Age estimation based on DNA methylation in teeth using real-time methylation-specific PCR

Ayano Ogata
Nihon University Graduate School of Dentistry
Major in Legal Medicine
(Directors: Prof. Naoto Suzuki, Project Prof. Hirofumi Aboshi,
and Assoc. Prof. Masahiro Kondo)

Index

This thesis is based on the following article and additional results in terms of age estimation using dental pulp DNA methylation (Table. 1c, Fig. 6,7):

Ogata A, Kondo M, Yoshikawa M, Okano M, Tsutsumi T, Aboshi H (2022) Dental age estimation based on DNA methylation using real-time methylation-specific PCR. Forensic Sci Int, 340:111445.

Abstract

Age estimation is crucial for reconstructing the biological profiles of deceased victims in the forensic field. DNA methylation, which varies in an age-dependent manner in specific genes, is a candidate biomarker for estimating chronological age. DNA methylation-based models for estimating age have been developed using various technologies such as pyrosequencing. Recently the methylation levels of elongation of very long chain fatty acids protein 2 (ELOVL2) in teeth were quantified using real-time methylation-specific polymerase chain reaction (RT-MSP) to rapidly assess the methylation value of CpG sites within a CpG island. The methylation levels of *ELOVL2* were moderately correlated with chronological age, suggesting the usefulness of RT-MSP for age estimation. In this study, eight and five new primer sets for *ELOVL2* and ectodysplasin A receptor-associated death domain (*EDARADD*), respectively, were designed, and the best primer set was selected. The DNA methylation level was analyzed in 59 tooth samples using the selected primer set. The *ELOVL2* methylation value positively correlated with age (R^2 = 0.50), whereas the *EDARADD* methylation value negatively correlated with age (R^2 = 0.44). A multiple regression model combining *ELOVL2* and *EDARADD* showed high accuracy [mean absolute error (MAE) = 6.69], which was verified using 40 test samples (MAE = 8.28). Additionally, the MAE of three age groups showed no significant difference. These results indicate that the multiple regression model based on the two genes is useful for accurate age estimation across the human lifespan.

Keywords: age estimation, CpG island, DNA methylation, teeth, real-time PCR

Introduction

Age estimation is fundamental for building the biological profile of unidentified human remains recovered from different forensic contexts and aiding in individual identification from a list of potential candidates. Several methods have been proposed for estimating age based on biochemical markers, such as racemization of aspartic acid in dentin [1], radiocarbon dating of tooth enamel [2], measurement of the telomere length [3], and determination of the number of mitochondrial DNA mutations [4,5]. However, most of these methods have not been widely used by the forensic science community because of their low accuracy, complicated techniques, and/or limited applicability.

DNA methylation plays an essential role in various biological processes such as embryonic development, cellular differentiation, and gene expression regulation [6]. Cytosine methylation of CpG islands upstream of specific genes is developmentally regulated in a tissue-specific manner [7,8]. Over the last decade, several groups have independently evaluated the methylation levels of DNA extracted from blood and demonstrated a high correlation between DNA methylation levels and chronological age [9–16]. Some studies focused on the methylation of dental DNA to estimate age [10,17–21]. For example, Bekaert et al. [10] investigated the correlation between the methylation levels of seven CpGs in three genes [elongation of very long chain fatty acids protein 2 (*ELOVL2*), phosphodiesterase 4C (*PDE4C*), and ectodysplasin A receptor-associated death domain (*EDARADD*)] and chronological age using DNA extracted from 29 dentin samples and developed an age estimation model ($R^2 = 0.74$) with a mean absolute deviation (MAD) of 4.86 years between chronological and predicted ages. Similarly, using 65 tooth samples, Márquez-Ruiz et al. [18] developed an age estimation model based on analysis of DNA methylation of nine CpGs in two genes (*ELOVL2* and *PDE4C*) with a mean absolute error (MAE) of 5.04 years. Recently, Zapico et al. [20] reported multivariate regression models for CpGs in *ELOVL2*, four and a half LIM domains 2 (*FHL2*), neuronal pentraxin 2 (*NPTX2*), Krüppel-like factor 14 (*KLF14*), and secretagogin (*SCGN*) from 20 dental pulp samples of wisdom teeth with an MAE of 1.5–2.13 years. Dias et al. [21] reported an accurate age prediction model for CpGs in *ELOVL2* and *KLF14*, explaining 76.4% of age variation with an MAD of 7.07 years. Overall, various combinations of genes have been used to develop an optimized age estimation model. Some specific genes including *ELOVL2* that show a strong correlation with age estimation and wide range of changes in methylation during aging are promising marker candidates for age estimation. Thus, evaluation of DNA methylation has received considerable attention from forensic scientists as a novel method for age estimation [22].

Pyrosequencing is the most commonly used method in this field of investigation [23]. A major advantage of pyrosequencing is that quantitative DNA methylation data can be obtained with high accuracy

through direct sequencing of polymerase chain reaction (PCR) products [24]. However, this method has some limitations, such as the low penetration rate of the pyrosequencer and high costs.

Methylation-specific PCR (MSP), which was developed by Herman et al. [25], is a cost-effective method for rapidly assessing the methylation status of CpG sites within a CpG island. In particular, MSP analysis can be performed using a conventional real-time PCR machine.

Kondo et al. [26] recently used the real-time MSP (RT-MSP) method to quantify the methylation levels of three CpGs in *ELOVL2* extracted from 29 tooth samples. The methylation levels of *ELOVL2* showed a moderate correlation with chronological age, suggesting that RT-MSP can be applied in age estimation. The aim of the present study was to measure the methylation levels of two genes, *ELOVL2* and *EDARADD*, using RT-MSP and to develop a multiple regression model for age estimation.

Materials and methods

Sample selection

One hundred twenty-one teeth were extracted from Japanese individuals (aged 20–85 years); these teeth had been stored in the laboratory and were used in this study after obtaining approval from the Ethics Committee of Nihon University School of Dentistry (approval #EP19D007).

Of the 121 teeth, 99 were used for regression analysis using whole-tooth DNA. Fifty-nine of these 99 teeth were selected as training samples, and the remaining 40 were used as test samples. The remaining 22 teeth were used for regression analysis using DNA derived from dental pulp. All of these teeth were used as training samples. The sample distribution according to age group is shown in Table 1. All samples were permanent teeth (109 molars, 7 premolars, and 5 anterior teeth) that had been stored in a paper bag under dry conditions following extraction and preserved in a container to avoid shrinkage due to changes in temperatures and humidity. Samples with large caries, metal prostheses, and root canal treatment, as well as split teeth, were excluded.

