

BMB Reports – Manuscript Submission

Manuscript Draft

Manuscript Number: BMB-16-070

Title: Forensic DNA methylation profiling to obtain investigative leads from evidence material

Article Type: Mini Review

Keywords: DNA methylation; forensic genetics; investigative lead; age prediction; body fluid identification

Corresponding Author: Hwan Young Lee

Authors: Hwan Young Lee^{1,*}, Soong Deok Lee¹, Kyoung-Jin Shin¹

Institution: ¹Department of Forensic Medicine, Yonsei University College of Medicine, 50-1 Yonsei-Ro, Seodaemun-Gu, Seoul 03722, Korea,

²Department of Forensic Medicine, Seoul National University College of Medicine, 28 Yongon-Dong, Chongno-Gu, Seoul 03080, Korea,

Manuscript Type: Mini Review

Title: Forensic DNA methylation profiling to obtain investigative leads from evidence material

Author's name: Hwan Young Lee^{1,*}, Soong Deok Lee² & Kyoung-Jin Shin¹

*Corresponding author

Affiliation:

¹Department of Forensic Medicine, Yonsei University College of Medicine, 50-1 Yonsei-Ro, Seodaemun-Gu, Seoul 03722, Korea

²Department of Forensic Medicine, Seoul National University College of Medicine, 28 Yongon-Dong, Chongno-Gu, Seoul 03080, Korea

Running Title: Forensic DNA methylation profiling

Keywords: DNA methylation, forensic genetics, investigative lead, age prediction, body fluid identification

Corresponding Author's Information:

Tel.: +82-2-2228-2482

Fax: +82-2-362-0860

E-mail: hylee192@yuhs.ac

ABSTRACT

DNA methylation is emerging as an attractive marker that can provide investigative leads to solve crimes in forensic genetics. Identification of body fluids that utilizes tissue-specific DNA methylation can contribute to solving crimes by allowing predicting activity that led to the deposition of the evidence material. DNA methylation-based age estimation is expected to be useful in helping to reduce the number of potential suspects, when the DNA profile from the evidence does not match that of any known person, including those stored in the forensic database. In addition, DNA methylation variation implicates environmental exposure, such as cigarette smoking and alcohol consumption, thereby suggesting the possibility to be used as a marker for predicting the lifestyle of potential suspect. This review will describe recent advances in our understanding of DNA methylation variations and discuss the utility of DNA methylation as a forensic marker to get more investigative leads from evidence materials.

INTRODUCTION

Currently, forensic DNA typing mainly focuses on matching a suspect with the evidence by testing a set of short tandem repeat (STR) markers. With advances in DNA detection technology and the increased number of available markers, useful DNA profiles are obtainable even with highly damaged evidence materials. However, DNA profiles can often fail to identify persons when there is no suspect available and the evidence DNA profile does not match that of any person in the forensic DNA database (1, 2). When no other evidence is available, hundreds to thousands of volunteers can be invited to provide their samples for DNA mass screening, but such DNA dragnets without specific cause and evidence to ask volunteers often face criticisms and are legally forbidden in some countries due to ethical concerns (1, 3).

Nowadays, forensic phenotyping, which aims to infer the unknown sample donor's appearance from DNA, is expected to be useful in helping to reduce the number of potential suspects (4, 5). To date, the most well studied externally visible characteristics (EVC) markers are the single nucleotide polymorphism (SNP) markers associated with pigmentation, e.g., variations in the coloration of the iris, hair and skin (6-8). However, such pigmentation markers may not be useful in certain populations, such as Asians and Africans that have little variation in coloration. More recently, age has been suggested as an EVC that can be used regardless of ethnicity to predict an individual's appearance, thereby providing an investigative lead with which to track an unknown suspect or to identify a missing person. Recently, a DNA test based on sjTREC DNA quantification has been introduced in the forensic field. This method was based on the former knowledge of the decrease in T-cells and a particular T-cell DNA rearrangement in blood with increased age (9), and showed relatively high prediction accuracy with an error of 9 years. However, this method cannot be applied to

the age prediction of other body fluids that do not include T-cells, and needs to be tested using patients' blood samples due to its immunity-dependent characteristics. In the meantime, the advancement of epigenetics led to the identification of a number of CpG markers which show age-associated DNA methylation changes in various types of tissues and cells. Several researchers reported age-predictive models that could be applied across broad spectrum of tissues as well as those based on the use of blood or saliva with a considerable prediction accuracy (<5 years) (10-13). To date, DNA methylation is regarded as the most promising age-predictive biomarker.

By the way, even when the evidence profile matches a suspect's DNA, it does not always prove the suspect guilty. It is because the matching DNA profiles only help address the issue at source or sub-source level. The probative force of matching DNA profile for sub-source and source level issues, in most cases, does not transfer directly to the probative force at activity level or offence level (14). Therefore, the activity that led to the deposition of the evidence material is more and more assessed by forensic experts (15, 16); offence level issues are normally dealt with by the court. To evaluate the evidence at activity level, knowledge about the cell types residing in an evidence material as well as the investigation of fingerprints and bloodstain pattern analysis could help (17). For example, detection of semen from vaginal swabs can indicate the involvement of some form of sexual encounter or assault, and blood stains can indicate some form of physical struggle, assault or murder whereas menses may be proposed as an alternative inoffensive scenario for a blood stain in an alleged violent assault (17). For the past decades, numerous types of analysis methods including chemical tests, immunological tests, protein catalytic activity tests, spectroscopic methods, and microscopy have been developed for investigation of forensically relevant body fluids (18). However, most of these methods suffer from limitations, such as low specificity, lack of sensitivity, sample destruction, instability of biomolecule assayed, or incompatibility with

downstream STR analysis (19). Recently, molecular approaches which involve the detection of specific messenger RNA (mRNA) and micro RNA (miRNA) expressions and differential DNA methylation patterns have therefore been intensively researched (17). Among these methods, DNA-methylation based assays, which identify differential DNA methylation profiles of different cell or tissue types, have been proposed as a promising new method for distinguishing among different types of body fluids because of their high specificity, DNA-based testing characteristics, fit with current forensic casework application, and applicability to old cases which have only DNA extracts available.

DNA methylation has only recently come into focus in the forensic field, but its applicability is being highly estimated among an increasing number of forensic investigators. Just like epigenetics to genetics, DNA methylation analyses are now expected to add more informative layers to the forensic genetic analyses of the evidence materials.

BIOLOGICAL BACKGROUND OF DNA METHYLATION

DNA methylation

Methylation of nucleotides provides a molecular mean to reversibly mark genomic DNA (20). DNA methylation is involved in immune recognition in bacteria but regulates the structure and expression of the genome in complex higher eukaryotes (21). In eukaryotes, methylation occurs only at cytosine residue, which is different from in bacteria that can methylate adenosine or cytosine (20). Moreover, vertebrates are unique in that cytosine methylation occurs throughout the entire genome, while plants and invertebrates show mosaic methylation patterns with only specific genomic elements targeted (20).

Despite the widespread methylation in vertebrate genomes, CpG islands (CGIs), which often overlap with promoter regions, generally remain unmethylated, whereas CG-poor

promoters are methylated when not active (20). Since DNA methylation blocks the start of transcription initiation not elongation (22), CGI methylation at transcription start site is also associated with long-term silencing, e.g., X-chromosome inactivation, imprinting, and genes expressed dominantly in germ cells and some tissue-specific genes (22). On the other hand, CGIs in gene bodies are known to be sometimes methylated in a tissue-specific manner and non-CGI methylation to be more dynamic and more tissue-specific than CGI methylation (22).

