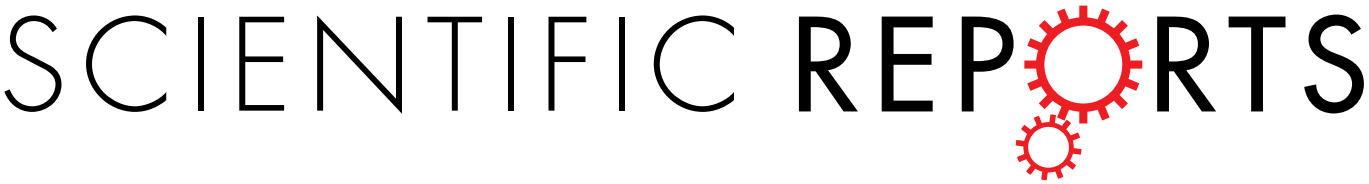

OPEN

Received: 11 May 2017
Accepted: 14 August 2017
Published online: 05 September 2017

Forensic age prediction for saliva samples using methylation-sensitive high resolution melting: exploratory application for cigarette butts

Yuya Hamano[1,2], Sho Manabe[1], Chie Morimoto[1], Shuntaro Fujimoto[1] & Keiji Tamaki[1]

There is high demand for forensic age prediction in actual crime investigations. In this study, a novel age prediction model for saliva samples using methylation-sensitive high resolution melting (MS-HRM) was developed. The methylation profiles of *ELOVL2* and *EDARADD* showed high correlations with age and were used to predict age with support vector regression. *ELOVL2* was first reported as an age predictive marker for saliva samples. The prediction model showed high accuracy with a mean absolute deviation (MAD) from chronological age of 5.96 years among 197 training samples. The model was further validated with an additional 50 test samples (MAD = 6.25). In addition, the age prediction model was applied to saliva extracted from seven cigarette butts, as in an actual crime scene. The MAD (7.65 years) for these samples was slightly higher than that of intact saliva samples. A smoking habit or the ingredients of cigarettes themselves did not significantly affect the prediction model and could be ignored. MS-HRM provides a quick (2 hours) and cost-effective (95% decreased compared to that of DNA chips) method of analysis. Thus, this study may provide a novel strategy for predicting the age of a person of interest in actual crime scene investigations.

In forensic science, predicting the age of a victim or a suspect can trigger the quick solution of a crime. Nonetheless, forensic scientists have had few options for estimating the age of the person of interest in actual practice, such as examining bones morphologically[1] or analysing the amino acid racemization of teeth[2]. These techniques are not versatile methods, as they limit sample sources. In addition, biological fluids, which are more commonly found at crime scenes, cannot be analysed with these morphological techniques. For this reason, forensic scientists have begun to apply knowledge of genetics to forensic cases, *e.g.* signal joint T-cell receptor excision circles (sjTREC)[3], telomere length[4], and somatic gene arrangements[5]. However, these genetic biomarkers exhibit relatively low accuracy or are severely influenced by the degradation of DNA collected from evidentiary materials found in actual crime scenes.

Epigenetics have recently come to play an important role in forensic age prediction. Cytosine methylation at CpG sites has been well investigated as a novel epigenetic marker of chronological age[6–17]. Hannum *et al.* built a predictive model for aging blood with 71 methylation markers selected from the Illumina Infinium HumanMethylation450 BeadChip, resulting in an error of 4.89 years[6]. Huang *et al.* also developed a predictive model for bloodstains using 5 CpG sites analysed by pyrosequencer with a mean absolute deviation (MAD) of 7.98 years[7]. Although these methods are novel, none are routinely applied for actual criminal investigations currently, likely because of their high cost and time requirements.

Traditional polymerase chain reaction (PCR), which is a universal and cost- and time-effective method, may be the key technique for the realization of forensic age prediction in actual crime investigations owing to its many advantages[18]. Recently, Mawlood *et al.* developed a useful age prediction method based on a qPCR system[19]. We have newly developed a novel age prediction model that involves the use of methylation-sensitive

[1]Department of Forensic Medicine, Graduate School of Medicine, Kyoto University, Kyoto, Japan. [2]Forensic Science Laboratory, Kyoto Prefectural Police Headquarters, Kyoto, Japan. Correspondence and requests for materials should be addressed to K.T. (email: ktamaki@fp.med.kyoto-u.ac.jp)

Target marker		Sequence
ELOVL2	Fw	CGATTTGTAGGTTTAGT
	Rv	ACTACCAATCTAAACAA
EDARADD	Fw	AGAAGGTTTGATTTTGGTTAGAT
	Rv	CCTCTCCCCATCTATTTAAT

Table 1. Sequences of PCR primers for *ELOVL2* and *EDARADD*.

high resolution melting (MS-HRM) for blood samples[20]. Antunes *et al*. also described the application potential of MS-HRM for forensic use[21]. MS-HRM is a method that measures methylation profiles easily, quickly, and cost-effectively, where the PCR amplification of bisulphite-treated DNA is followed by melting analysis[22–24]. In bisulphite-treated DNA analysis, unmethylated cytosines are converted into uracils by bisulphite conversion while methylated cytosines are kept intact. Therefore, the methylation status of each cytosine is directly converted into the sequence, where it alters the thermodynamic stability of double-stranded DNA. Thus, a novel age prediction model that is suitable for actual crime investigations using MS-HRM has been developed.

However, in most of the studies performed previously[6–8, 10–17, 19, 20], the research object has been limited to blood samples. DNA methylation profiles can differ depending on the cell type from which the DNA is derived[25–27]. Therefore, an age prediction model established from blood DNA may not be applicable for DNA derived from other biological fluids. To the best of our knowledge, only Bocklandt *et al*.[28] and Horvath[9] have investigated saliva samples, which are also commonly found at crime scenes. During the writing of this manuscript, Hong *et al*. also developed an age estimation model for saliva samples[29]. However, all of these methods suffer from the abovementioned difficulties in practical use.

Here, we report a practical age prediction method that involves analysing the methylation status of *ELOVL2* and *EDARADD* via MS-HRM of saliva samples. *ELOVL2* is newly reported to correlate with chronological age in saliva samples. In this study, 197 saliva samples were analysed to develop an age prediction model, and the model was further validated using 50 additional samples. The cost and time required for analysis were dramatically reduced with this method. In addition, saliva DNA was extracted from cigarette butts, and then age prediction was performed as in an actual crime scene for the first time ever. This HRM-based method has great potential for predicting age and is quite useful, especially when DNA data for the person of interest are not recorded in criminal databases.

Results

Identification of optimal age markers for saliva samples with MS-HRM. In previous work, we developed an age prediction model for blood samples by analysing methylation profiles of the promoter regions of *ELOVL2* and *FHL2*[20]. The degrees of methylation for both these markers increased with chronological age in blood samples. Therefore, we first investigated whether these markers could be applied for the analysis of saliva samples with MS-HRM. The methylation profile of *ELOVL2* clearly correlated with the age of the saliva samples, while that of *FHL2* exhibited no correlation with chronological age in the preliminary test (Supplementary Fig. 1).

To identify another methylation marker for MS-HRM, the top five markers positively correlated with age (*KCNG3*, *NPTX2*, *GREM*, *VGF*, and *PDE4C*) and the top five negatively correlated with age (*ASPA*, *Bles03*, *EDARADD*, *TCEA2*, and *ELN*) were selected from the study of Bocklandt *et al*., in which Illumina HumanMethylation27 microarrays were used to analyse saliva samples[28]. Bisulphite PCR primers were newly designed for these 10 markers for HRM, though only *EDARADD* showed site-specific bisulphite PCR amplification due to the sequence simplicity of bisulphite-modified DNA (*i.e*. most cytosines are converted to uracils, which act as thymines in the PCR amplification process). Thus, *ELOVL2* and *EDARADD* were selected as age prediction candidate markers for use with MS-HRM of saliva samples. The sequences of the PCR primers used in this study are shown in Table 1.

PCR bias often occurs when amplifying bisulphite-treated DNA[30, 31], since unmethylated DNA tends to be amplified more efficiently than methylated DNA. To analyse methylation profiles accurately, therefore, an interpolation line or curve must be obtained before measuring unknown methylated samples with MS-HRM. Thus, a standard line and curve were first established (Fig. 1). The promoter region of *ELOVL2* showed some PCR bias, as expected[20]. In contrast, *EDARADD* showed no PCR bias; thus, the standard line was linear. The maximum absolute relative signal difference values (Df values) obtained following HRM analysis of each sample were plotted, and a non-linear regression model was developed for *ELOVL2*, as depicted in Eq. (1):

$$\frac{0.054 \times M_1}{100 - M_1} = \frac{Df}{Df_{max} - Df} \quad (1)$$

where M_1 is the methylation score of *ELOVL2* and Df_{max} is the Df value of a 100% methylated control sample. For *EDARADD*, a simple linear regression model was developed, as depicted in eq. (2):

$$M_2 = 1.765 + 0.737 \times Df \quad (2)$$

where M_2 is the methylation score of *EDARADD*. Thereafter, the methylation scores of *ELOVL2* and *EDARADD* were calculated by substituting the Df value into the corresponding regression model.