DNA extraction and bisulfite conversion

The tooth surface was cleaned with chlorine bleach to remove stains. Tartar and soft tissue were carefully removed using a probe or scaler. Further, DNA was extracted from the 99 teeth used for regression analysis using whole-tooth DNA (i.e., without isolating any part of the tooth, or splitting or grinding of the tooth) using a DNA extraction kit for hard tissue (TBONE EX Kit; DNA Chip Research Inc., Tokyo, Japan) according to the manufacturer's protocol. The DNA was eluted with 50 μL of EB buffer (QIAamp DNA Mini Kit, Qiagen, Hilden, Germany). The dental pulp was collected from the remaining 22 teeth by splitting them under water using a diamond point. Genomic DNA was extracted from the dental pulp using the QIAamp DNA Investigator kit (Qiagen) or NucleoSpin® DNA Forensic kit (TAKARA Bio, Shiga, Japan). All of the DNA was quantified using a spectrophotometer (NanoDrop, Thermo Fisher Scientific, Waltham, MA, USA). The DNA extracts were bisulfite-converted using an EZ DNA Methylation kit (ZYMO Research, Irvine, CA, USA). The final concentration was adjusted to 20 ng/μL, and 50 μL of elution buffer from the kit was used for elution.

Percentage of methylated reference measured using RT-MSP

Methylation-specific primers were designed for CpG sites upstream of each target gene (*ELOVL2* and *EDARADD*) by referring to the MethPrimer website (http://www.urogene.org/cgi-bin/methprimer/methprimer.cgi); the primer sequences are listed in Table 2. The primers targeting *ELOVL2* and *EDARADD* contained six CpG sites (*ELOVL2*: chr6:11044644, chr6:11044647, chr6:11044655, chr6:11044661, chr6:11044711, and chr6:11044727 of GRCh38/hg38; *EDARADD*: chr1:236348188, chr1:236348190,

chr1:236348193, chr1:236348272, chr1:236348276, and chr1:236348280 of GRCh38/hg38). The human Alu sequence was used as the reference primer [27,28]. The bisulfite-converted DNA was amplified using an EpiScope® MSP Kit (TAKARA Bio) and 0.3 μM primers. RT-MSP was performed as previously described [26]. Briefly, the mean Ct value for each sample was calculated using a standard curve. The standard solutions, 100% methylated human DNA (EpiScope® Methylated HCT116 gDNA, TAKARA Bio), and 100% unmethylated human DNA (EpiScope® Unmethylated HCT116 DKO gDNA, EpiScope® Unmethylated HCT116 DKO gDNA, TAKARA Bio) were bisulfite-treated in the same manner as the sample, and the methylation rates were adjusted to 0%, 10%, 25%, 50%, 75%, 90%, and 100%. PCR was performed for 45 cycles using a Thermal Cycler Dice Real Time System II (TP900; TAKARA Bio). The annealing temperatures were 55 ℃ for *ELOVL2* and 60 ℃ for *EDARADD*. After the reaction, the percentage of methylated reference (PMR) was calculated from the fluorescence level (relative quantity) of each sample using the following equation:

PMR (%) = {[(target gene†/Alu) mean value sample]/[(target gene†/Alu) mean value universal methylated human DNA]} × 100

† represents target *ELOVL2* or *EDARADD*.

Statistical analysis

The correlations between PMR and chronological age were examined using 59 teeth samples from the training samples, and regression equations were obtained (Microsoft Excel 2019, Microsoft, Redmond, WA, USA). To compare the residuals of the regression equations among age groups, the samples were divided into three age groups (young adults: 20–34 years; middle-aged adults: 35–54 years; and older adults: ≥55 years), and the MAE of each group was calculated and tested for significant differences between groups [one-way analysis of variance (ANOVA) with post-hoc Tukey–Kramer test in GraphPad Prism 9 software (GraphPad, Inc., San Diego, CA, USA). Multiple regression analysis was performed between chronological age and the PMR of *ELOVL2* and *EDARADD* (Microsoft Excel 2019). The training samples were divided into men and women, and multiple regression equations were obtained for each sex. The significance of the difference between the correlation coefficients was analyzed using VassarStats (http://vassarstats.net/). Forty teeth were used to verify the accuracy of the multiple regression equations. To compare the test sample residuals from the multiple regression equation among age groups, the MAE of each age group was calculated, and a significance test among age groups was performed (one-way ANOVA with post-hoc Tukey–Kramer test, GraphPad Prism 9). Statistical significance was set at $p < 0.05$.

Finally, 22 dental pulps were used to examine the correlation between PMR and age for the two genes（*ELOVL2* and *EDARADD*）and single or multiple regression equations were obtained.

Results

Correlation between age and PMR of ELOVL2 *and* EDARADD *derived from whole teeth*

The correlation between the PMR of the two genes (*ELOVL2* and *EDARADD*) and chronological age was assessed for the 59 whole-tooth training samples (Fig. 1). The PMR of *ELOVL2* was strongly positively correlated with age, whereas that of *EDARADD* was negatively correlated with age. A quadratic regression was a better fit for the relationship between age and PMR of *ELOVL2* ($R^2 = 0.50$) than a linear ($R^2 = 0.42$) or logarithmic ($R^2 = 0.42$) regression. For *EDARADD*, logarithmic regression was a better fit ($R^2 = 0.44$) than linear ($R^2 = 0.40$) or quadratic ($R^2 = 0.42$) regression. The regression equations were as follows: $Y = -0.0784X^2 + 3.3587X + 23.529$ for *ELOVL2* and $Y = -13.91\ln(X) + 92.259$ for *EDARADD*.

Validation of regression equations using the training sample

Each regression equation yielded MAEs of 9.59 years (*ELOVL2*) and 10.12 years (*EDARADD*) in the training sample. In the regression equations for the PMR of *ELOVL2* and *EDARADD*, the residuals were biased toward positive values for individuals in their 20s and 30s and toward negative values for those in their 60s (Fig. 2).