Such pattern of DNA methylation is established by de novo DNA methyltransferases DNMT3A and DNMT3B in combination with DNMT3L, and then is faithfully maintained through cell division by the maintenance methyltransferase DNMT1 and associated proteins (23). DNA methylation has long been considered to be lost passively through imperfect maintenance, but the recent discovery of mammalian ten-eleven translocation family of proteins showed a convincing path for catalyzed active demethylation (24).

DNA methylation variations

DNA methylation at a single CpG site within a single DNA strand is a binary trait; the site is either methylated or not (25). However, experimental samples contain a large number of DNA strands, and therefore, DNA methylation, which is recorded as a fraction between zero and one, represents the frequency of methylation at a given CpG site across the population of cells in the sample (11).

DNA methylation changes during development and aging. The full range of DNA methylation variation is potentially enormous considering that the diploid human genome contains more than 10^7 CpGs that may all potentially vary (26). The most common features in which DNA methylation varies appear as methylation variable position (MVP), variably methylated region (VMR) or differentially methylated region (DMR) (27-29). DNA

methylation at a single CpG site is known as a MVP, and the region defined by increased variability rather than gain or loss of DNA methylation refer to as VMR. DMR is a region of the genome at which multiple adjacent CpG sites show differential methylation, and can occur in many different contexts, e.g., iDMR (imprinting-specific DMR), tDMR (tissue-specific DMR), rDMR (reprogramming-specific DMR), cDMR (cancer-specific DMR) and aDMR (aging-specific DMR).

The factors underlying variable DNA methylation include cell differentiation, aging and environmental exposures as well as genetic factors. During early embryo development, cell-specific DNA methylation patterns develop to aid in cell differentiation, thereby enabling cells to have specific structures and functions (30). The established DNA methylation patterns are maintained through consecutive cell divisions, and are relatively stable throughout life (25). However, aging can modify DNA methylation through epigenetic drift and epigenetic clock; epigenetic drift is caused by the stochastic accumulation of small errors in transmitting and maintaining DNA methylation (31) and epigenetic clock refers to the phenomenon that specific sites in the genome undergo DNA methylation changes with age that are progressive and common across individuals and sometimes even tissues (11, 12). In addition, environmental exposures, such as diet, stress or smoking, can alter DNA methylation at various stages of human development (25). Besides, twin and family-based studies suggested that a considerable amount of inter-individual variation in DNA methylation is in part determined genetically (32, 33). The genome-wide association studies which test genotype-phenotype associations identified a large number of SNPs associated with the DNA methylation levels at various CpGs (34, 35). Therefore, DNA methylation at a certain CpG site is the result of multiple processes that are regulated both by genetic and environmental factors.

FORENSIC APPLICATION OF DNA METHYLATION

Tissue-specific DNA methylation changes and forensic body fluid identification

A lot of previous studies have reported tissue-specific DNA methylation and gene expression variations (17, 18). In fact, tissue of origin is the primary difference in DNA methylation profiles from different samples, regardless of whether they originate from the same or different individuals (36-39). Genome-wide DNA methylation analysis of different tissues demonstrated that numerous tDMRs exist in the mammalian genome and that DNA methylation patterns were more consistent between the same tissues from different individuals than between different tissues from the same individual (36).

In the forensic field, the potential of tissue-specific differential DNA methylation has been examined for the identification of body fluids that are frequently observed in crime scenes (40-45). Frumkin *et al.* (40) first reported genomic loci that are differentially methylated among blood, saliva, semen, skin epidermis, urine and vaginal secretion using a method based on methylation-sensitive restriction enzyme-PCR (MSRE-PCR). Lee *et al.* (41) examined the potential of tDMRs for forensic body fluid identification using a bisulfite sequencing method. They suggested two testis-specific tDMRs located at the *DACT1* and *USP49* genes as a semen-specific marker and the *PFN3* tDMR as a vaginal fluid-specific marker. Madi *et al.* (42) also examined a few genomic loci using bisulfite modification and pyrosequencing, and then reported that the methylation patterns at the *ZC3H12D* and *FGF7* loci can differentiate sperm from other biological samples while the *C20orf117* locus and the *BCAS4* locus can differentiate blood and saliva from other samples, respectively.

Recently, rapid advances in epigenetics made genome-wide DNA methylation profiling accessible to many researchers. In particular, since Illumina's HumanMethylation450 (450K) BeadChip array provides DNA methylation profiles at more than 450,000 CpG sites using

only as little as 0.5 ug of genomic DNA, several researchers have reported CpG markers that show differential DNA methylation patterns in different types of body fluids based on the 450K BeadChip array results. Park *et al.* (43) investigated the methylome data from 16 samples of blood, saliva, semen, and vaginal secretions, and identified eight CpG sites as forensically relevant DNA methylation markers: cg06379435 and cg08792630 for blood, cg26107890 and cg20691722 for saliva, cg23521140 and cg17610929 for semen, and cg01774894 and cg14991487 for vaginal secretions. In a subsequent validation with pyrosequencing analysis, the blood and semen markers were confirmed to have high specificity for identification of the target body fluid, but their vaginal secretion and saliva markers did not show sufficient methylation difference from other body fluids for acceptable specificity. Lee *et al.* (44) also identified markers using the same approach of generating the 450K BeadChip array data for 42 body fluids including blood, saliva, semen, vaginal fluid and menstrual blood. Further validation with bisulfite sequencing and methylation SNaPshot produced a total of 8 CpGs located at or adjacent to the sites investigated in the BeadChip array as body fluid-specific markers: cg17610929, cg26763284 and cg17621389 for semen, cg06379435 and cg01543184 for blood, cg09765089-231d (a CpG located 231 bp downstream of cg09765089) and cg26079753-7d (a CpG located 7 bp downstream of cg26079753) for vaginal fluid, and cg09652652-2d (a CpG located 2 bp downstream of cg09652652) for saliva. From both of the two studies, cg17610929 has been suggested as a semen-specific marker, and cg06379435 has been suggested as a blood-specific marker. In addition, cg08792630 that had been proposed as a blood-specific marker by Park *et al.* was confirmed to show methylation signal only in blood even when examined using the array data generated by Lee *et al.* (unpublished data). More recently, Forat *et al.* (45) selected body fluid-specific CpG marker candidates from HumanMethylation27 (27K) and 450K BeadChip array data of pooled DNA samples from various body fluids and tissues: cg26285698 and

cg03363565 for blood, cg09696411 for menstrual blood, cg21597595 and cg15227982 for saliva, and cg14991487 and cg03874199 for vaginal fluid, and cg22407458 and cg05656364 for semen. Through further validation with methylation SNaPshot called methylation-sensitive single nucleotide primer extension (Ms-SNuPE), they suggested body fluid-specific CpG markers that were not identical with the ones investigated in the BeadChip array but were adjacent to them with 0 to 288 bp apart. However, no marker overlaps between Forat *et al.* (45) and the two previous studies, which suggests the need to further evaluate those markers.