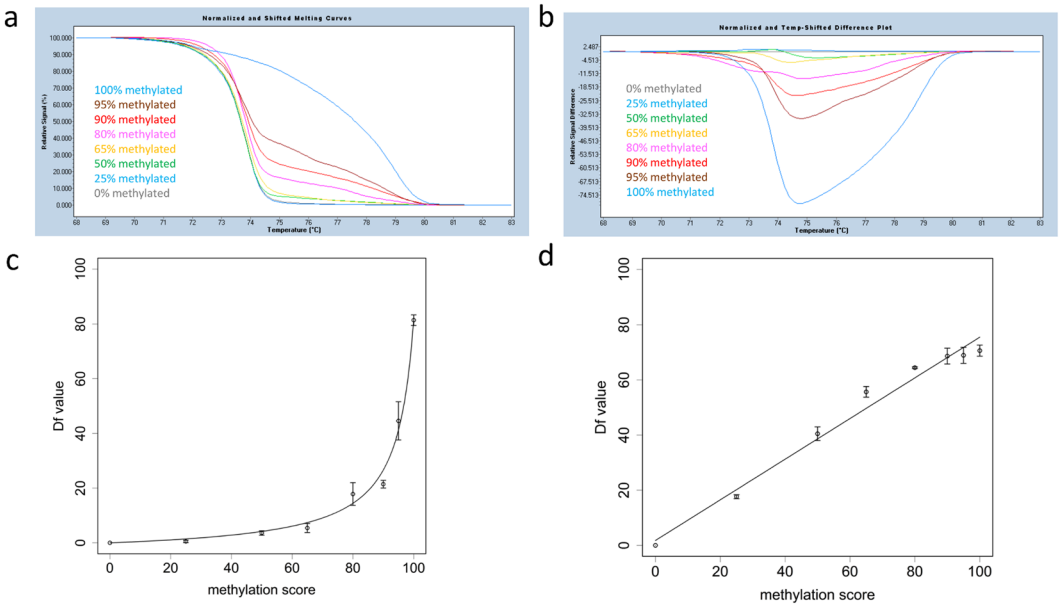

Figure 1. MS-HRM analysis of DNA methylation. (**a**) Schematic representation of MS-HRM. Normalized melting curve. Control DNA samples were mixed and adjusted to 0%, 25%, 50%, 65%, 80%, 90%, 95%, and 100% methylated. (**b**) Normalized difference plot of control DNA samples. Melting data of 0% methylated standard sample was set to baseline (grey). (**c**) Standard curve of *ELOVL2*. Error bars represent standard errors. (**d**) Standard line of *EDARADD*. Error bars represent standard errors.

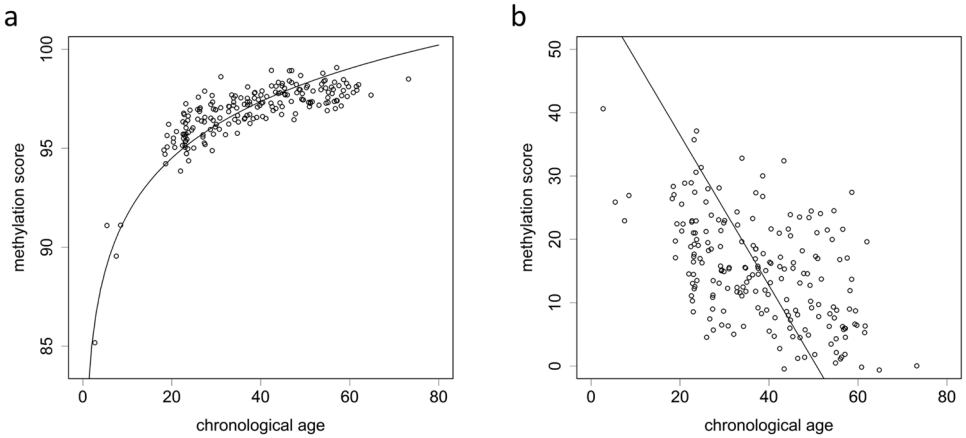

Figure 2. Relationship between age and methylation score for (**a**) *ELOVL2* and (**b**) *EDARADD*.

	Training set		Test set	
	Male	Female	Male	Female
under 20	5	5	2	3
20–39	51	49	10	16
40–59	45	36	13	6
over 60	5	1		

Table 2. Age and gender information for 197 training and 50 test samples used in this study.

Developing an age prediction model. Next, we analysed the methylation scores of *ELOVL2* and *EDARADD* in 197 saliva samples with MS-HRM (Fig. 2). Detailed information for the samples is shown in Table 2. *ELOVL2* was positively correlated with the logarithm of chronological age (Pearson's correlation coefficient r = 0.868), while *EDARADD* showed a negative correlation (r = −0.519). The relationship between the

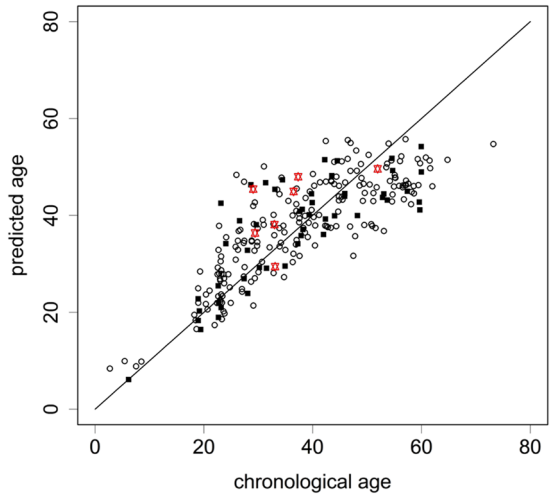

Figure 3. Correlation between predicted age and chronological age. In total, 197 training set samples plotted as white circles, 50 test set samples plotted as black squares, and seven cigarette butts plotted as red stars. The black line represents the y = x diagonal line.

methylation score and the chronological age fit the logarithmic curve well for *ELOVL2*. The methylation score of *EDARADD* showed a linear decrease with chronological age. No statistically significant difference was observed between male and female samples for either of the two markers when performing analysis of co-variance (ANCOVA) (Supplementary Fig. 2; $p = 0.849$ and 0.382 for *ELOVL2* and *EDARADD*, respectively). Subsequently, a final age prediction model was developed with support vector regression[16] using information from both markers (Fig. 3). The MAD was 5.96 years for the training set (adjusted $R^2 = 0.69$). Then, an additional independent set of 50 saliva samples was analysed to validate this model. The accuracy of the age prediction model was demonstrated with a MAD of 6.25 years for the test set (adjusted $R^2 = 0.60$). However, the MAD was smaller for younger individuals than for seniors (Supplementary Table 1).

Exploratory application. Until now, a few groups had developed age prediction models for saliva samples[9,28,29] and buccal epithelial cells[9,32,33]. However, no researcher have yet examined the utility of these methods for the analysis of forensic trace evidence, such as cigarette butts. In this study, we extracted DNA from seven cigarette butts and performed age estimation as an exploratory application (Fig. 3). The applicability of our model to cigarette butts was thus demonstrated, although the MAD of 7.65 years was slightly higher than that of intact saliva.

Based on this, the effects of smoking habits and the ingredients in the cigarettes themselves were further examined. Tsaprouni *et al*. investigated the effect of a smoking habit on genome-wide DNA methylation and found some significant smoking-related markers[34]. The methylation statuses of 54 people (50 ± 1 years old) were retrieved from publicly available data sets (GSE50660), and the effect of a smoking habit was analysed for cg16867657 (*ELOVL2*) and cg09809672 (*EDARADD*) (Supplementary Fig. 3). No statistically significant differences were observed among non-, former, or current smokers according to analysis of variance (ANOVA; $p = 0.075$ and 0.332 for *ELOVL2* and *EDARADD*, respectively). Moreover, we collected nine cigarette butts and nine saliva samples from the same volunteers for use as smokers' samples, as well as seven saliva samples from non-smokers. All of the sample donors were 40 years old. For these 25 samples (nine cigarette butts, nine smokers' saliva samples, and seven non-smokers' saliva samples), we analysed the methylation scores of *ELOVL2* and *EDARADD* with MS-HRM (Supplementary Fig. 4). No statistically significant differences in methylation scores were observed among cigarette butts, smokers' saliva, or non-smokers' saliva for *EDARADD* (ANOVA; $p = 0.072$). For *ELOVL2*, a statistically significant difference was observed ($p = 0.012$), but the difference was very slight. Subsequently, age predictions were successfully performed on these samples, resulting in MADs of 4.07, 2.56, and 2.79 years for cigarette butts, smokers' saliva, and non-smokers' saliva, respectively (Fig. 4). No statistically significant difference in prediction was observed among these categories according to ANOVA ($p = 0.22$). This demonstrates that the effect of a smoking habit and the contents of cigarettes themselves can be ignored when performing age prediction using the method developed in this study.

Discussion

Age prediction has long been one of the most practically important goals for forensic scientists. Recently, novel age estimation models were developed by analysing the methylation degrees of some CpG markers for blood samples[6–8,10–17]. However, none of these methods has been applied in actual crime investigations due to the high cost and extended length of time required for analysing DNA chips or pyrosequencing. In addition, only blood samples have been well investigated; thus, other forensically relevant body fluids—such as saliva—have been less discussed. The current study represents an age prediction model for saliva samples using MS-HRM, and it may solve the abovementioned problems of age prediction analysis.