In the regression equations for *ELOVL2*, the largest residual was 12.67 years, which was observed in older adults (≥55 years), and the smallest residual was 8.02 years, which was observed in young adults (20–34 years; Fig. 3a). In contrast, in the regression equations for *EDARADD*, the largest residual was 12.20 years, which was observed in older adults (≥55 years), and the smallest residual was 8.00 years, which was observed in middle-aged adults (35–54 years; Fig. 3b). There was no significant difference between any of the age groups for *ELOVL2* and *EDARADD* (one-way ANOVA with post-hoc Tukey–Kramer test).

Multiple regression analysis using the PMR of ELOVL2 *and* EDARADD *and accuracy verification using a test sample*

PMR values for *ELOVL2* and *EDARADD* were used in a multivariate regression model to obtain an age estimation formula. The multiple regression equation was as follows: $Y = -0.041X_1^2 + 2.243X_1 - 10.815\ln(X_2) + 66.538$ ($R^2 = 0.74$; X_1: PMR of *ELOVL2*, X_2: PMR of *EDARADD*). Age was estimated using the multiple regression equation (Fig. 4). The model produced an MAE between the estimated and chronological age of 6.69 years for the training samples (n = 59). The correlation between the PMR and chronological age did not significantly differ between men and women (two correlation coefficients test, $Z = -1.18$, $P = 0.24$, two-tailed). In addition, the multiple regression equation was validated using a test sample (n = 40), yielding an MAE of 8.28 years. The MAE was 7.61 years for young adults (n = 14), 7.48 years for middle-aged adults (n =

15), and 10.21 years for older adults (n = 11) (Fig. 5). No significant differences were observed between any age of the groups (one-way ANOVA with post-hoc Tukey–Kramer test).

Age estimation based on PMR of ELOVL2 *and* EDARADD *derived from dental pulp*

The correlation between the PMR of two genes (*ELOVL2* and *EDARADD*) and chronological age was assessed for the 22 dental pulp samples (Fig. 6). Similar to those observed in the whole tooth, the PMR of *ELOVL2* showed strong positive correlation with age (Fig. 6a), whereas that of *EDARADD* showed weak negative correlation with age (Fig. 6b). The calculated regression equations were as follows: $Y = -0.1709X^2 + 5.6239X + 14.8$ ($R^2 = 0.64$) for *ELOVL2*, and $Y = -18.04\ln(X) + 115.27$ ($R^2 = 0.12$) for *EDARADD*. The single regression analysis revealed that the PMR of dental pulp-derived *ELOVL2* exhibited a stronger correlation with age than that of whole-tooth-derived *ELVOL2*. In contrast, the PMR of *EDARADD* exhibited a stronger correlation with chronological age in the whole teeth than in the dental pulp. The MAEs between estimated and chronological age were 6.42 years (for *ELOVL2*) and 11.42 years (for *EDARADD*) for the 22 dental pulp training samples.

Based on the PMR of *ELOVL2* and *EDARADD*, the multiple regression equation was calculated as follows: $Y = -0.162(X_1^2) + 5.351(X_1) - 11.847\ln(X_2) + 64.701$ (X_1: PMR of *ELOVL2*, X_2: PMR of *EDARADD*). The multiple regression equation showed high accuracy ($R^2 = 0.69$), with an MAE between the estimated and chronological age of 5.86 years (Fig. 7).

Discussion

Over the last three decades, various molecular markers for age estimation have been identified for use in the forensic field. Particularly, racemization of aspartic acid in a dental protein was tested for forensic applications [1]. However, aspartic acid racemization is affected by temperature, which is a limitation of this analysis. Alternatively, DNA methylation is relatively stable at different temperatures, suggesting that it can be used to estimate the chronological age of burnt remains.

Epigenetic age estimation using pyrosequencing [10,11,20], EpiTyper [17,29], SNaPshot [19, 21, 30,31], and methylation-sensitive high-resolution melting (MS-HRM) [13] have been tested for practical applications in the forensic field [32].MS-HRM, a PCR-based method, enables rapid and relatively inexpensive quantification of methylation levels. DNA methylation levels are estimated by comparing the melting profiles of an unknown PCR product with those of a PCR product derived from a standard material with a known methylation ratio [33]. Recently, Hamano et al. [13] used MS-HRM to quantify the methylation levels of *ELOVL2* and *FHL2* in blood-derived DNA and developed a multiple regression model with an MAD of 7.44 years.

RT-MSP, a PCR-based methylation analysis similar to MS-HRM, has been widely used in medical sciences to diagnose cancer and some syndromes [25,34]. This method can also be used to rapidly assess the methylation values of virtually any group of CpG sites within a CpG island using methylation-specific primer sets for the sequence of interest. The greatest advantage of this method is that the analysis can be performed using a conventional real-time PCR machine, without requiring specific equipment such as a pyrosequencer. Therefore, this method can be used in primary screening for personal identification, such as following a large-scale disaster, where a large number of cadavers must be analyzed immediately. This study showed that RT-MSP can be used to quantify CpG methylation levels and estimate chronological age.

The selection of appropriate genes is crucial for developing a highly accurate age estimation model. To date, at least 10 genes derived from teeth have been considered as targets for developing age estimation models [5,10,14,17–21]. *ELOVL2* is the most promising gene for creating an age estimation model because its methylation values are strongly positively correlated with age [10–12,15]. In addition, this correlation is highly conserved in multiple tissues [35]. In contrast to *ELOVL2* methylation, *EDARADD* methylation is negatively correlated with age [10,36,37]. *EDARADD* has also been used for methylation-based age estimation in tooth samples [10,19]. Therefore, *ELOVL2* and *EDARADD* were used to develop an age estimation model.

Eight and five sets of primers were designed for *ELOVL2* and *EDARADD*, respectively. Each primer contained 2–4 CpGs of 161 (*ELOVL2*) and 136 (*EDARADD*) CpG sites on CpG islands, with each primer set consisting of forward and reverse primers, containing 4–8 CpGs with various combinations. The best primer set

for each gene was selected based on a single peak in the dissociation curve and correct calibration curve (R^2 = 0.89, *ELOVL2*; R^2 = 0.89, *EDARADD*).

The final primer sets for *ELOVL2* and *EDARADD* each contained six CpGs. Single regression analyses of individual genes revealed moderate correlations (R^2 = 0.50, *ELOVL2*; R^2 = 0.44, *EDARADD*). Moreover, the final multiple regression model combining six CpGs in *ELOVL2* and six CpGs in *EDARADD* (i.e., 12 CpGs in total) showed a high correlation (R^2 = 0.74) and an MAE of 6.69 years.