Nevertheless, because body fluid-specific hypo- or hypermethylation status, which does not give “on or off” signal, has attracted criticism when used for analysis of mixed samples, recent three publications (43-45) are worth noting in that they reported a set of CpG markers that show a methylation signal only in the target body fluids through genome-wide methylation profiling and gene-specific confirmation. Overall, the identification and use of semen-specific CpG markers is not a problem due to the significant difference in DNA methylation between somatic cells and germ cells. However, since most somatic tissues and body fluids possess similar DNA methylation patterns across genome, CpGs that have been suggested to be specific to blood, saliva, vaginal fluid and menstrual blood would need to be carefully validated for better application to forensic caseworks.

Age-associated DNA methylation changes and forensic age estimation

Age-associated DNA methylation changes have been studied both globally and at specific regions by multiple studies (46). These studies indicated a gain in DNA methylation in early life and gradual loss in later life across the genome (46). After birth, average DNA methylation levels increase in blood throughout the first year of life (47, 48), and after the first year, the median global DNA methylation levels are relatively stable except some

regions that frequently gain methylation (47). After reaching adulthood, DNA methylation decreases generally across the genome and specifically at repetitive elements (46).

With the advent of microarray technology, more specific DNA methylation changes at certain genes or genomic regions have been reported to show high association with age (46, 49, 50). These age-associated CpGs have been found both within a specific tissue and across tissues because DNA methylation profiles are highly divergent in different tissues. In earlier studies (51-53), Polycomb group protein target genes (*PCGTs*) known to have roles in the chromatin remodeling related with gene silencing were reported to gain methylation in blood and other tissues during aging. Bocklandt *et al.* (10) identified three age-associated CpG sites from the promoters of *EDARADD*, *TOM1L1* and *NPTX2* in saliva based on the 27K BeadChip array results, and built a regression model that could predict the age of an individual with an average accuracy of 5.2 years. Koch and Wagner (49) also using the 27K array results on various tissues and cells suggested an age-predictive model composed of five CpGs (associated with the genes *NPTX2*, *TRIM58*, *GRIA2*, *KCNQ1DN* and *BIRC4BP*), which was applicable for many tissues but the average absolute difference between predicted and chronological age was about 11 years. Of the studies using the 450K BeadChip array, Garagnani *et al.* (54) demonstrated a high association between age and DNA methylation at three CpG sites of the genes *ELOVL2*, *FHL2* and *PENK*, and suggested *ELOVL2* as the most promising age predictive maker in blood. A quantitative aging model built by Hannum *et al.* (11) used 71 CpG sites of the 450K array and showed very high prediction accuracy with an error of 3.9 years in blood. Horvath (12) developed a multi-tissue predictor of age using 8000 samples from the 27K and 450K array data sets. The model was applicable across a broad spectrum of tissues and showed high accuracy with an error of 3.6 years.

However, the genome-wide DNA methylation analysis requires substantial amount of DNA, time and efforts; therefore, an assay utilizing only a few CpG sites was expected to be

more appealing to forensic investigators if it could provide accuracy comparable to that provided by genome-wide methylation profiling. The model by Weidner *et al.* (13) was notable in that it was based on the pyrosequencing data of only three CpGs at the genes *ITGA2B*, *ASPA* and *PDE4C* for age prediction of blood and the accuracy was high with the average absolute difference between predicted and chronological age of 4.3 years. In the forensic field, Zbiec-Piekarska *et al.* (55) reported an age predictive model for blood using two CpGs in the *ELOVL2* gene, which had prediction error of 6.85 years and a mean absolute deviation (MAD) from chronological age of 5.03 years. Later, they also demonstrated that a model composed of five CpGs of the genes *ELOVL2*, *C1orf132*, *TRIM59*, *KLF14* and *FHL2* had an improved prediction accuracy in blood with a MAD from chronological age of 3.9 years (56). Recently, Park *et al.* (57) also reported an age predictive model for blood using three CpGs of the genes *ELOVL2*, *ZNF423* and *CCDC102B*. These CpGs were evaluated in more than 760 blood samples based on a pyrosequencing platform, and the model provided very high prediction accuracy with a MAD from chronological age of 3.4 years. In this paper, the authors mentioned that the DNA methylation at *KLF14* and *FHL2* were significantly associated with age, but primer design for pyrosequencing of these genes failed. Because strong age correlation of DNA methylation at the genes *ELOVL2*, *KLF14* and *FHL2* have been repeatedly observed in many independent studies with blood, they are considered to be some of the most promising age-predictive markers for blood.

Recently, Lee *et al.* (58) identified age-associated CpGs for semen, which is a particularly relevant body fluid in forensic analyses, using the 450K array and subsequent methylation SNaPshot analyses. Although the model by Horvath is applicable to age prediction with various kinds of tissues and cells, the age prediction values for sperm were significantly lower than the chronological age of the donors (12). The model by Lee *et al.* was composed of three CpGs (cg06304190 in the *TTC7B* gene, cg06979108 in the *NOX4*

gene and cg12837463), and showed a high correlation between the predicted age and the chronological age with a MAD from chronological age of approximately 5 years. In addition, the region around the *TTC7B* gene has been reported to show age-related DNA methylation alteration in the sperm methylome of two samples collected from each individual 9–19 years apart (59), suggesting *TTC7B* as one of the most promising age predictive marker for semen.

Besides, Bekaert *et al.* (60) constructed a model with four age-associated makers suggested for blood (*ASPA*, *PDE4C*, *ELOVL2* and *EDARADD*), and demonstrated that the model was also capable of producing highly accurate age predictions for teeth samples with a MAD from chronological age of 4.9 years. Giuliani *et al.* (61) also showed that the previously reported age-associated markers for blood, i.e., CpGs located in the *ELOVL2*, *FHL2* and *PENK* genes (54), could be a powerful tool to predict age in teeth, but the MAD from chronological age varied (1.2–7.1 years) depending on the part of the tooth from which DNA was extracted.

Exposure-related DNA methylation changes and their potential of forensic application

Environmental factors, such as cigarette smoking and alcohol consumption, may also alter DNA methylation (62). Environments affecting during early embryogenesis may induce extensive, soma-wide modifications of DNA methylation, whereas environments affecting later life are more likely to induce less extensive, tissue-specific modifications of DNA methylation (25).

Cigarette smoke is considered one of the most powerful environmental modifiers of DNA methylation (63), and is implicated in both effects during early embryogenesis and later life. Cigarette smoke may modulate DNA methylation through the mechanisms related with carcinogen-induced DNA damage and repair (64), down regulating effect of nicotine on DNMTs (65) and hypoxia (66).

Breitling *et al.* (63) generated DNA methylation profiles of peripheral lymphocytes from smokers, ever-smokers and never-smokers with the use of 27K BeadChip array, and reported that a single locus, cg03636183 located in *F2RL3* gene is differentially methylated between smokers and non-smokers. This CpG showed significantly lower methylation in smokers than non-smokers (% methylation difference = 12%), which has been also observed in other independent studies (67). Later, several epigenome-wide studies have been conducted using the 450K BeadChip array to identify smoking-associated DNA methylation changes (67, 68). These studies replicated the previous findings related with *F2RL3*, and revealed additionally that DNA methylation at multiple CpGs located in the *AHRR* were significantly lower in smokers than non-smokers. These sites (cg23576855 and cg05575921) showed smoking-related hypomethylation in the lungs as well as in the peripheral lymphocytes, but the % methylation difference between smokers and non-smokers in the lungs was much larger than in the peripheral lymphocytes (34% vs. 17%). Another study implicating more than 2,000 whole bloods also demonstrated significantly low DNA methylation at these sites in smokers (% methylation difference = 24%) (69). The protein encoded by *AHRR* (aryl hydrocarbon receptor repressor) participates in the aryl hydrocarbon receptor signaling cascade, which mediates detoxification of environmental pollutants (70). Thereby, cigarette smoking-induced decreases in *AHRR* DNA methylation and related increases in *AHRR* expression may contribute to removing harmful environmental chemicals such as polycyclic aromatic hydrocarbons contained in cigarette smoking (25).