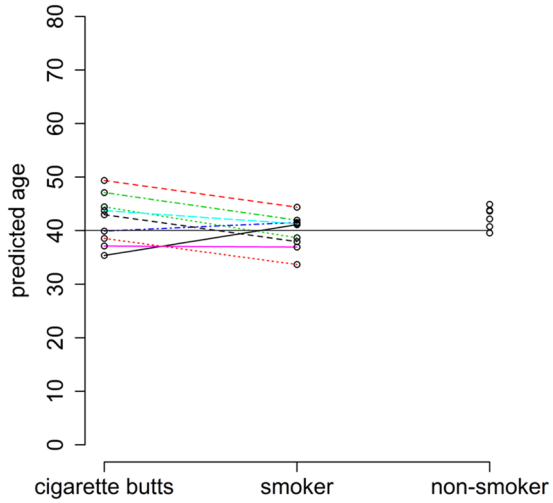

Figure 4. The results of age prediction for nine cigarette butts, nine smokers' saliva samples, and seven non-smokers' saliva samples. All sample donors were 40 years old. Cigarette butts and smokers' saliva samples were collected from the same nine individuals (connected by straight lines).

MS-HRM is a real-time PCR-based technique that measures the integrated methylation statuses of multiple CpG sites in a single assay that is quick (2 hours) and cost-effective (approximately £3 for age prediction based on two markers). According to Mawlood *et al.*, 35 hours are essential for pyrosequencing and next generation sequencing (NGS)[19], which cost £75 and £90, respectively. Therefore, many other groups have begun to use MS-HRM for various aspects of forensic research, such as differentiating monozygotic twins[35], identifying body fluids[21], and discriminating between tigers[36]. Notably, Migheli *et al.* showed that MS-HRM gave estimates of *APC* and *CDKN2A* gene methylation that were similar to those obtained by pyrosequencing[37]. Amornpisutt *et al.* also referred to the presence of a significant agreement between MS-HRM and pyrosequencing[38]. However, MS-HRM has some disadvantages. The biggest may be that individual methylation rates cannot be measured by MS-HRM. For 427 blood samples, Zbieć-Piekarska *et al.* investigated the methylation rates of seven CpG sites in *ELOVL2* with pyrosequencing (from C1 to C7 in their study)[12], which are also included in our analysing region with MS-HRM. The MAD of their model was 5.03 and 5.75 years for 303 training set and 124 test set, respectively. To evaluate the ability of MS-HRM in age prediction, another model was generated by performing support vector regression using the methylation score of *ELOVL2* only (the methylation score of *EDARADD* was not used). The MAD of this model (6.59 and 6.83 years for training set and test set, respectively) was a little higher than that of Zbieć-Piekarska's model (Supplementary Table 2). It is important to note that there is the difference in body fluids; they investigated blood samples, but we analysed saliva samples.

In the study of Zbieć-Piekarska *et al.*, the methylation rates of all seven CpG sites showed nearly the same correlation with chronological age ($r = 0.798–0.913$)[12]. Likewise, Garagnani *et al.* indicated that the methylation rates of CpG sites neighbouring an age-related CpG site were also associated with chronological age in *ELOVL2*[8]. Moreover, Day *et al.* investigated the effect of age-related CpG sites to methylation on neighbouring CpG sites in detail[39]. In his research, age-related CpG sites that were proximal to the same gene region showed a ~91% overlap in association with age. These findings are consistent with our results that a certain level of accurate age prediction can be performed with MS-HRM. As mentioned previously, MS-HRM has its advantages in time and cost required for analysis. While less information is obtained with MS-HRM as compared to other techniques measuring individual CpG methylation rates; however, our model has a potential to provide scientists with another option to predict a subject's age in an actual crime investigation and maybe useful to screen samples.

Another disadvantage may be the issue of PCR bias. In this study, the interpolation curve for *ELOVL2* showed non-linearity, indicating the presence of PCR bias, while *EDARADD* exhibited little PCR bias. Warnecke *et al.* proposed that the presence of PCR bias depends on the sequence of the bisulphite-treated DNA[30]. Thus, an interpolation curve must be obtained for each marker before analysing the methylation profile with MS-HRM, even when adapting the strategy for reversing PCR bias[31].

The prediction accuracy of our model (MAD = 6.25 years) was a little lower than that of Bocklandt *et al.* (MAD = 5.2 years)[28]. As for blood samples, increasing the number of target sites tends to improve the age prediction accuracy. For example, Weidner *et al.* developed a prediction model with three CpG markers (MAD = 5.4 years), while a more accurate model required 102 markers (MAD = 3.34 years)[10]. Park *et al.* investigated the relationship between the age prediction accuracy and the number of target sites and suggested that the most preferable number of target sites might be three for practical reasons[40]. In this study, two markers were used to predict age; however, additional markers may improve the prediction accuracy. We initially selected 10 candidate CpG sites for age estimation using data from Illumina HumanMethylation27, which assesses 27,578 CpG sites. HumanMethylation450, which assesses > 450,000 CpG sites, may result in better candidate markers for enhancing prediction accuracy. Thus, further studies may be required to incorporate at least one more marker to establish a useful model for practical application.

The MAD was smaller for younger individuals than for seniors, which is consistent with the results of a study by Branicki[13]. In our study, the speed of methylation change of *ELOVL2* was significantly higher in youth. Thus, the prediction is more precise in young people. Addition of another CpG site that undergoes a change its methylation profile in older individuals will improve the accuracy of the predictive model in the senior segment of the population. *ELOVL2* is a promising age marker for blood samples[8,12], but methylation profiles for many CpG markers can change dramatically depending on cell type[41]. This is the first report to demonstrate the utility of *ELOVL2* in the determination of age using saliva samples. In our study, *EDARADD* ($r = -0.519$) showed a modest correlation coefficient compared to that of *ELOVL2* ($r = 0.868$). Huang *et al.* developed an age prediction model[7] with four CpG markers ranging in absolute correlation coefficient ($|r|$) from 0.409 to 0.857. Higher marker correlations will also improve the age prediction model.

Individual lifestyles can cause changes in DNA methylation. The effect of a smoking habit on DNA methylation profiles has been particularly well investigated[34,42]. According to previous studies, some loci (*AHRR*, *F2RL3*, etc.) showed significant differences in methylation between smokers and non-smokers. To the best of our knowledge, none of these smoking-associated markers were also identified as age-predictive markers. *ELOVL2* and *EDARADD* showed almost no relationship with smoking habit in this study; however further study might be required due to the small sample size of this study. Notably, the smoking habit did not significantly affect the accuracy of age prediction in our study. Thus, we conclude that when performing age prediction with saliva samples extracted from cigarette butts, any effects of a smoking habit or of the ingredients of cigarettes themselves can be ignored. Age prediction with nine cigarette butts from 40-year-old donors resulted in accurate predictions (MAD = 4.07 years), though the MAD of seven cigarette butts from volunteers ranging in age from 29 to 51 years was higher (MAD = 7.65 years). This difference may be attributed to the small sample size. In total, the MAD was 5.64 for 16 cigarette butts analysed in this study, although further research is necessary to support these findings. Saliva consists mainly of leucocytes and epithelial cells[43]. According to Weidner *et al.*[33], a smaller MAD may be achieved by adding cell type markers to the prediction model.

In conclusion, a novel age prediction model for saliva samples using MS-HRM was developed in this study. There are three major points of caution before applying this method to actual forensic investigations. First, interpolation curves must be established for each instrument or reaction reagent, as the methylation score is affected by these conditions. Second, body fluid identification must be performed prior to age prediction. It is not appropriate to apply an age prediction model developed for saliva samples to blood or mixed samples. Third, since forensic samples are left in various conditions, the effect of prolonged storage and sample preservation methods must be investigated before applying this model to practice. When these requirements are fulfilled, the analysis of the methylation profiles of saliva samples with MS-HRM offers great potential for predicting age in actual crime scene investigations.

Methods

Ethic statement. All samples in this study were collected with permission for research use from the ethical committee of the Graduate School of Medicine of Kyoto University with approval number G1036. All experiments of this study were carried out in accordance with the Japanese ethical guidelines for human genome/gene analysis research, Ministry of Health, Labour and Welfare of Japan.

Sample collection, DNA extraction, and bisulphite conversion. Saliva samples from 263 healthy donors ranging in age from 1 to 73 years were collected using plastic tubes. Cigarette butts were collected from 16 volunteers. All samples were immediately stored in a −30 °C freezer until use. All donors or their parents signed written consent forms including specific consent to publish the images in an online open-access publication prior to donation. Ethical approval was received from the ethical committee of the Graduate School of Medicine of Kyoto University. We obtained participants' informed consent for all samples collected. For these samples, DNA was extracted and bisulphite-modified according to our previously published methods[20].