In the narrow *ELOVL2* upstream region, dozens of consecutive CpG sites have been identified to be specifically correlated with age [11,38]. Interestingly, the CpG showing the highest correlation with age differed in each analysis, even when the same tissue (blood) and same method (pyrosequencing) were used. For example, Zbiec-Piekarska et al. [11] found that two of the seven CpGs (chr6:11,044,642 and chr6:11,044,634) were strongly correlated with chronological age; they also developed an age estimation model (R^2 = 0.859, MAD = 5.03) using blood samples. In contrast, Bekaert et al. [10] showed that three CpGs (chr6:11,044,661, chr6:11,044,640, and chr6:11,044,625) in *ELOVL2* were correlated in their age estimation model, with chr6: 11,044,640 as the most informative CpG in the blood sample. Park et al. [15] found a strong correlation between age and the CpG (chr 6: 11,044,894) in *ELOVL2*. Similar to the results observed in blood, CpG sites showing the best correlation with age differed in each model in teeth [10,18,21]. Bekaert et al. [10], Marquez-Ruiz et al. [18], and Dias et al. [21] used five CpGs (chr6:11,044,655, chr6:11,044,640, chr6:11,044,634, chr6:11,044,628, and chr6:11,044,625), six CpGs (chr6:11,044,617; chr6:11,044,631; chr6:11,044,642; chr6:11,044,644; chr6:11,044,649; chr6:11,044,661), and one CpG (chr6:11,044,628), respectively. The difference in the optimal CpG sites in each model may be attributed to the stochastic phenomenon of DNA methylation. In other words, the best CpG site may vary when samples are added to the analysis. Three (chr6:11,044,655, chr6:11,044,647, chr6:11,044,644) of the six CpGs contained in the primer set for *ELOVL2* used in this study were previously used for age estimation analysis, suggesting the usefulness of the region containing nine CpGs.

In contrast, a minimum of six CpGs in *EDARADD* has been used to develop epigenetic age estimation models [10,36,37]. To the best of my knowledge, the six CpGs in my *EDARADD* primers have not been previously used for age estimation analyses.

Blood is relatively easy to obtain and has been used for DNA methylation-based age estimation [5]. Compared with blood, teeth are a stable source of DNA, even in severely damaged cadavers (e.g., highly decomposed corpses, burned corpses, and white skeletons). The molars are the most protected teeth in the jaw, and thus are useful for assessing age by analyzing their methylation levels.

In this study, teeth that had not been crushed or split were used to extract DNA, which is required for rapid processing in the forensic field. Therefore, the origin of the DNA is expected to be from dentin

(odontoblasts), dental pulp (odontoblasts, fibroblasts, defense cells, and undifferentiated mesenchymal cells), and cementum (cementocytes) [39].

Giuliani et al. [17] divided the tooth into three parts, dentin, pulp, and cementum, and built an age estimation model for each. The MAD between the estimated age and chronological age was better in the model from the pulp (MAD = 2.25) or cementum (MAD = 2.45) compared to that from the dentin (MAD = 7.07). Recently, Zapico et al. [20] developed excellent multiple regression models (MAE = 1.5−2.13 years) by quantifying DNA methylation in dental pulp. The results suggested that using dental pulp can provide the most accurate estimate of age.

To validate the results, a regression model based on altering the methylation of dental pulp-derived DNA (*ELOVL2* and *EDARADD*) was also developed by measuring methylation using the RT-MSP method. Consistent with previous reports, the age estimation model based on dental pulp-derived *ELOVL2* methylation levels showed higher accuracy compared with that based on whole-tooth-derived *ELOVL2* methylation levels. However, the accuracy of the age estimation model based on the *EDARADD* methylation level was in conflict with that of *ELOVL2*.

The use of dental pulp as an origin of DNA to develop multiple regression models was only slight improvement compared with the use of whole teeth. In the multiple regression analysis based on the PMR of the two genes from dental pulp, the MAE was slightly smaller than that from whole teeth, but the coefficient of determination was lower. Therefore, it was difficult to determine whether limiting the DNA source to dental pulp would improve the accuracy of age estimation.

In contrast, when collecting DNA, a certain amount of time is required to divide the teeth. Given the rapid processing needs in the forensic field, such as following large-scale disasters, DNA extraction from the whole tooth is a suitable method. In addition, as the small amount of DNA commonly found in forensic cases increases the margin of error when determining DNA methylation levels, extracting DNA from the whole tooth may increase the total amount of DNA collected. Thus, the use of whole tooth-derived DNA for age estimation has many advantages in forensic practice.

Furthermore, to develop a highly accurate regression model, it is important to consider a sufficient sample size or population variability in each age group. However, in most previously reported experiments, the number of tooth samples used was only 20–30, and the sample distribution in each age group was not always uniform. In contrast, I used 99 samples (59 training samples and 40 test samples) and observed a low level of sample variability in each age group, indicating that my experimental conditions are suited for developing an age estimation model.

Both highly accurate age estimation models and homogeneous estimation accuracy across generations are important in the forensic field. Previous studies showed that age estimation accuracy decreases in an age-

dependent manner [10,18]. In the present study, no statistically significant difference was observed in the MAE of each age group using the multiple regression model for age estimation (Fig. 5). These results suggest that my model is stable across the human lifespan. Therefore, this method may be useful for identifying missing people (e.g., in criminal cases), which mostly includes adults and the elderly.

Age estimation, which is based on various biological or molecular characteristics including DNA methylation, is a scoring method used as part of individual identification. The final personal identification is performed by comparing postmortem and ante mortem dental charts and/or through DNA identification. In this study, the multiple regression model for age estimation was developed using RT-MSP to evaluate DNA methylation in tooth samples. This model showed a slightly higher MAE and was less accurate than other age estimation methods. However, as described above, the method can be easily performed at low costs using conventional real-time PCR. Therefore, this method can be used in primary screening for personal identification in situations such as large-scale disasters, where large numbers of cadavers must be analyzed immediately.

Conclusions

I assessed the methylation of *ELOVL2* and *EDARADD* in 59 tooth samples using RT-MSP and developed a multiple regression model, the MAE value of which was 6.69 years. The accuracy of the age estimation model was verified using an additional 40 extracted teeth; the obtained MAE value was 8.28 years, which was close to that obtained for the training samples. This result indicates that my multiple regression model can be used for dental age estimation.