However, it should be noted that such DNA methylation changes differ between tissues; Wu *et al.* (71) showed that a significant difference existed in DNA methylation between blood and saliva even from non-smokers. Therefore, potential tissue variability should also be examined to predict the effect of cigarette smoking with various tissues. In addition, prenatal exposure to cigarette smoke induces DNA methylation change at the *AHRR* and

CYP1A1 genes in cord blood and placental DNA (72). At these sites, DNA methylation differences between exposed and non-exposed newborns, however, were less than 10%, and potential DNA methylation change after that time should be investigated in various tissues.

Other studies have shown that alcohol consumption may induce DNA methylation changes in bloods (73). Global DNA methylation increases in peripheral bloods of patients with alcoholism (74, 75). Early candidate gene-based investigations have generally focused on the genes related to neurotransmitter systems, and reported small but significant changes associated with alcohol dependence (AD) (73). However, the loci of these studies failed to replicate previous association or found association only in subgroups (76-78). Epigenome-wide analyses using the 27K and 450K BeadChip arrays reported many significant probes, which are mostly hypomethylated by alcohol consumption (79-81). Zhang *et al.* (79) reported that the differentially methylation regions in AD patients include dehydrogenases 1A, *ADH7*, *ADH3B2*, *CYP2A13* and five loci (*C8orf4*, *HCRT1*, *FLJ38379*, *HSA277841* and *TSC2*) with methylation difference over 40%. Zhao *et al.* (80) produced DNA methylation profiles of AD-discordant siblings using the 450K BeadChip array, and reported 865 hypomethylated and 716 hypermethylated loci with a methylation difference of more than 20%; the most hypomethylated CpG was located in the promoter of *SSTR4* and the most hypermethylated CpG was *GABRP*. Philibert *et al.* (81) also investigated alcohol-associated methylation changes using the 450K BeadChip array, but they selected active heavy alcohol consumers entering and exiting treatment for alcohol use disorders and compared them to community controls. In this report, significant changes within individuals were not detected at 4 weeks following treatment entry, but 8626 CpGs were found to be differentially methylated between current heavy drinkers and controls. However, methylation differences at these loci were generally less than 10%. Moreover, almost all alcohol users (85%) were daily smokers, and the cg09935388 at *GFII* gene with the largest value of 15% has been reported in a number of

smoking-related studies, thereby suggesting possible confounding effect at this site (73). In contrast to smoking, numerous limitations were found both in terms of quality and quantity of studies on alcohol (73). The primary weakness is the lack of replicated findings, which may attribute to the difference in study design and population investigated. In addition, the effect size is relatively small with the largest methylation difference frequently less than 10% (73).

In an effort to establish tools to aid diagnosis and monitor treatment response on environmental exposures, Endo *et al.* (82) developed an assay to measure DNA methylation at smoking and alcohol consumption-associated CpGs using the MethyLight method that utilizes fluorescence-based real-time PCR technology. They selected cg23576855 in the *AHRR* gene as smoking and cg02583484 in the *HNRNPA1* gene as alcohol consumption biomarkers, and tested their specificity in 33 bloods from healthy donors. The methylation rate at cg23576855 was significantly different between the current and never smokers (never smokers: 72.0 ± 9.5 , past smokers: 65.8 ± 8.0 , current smokers: 44.4 ± 15.1), and the AUC for the DNA methylation rate was 0.955 for current smokers. The cg23576855 was selected from the study by Philibert *et al.* (81), and the methylation rate at this site could differentiate never and habitual alcohol consumers (never drinkers: 58.5 ± 13.2 , habitual alcohol drinkers: 48.1 ± 9.1). In the ROC analysis, the AUC for the DNA methylation rate was 0.746 for habitual alcohol consumers. Although the DNA methylation rates of *AHRR* and *HNRNPA1* did not correlate with the frequency of smoking and alcohol consumption (82) and further analyses with more various types of tissues are needed, DNA methylation at these sites seem to have the potential to monitor the lifestyle of the blood sample donors. Once the specificity and the effect size of selected markers are confirmed to be suitable as biomarkers, cigarette smoking and alcohol consumption-associated DNA methylation changes could be used for the prediction of individual's life style, which will also be useful in helping to reduce the number of potential suspects.

Assay design for DNA methylation measurement in forensic evidence samples

Because DNA methylation patterns are erased by PCR amplification, an extra step is needed to convert DNA methylation information into readily assayable DNA sequence information (83). Above all, chemical treatment of the DNA with sodium bisulfite gives rise to methylation-specific sequence variants, which can be mapped and quantified by various epigenome-wide technologies and locus-specific analysis.

Assay design means development of locus-specific analysis for detection of single or a few genes that are already established as predictive biomarkers. Therefore, it is essential to select relevant markers to allow detection of desired characteristics. In general, relevant markers can be selected from studies employing epigenome-wide technologies. Whole genome bisulfite sequencing, bisulfite microarray and enrichment methods are the current most useful and popular approaches (83), and among them, bisulfite microarray is the most public and easily accessible. The latest 450K BeadChip array allows for the high resolution, genome-wide DNA methylation profiling of human samples to be carried out, covering 99% of all RefSeq genes and approximately 450,000 CpGs overall (84). In addition, its low cost allows the profiling of a large number of samples, thereby now enormous data can be browsed and downloaded from public databases, e.g., Gene Expression Omnibus (GEO) database of the US National Center for Biotechnology Information (83).

Once the markers to be included in the assay were determined, locus-specific analysis can be designed using several approaches. Although an approach based on the use of MSRE has been reported in the forensic body fluid identification, sequencing-based technologies such as direct bisulfite pyrosequencing or alternatives such as methylation-sensitive single nucleotide primer extension-based analysis are more suitable for reliably assessing DNA methylation level in term of reproducibility and resolution. Direct bisulfite pyrosequencing

has been most widely used in forensic field as well as in other fields. This approach has been reported to provide considerable reproducibility with more than 10 ng of DNA (56, 85). Tost and Gut (86) reported reproducibility of the pyrosequencing method to be very high with differences in methylation results of only 2% in the case of the same PCR reaction and 5% in the case of different bisulfite treatments and/or separate PCR reactions. However, since this assay does not allow multiplexing of markers, a lot of time, cost and efforts may be required when multiple markers are analyzed in a large number of samples.

