genome Res* 13(5):838–844. DOI: 10.1186/1756-8935-6-26. DOI: <https://doi.org/10.1101/gr.1152803>.
87. Chatterjee S, Pal JK (2009) Role of 5'- and 3'-untranslated regions of mRNAs in human diseases. *Biol Cell* 101(5):251–262. <https://doi.org/10.1042/BC20080104>
88. Reamon-Buettner SM, Cho SH, Borlak J (2007) Mutations in the 3'-untranslated region of GATA4 as molecular hotspots for congenital heart disease (CHD). *BMC Medical Genetics* 8:38. <https://doi.org/10.1186/1471-2350-8-38>
89. Rogozin IB, Pavlov YI (2003) Theoretical analysis of mutation hotspots and their DNA sequence context specificity. *Mutat Res* 544(1):65–85
90. Baer CF, Miyamoto MM, Denver DR (2007) Mutation rate variation in multicellular eukaryotes: Causes and consequences. *Nat Rev Genet* 8:619–631. <https://doi.org/10.1038/nrg2158>
91. Duret L (2009) Mutation patterns in the human genome: More variable than expected. *PLoS Biology* 7(2):e1000028. <https://doi.org/10.1371/journal.pbio.1000028>
92. Arndt PF, Hwa T, Petrov DA (2005) Substantial regional variation in substitution rates in the human genome: Importance of GC content, gene density, and telomere-specific effects. *J Mol Evol* 60:748–763. <https://doi.org/10.1007/s00239-004-0222-5>
93. Duret L, Arndt PF (2008) The impact of recombination on nucleotide substitutions in the human genome. *PLoS Genetics* 4(5):e1000071. <https://doi.org/10.1371/journal.pgen.1000071>
94. Nabel CS, Manning SA, Kohli RM (2012) The Curious Chemical Biology of Cytosine: Deamination, Methylation and Oxidation as Modulators of Genomic Potential. *ACS Chem Biol* 7(1):20–30. <https://doi.org/10.1021/cb2002895>
95. Frederico LA, Kunkel TA, Shaw BR (1993) Cytosine deamination in mismatched base pairs. *Biochemistry* 32:6523–6530. <https://doi.org/10.1021/bi00077a005>
96. Fryxell KJ, Moon WJ (2005) CpG mutation rates in the human genome are highly dependent on local GC content. *Mol Biol Evol* 22(3):650–658. <https://doi.org/10.1093/molbev/msi131>
97. Duret L, Semon M, Piganeau G, Mouchiroud D, Galtier N (2002) Vanishing GC-rich isochores in mammalian genomes. *Genetics* 162:1837–1847 PMC1462357
98. Keightley PD, Eyre-Walker A (2000) Deleterious mutations and the evolution of sex. *Science* 5490:331–333. <https://doi.org/10.1126/science.290.5490.331>
99. Martinez-Arias R, Calafell F, Mateu E, Comas D, Andres A, Bertranpetit J (2001) Sequence variability of a human pseudogene. *Genome Res* 11:1071–1085. <https://doi.org/10.1101/gr.167701>
100. Nachman MW, Crowell SL (2000) Estimate of the mutation rate per nucleotide in humans. *Genetics* 156:297–304 PMC1461236
101. Chen FC, Vallender EJ, Wang H, Tzeng CS, Li WH (2001) Genomic divergence between human and chimpanzee estimated from large-scale alignments of genomic sequences. *J Hered* 92:481–489. <https://doi.org/10.1093/jhered/92.6.481>
102. Mathews DJ, Kashuk C, Brightwell G, Eichler EE, Chakravarti A (2001) Sequence variation within the fragile X locus. *Genome Res* 11:1382–1391. <https://doi.org/10.1101/gr.172601>
103. Yu N, Zhao Z, Fu YX, Sambuughin N, Ramsay M, Jenkins T, Leskinen E, Patthy L, Jorde LB, Kuromori T, Li WH (2001) Global patterns of human DNA sequence variation in a 10-kb region on chromosome 1. *Mol Biol Evol* 18:214–222. <https://doi.org/10.1093/oxfordjournals.molbev.a003795>
104. Botstein D, Risch N (2003) Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. *Nat Genet* 33:228–237. <https://doi.org/10.1038/ng1090>