Acknowledgments

I am grateful to Prof. Naoto Suzuki and Project Prof. Hirofumi Aboshi for the opportunity to perform this study, Assoc. Prof. Masahiro Kondo for their instructions of this study, and colleagues in Department of Legal Medicine for their technical advice and assistance.

References

[1] S. Ohtani, K. Yamamoto, Estimation of age from a tooth by means of racemization of an amino acid, especially aspartic acid - Comparison of enamel and dentin, J. Forensic Sci. 37(4) (1992) 1061–1067. https://www.astm.org/jfs13292j.html.

[2] K.L. Spalding, B.A. Buchholz, L.E. Bergman, H. Druid, J. Frisén, Forensics: age written in teeth by nuclear tests, Nature. 437(7057) (2005) 333–334. https://doi.org/10.1038/437333a.

[3] A.O. Karlsson, A. Svensson, A. Marklund, G. Holmlund, Estimating human age in forensic samples by analysis of telomere repeats, Forensic Sci. Int. Genet. Suppl. Ser. 1(1) (2008) 569–571. https://doi.org/10.1016/j.fsigss.2007.10.153.

[4] C. Meissner, N. von Wurmb, M. Oehmichen, Detection of the age-dependent 4977 bp deletion of mitochondrial DNA, A pilot study. Int. J. Legal Med. 110(5) (1997) 288–291. https://doi.org/10.1007/s004140050089.

[5] A. Vidaki, B.Daniel, D.S. Court, Forensic DNA methylation profiling - Potential opportunities and challenges, Forensic Sci. Int. Genet. 7(5) (2013) 499–507. https://doi.org/10.1016/j.fsigen.2013.05.004.

[6] A.M. Deaton, A. Bird, CpG islands and the regulation of transcription, Genes Dev. 25(10) (2011) 1010–1022. https://doi.org/10.1101/gad.2037511.

[7] S. Horvath, DNA methylation age of human tissues and cell types, Genome Biol. 14(10) (2013) R115. https://doi.org/10.1186/gb-2013-14-10-r115.

[8] D. Schübeler, Function and information content of DNA methylation, Nature. 517(7534) (2015) 321–326. https://doi.org/10.1038/nature14192.

[9] 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(2) (2014) R24. https://doi.org/10.1186/gb-2014-15-2-r24.

[10] B. Bekaert, A. Kamalandua, S.C. Zapico, W. Van de Voorde, R. Decorte, Improved age determination of blood and teeth samples using a selected set of DNA methylation markers, Epigenetics. 10(10) (2015) 922–930. https://doi.org/10.1080/15592294.2015.1080413.

[11] R. Zbieć-Piekarska, M. Spólnicka, T. Kupiec, Z. 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. https://doi.org/10.1016/j.fsigen.2014.10.002.

[12] R. Zbieć-Piekarska, M. Spólnicka, T. Kupiec, A. Parys-Proszek, Z. 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. https://doi.org/10.1016/j.fsigen.2015.05.001.

[13] Y. Hamano, S. Manabe, C. Morimoto, S. Fujimoto, M. Ozeki, K. Tamaki, Forensic age prediction for dead or living samples by use of methylation-sensitive high resolution melting, Leg. Med. (Tokyo). 21 (2016) 5–10. https://doi.org/10.1016/j.legalmed.2016.05.001.

[14] H.Y. Lee, S.D. Lee, K.J. Shin, Forensic DNA methylation profiling from evidence material for investigative leads, BMB Rep. 49(7) (2016) 359–369. https://doi.org/10.5483/bmbrep.2016.49.7.070.

[15] J.L. Park, J.H. Kim, E. Seo, D.H. Bae, S.Y. Kim, H.C. Lee, K.M. Woo, Y.S. Kim, Identification and evaluation of age-correlated DNA methylation markers for forensic use, Forensic Sci. Int. Genet. 23 (2016) 64–70. https://doi.org/10.1016/j.fsigen.2016.03.005.

[16] Z. Thong, X.L.S. Chan, J.Y.Y. Tan, E.S. Loo, C.K.C. Syn, Evaluation of DNA methylation-based age prediction on blood, Forensic Sci. Int. Genet. Suppl. Ser. 6 (2017) e249–e251. https://doi.org/10.1016/J.FSIGSS.2017.09.095.

[17] C. Giuliani, E. Cilli, M.G. Bacalini, C. Pirazzini, M. Sazzini, G. Gruppioni, C. Franceschi, P. Garagnani, D. Luiselli, Inferring chronological age from DNA methylation patterns of human teeth, Am. J. Phys. Anthropol. 159(4) (2016) 585–595. https://doi.org/10.1002/ajpa.22921.

[18] A.B. Márquez-Ruiz, L. González-Herrera, J.D. Luna, A. Valenzuela, DNA methylation levels and telomere length in human teeth: usefulness for age estimation, Int. J. Legal Med. 134(2) (2020) 451–459. https://doi.org/10.1007/s00414-019-02242-7.

[19] H.C. Dias, L. Manco, F. Corte-Real, E. Cunha, A blood-bone-tooth model for age prediction in forensic contexts, Biology (Basel). 10(12) (2021) 1312. https://doi.org/10.3390/biology10121312.

[20] S.C. Zapico, Q. Gauthier, A. Antevska, B.R. McCord, Identifying methylation patterns in dental pulp aging: Application to age-at-death estimation in forensic anthropology, Int. J. Mol. Sci. 22(7) (2021) 3717. https://doi.org/10.3390/ijms22073717.

[21] H.C. Dias, F. Corte-Real, E. Cunha, L. Manco, DNA methylation age estimation from human bone and teeth, Aust. J. Forensic Sci. 54(2) (2022) 163–176. https://doi.org/10.1080/00450618.2020.1805011.

[22] W. Parson, Age estimation with DNA: From forensic DNA fingerprinting to forensic (epi)genomics: A mini-review, Gerontology. 64(4) (2018) 326–332. https://doi.org/10.1159/000486239.

[23] A. Freire-Aradas, C. Phillips, M.V. Lareu, Forensic individual age estimation with DNA: From initial approaches to methylation tests, Forensic Sci. Rev. 29(2) (2017) 121–144.