The second most popular approach used in the forensic field is methylation-sensitive single nucleotide primer extension-based approach, generally called methylation SNaPshot. This method enables simultaneous analysis of multiple markers by designing multiplex methylation SNaPshot, and methylation levels at multiple CpG sites can conveniently be seen at a glance in the electropherogram. In addition, it employs the same workflow used in forensic SNP analysis except that bisulfite-converted DNA is used as a template, thus special training is not required for an assay. However, there exists the difference in signal strength between fluorescence dyes in the electropherograms, and therefore, this method is most favored when detecting “on-off” signal rather than measuring hyper or hypomethylation signal. As such, this method has been consistently used for assaying markers for forensic body fluid identification (Fig. 1) (44, 87). The sensitivity varied from 0.01 to 0.5 ng of bisulfite-converted DNA (0.05 to 10 ng of genomic DNA) depending on the marker (44, 85, 87), which indicates the importance of primer design and sequence analyzed in the assay. Alternatives such as MethyLight and Epityper methods employing real time-PCR or MALDI-TOF mass spectrometry could also be used for assay design, but they do not provide 1 bp resolution (61, 82).

Because each method has different limitations, the assay should be designed in consideration of markers to be analyzed. For example, bisulfite pyrosequencing can be more

suitable for the analysis of CpG islands than methylation SNaPshot, because SBE reaction of methylation SNaPshot may be easily affected by CpGs adjacent to the target CpG site. In addition, if the assay needs precise quantitative measurement of DNA methylation such as age-prediction analysis, determination of the lowest DNA concentration that guarantees reproducibility will be much more crucial during assay validation, which is because methylation rate represents the frequency of methylation at a given CpG site across the population of cells taken from a sample vial and would be influenced by the number of DNA strands included.

CONCLUDING REMARKS

Recent advances in epigenetics have suggested using DNA methylation markers for adding more informative layers to the evidence in forensic analysis. DNA methylation profiling can not only provide information about tissue origin of evidence sample, but it can also provide information about the age and history of environmental exposure of an individual. The ability to infer the unknown sample donor's age or lifestyle such as smoking and alcohol consumption habits can guide police investigations in cases without known suspects, thereby allowing forensic use of DNA methylation for investigation, not in the courtroom. On the other hand, body fluid identification that allows predicting activity that led to the deposition of the evidence material can play an important role to evaluate the evidence at activity level in the courtroom. However, since cell lineage and tissue of origin are major determinants of DNA methylation, identification of the type of tissues or body fluids is an essential prerequisite to get investigative leads such as predicted age and lifestyle using DNA methylation profiling of evidence materials. Therefore, identification of CpG markers for more kinds of tissues and body fluids and following assay development to detect them need

to be more encouraged. It's good that more and more DNA methylation markers are being suggested as age-specific, smoking-specific and alcohol consumption-specific markers, but some of them failed to replicate previous findings and require further validation with various kinds of tissues and cells. In addition, since forensic evidence materials often possess only a scarce amount of DNA, determination of the lowest DNA concentration that guarantees reliable quantitative results will be important. Moreover, some methylation markers show methylation changes of a mere of 10% or less, careful choice should be made in assay design; the analysis platform employed in an assay design should be able to detect the methylation change at least equivalent to the effect size of the marker. Nevertheless, DNA methylation has great potential to provide additional useful information to current forensic DNA profiling. In addition, DNA methylation profiling procedures fit well with current forensic workflow, and accordingly, could be easily integrated into forensic standardized procedures. In the near future, forensic investigations, therefore, should improve a great deal with the continued advances in epigenetics that allow the extraction of more investigative leads from forensic evident materials.

ACKNOWLEDGEMENTS

This work was supported by the Bio & Medical Technology Development Program of the National Research Foundation of Korea (NRF) funded by the Korean government (NRF-2014M3A9E1069992).

FIGURE LEGENDS

Figure 1. Representative electropherograms of body fluid identification using multiplex methylation SNaPshot. (A) Semen, (B) blood, (C) vaginal fluid, (D) menstrual blood, and (E) saliva. SE1, SE2, SE3, BL1, BL2, VF1, VF2, and SA1 represent cg17610929, cg26763284, cg17621389, cg06379435, cg01543184, cg09765089-231d, cg26079753-7d, and cg09652652-2d, respectively. In a DNA methylation profile produced by the multiplex methylation SNaPshot, a green peak represented the nucleotide A as an unmethylation signal and a blue peak represented the nucleotide G as a methylation signal. Reprinted from Lee *et al.* (44) with kind permission from Elsevier.

REFERENCES

1. Kayser M (2015) Forensic DNA Phenotyping: Predicting human appearance from crime scene material for investigative purposes. *Forensic Sci Int Genet* 18, 33-48
2. Santos F, Machado H and Silva S (2013) Forensic DNA databases in European countries: is size linked to performance? *Life Sci Soc Policy* 9, 12
3. European Network of Forensic Science Institutes DNA Working Group, Group DW. Report on Criminal Cases in Europe solved by ILS (DNA Mass Testing). http://www.enfsi.eu/sites/default/files/documents/enfsi_report_on_criminal_cases_in_europe_solved_by_ils_0.pdf, 2006
4. Kayser M and Schneider PM (2009) DNA-based prediction of human externally visible characteristics in forensics: motivations, scientific challenges, and ethical considerations. *Forensic Sci Int Genet* 3, 154-161
5. Kayser M and de Knijff P (2011) Improving human forensics through advances in genetics, genomics and molecular biology. *Nat Rev Genet* 12, 179-192
6. Sulem P, Gudbjartsson DF, Stacey SN et al (2007) Genetic determinants of hair, eye and skin pigmentation in Europeans. *Nat Genet* 39, 1443-1452
7. Walsh S, Lindenbergh A, Zuniga SB et al (2011) Developmental validation of the IrisPlex system: determination of blue and brown iris colour for forensic intelligence. *Forensic Sci Int Genet* 5, 464-471
8. Walsh S, Chaitanya L, Clarisse L et al (2014) Developmental validation of the HIrisPlex system: DNA-based eye and hair colour prediction for forensic and anthropological usage. *Forensic Sci Int Genet* 9, 150-161
9. Zubakov D, Liu F, van Zelm MC et al (2010) Estimating human age from T-cell DNA rearrangements. *Curr Biol* 20, R970-R971
10. Bocklandt S, Lin W, Sehl ME et al (2011) Epigenetic predictor of age. *PLoS ONE* 6,

e14821.

11. Hannum G, Guinney J, Zhao L et al (2013) Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol Cell* 49, 359-367
12. Horvath S (2013) DNA methylation age of human tissues and cell types. *Genome Biol* 14, R115
13. Weidner CI, Lin Q, Koch CM et al (2014) Aging of blood can be tracked by DNA methylation changes at just three CpG sites. *Genome Biol* 15, R24
14. Jackson G (2013) The impact of commercialization on the evaluation of DNA evidence. *Front Genet* 4, 227
15. McKenna L (2013) Understanding DNA results within the case context: importance of the alternative proposition. *Front Genet* 4, 242
16. Gunn P, Walsh S, and Roux C (2014) The nucleic acid revolution continues, will forensic biology become forensic molecular biology? *Front Genet* 5, 44
17. Sijen T (2015) Molecular approaches for forensic cell type identification: On mRNA, miRNA, DNA methylation and microbial markers. *Forensic Sci Int Genet* 18, 21-32
18. An JH, Shin KJ, Yang WI and Lee HY (2012) Body fluid identification in forensics. *BMB Rep* 45, 545-553.
19. Virkler K and 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
20. Schübeler D (2015) Function and information content of DNA methylation. *Nature* 517, 321-326
21. Bestor TH (1990) DNA methylation: evolution of a bacterial immune function into a regulator of gene expression and genome structure in higher eukaryotes. *Philos Trans R Soc Lond B Biol Sci* 326, 179-187