105. Ng PC, Levy S, Huang J, Stockwell TB, Walenz BP, Li K, Axelrod N, Busam DA, Strausberg RL, Venter JC (2008) Genetic Variation in an Individual Human Exome. *PLoS Genetics* 4(8):e1000160
106. Daca-Roszak P, Pfeifer A, Zebracka-Gala J, Rusinek D, Szybinska A, Jarzab B et al (2015) Impact of SNPs on methylation readouts by Illumina Infinium HumanMethylation450 BeadChip Array: implications for comparative population studies. *BMC Genomics* 16:1003. <https://doi.org/10.1186/s12864-015-2202-0>
107. Galanter JM, Gignoux CR, Oh SS, Pino-Yanes M, Thakur N, Eng C et al (2017) Differential methylation between ethnic sub-groups reflects the effect of genetic ancestry and environmental exposures. *eLife* 6:e20532. <https://doi.org/10.7554/eLife.20532>
108. Dunn J, Thabet S, Jo H (2015) Flow-dependent epigenetic DNA methylation in endothelial gene expression and atherosclerosis. *Arterioscler Thromb Vasc Biol* 35(7):1562–1569. <https://doi.org/10.1161/ATVBAHA.115.305042>
109. Tost J (2010) DNA Methylation: An Introduction to the Biology and the Disease-Associated Changes of a Promising Biomarker. *Mol Biotech* 44:71–81. https://doi.org/10.1007/978-1-59745-522-0_1
110. Giuliani C, Sazzini M, Bacalini MG, Pirazzini C, Marasco E, Fontanesi E (2016) Epigenetic Variability across Human Populations: A Focus on DNA Methylation Profiles of the KRTCAP3, MAD1L1 and BRSK2 Genes. *Genome Biol Evol* 8(9):2760–2773. <https://doi.org/10.1093/gbe/evw186>
111. Terry MB, Ferris JS, Pilsner R, Flom JD, Tehranifar P, Sentellar RM et al (2008) Genomic DNA Methylation among Women in a Multiethnic New York City Birth Cohort. *Cancer Epidemiol Biomarkers Prev* 17:2306–2310. <https://doi.org/10.1158/1055-9965.EPI-08-0312>

112. Zaghlool SB, Al-Shafai M, Al Muftah WA, Kumar P, Falchi M, Suhre K (2015) Association of DNA methylation with age, gender, and smoking in an Arab population. *Clin Epigenetics* 7(1):6. <https://doi.org/10.1186/s13148-014-0040-6>
113. Reed K, Poulin ML, Yan L, Parissenti AM (2010) Comparison of bisulfite sequencing PCR with pyrosequencing for measuring differences in DNA methylation. *Anal Biochem* 397(1):96–106. <https://doi.org/10.1016/j.ab.2009.10.021>
114. Wojdacz TK, Møller TH, Thestrup BB, Kristensen LS, Hansen LL (2010) Limitations and advantages of MS-HRM and bisulfite sequencing for single locus methylation studies. *Expert Rev Mol Diagn* 10(5):575–580. <https://doi.org/10.1586/erm.10.46>
115. Wreczycka K, Godschan A, Yusuf D, Grüning B, Assenov Y, Akalin A (2017) Strategies for analyzing bisulfite sequencing data. *J Biotechnol* 261:105–115. <https://doi.org/10.1016/j.jbiotec.2017.08.007>
116. Bell JT, Pai AA, Pickrell JK, Gaffney DJ, Pique-Regi R, Degner JF et al (2011) DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. *Genome Biol* 12:1–13. <https://doi.org/10.1186/gb-2011-12-1-r10>
117. Schalkwyk LC, Meaburn EL, Smith R, Dempster EL, Jeffries AR, Davies MN, Plomin R, Mill J (2010) Allelic skewing of DNA methylation is widespread across the genome. *Am J Hum Genet* 86(2):196–212. <https://doi.org/10.1016/j.ajhg.2010.01.014>
118. Cooper DN (2010) Functional intronic polymorphisms: buried treasure awaiting discovery within our genes. *Hum Genomics* 4:284. <https://doi.org/10.1186/1479-7364-4-5-284>
119. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorf LA, Hunter DJ, McCarthy M, Ramos EM, Cardon LR, Chakravarti A, Cho JH, Guttmacher AE, Kong A, Kruglyak L, Mardis E, Rotimi CN, Slatkin M, Valle D, Whittemore AS, Boehnke M, Clark AG, Eichler EE, Gibson G, Haines JL, Mackay TF, McCarroll S, Visscher PM (2009) Finding the missing heritability of complex diseases. *Nature* 461:747–753. <https://doi.org/10.1038/nature08494>
120. Gervin K, Hammero M, Akselsen HE, Moe R, Nygard H, Brandt I et al (2011) Extensive variation and low heritability of DNA methylation identified in a twin study. *Genome Res* 21:1813–1821. <https://doi.org/10.1101/gr.119685.110>
121. Bell JT, Tsai PC, Yang TP, Pidsley R, Nisbet J, Glass D, Mangino M, Zhai G, Zhang F, Valdes A, Shin SY, Dempster EL, Murray RM, Grundberg E, Hedman AK, Nica A, Small KS, MuTHER Consortium, Dermitzakis ET, McCarthy M, Mill J, Spector TD, Deloukas P (2012) Epigenome-wide scans identify differentially methylated regions for age and age-related phenotypes in a healthy ageing population. *PLoS Genetics* 8(4):e1002629. <https://doi.org/10.1371/journal.pgen.1002629>
122. Boks MP, Derks EM, Weisenberger DJ, Strengman E, Janson E, Sommer IE et al (2009) The relationship of DNA methylation with age, gender and genotype in twins and healthy controls. *PLoS One* 4:6767. <https://doi.org/10.1371/journal.pone.0006767>
123. Breton CV, Salam MT, Gilliland FD (2011) Heritability and role for the environment in DNA methylation in AXL receptor tyrosine kinase. *Epigenetics* 6(7):895–898. <https://doi.org/10.4161/epi.6.7.15768>
124. Gibbs JR, van der Brug MP, Hernandez DG, Traynor BJ, Nalls MA, Lai SL, Arepalli S, Dillman A, Rafferty IP, Troncoso J, Johnson R, Zielke HR, Ferrucci L, Longo DL, Cookson MR, Singleton AB (2010) Abundant quantitative trait loci exist for DNA methylation and gene expression in human brain. *PLoS Genetics* 6(5):e1000952. <https://doi.org/10.1371/journal.pgen.1000952>
125. Quon G, Lippert C, Heckerman D, Listgarten J (2013) Patterns of methylation heritability in a genome-wide analysis of four brain regions. *Nucleic Acids Res* 41(4):2095–2104. <https://doi.org/10.1093/nar/gks1449>
126. Rowlatt A, Hernández-Suárez G, Sanabria-Salas MC, Serrano-López M, Rawlik K, Hernandez-Illan E, Alenda C, Castillejo A, Soto JL, Haley CS, Tenesa A (2016) The heritability and patterns of DNA methylation in normal human colorectum. *Hum Mol Genet* 25(12):2600–2611. <https://doi.org/10.1093/hmg/ddw072>
127. Zbiec-Piekarska R, Spolnicka M, Kupiec T, Makowska Z, Spas A, Parys-Proszek A et al (2015) Examination of DNA methylation status of the ELOVL2 marker may be useful for human age prediction in forensic science. *Forensic Sci Int Genet* 14:161–167. <https://doi.org/10.1016/j.fsigen.2014.10.002>
128. Thong Z, Chan XLS, Tan JYY, Loo ES, Syn CKC (2017) Evaluation of DNA methylation-based age prediction on blood. *Forensic Sci Int Genet Suppl Series* 6:e249–e251. <https://doi.org/10.1016/j.fsigs.2017.09.095>
129. Park J-L, Kim JH, Seo E, Bae DH, Kim S-Y, Lee H-C et al (2016) Identification and evaluation of age-correlated DNA methylation markers for forensic use. *Forensic Sci Int Genet* 23:64–70. <https://doi.org/10.1016/j.fsigen.2016.03.005>
130. Vidaki A, Ballard D, Aliferi A, Miller TH, Barron LP, Syndercombe Court D (2017) DNA methylation-based forensic age prediction using artificial neural networks and next generation sequencing. *Forensic Sci Int Genet* 28:225–236. <https://doi.org/10.1016/j.fsigen.2017.02.009>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.