[24] J. Tost, I.G. Gut, Analysis of gene-specific DNA methylation patterns by pyrosequencing technology, Methods Mol. Biol. 373 (2007) 89–102. https://doi.org/10.1385/1-59745-377-3:89.

[25] J.G. Herman, J.R. Graff, S. Myöhänen, B.D. Nelkin, S.B. Baylin, Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands, Proc. Natl. Acad. Sci. U. S. A. 93(18) (1996) 9821–9826. https://doi.org/10.1073/pnas.93.18.9821.

[26] M. Kondo, H. Aboshi, M. Yoshikawa, A. Ogata, R. Murayama, M. Takei, S. Aizawa, A newly developed age estimation method based on CpG methylation of teeth-derived DNA using real-time methylation-specific PCR, J. Oral Sci. 63(1) (2020) 54–58. https://doi.org/10.2334/josnusd.20-0138.

[27] D.J. Weisenberger, M. Campan, T.I. Long, M. Kim, C. Woods, E. Fiala, M. Ehrlich, P.W. Laird, Analysis of repetitive element DNA methylation by MethyLight, Nucleic Acids Res. 33(21) (2005) 6823–6836. https://doi.org/10.1093/nar/gki987.

[28] T. Niwa, T. Tsukamoto, T. Toyoda, A. Mori, H. Tanaka, T. Maekita, M. Ichinose, M. Tatematsu, T. Ushijima, Inflammatory processes triggered by Helicobacter pylori infection cause aberrant DNA methylation in gastric epithelial cells, Cancer Res. 70(4) (2010) 1430–1440. https://doi.org/10.1158/0008-5472.CAN-09-2755.

[29] S. Kunze, Quantitative region-specific DNA methylation analysis by the EpiTYPER™ technology, Methods Mol. Biol. 1708 (2018) 515–535. https://doi.org/10.1007/978-1-4939-7481-8_26.

[30] H.Y. Lee, S.E. Jung, Y.N. Oh, A. Choi, W.I. Yang, K.J. Shin, Epigenetic age signatures in the forensically relevant body fluid of semen: a preliminary study, Forensic Sci. Int. Genet. 19 (2015) 28–34. https://doi.org/10.1016/j.fsigen.2015.05.014.

[31] S.E. Jung, S.M. Lim, S.R. Hong, E.H. Lee, K.J. Shin, H.Y. Lee, DNA methylation of the ELOVL2, FHL2, KLF14, C1orf132/MIR29B2C, and TRIM59 genes for age prediction from blood, saliva, and buccal swab samples, Forensic Sci. Int. Genet. 38 (2019) 1–8. https://doi.org/10.1016/j.fsigen.2018.09.010.

[32] A. Vidaki, M. Kayser, Recent progress, methods and perspectives in forensic epigenetics, Forensic Sci. Int. Genet. 37 (2018) 180–195. https://doi.org/10.1016/j.fsigen.2018.08.008.

[33] T.K. Wojdacz, A. Dobrovic, L.L. Hansen, Methylation-sensitive high-resolution melting, Nat. Protoc. 3(12) (2008) 1903–1908. https://doi.org/10.1038/nprot.2008.191.

[34] Y. Pan, G. Liu, F. Zhou, B. Su, Y. Li, DNA methylation profiles in cancer diagnosis and therapeutics, Clin. Exp. Med. 18(1) (2018) 1–14. https://doi.org/10.1007/s10238-017-0467-0.

[35] R.C. Slieker, C.L. Relton, T.R. Gaunt, P.E. Slagboom, B.T. Heijmans, Age-related DNA methylation changes are tissue-specific with ELOVL2 promoter methylation as exception, Epigenetics Chromatin. 11(1) (2018) 25. https://doi.org/10.1186/s13072-018-0191-3.

[36] S. Bocklandt, W. Lin, M.E. Sehl, F.J. Sánchez, J.S. Sinsheimer, S. Horvath, E. Vilain, Epigenetic predictor of age, PLoS One. 6(6) (2011) e14821. https://doi.org/10.1371/journal.pone.0014821.

[37] A. Freire-Aradas, C. Phillips, L. Girón-Santamaría, A. Mosquera-Miguel, A. Gómez-Tato, M.A. Casares de Cal J. Álvarez-Dios, M.V. Lareu, Tracking age-correlated DNA methylation markers in the young, Forensic Sci. Int. Genet. 36 (2018) 50–59. https://doi.org/10.1016/j.fsigen.2018.06.011.

[38] P. Sukawutthiya, T. Sathirapatya, K. Vongpaisarnsin, A minimal number CpGs of ELOVL2 gene for a chronological age estimation using pyrosequencing, Forensic Sci. Int. 318 (2021) 110631. https://doi.org/10.1016/j.forsciint.2020.110631.

[39] D. Higgins. J.J. Austin, Teeth as a source of DNA for forensic identification of human remains: a review, Sci. Justice. 53(4) (2013) 433–441. https://doi.org/10.1016/j.scijus.2013.06.001.

Tables

Table 1. Number of samples and sex distribution per age group.

a

Age (years)	Frequency	Men	Women
20–29	14	6	8
30–39	12	6	6
40–49	12	8	4
50–59	8	7	1
60–69	7	5	2
≥70	6*	1	4
Total	59	33	25

b

Age (years)	Frequency	Men	Women
20–29	8	3	5
30–39	9	4	5
40–49	8	7	1
50–59	6	3	3
60–69	5	4	1
≥70	4	3	1
Total	40	24	16

c

Age (years)	Frequency	Men	Women
20–29	6	4	2
30–39	4	1	3
40–49	5	3	2
50–59	4	1	3
60–69	3	1	2
≥70	0	0	0
Total	22	10	12

a: training samples of whole teeth (n = 59). **b**: test samples of whole teeth (n = 40). **c**: training samples of dental pulp (n = 22) *One sample in the ≥70 group was of unknown sex.

Table 2. Sequences of methylation-specific and reference primers used in methylation-specific polymerase chain reaction.

Target	Primer	Sequence (5′-3′)
ELOVL2	ELOVL2-F	GCGGCGGTTTAACGTTTAC
	ELOVL2-R	CACGATACTACTTCTCCCCG
EDARADD	EDARADD-F	GTAGATGTTAGGCGCGGC
	EDARADD-R	CCCTACCTTACGATCGTCCG
Alu	ALU-F	GGTTAGGTATAGTGGTTTATATTTGTAATTTTAGTA
	ALU-R	ATTAACTAAACTAATCTTAAACTCCTAACCTCA

*Methylation specific primers: *ELOVL2*-F, *ELOVL2*-R, *EDARADD*-F, and *EDARADD*-R; reference primers: ALU-F and ALU-R.