High resolution melting. PCR primers were designed with either BiSearch[44,45] or manually. For *ELOVL2*, the amplicon is 91 bp long and includes 10 CpG markers between primer binding sites (chr6: 11,044,611–11,044,701; UCSC Genome Browser GRCh38). For *EDARADD*, the amplicon is 139 bp long and includes four CpG sites (chr1: 236,394,341–236,394,480). PCR amplification was carried out with a Roche LightCycler 480 Instrument II (Roche Diagnostics GmbH, Mannheim, Germany) equipped with the Gene Scanning Software (version 1.5.1.62 SP2) in a 25 μL total volume containing 1 × EpiTect HRM PCR Master Mix, 250 nM of each primer, and 10 ng of bisulphite-modified template. When HRM analysis was performed, we set the pre-melt temperature region to 68–69 °C and the post-melt temperature region to 82–83 °C for *ELOVL2*. For *EDARADD*, these were set to 65–66 °C and 80–81 °C, respectively. In total, 263 saliva samples (197 in the training set, 50 in the test set, and 16 to examine the effect of smoking) were analysed using HRM in duplicate. Other variables were set appropriately according to our previous methods[20].

Calculating methylation scores. Fully methylated control DNA and fully unmethylated DNA were purchased from Qiagen (Hilden, Germany) and mixed in appropriate ratios to make 0%, 25%, 50%, 65%, 80%, 90%, 95%, and 100% methylated control DNA. The Df value of each sample obtained by HRM was plotted, and a non-linear regression model was developed for *ELOVL2* with R (version 3.2.2)[46] using the "nls" command. For *EDARADD*, a simple linear regression model was developed with R using the "lm" command. HRM measurements were performed in triplicate to obtain the interpolation curve or line. We newly defined the methylation score, since HRM provides the overall methylation profile of PCR-amplified products rather than the methylation rates of the individual CpG markers. The methylation rates of all CpG markers present in the region of interest

were integrated to determine the value of the methylation score and analysed with one pair of PCR primers in one measurement.

Developing an age prediction model. First, to predict age, a non-linear regression model for *ELOVL2* was built from 197 saliva samples with R using the "nls" command. For *EDARADD*, a linear regression model was built using the "lm" command. Secondly, ANCOVAs were performed with IBM SPSS Statistics 20 to determine whether gender affected the regression models ($p < 0.05$ was considered statistically significant). Finally, a support vector regression model was built using the "e1071" package[47]. Support vector regression parameters were optimized with "tune.svm" command and set as "Cost = 1.1, gamma = 0.1". The final model was further validated using an additional set of 50 test samples.

Assessing the impact of smoking. The methylation profiles of 54 people ranging in age from 49 to 51 years were retrieved from a publicly available dataset (GSE50660)[34]. They were categorized into three groups by their smoking habits (non-smokers, former smokers, and current smokers) according to Tsaprouni[34]. ANOVA was performed for those data with R using the "anova" and "aov" commands ($p < 0.05$ was considered statistically significant). In addition, to evaluate if a smoking habit or the ingredients of the cigarettes themselves affected the methylation score, we collected nine cigarette butts and nine saliva samples from the same volunteers for use as smokers' samples, as well as seven saliva samples from non-smokers. All sample donors were 40 years old. ANOVAs were performed on the methylation scores and the predicted ages of these samples with R using the "anova" and "aov" commands with default settings.

Availability of data and material. The datasets generated during and/or analysed during the current study are not publicly available due to protecting participant confidentiality but are available from the corresponding author on reasonable request.

References

1. Ritz-Timme, S. *et al.* Age estimation: The state of the art in relation to the specific demands of forensic practise. *Int. J. Legal Med.* **113**, 129–136 (2000).
2. Ohtani, S. & Yamamoto, T. Age estimation by amino acid racemization in human teeth. *J. Forensic Sci.* **55**, 1630–1633 (2010).
3. Ou, X. L. *et al.* Predicting human age with bloodstains by sjTREC Quantification. *PLoS ONE* **7**, e42412 (2012).
4. Benetos, A. *et al.* Telomere length as an indicator of biological aging the gender effect and relation with pulse pressure and pulse wave velocity. *Hypertension* **37**, 381–385 (2001).
5. Zubakov, D. *et al.* Estimating human age from T-cell DNA rearrangements. *Curr. Biol.* **20**, R970–R971 (2010).
6. Hannum, G. *et al.* Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol. Cell* **49**, 359–367 (2013).
7. Huang, Y. *et al.* Developing a DNA methylation assay for human age prediction in blood and bloodstain. *Forensic Sci. Int. Genet.* **17**, 129–136 (2015).
8. Garagnani, P. *et al.* Methylation of ELOVL2 gene as a new epigenetic marker of age. *Aging Cell* **11**, 1132–1134 (2012).
9. Horvath, S. DNA methylation age of human tissues and cell types. *Genome Biol.* **14**, R115 (2013).
10. Weidner, C. I. *et al.* Aging of blood can be tracked by DNA methylation changes at just three CpG sites. *Genome Biol.* **14**, R24 (2014).
11. Florath, I., Butterbach, K., Müller, H., Hudler, M. B. & Brenner, H. 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.* **14**, 1186–1201 (2014).
12. Zbieć-Piekarska, R. *et al.* 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 (2015).
13. Zbieć-Piekarska, R. *et al.* Development of a forensically useful age prediction model based on DNA methylation analysis. *Forensic Sci. Int. Genet.* **17**, 173–179 (2015).
14. Bekaert, B., Kamalandua, A., Zapico, S. C., Voorde, W. V. & Decorte, R. Improved age determination of blood and teeth samples using a selected set of DNA methylation markers. *Epigenetics* **10**, 922–930 (2015).
15. Freire-Aradas, A. *et al.* Development of a methylation marker set for forensic age estimation using analysis of public methylation data and the Agena Bioscience EpiTYPER system. *Forensic Sci. Int. Genet.* **24**, 65–74 (2016).
16. Xu, C. *et al.* A novel strategy for forensic age prediction by DNA methylation and support vector regression model. *Sci. Rep.* **5**, 17788 (2015).
17. Vidaki, A. *et al.* DNA methylation-based forensic age prediction using artificial neural network and next generation sequencing. *Forensic Sci. Int. Genet.* **28**, 225–236 (2017).
18. Hernández, H. G., Tse, M. Y., Pang, S. C., Arboleda, H. & Forero, D. A. Optimizing methodologies for PCR-based DNA methylation analysis. *BioTechniques* **55**, 181–197 (2013).
19. Mawlood, S. K. *et al.* The EpiTect Methyl qPCR Assay as novel age estimation method in forensic biology. *Forensic Sci. Int.* **264**, 132–138 (2016).
20. Hamano, Y. *et al.* Forensic age prediction for dead or living samples by use of methylation-sensitive high resolution melting. *Leg. Med.* **21**, 5–10 (2016).
21. Antunes, J. *et al.* High-resolution melt analysis of DNA methylation to discriminate semen in biological stains. *Anal. Biochem.* **494**, 40–45 (2016).
22. Wojdacz, T. K. & Dobrovic, A. Methylation-sensitive high resolution melting (MS-HRM): a new approach for sensitive and high-throughput assessment of methylation. *Nucleic Acids Res.* **35**, e41 (2007).
23. Wojdacz, T. K., Dobrovic, A. & Hansen, L. L. Methylation-sensitive high-resolution melting. *Nat. Protoc.* **3**, 1903–1908 (2008).
24. Tse, M. Y. *et al.* A refined, rapid reproducible high resolution melt (HRM)-based method suitable for quantification of global LINE-1 repetitive element methylation. *BMC Res. Notes* **4**, 565 (2011).
25. Forat, S. *et al.* Methylation markers for the identification of body fluids and tissues from forensic trace evidence. *PLoS ONE* **11**, e0147973 (2016).
26. Park, J. L. *et al.* Identification of body fluid-specific DNA methylation markers for use in forensic science. *Forensic Sci. Int. Genet.* **13**, 147–153 (2014).
27. Watanabe, K., Akutsu, T., Takamura, A. & Sakurada, K. Evaluation of a blood-specific DNA methylated region and trial for allele-specific blood identification from mixed body fluid DNA. *Leg. Med.* **22**, 49–53 (2016).
28. Bocklandt, S. *et al.* Epigenetic predictor of age. *PLoS ONE* **6**, e14821 (2011).
29. Hong, S. R. *et al.* DNA methylation-based age prediction from saliva: high age predictability by combination of 7 CpG markers. *Forensic Sci. Int. Genet.* **29**, 118–125 (2017).
30. Warnecke, P. M. *et al.* Detection and measurement of PCR bias in quantitative methylation analysis of bisulfite-treated DNA. *Nucleic Acids Res.* **25**, 4422–4426 (1997).