22. Jones PA (2012) Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet* 13, 484-492
23. Chen T, Ueda Y, Dodge JE, Wang Z and Li E (2003) Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. *Mol Cell Biol* 23, 5594-5605
24. Tahiliani M, Koh KP, Shen Y et al (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* 324, 930-935
25. Lee KW and Pausova Z (2013) Cigarette smoking and DNA methylation. *Front Genet* 4, 132
26. Rakyan VK, Down TA, Balding DJ and Beck S (2011) Epigenome-wide association studies for common human diseases. *Nat Rev Genet* 12, 529-541
27. Rakyan VK, Hildmann T, Novik KL et al (2004) DNA methylation profiling of the human major histocompatibility complex: a pilot study for the human epigenome project. *PLoS Biol* 2, e405
28. Frigola J, Song J, Stirzaker C, Hinshelwood RA, Peinado MA and Clark SJ (2006) Epigenetic remodeling in colorectal cancer results in coordinate gene suppression across an entire chromosome band. *Nature Genet* 38, 540-549
29. Irizarry RA, Ladd-Acosta C, Wen B et al (2009) The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nature Genet* 41, 178-186
30. Straussman R, Nejman D, Roberts D et al (2009) Developmental programming of CpG island methylation profiles in the human genome. *Nat Struct Mol Biol* 16, 564-571
31. Martin GM (2005) Epigenetic drift in aging identical twins. *Proc Natl Acad Sci USA* 102, 10413-10414.
32. Bjornsson HT, Sigurdsson MI, Fallin MD et al (2008) Intra-individual change over

time in DNA methylation with familial clustering. *JAMA* 299, 2877-2883.

33. Kaminsky ZA, Tang T, Wang SC et al (2009) DNA methylation profiles in monozygotic and dizygotic twins. *Nat Genet* 41, 240-245
34. Bell JT, Pai AA, Pickrell JK et al (2011) DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. *Genome Biol* 12, R10
35. Numata S, Ye T, Hyde TM et al (2012) DNA methylation signatures in development and aging of the human prefrontal cortex. *Am J Hum Genet* 90, 260-272
36. Byun HM, Siegmund KD, Pan F et al (2009) Epigenetic profiling of somatic tissues from human autopsy specimens identifies tissue- and individual-specific DNA methylation patterns. *Hum Mol Genet* 18, 4808-4817
37. Davies MN, Volta M, Pidsley R et al (2012) Functional annotation of the human brain methylome identifies tissue-specific epigenetic variation across brain and blood. *Genome Biol* 13, R43
38. Ziller MJ, Gu H, Müller F et al (2013) Charting a dynamic DNA methylation landscape of the human genome. *Nature* 500, 477-481
39. Jiang R, Jones MJ, Chen E et al (2015) Discordance of DNA methylation variance between two accessible human tissues. *Sci Rep* 5, 8257
40. Frumkin D, Wasserstrom A, Budowle B and Davidson A (2011) DNA methylation-based forensic tissue identification. *Forensic Sci Int Genet* 5, 517-524
41. Lee HY, Park MJ, Choi A, An JH, Yang WI and Shin KJ (2012) Potential forensic application of DNA methylation profiling to body fluid identification. *Int J Legal Med* 126, 55-62
42. Madi T, Balamurugan K, Bombardi R, Duncan G and McCord B (2012) The determination of tissue-specific DNA methylation patterns in forensic biofluids using bisulfite modification and pyrosequencing. *Electrophoresis* 33, 1736-1745

43. Park JL, Kwon OH, Kim JH et al (2014) Identification of body fluid-specific DNA methylation markers for use in forensic science. *Forensic Sci Int Genet* 13, 147-153
44. Lee HY, An JH, Jung SE et al (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
45. Forat S, Huettel B, Reinhardt R et al (2016) Methylation Markers for the Identification of Body Fluids and Tissues from Forensic Trace Evidence. *PLoS One* 11, e0147973
46. Jones MJ, Goodman SJ and Kobor MS (2015) DNA methylation and healthy human aging. *Aging Cell* 14, 924-932
47. Martino DJ, Tulic MK, Gordon L et al (2011) Evidence for age-related and individual-specific changes in DNA methylation profile of mononuclear cells during early immune development in humans. *Epigenetics* 6, 1085-1094
48. Jenkins TG, Aston KI, Cairns BR and Carrell DT (2013) Paternal aging and associated intraindividual alterations of global sperm 5-methylcytosine and 5-hydroxymethylcytosine levels. *Fertil Steril* 100, 945-951
49. Koch C and Wagner W (2011) Epigenetic-aging-signature to determine age in different tissues. *Aging* 3, 1018-1027
50. Bell JT, Tsai PC, Yang TP et al (2012) Epigenome-wide scans identify differentially methylated regions for age and age-related phenotypes in a healthy ageing population. *PLoS Genet* 8, e1002629
51. Viré E, Brenner C, Deplus R et al (2006) The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* 439, 871-874
52. Teschendorff AE, Menon U, Gentry-Maharaj A et al (2010) Age-dependent DNA methylation of genes that are suppressed in stem cells is a hallmark of cancer. *Genome Res*

20, 440-446

53. Heyn H, Li N, Ferreira HJ et al (2012) Distinct DNA methylomes of newborns and centenarians. *Proc Natl Acad Sci USA* 109, 10522-10527
54. Garagnani P, Bacalini MG, Pirazzini C et al (2012) Methylation of ELOVL2 gene as a new epigenetic marker of age. *Aging Cell* 11, 1132-1134
55. Zbieć-Piekarska R, Spólnicka M, Kupiec T 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
56. Zbieć-Piekarska R, Spólnicka M, Kupiec T et al (2015) Development of a forensically useful age prediction method based on DNA methylation analysis. *Forensic Sci Int Genet* 17, 173-179
57. Park JL, Kim JH, Seo E et al (2016) Identification and evaluation of age-correlated DNA methylation markers for forensic use. *Forensic Sci Int Genet* 23, 64-70
58. Lee HY, Jung SE, Oh YN, Choi A, Yang WI and Shin KJ (2015) Epigenetic age signatures in the forensically relevant body fluid of semen: a preliminary study. *Forensic Sci Int Genet* 19, 28-34
59. Jenkins TG, Aston KI, Pflueger C, Cairns BR and Carrell DT (2014) Age-associated sperm DNA methylation alterations: possible implications in offspring disease susceptibility. *PLoS Genet* 10, e1004458
60. Bekaert B, Kamalandua A, Zapico SC, Van de Voorde W and Decorte R (2015) Improved age determination of blood and teeth samples using a selected set of DNA methylation markers. *Epigenetics* 10, 922-930
61. Giuliani C, Cilli E, Bacalini MG et al (2016) Inferring chronological age from DNA methylation patterns of human teeth. *Am J Phys Anthropol* 159, 585-595
62. Terry MB, Delgado-Cruzata L, Vin-Raviv N, Wu HC and Santella RM (2011) DNA

methylation in white blood cells: association with risk factors in epidemiologic studies. *Epigenetics* 6, 828-837