†*ELOVL2*, elongation of very long-chain fatty acids protein 2; *EDARADD*, ectodysplasin A receptor-associated death domain.

Figures

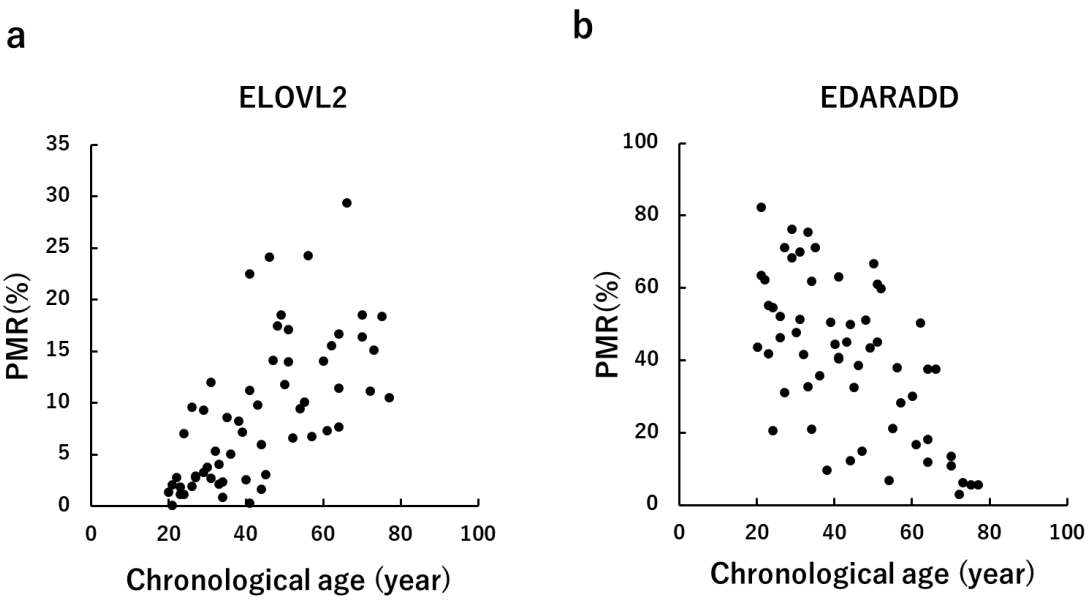

Fig. 1. Correlation between the percentage of methylated reference (PMR) and chronological age in the training samples (n = 59).

a: PMR of *ELOVL2* positively correlated with age. **b**: PMR of *EDARADD* negatively correlated with age.

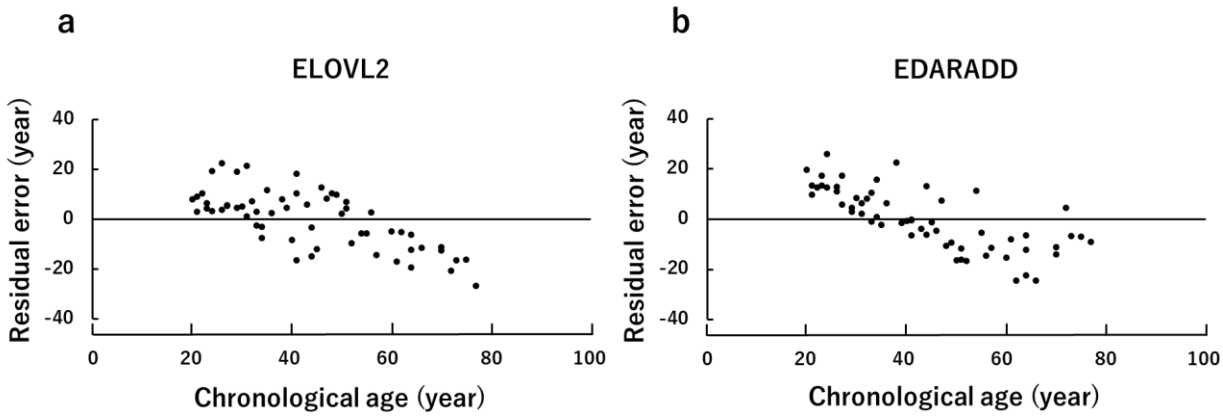

Fig. 2. Residuals of the training sample (n = 59) using a simple regression equation.

a: Mean absolute deviation (MAE) was 9.59 years. **b**: MAE was 10.12 years.

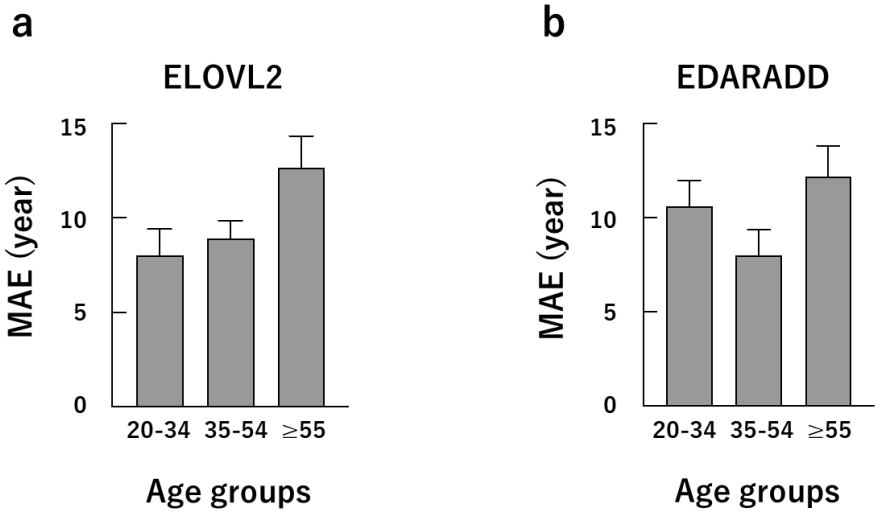

Fig. 3. Mean absolute deviation (MAE) per age group for the training sample (n = 59) using the simple regression equation (young adults, middle-aged adults, and older adults; n = 22, 21, and 16, respectively).

a: No significant difference was observed in *ELOVL2* between any age groups. **b**: No significant difference was observed in *EDARADD* between any age groups.