31. Wojdacz, T. K. & Hansen, L. L. Reversal of PCR bias for improved sensitivity of the DNA methylation melting curve assay. *BioTechniques* **41**, 274–278 (2006).
32. Bispo, D. S. *et al*. Evaluation of DNA methylation markers and their potential to predict human aging. *Electrophoresis* **36**, 1775–1780 (2015).
33. Eipel, M. *et al*. Epigenetic age predictions based on buccal swabs are more precise in combination with cell type-specific DNA methylation signatures. *Aging* **8**, 1034–1047 (2016).
34. Tsaprouni, L. G. *et al*. Cigarette smoking reduces DNA methylation levels at multiple genomic loci but the effect is partially reversible upon cessation. *Epigenetics* **9**, 1382–1396 (2014).
35. Leander, S., Neil, E., Kimberley, J. B., Dieudonne, J. M. & Graham, A. W. Differentiating between monozygotic twins through DNA methylation-specific high-resolution melt curve analysis. *Anal. Biochem.* **476**, 36–39 (2015).
36. Buddhakosai, W. *et al*. Discrimination of tiger using a novel high resolution melting (HRM) and multiplex SNP specific HRM (MSS-HRM) technique. *Forensic Sci. Int. Genet.* **13**, 30–33 (2014).
37. Migheli, F. *et al*. Comparison study of MS-HRM and pyrosequencing techniques for quantification of APC and CDKN2A gene methylation. *PLoS ONE* **8**, e52501 (2013).
38. Amornpisutt, R., Sriraksa, R. & Limpaiboon, T. Validation of methylation-sensitive high resolution melting for the detection of DNA methylation in cholangiocarcinoma. *Clin. Biochem.* **45**, 1092–1094 (2012).
39. Day, K. *et al*. Differential DNA methylation with age displays both common and dynamic features across human tissues that are influenced by CpG landscape. *Genome Biol.* **14**, R102 (2013).
40. Park, J. L. *et al*. Identification and evaluation of age-correlated DNA methylation markers for forensic use. *Forensic Sci. Int. Genet.* **23**, 64–70 (2016).
41. Lee, H. Y. *et al*. Epigenetic age signatures in the forensically relevant body fluid of semen: a preliminary study. *Forensic Sci. Int. Genet.* **19**, 28–34 (2015).
42. Zaghlool, S. B. *et al*. Association of DNA methylation with age, gender, and smoking in an Arab population. *Clin. Epigenetics* **7**, 6 (2015).
43. Thiede, C., Prange-Krex, G., Freiberg-Richter, J., Bornhäuser, M. & Ehninger, G. Buccal swabs but not mouthwash samples can be used to obtain pretransplant DNA fingerprints from recipients of allogeneic bone marrow transplants. *Bone Marrow Transplant.* **25**, 575–577 (2000).
44. Tusnády, G. E., Simon, I., Váradi, A. & Arányi, T. BiSearch: primer-design and search tool for PCR on -treated genomes. *Nucleic Acids Res.* **33**, e9 (2005).
45. Arányi, T., Váradi, A., Simon, I. & Tusnády, G. E. The BiSearch web server. *BMC Bioinformatics* **7**, 431 (2006).
46. R Development Core Team. R: a language and environment for statistical computing. Austria: R Foundation for Statistical Computing. Available from: http://www.R-project.org (2006).
47. Chang, C. C. & Lin, C. J. LIBSVM: a library for support vector machines. *ACM. T. Intel. Tec.* **2**, 27 (2011).

Acknowledgements

The authors would like to offer special thanks to Prof. Ryosuke Takahashi (Department of Neurology, Graduate School of Medicine, Kyoto University) for his technical support.

Author Contributions

K.T. designed the research. Y.H. carried out the experiments and wrote the manuscript. S.M., C.M., and S.F. contributed the model formulation and reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper at doi:10.1038/s41598-017-10752-w

Competing Interests: The authors declare that they have no competing interests.

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

Legal Medicine 21 (2016) 5–10

Contents lists available at ScienceDirect

Legal Medicine

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

Forensic age prediction for dead or living samples by use of methylation-sensitive high resolution melting

Yuya Hamano [a,b], Sho Manabe [a], Chie Morimoto [a], Shuntaro Fujimoto [a], Munetaka Ozeki [a], Keiji Tamaki [a,*]

[a] Department of Forensic Medicine, Graduate School of Medicine, Kyoto University, Sakyo-ku, Kyoto, Japan
[b] Forensic Science Laboratory, Kyoto Prefectural Police Headquarters, Kamigyo-ku, Kyoto, Japan

ARTICLE INFO

Article history:
Received 13 January 2016
Received in revised form 15 April 2016
Accepted 5 May 2016
Available online 7 May 2016

Keywords:
Age prediction
Forensic science
MS-HRM
DNA methylation

ABSTRACT

Age prediction with epigenetic information is now edging closer to practical use in forensic community. Many age-related CpG (AR-CpG) sites have proven useful in predicting age in pyrosequencing or DNA chip analyses. In this study, a wide range methylation status in the *ELOVL2* and *FHL2* promoter regions were detected with methylation-sensitive high resolution melting (MS-HRM) in a labor-, time-, and cost-effective manner. Non-linear-distributions of methylation status and chronological age were newly fitted to the logistic curve. Notably, these distributions were revealed to be similar in 22 living blood samples and 52 dead blood samples. Therefore, the difference of methylation status between living and dead samples suggested to be ignorable by MS-HRM. Additionally, the information from *ELOVL2* and *FHL2* were integrated into a logistic curve fitting model to develop a final predictive model through the multivariate linear regression of logit-linked methylation rates and chronological age with adjusted $R^2 = 0.83$. Mean absolute deviation (MAD) was 7.44 for 74 training set and 7.71 for 30 additional independent test set, indicating that the final predicting model is accurate. This suggests that our MS-HRM-based method has great potential in predicting actual forensic age.

1. Introduction

Although age is one of the most important pieces of information for criminal investigations, there are few techniques available to predict age in actual practice, such as examining bones or teeth morphologically. These techniques require expert medical experience, but the result of prediction might not be "objective". Moreover, these are not versatile methods and are limited to samples such as bones or teeth in actual practice.

Age-related changes in cytosine methylation have been recently reported by many groups [1–7]. For example, Hannum et al. built a predictive model of aging blood with the use of 71 methylation markers selected from the Illumina Infinium HumanMethylation450 BeadChip, which measures more than 450,000 CpG markers [8]. Branicki et al. investigated the usefulness of CpGs located in the promoter region of *ELOVL2* with pyrosequencing [9,10]. The promoter region of *FHL2* has also been identified as a useful age-predictive marker in many studies [4,10]. Owing to these studies, knowledge on the relationships between methylation patterns and chronological age has accumulated. However, the BeadChip method requires specialized instruments and analyzing machines followed by complex bioinformatic analysis for age prediction. The pyrosequencing method also requires specialized instruments. In general, very few forensic laboratories are equipped with these kinds of machines. Even if so, high costs has prevented these methods from being routinely used in criminal investigations.

Methylation-sensitive high resolution melting (MS-HRM) is a method that measures methylation statuses easily, quickly and cost effectively, where bisulfite-treated DNA is PCR amplified followed by melting analysis [11–15]. In bisulfite-treated DNA analyses, unmethylated cytosines are converted to uracil by bisulfite conversion while methylated cytosines are kept intact. Therefore, the information of methylation status is directly converted to the sequence, where it alters the thermodynamic stability of double-stranded DNA, enabling quantitative methylation assessment. The unique characteristic of MS-HRM is that it measures the overall methylation status of amplified PCR products, rather than the individual CpG marker. As a result, the information of many CpG markers present in the region of interest can be integrated and analyzed with one pair of PCR primers in one measurement.

Abbreviations: MS-HRM, methylation-sensitive high resolution melting; MAD, mean absolute deviation.

* Corresponding author at: Department of Forensic Medicine, Graduate School of Medicine, Kyoto University, Sakyo-ku, Kyoto 606-8501, Japan.
E-mail address: ktamaki@fp.med.kyoto-u.ac.jp (K. Tamaki).

http://dx.doi.org/10.1016/j.legalmed.2016.05.001
1344-6223/

We have to consider the possibility that post-mortem changes alter the methylation status when performing age prediction in actual cases. For example, a forensic scientist is not always cognizant of whether a victim is alive or deceased, as in abduction cases. Before applying this technology to actual cases, we must investigate the effect of post-mortem changes on forensic age prediction. To our best knowledge, no one has focused on this point, which might become a more significant issue when performing age prediction for actual forensic cases.

Here, we report on a labor-, time-, and cost-effective method of forensic age prediction using MS-HRM for the *ELOVL2* and *FHL2*. The analysis of 74 blood samples from 22 living and 52 dead donors who varied in age from 0 to 95 years yielded a logistic curve model. While the majority of previous studies constructed simple linear models for this analysis, such models were not rational for the purposes of our study. Finally, 30 independent dead blood samples were used to test the prediction accuracy of the model.

2. Materials and methods

2.1. Sample collection and DNA extraction

Blood samples from 19 healthy donors were collected at the same time of health checking. Blood samples from three children were collected from epistaxis caused by daily life hurt rather than performing any operation. For these blood samples, DNA was extracted with QIAamp DNA Investigator Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol. All donors or their parents signed written consent form prior to donation. Additionally, cadaver blood samples were collected from 82 autopsies performed during 2006–2009 at Kyoto University, Kyoto, Japan. Cadaver blood samples were collected in cases of extrinsic death—such as burn cases or suicides—and all autopsies were performed within 10 days of death. All dead bodies had no evidence of disease (e. g. cancer) which affects the methylation status. DNA from cadaver blood was extracted using the QIAprep DNA Blood Kit and stored at -20 °C until use. All samples in this study were used with permission for research use from the ethical committee of Graduate School of Medicine of Kyoto University.