63. Breitling LP, Yang R, Korn B, Burwinkel B and Brenner H (2011). Tobacco-smoking-related differential DNA methylation: 27K discovery and replication. *Am J Hum Genet* 88, 450-457

64. Huang J, Okuka M, Lu W et al (2012) Telomere shortening and DNA damage of embryonic stem cells induced by cigarette smoke. *Reprod Toxicol* 35, 89-95

65. Satta R, Maloku E, Zhubi A et al (2008) Nicotine decreases DNA methyltransferase 1 expression and glutamic acid decarboxylase 67 promoter methylation in GABAergic interneurons. *Proc Natl Acad Sci USA* 105, 16356-16361

66. Liu Q, Liu L, Zhao Y et al (2011) Hypoxia induces genomic DNA demethylation through the activation of HIF-1alpha and transcriptional upregulation of MAT2A in hepatoma cells. *Mol Cancer Ther* 10, 1113-1123

67. Shenker NS, Polidoro S, Van Veldhoven K et al (2012) Epigenome-wide association study in the European Prospective Investigation into Cancer and Nutrition (EPIC-Turin) identifies novel genetic loci associated with smoking. *Hum Mol Genet* 22, 843-851

68. Monick MM, Beach SR, Plume J et al (2012) Coordinated changes in AHRR methylation in lymphoblasts and pulmonary macrophages from smokers. *Am J Med Genet B Neuropsychiatr Genet* 159B, 141-151

69. Zeilinger S, Kühnel B, Klopp N et al (2013) Tobacco smoking leads to extensive genome-wide changes in DNA methylation. *PLoS One* 8, e63812

70. Opitz CA, Litzenburger UM, Sahm F et al (2011) An endogenous tumour-promoting ligand of the human aryl hydrocarbon receptor. *Nature* 478, 197-203

71. Wu HC, Wang Q, Chung WK et al (2014) Correlation of DNA methylation levels in blood and saliva DNA in young girls of the LEGACY Girls study. *Epigenetics* 9, 929-933

72. Joubert BR, Haberg SE, Nilsen RM et al (2012) 450K epigenome-wide scan identifies differential DNA methylation in newborns related to maternal smoking during pregnancy. *Environ Health Perspect* 120, 1425-1431
73. Andersen AM, Dogan MV, Beach SR and Philibert RA (2015) Current and Future Prospects for Epigenetic Biomarkers of Substance Use Disorders. *Genes* 6, 991-1022
74. Bonsch D, Lenz B, Reulbach U, Kornhuber J and Bleich S (2004) Homocysteine associated genomic DNA hypermethylation in patients with chronic alcoholism. *J Neural Transm* 111, 1611-1616
75. Semmler A, Heese P, Stoffel-Wagner B et al (2015) Alcohol abuse and cigarette smoking are associated with global DNA hypermethylation: Results from the german investigation on neurobiology in alcoholism (GINA). *Alcohol* 49, 97-101
76. Nieratschker V, Grosshans M, Frank J et al (2014) Epigenetic alteration of the dopamine transporter gene in alcohol-dependent patients is associated with age. *Addict Biol* 19, 305-311
77. Jasiewicz A, Rubis B, Samochowiec J et al (2015) DAT1 methylation changes in alcohol-dependent individuals vs. controls. *J Psychiatr Res* 64, 130-133
78. Philibert RA, Gunter TD, Beach SR, Brody GH and Madan A (2008) MAOA methylation is associated with nicotine and alcohol dependence in women. *Am J Med Genet B Neuropsychiatr Genet* 147B, 565-570
79. Zhang H, Wang F, Kranzler HR, Zhao H and Gelernter J (2013) Profiling of childhood adversity-associated DNA methylation changes in alcoholic patients and healthy controls. *PLoS ONE* 8, e65648
80. Zhao R, Zhang R, Li W et al (2013) Genome-wide DNA methylation patterns in discordant sib pairs with alcohol dependence. *Asia Pac Psychiatry* 5, 39-50
81. Philibert RA, Penaluna B, White T et al (2014) A pilot examination of the genome-

wide DNA methylation signatures of subjects entering and exiting short-term alcohol dependence treatment programs. *Epigenetics* 9, 1212-1219

82. Endo K, Li J, Nakanishi M et al (2015) Establishment of the MethyLight Assay for Assessing Aging, Cigarette Smoking, and Alcohol Consumption. *Biomed Res Int* 2015, 451981

83. Bock C (2012) Analysing and interpreting DNA methylation data. *Nat Rev Genet* 13, 705-719

84. Sandoval J, Heyn H, Moran S et al (2011) Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. *Epigenetics* 6, 692-702

85. Silva DS, Antunes J, Balamurugan K, Duncan G, Alho CS and 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

86. Tost J and Gut IG (2007) DNA methylation analysis by pyrosequencing. *Nat Protoc* 2, 2265-2275

87. An JH, Choi A, Shin KJ, Yang WI and Lee HY (2013) DNA methylation-specific multiplex assays for body fluid identification. *Int J Legal Med* 127, 35-43

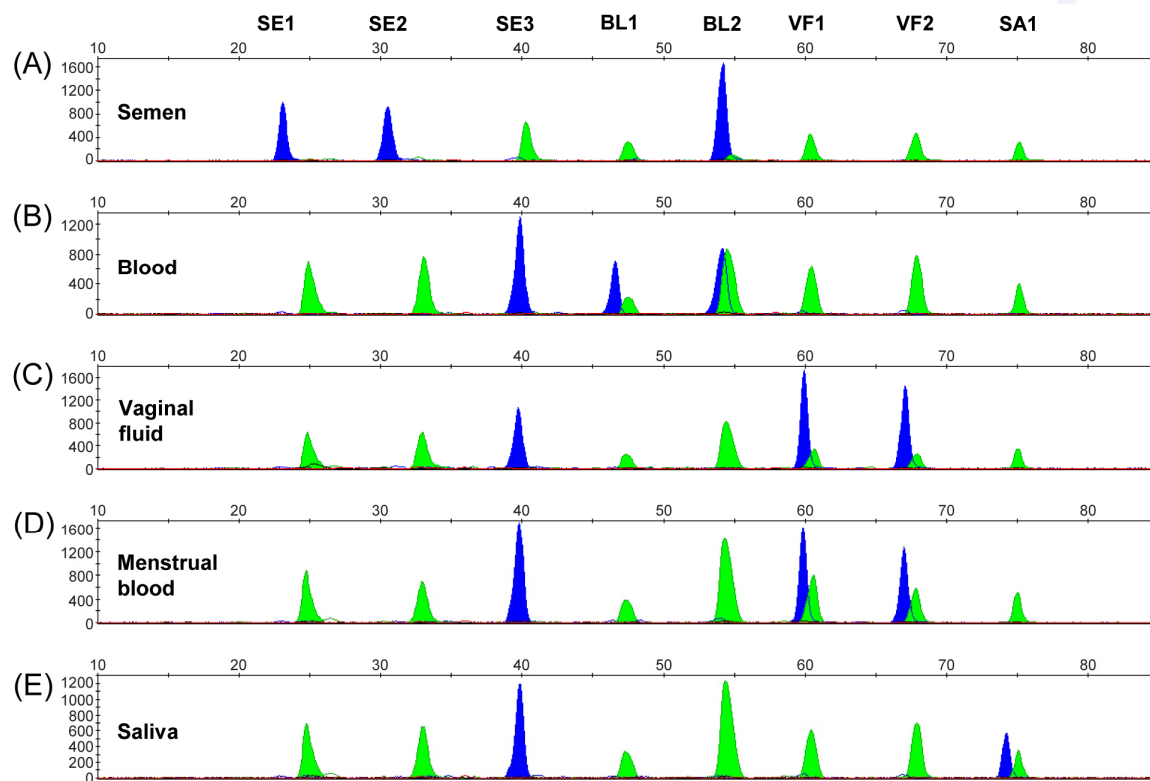
Table 1. DNA methylation-based age predictive models

Models	Age signatures ^a	Gene	Tissue	Analysis platform	Accuracy ^b
Bocklandt <i>et al.</i> [10]	cg09809672 cg27210390 cg12799895	EDARADD TOM1L1 NPTX2	Saliva	27K array	5.2 years
Koch & Wagner [49]	cg12799895 cg07533148 cg25148589 cg01530101 cg23571875	NPTX2 TRIM58 GRIA2 KCNQ1DN BIRC4BP	Various tissues	27K array	11 years
Hannum <i>et al.</i> [11]	71 CpG sites		Blood	450K array	3.9 years ^c
Horvath [12]	353 CpG sites		Various tissues	27K & 450K array	3.6 years ^c
Weidner <i>et al.</i> [13]	cg25809905 cg02228185 cg17861230 (-14 bp)	ITGA2B ASPA PDE4C	Blood	Pyrosequencing	4.3 years
Zbieć-Piekarska <i>et al.</i> [55]	cg16867657 (-2 bp) cg16867657 (-10 bp)	ELOVL2	Blood	Pyrosequencing	5.0 years
Zbieć-Piekarska <i>et al.</i> [56]	cg16867657 (-10 bp) cg10501210 (+6 bp) cg07553761 (+10 bp) cg14361627 cg06639320 (+6 bp)	ELOVL2 C1orf132 TRIM59 KLF14 FHL2	Blood	Pyrosequencing	3.9 years
Park <i>et al.</i> [57]	cg21572772 cg04208403 cg19283806	ELOVL2 ZNF423 CCDC102B	Blood	Pyrosequencing	3.4 years
Lee <i>et al.</i> [58]	cg06304190 cg06979108 cg12837463	TTC7B NOX4	Semen	Methylation SNaPshot	5.4 years
Bekaert <i>et al.</i> [60]	cg02228185 cg17861230 (-14 bp) cg16867657 (-4 bp) cg09809672	ASPA PDE4C ELOVL2 EDARADD	Blood, teeth	Pyrosequencing	4.9 years
Giuliani <i>et al.</i> [61]	4 CpG units	ELOVL2	Cementum,	EpiTYPER	1.2 years ^d
	4 CpG units	FHL2	dental pulp		
	5 CpG units	ELOVL2	Dental pulp		
	5 CpG units	FHL2			2.3 years ^d
	3 CpG units	PENK			
	7 CpG units	ELOVL2	Cementum		
	4 CpG units	FHL2			2.5 years ^d
	1 CpG units	PENK			
	3 CpG units	ELOVL2	Dentin		
1 CpG units	FHL2	7.1 years ^d			
1 CpG units	PENK				

^aAge signatures indicate the CpGs used for model construction. When CpG differs from that investigated in the BeadChip array, distance is indicated in parenthesis. + and – indicates downstream and upstream, respectively.

^bAccuracy is indicated by MAD from chronological age. When indicated by root-mean-square error or median of absolute difference, accuracy is marked by superscript c or d, respectively.

Figure 1



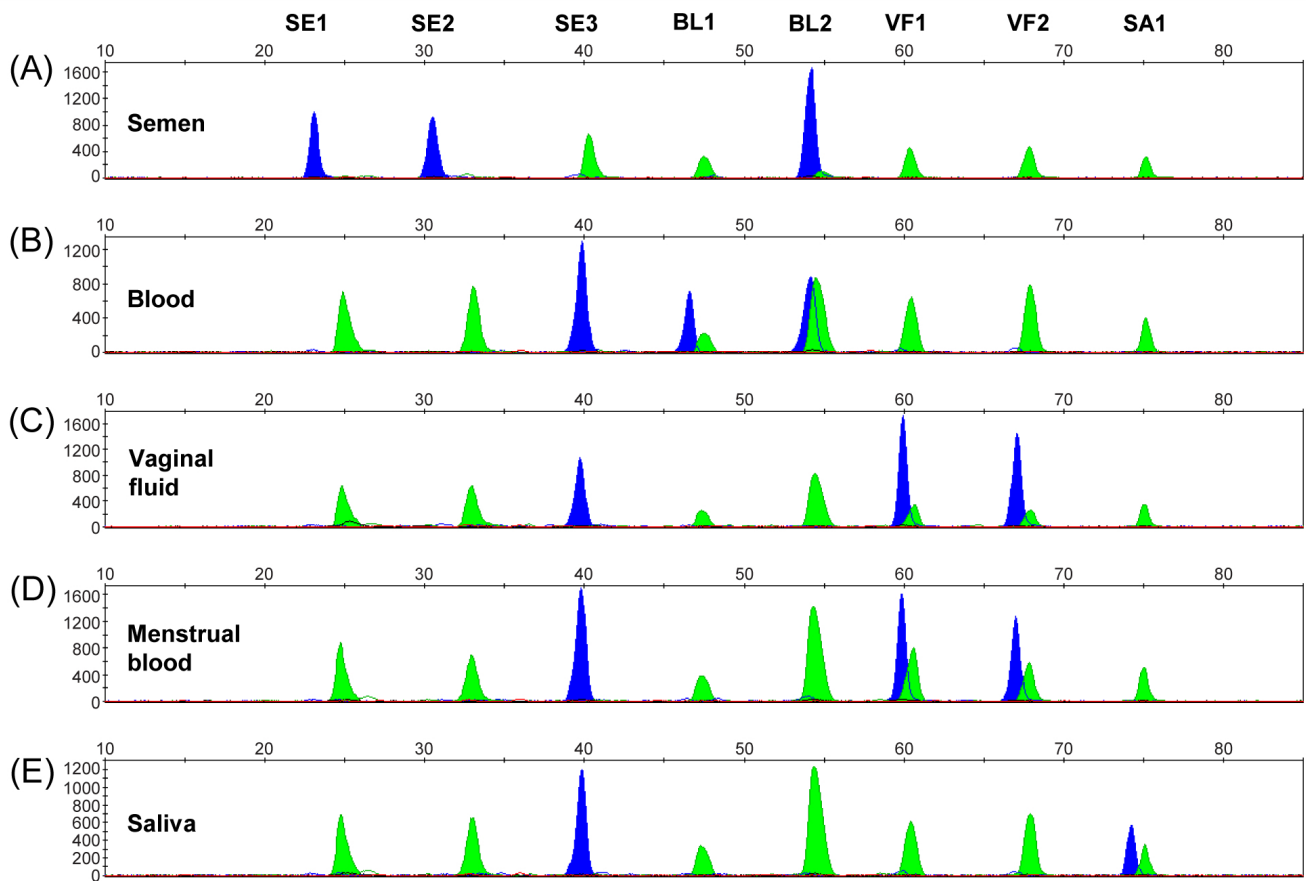


Fig. 1