2.2. Bisulfite modification and control DNA

All DNA extracted from blood was treated with EpiTect Fast Bisulfite Conversion Kit (Qiagen) and bisulfite-converted DNA were eluted with Buffer EB (10 mM Tris-Cl, pH 8.5). The concentration of eluted DNA was then measured with the Nano Vue Plus (GE Healthcare, Amersham, England) and subsequently adjusted to 10 ng/μL with Buffer EB. As a positive (fully methylated) or negative (fully unmethylated) control, we used "EpiTect Control DNA (human), methylated/unmethylated and bisulfite converted (Qiagen)" respectively. Control DNA was stored in Buffer EB and adjusted to 10 ng/μL.

2.3. High resolution melting step

PCR primers were designed with BiSearch [16,17] according to Table 1. For *ELOVL2*, the amplicon is 91 bp long and includes 10 CpG markers between primer binding sites (chr6: 11,044,611–11,044,701; Genome browser UCSC GRCh38, Fig. 1). For *FHL 2*, the amplicon is 133 bp long and includes 14 CpG markers (chr2: 105,399,228–105,399,360). PCR amplification was carried out with a Roche LightCycler 480 Instrument II (Roche Diagnostics GmbH, Mannheim, Germany) equipped with the Gene Scanning Software (version 1.5.1.62 SP2) in a 25 μL total volume containing: 1× EpiTect HRM PCR Master Mix (EpiTect HRM PCR Kit, Qiagen), 250 nM of each primer and 20 ng of bisulfite modified template. First, polymerase was activated at 95 °C for 5 min, followed by 45 cycles of 95 °C for 10 s, 50 °C for 30 s, and 72 °C for 10 s. After the amplification, HRM analysis was initiated by denaturing all products at 95 °C for 1 min, followed by re-annealing at 40 °C for 1 min. Subsequently, the samples were quickly warmed to 50 °C and heated to 95 °C at 0.1 °C/s. Fluorescence intensity was measured at 25 acquisitions/s. All reactions were performed in duplicate.

Table 1
The sequence of PCR primers.

Primers	Primer sequence (5′ to 3′)
ELOVL2-Fw	CGATTTGTAGGTTTAGT
ELOVL2-Rv	ACTACCAATCTAAACAA
FHL2-Fw	TTTACCAAAACTCCTTTCTT
FHL2-Rv	GTGGGTAGATTTTTGTTATT

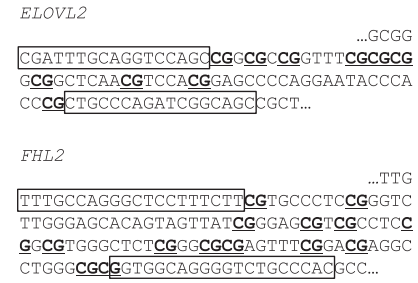

Fig. 1. Sequences of PCR target sites in this study (before bisulfite conversion). PCR primer binding sites are boxed. CpG markers that can be analyzed by MS-HRM are emphasized and underlined.

When HRM analysis was performed, Gene Scanning Software first normalized raw melt curves so that different samples can be compared. In this normalizing process, we set the pre-melt temperature region to 68–69 °C and the post-melt temperature region to 82–83 °C. Although, the temperature shift process is often run when the software is used for analyzing heterozygous mutant, no adjustment was performed in this study by setting the threshold to zero, because the shape of melt curve itself was important in analysis of the overall methylation status of the amplicon. If the temperature shift process was performed, the shape of melt curve would be distorted. A difference curve was then derived from the first derivative of the melt curves, after setting the data of fully unmethylated sample as a baseline. Relative signal difference values were exported as .txt data, and the maximum absolute value were defined as "Df value" for each sample (Fig. 2B).

2.4. Methylation analysis

In general, PCR bias occurs when amplifying bisulfite-treated DNA [18,19], since unmethylated DNA (UG pair rich sequence after bisulfite modification) tends to be amplified more efficiently than methylated DNA (CG pair rich sequence). Therefore, a standard curve was first established for each target site to accurately measure methylation status. Fully methylated control DNA and fully unmethylated DNA were mixed in appropriate ratios to make 0%, 25%, 50%, 65%, 80%, 90%, and 100% methylated control DNA. For the 90% and 100% methylated standard sample of *ELOVL2*, 40 ng bisulfite-treated DNA was used as a template due to its small amplification efficiency caused by PCR bias. Df values of each control sample were plotted and a non-linear regression model was developed [18] with R (version 3.2.2) [20] depicted as Eq. (1) as follows

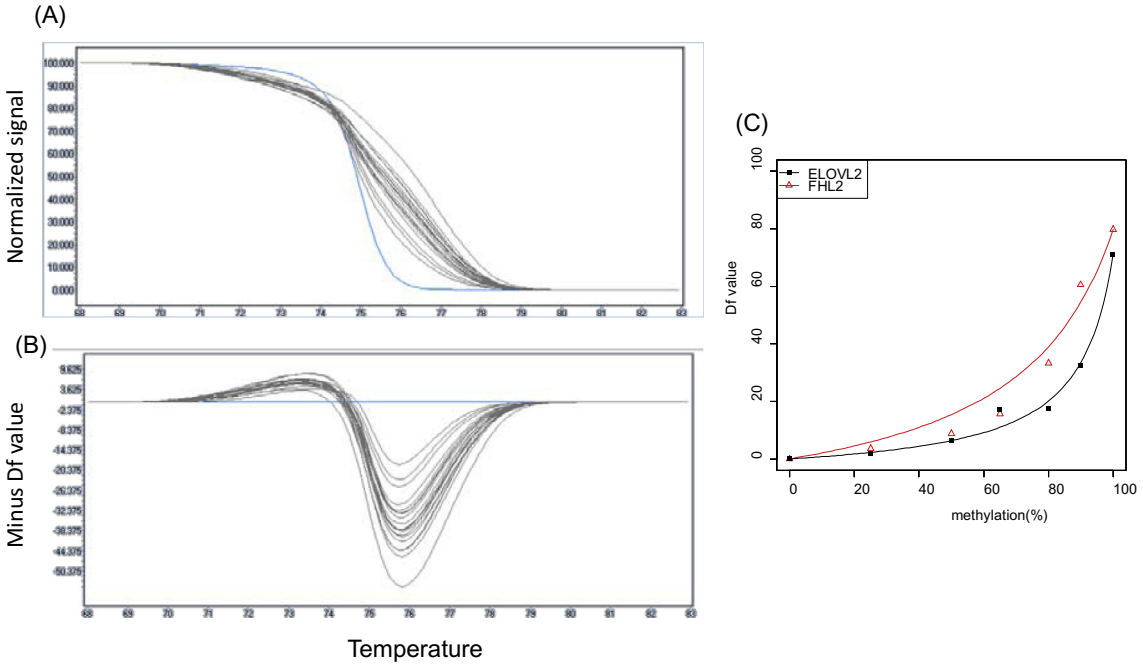

Fig. 2. (A) Representative normalized melting data. Melting data of 0% methylated standard sample is highlighted with light blue. B) Difference curves obtained by Gene Scanning software. The 0% methylated standard sample was set as a baseline. (C) Standard curves of the relationship between methylation rates and Df value for 0%, 25%, 50%, 65%, 80%, 90%, and 100% methylated standard samples.

$$\frac{a * M}{100 - M} = \frac{Df}{Df_{max} - Df} \quad (1)$$

where M is the proportion of a methylation status and Df_{max} is the Df value of 100% methylated control sample and "a" is a coefficient. Once the standard curve is established, the overall methylation status of the sample can be calculated by substituting the Df value to the Eq. (1). Therefore, Df values can be converted to methylation status.

2.5. Statistical analysis

First, logistic curve fittings were performed with R to determine the relationship between age and methylation status. In this fitting, the value of methylation status is converted to the logit form, and it was fitted to a line by ordinary least squares depicted as Eq. (2) for *ELOVL2* and *FHL2* each.

$$\text{Predicted age} = b + c * \ln\frac{M}{1 - M} \quad (2)$$

where "b, c" are coefficients. Secondly, ANCOVA were performed with IBM SPSS Statistics 20 to confirm whether live status (dead or alive) affects the regression line or not ($p < 0.05$ is considered as statistically significant).

Thirdly, a multivariate regression analysis was performed with 74 samples (22 living; 52 dead) as a training group to establish the final age predicting model depicted as Eq. (3) as follows

$$\text{Predicted age} = d + e * \ln\frac{M_E}{1 - M_E} + f * \ln\frac{M_F}{1 - M_F} \quad (3)$$

Where M_E and M_F stand for the proportion of a methylation status of *ELOVL2* and *FHL2,* respectively, and "d, e, f" are coefficients. The prediction accuracy of the regression model was assessed using the adjusted R^2. The mean absolute deviation (MAD) was also calculated. The final model was further validated using an additional set of 30 test samples (all dead).

3. Results

3.1. Methylation analysis with MS-HRM

Smooth melting curves were obtained from MS-HRM (Fig. 2A) and difference curves were also obtained with Gene Scanning Software using the 0% methylated data as a baseline (Fig. 2B). Table 2 shows the estimated "a" value of the Eq. (1) for *ELOVL2* and *FHL2*, meaning that unmethylated DNA exhibits a 5-fold or 5.6-fold amplification efficiently in MS-HRM due to PCR bias (Fig. 2C). In this study, methylation rate differences between 40% and 100% can be detected clearly by the differences of Df values both for *ELOVL2* and *FHL2*, while those between 0% and 30% were hard to detect.

3.2. Assessment of the methylation status difference between living and dead

Fig. 3A and D shows the methylation status of *ELOVL2* and *FHL2* for 74 training samples (22 alive; 52 dead). Coefficient values of the Eq. (2) ("b", "c") are listed in Table 2. Chronological age and logit-linked methylation status correlated well. Therefore, we decided to adopt logistic curve model for predicting age, while simple linear model is adopted in most of researches performed in past. Two simple linear regression lines of logit-linked methylation status and age derived from living or dead sample has no statistically significant difference ($p \geqslant 0.05$) in slope and intercept for each target site (Fig. 3B and E). New regression models explained nearly 80% of

Table 2
Coefficients calculated in this study.

	a ($\times 10^{-2}$)	b	c	d	e	f
ELOVL2	19.7	−28.8	22.9			
FHL2	9.8	−21.1	45.1			
Combined				−37.2	14.6	18.9

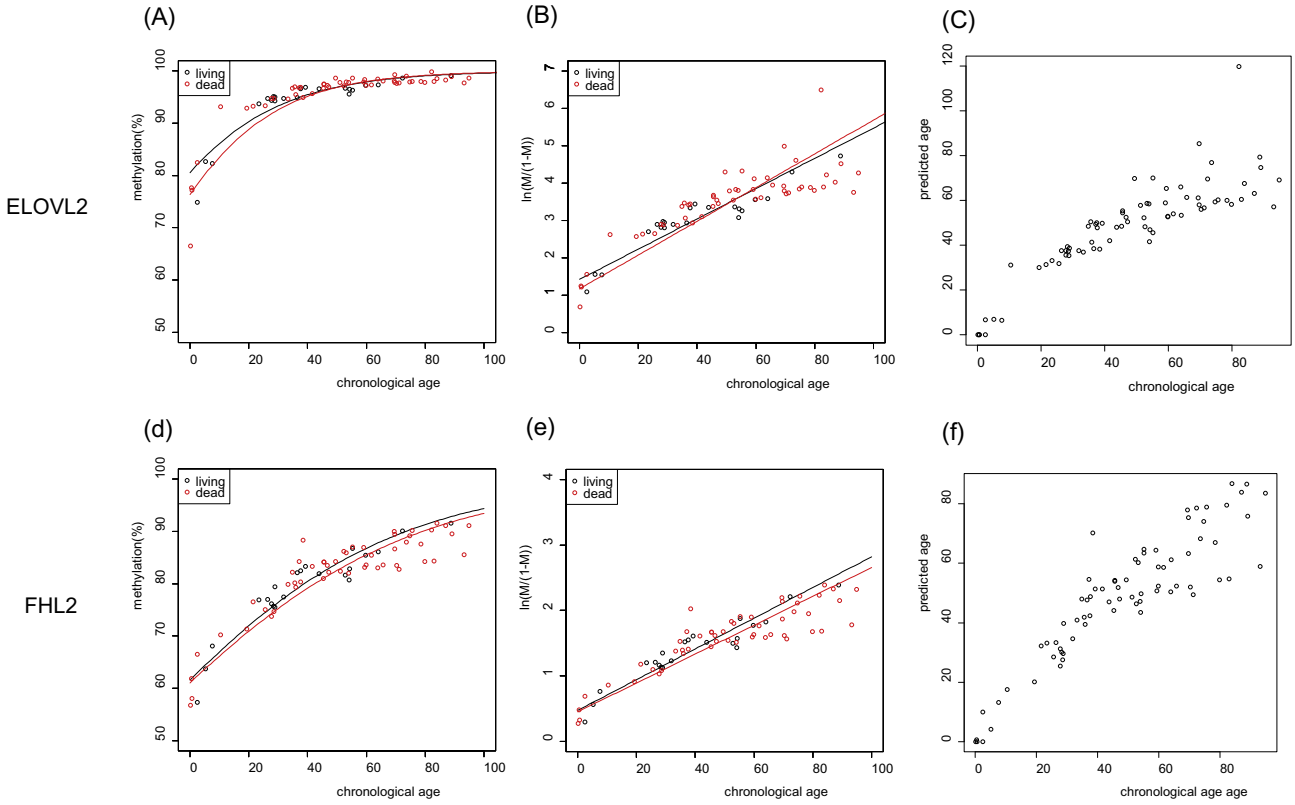

Fig. 3. (A, D) Distribution of methylation status and chronological age for living and dead blood samples. (B, E) Linear relation of chronological age and logit linked methylation status. (C, F) Accuracy of age prediction. Negative prediction values were set at 0.

the variation in age when combining living samples and dead samples (Fig. 3C and F; adjusted R^2 was 0.74 for *ELOVL2* and 0.81 for *FHL2*). Negative prediction values obtained for three or four young individuals were set at 0 [9,10]. The mean absolute deviation (MAD) was 9.67 for *ELOVL2* and 7.71 for *FHL2*. The prediction accuracy was good for the youth but a little poor for the old. No statistically significant difference in slope and intercept were observed based on gender either (Fig. S1). Thus, we decided to ignore differences due to gender to keep the prediction model simple.

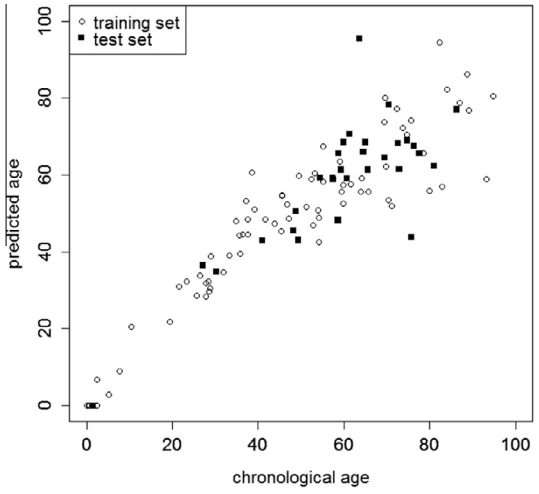

Fig. 4. Accuracy of the final age predicting model with combined information of the methylation status of ELOVL2 and FHL2 for 74 training set and 30 test set.

3.3. Developing final age prediction model and its validation

The final age prediction model was developed combining the methylation information of *ELOVL2* and *FHL2*. Estimated coefficient values of Eq. (3) ("d", "e", "f") are also listed in Table 2. This multivariate regression model showed further accuracy with MAD 7.44 (Fig. 4). In the end, holdout validation test was performed. The methylation status of 30 additional independent samples was analyzed and applied for the final model. The result is also shown in Fig. 4. MAD was slightly higher in this test group (7.71).

4. Discussion

Age prediction with the epigenetic techniques has attracted increasing attention from the forensic science community. For investigation of crimes, it is important to minimize the time required to obtain test results. In this study, MS-HRM was adopted to analyze the methylation status of the *ELOVL2* and *FHL2* promoter regions, which is able to return a test result within half a day after blood sample acquisition. The unique characteristic of this method is that it can detect the overall methylation status of the region of interest. Branicki et al. found that the methylation rates of CpG sites near AR-CpG correlate well with chronological age for *ELOVL2* (C1-C7 in his study [9]). The methylation status of many CpG sites can be detected with only one MS-HRM analysis, while DNA chip can detect only a limited number of CpG sites. On the other hand, MS-HRM has its limitations for practical usage. The biggest limitation might be the issue of PCR bias [18,19], where methylated templates (containing many Cs in the sequence) are less effectively amplified than unmethylated templates (contain many Us). Owing to the PCR bias, the methylation status

differences are hard to detect in lower methylated region (e.g. 0–30%). Therefore, if a researcher wanted to detect differences in the lowly methylated region with MS-HRM, he/she would have to identify a sequence without any PCR bias or design primers that reverse the PCR bias [19]. As for this study, PCR bias does not affect the detectability of methylation differences across samples for analyzing higher methylated regions. In addition, if CpG sites in a CpG island were intended to be analyzed, designing primers might be a little difficult because they should include as little CpG sites as possible.

The speed of the methylation change with aging has been less discussed. Most researchers have developed simple linear regression models for the methylation status and age (i.e. the speed of methylation change is constant through one's lifetime); however, these models have not accounted for pediatric specimens. Alisch et al. pointed out that the methylation changes accelerate in childhood for some CpG sites [21]. We found that changes of methylation status in the promoter region of *ELOVL2* and *FHL2* were not linear, but rather increased dramatically in youth and slowed down with increasing age. The methylation change for two target sites could be fitted well with a logistic or growth curve (adjusted $R^2 = 0.74$ and 0.81). It is reasonable to hypothesize that the methylation status plateaus with increasing age than to hypothesize that it has no upper limit.

The prediction accuracy depended on sample donor's chronological age in this study (Fig. 4), which was concordant with the research performed by Branicki et al. Age estimation for youths (0–20 years old) had little prediction error, but this error increased with chronological age. This might be explained by individual differences in the rate of methylation change. At first, individual differences may be slight, but they accumulate with age. Therefore, a prediction result must be handled carefully when the predicted age is high (>50 years old). Similarly, it is highly reliable when the predicted age is low (0–20 years old).

To our best knowledge, it is unknown whether death affects the methylation status. If post-mortem changes affect one's methylation status, forensic researchers cannot predict age without information about the sample donor's safety. In this study, all blood samples from dead bodies were collected within 10 days after death and the methylation status was analyzed. The distribution of age and methylation rate was similar regardless of the sample donor's life or death (Fig. 3A and D). Moreover, no statistically significant change was observed between living and dead blood samples. These observations may suggest that the difference of methylation status between living and dead samples is ignorable, though sample size is limited. It was difficult to collect more living blood samples especially for youths. Further analysis might be required to support our findings.

Additionally, differences in methylation status due to gender were considered (Fig. S1). No statistically significant change was observed between male and female genders, which is consistent with findings of other studies. Huang et al. developed an age prediction model for the gender combined case and two additional models for only male or only female cases; however, no statistically significant difference was observed between the gender-based models [6]. Branicki et al. decided not to include gender in prediction modeling (R^2 improved by 0.001 with age as a covariate) [10]. In our study, age prediction accuracy was minimally affected by distinguishing gender, and therefore, the simplicity of the combined model was preferred.

The multivariate regression model enhanced the prediction accuracy when combining the information about the methylation status of *ELOVL2* and *FHL2* in our study. It is well known that increasing the number of target sites enhances age prediction accuracy—Weidner et al. developed a prediction model with 3 target genes (MAD = 5.4), while more accurate model required 102 sites (MAD = 3.34) [2]. Additional target sites may also increase the prediction accuracy of our method. The authors are investigating other target sites with MS-HRM in order to improve the prediction accuracy now. However, analyzing more than 5 target sites might be too labor-intensive to perform age prediction in actual forensic cases.

Age prediction with epigenetic information has become popular in the forensic science community. This knowledge should now be used for actual criminal investigations. Most previous reports analyzed the methylation rate of CpG sites by using DNA chips or pyrosequencing. However, these techniques are too labor-, time-, and cost-intensive to apply to routine crime investigations. MS-HRM has the potential to be a gold standard for usual forensic test because of its convenience. However, there are a few causes for caution before applying this method to actual cases. First, the standard curve of methylation rate and Df value must be determined for each MS-HRM instrument and chemical because commercially available fully unmethylated DNA may not be perfectly unmethylated. Second, methylation analysis should be performed more than twice to validate the obtained data. Third, test samples are limited to blood samples. AR-CpG is considered to differ by tissue [3,5,7]. We are attempting to develop an age prediction model using saliva, semen, sweat, and bone samples with MS-HRM, but only the blood model has been fully developed. Huang et al. have shown that there was no statistically significant difference in age prediction results from blood samples and those from bloodstain [6]. When these points are considered, forensic age prediction for dead or living samples can be performed with MS-HRM conveniently. For instance, analysis of a bloodstain left at a crime scene could provide the approximate age of the suspect or victim. Thus, we believe that this study opens new possibilities for forensic DNA phenotyping.

5. Role of funding source

This study was supported by the operation grants of Department of Forensic Medicine, Graduate School of Medicine, Kyoto University.

6. Conflict of interest

The authors state that they have no conflicts of interest.

Acknowledgments

The authors would like to offer special thanks to Prof. Ryosuke Takahashi (Department of Neurology, Graduate School of Medicine, Kyoto University) for his technical support. We also thank Dr. Ken Watanabe (National Research Institute of Police Science, NRIPS) whose comments made an enormous contribution to this work.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.legalmed.2016.05.001.

References

[1] P. Garagnani, M.G. Bacalini, C. Pirazzini, D. Gori, C. Giuliani, D. Mari, A.M. Di Blasio, D. Gentilini, G. Vitale, S. Collino, S. Rezzi, G. Castellani, M. Capri, S. Salvioli, C. Franceschi, Methylation of ELOVL2 gene as a new epigenetic marker of age, Aging Cell 11 (2012) 1132–1134.

[2] C.I. Weidner, Q. Lin, C.M. Koch, L. Eisele, F. Beier, P. Ziegler, D.O. Bauerschlag, K. H. Jöckel, R. Erbel, T.W. Mühleisen, M. Zenke, T.H. Brümmendorf, W. Wagner,

Aging of blood can be tracked by DNA methylation changes at just three CpG sites, Genome Biol. 15 (2014) R24.
[3] K. Day, L.L. Waite, A. Thalacker-Mercer, A. West, M.M. Bamman, J.D. Brooks, R. M. Myers, D. Absher, Differential DNA methylation with age displays both common and dynamic features across human tissues that are influenced by CpG landscape, Genome Biol. 14 (2013) R102.
[4] I. Florath, K. Butterbach, H. Müller, M. Bewerunge-Hudler, H. Brenner, 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 (2014) 1186–1201.
[5] S. Horvath, DNA methylation age of human tissues and cell types, Genome Biol. 14 (2013) R115.
[6] Y. Huang, J. Yan, J. Hou, X. Fu, L. Li, Y. Hou, Developing a DNA methylation assay for human age prediction in blood and bloodstain, Forensic Sci. Int. Genet. 17 (2015) 129–136.
[7] H.Y. Lee, Sang-Eun Jung, Y.N. Oh, A. Choi, W.I. Yang, Kyoung-Jin Shin, Epigenetic age signatures in the forensically relevant body fluid of semen: a preliminary study, in: Forensic Sci. Int. Genet 19 (2015) 28–34.
[8] G. Hannum, J. Guinney, L. Zhao, L. Zhang, G. Hughes, S. Sadda, B. Klotzle, M. Bibikova, J.B. Fan, Y. Gao, R. Deconde, M. Chen, I. Rajapakse, S. Friend, T. Ideker, K. Zhang, Genome-wide methylation profiles reveal quantitative views of human aging rates, Mol. Cell 49 (2013) 359–367.
[9] R. Zbieć-Piekarska, M. Spólnicka, T. Kupiec, Ż. Makowska, A. Spas, A. Parys-Proszek, K. Kucharczyk, R. Płoski, W. Branicki, Examination of DNA methylation status of the ELOVL2 marker may be useful for human age prediction in forensic science, Forensic Sci. Int. Genet. 14 (2015) 161–167.
[10] R. Zbieć-Piekarska, M. Spólnicka, T. Kupiec, A. Parys-Proszek, Ż. Makowska, A. Pałeczka, K. Kucharczyk, R. Płoski, W. Branicki, Development of a forensically useful age prediction method based on DNA methylation analysis, Forensic Sci. Int. Genet. 17 (2015) 173–179.
[11] T.K. Wojdacz, A. Dobrovic, L.L. Hansen, Methylation-sensitive high-resolution melting, Nat. Protoc. 3 (2008) 1903–1908.
[12] T.K. Wojdacz, A. Dobrovic, Methylation-sensitive high resolution melting (MS-HRM): a new approach for sensitive and high-throughput assessment of methylation, Nucleic Acids Res. 35 (2007) e41.
[13] M.Y. Tse, J.E. Ashbury, N. Zwingerman, W.D. King, S.A. Taylor, S.C. Pang, A refined, rapid and reproducible high resolution melt (HRM)-based method suitable for quantification of global LINE-1 repetitive element methylation, BMC Res. Notes 4 (2011) 565.
[14] H.G. Hernandez, M.Y. Tse, S.C. Pang, H. Arboleda, D.A. Forero, Optimizing methodologies for PCR-based DNA methylation analysis, Biotechniques 55 (2013) 181–197.
[15] L. Stewart, N. Evans, K.J. Bexon, D.J. van der Meer, G.A. Williams, Differentiating between monozygotic twins through DNA methylation-specific high-resolution melt curve analysis, Anal. Biochem. 476 (2015) 36–39.
[16] G.E. Tusnady, I. Simon, A. Varadi, T. Aranyi, BiSearch: primer-design and search tool for PCR on bisulfite-treated genomes, Nucleic Acids Res. 33 (2005) e9.
[17] T. Aranyi, A. Varadi, I. Simon, G.E. Tusnady, The Bisearch web server, BMC Bioinformatics 7 (2006) 431.
[18] P.M. Warnecka, C. Stirzaker, J.R. Melki, D.S. Millar, C.L. Paul, S.J. Clark, Detection and measurement of PCR bias in quantitative methylation analysis of bisulphite-treated DNA, Nucleic Acids Res. 25 (1997) 4422–4426.
[19] T.K. Wojdacz, L.L. Hansen, Reversal of PCR bias for improved sensitivity of the DNA methylation melting curve assay, Biotechniques 41 (2006) 274–278.
[20] R Development Core Team. R: a language and environment for statistical computing. Vienna. Austria: R Foundation for Statistical Computing. 2006. Available from: http://www.R-project.org.
[21] R.S. Alisch, B.G. Barwick, P. Chopra, Age-associated DNA methylation in pediatric populations, Genome Res. 22 (2012) 623–632.