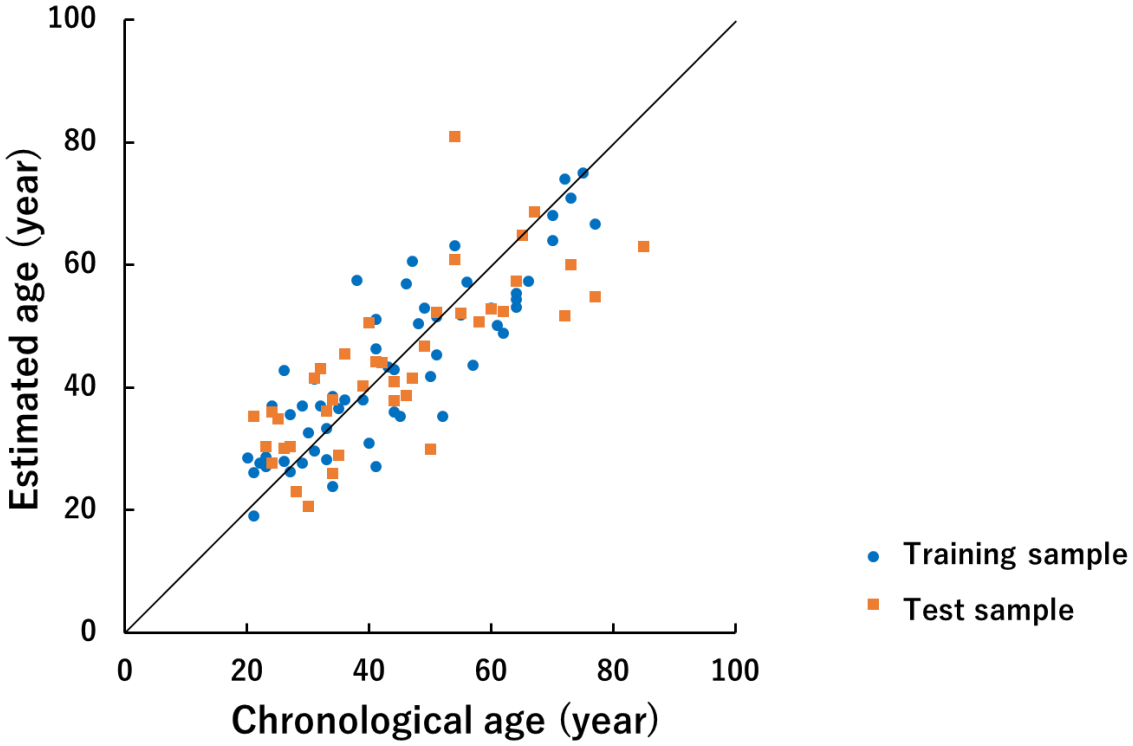

Fig. 4. Relationship between estimated and chronological age of the training and test samples using multiple regression equations.

In the training sample (n = 59), the mean absolute deviation (MAE) was 6.69 years. In the test sample (n = 40), the MAE was 8.28 years.

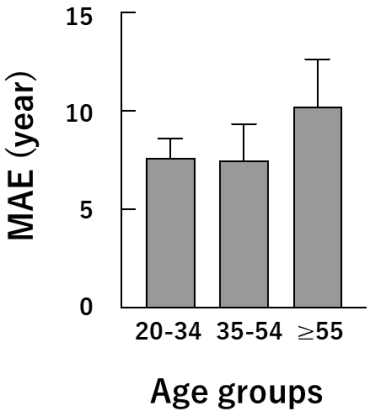

Fig. 5. Mean absolute deviation (MAE) per age group for the test sample (n = 40) using a multiple regression equation. No statistically significant difference was observed between any age groups (young adults, middle-aged adults, and older adults; n = 14, 15, and 11, respectively).

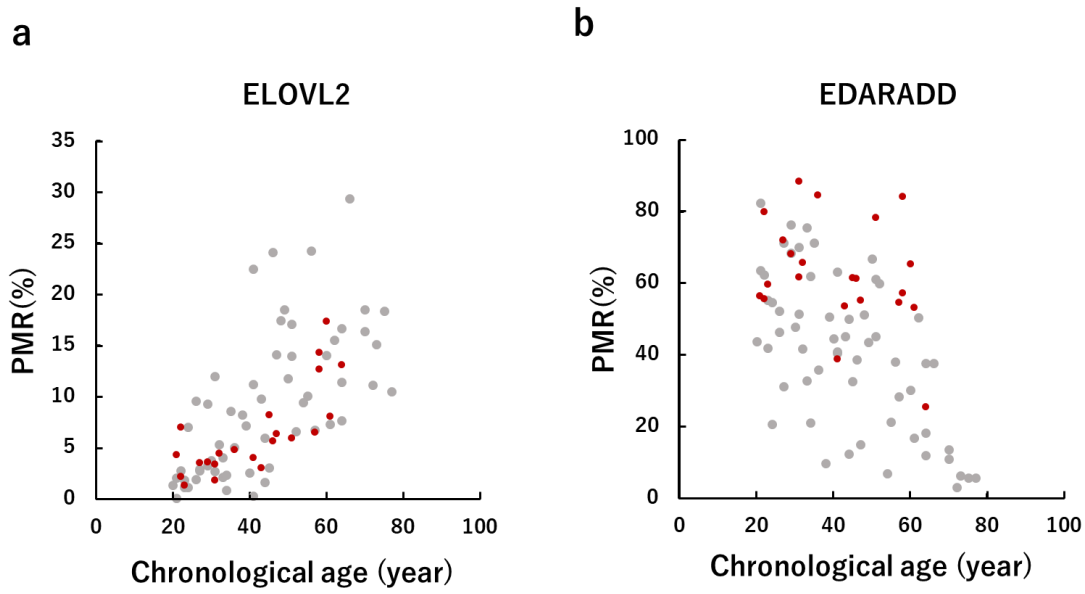

Fig. 6. Correlation between the percentage of methylated reference (PMR) and chronological age in the training samples of dental pulp (n = 22).

Red markers: dental pulp results, Gray markers: whole teeth results. **a**: PMR of *ELOVL2* positively correlated with age. **b**: PMR of *EDARADD* negatively correlated with age.

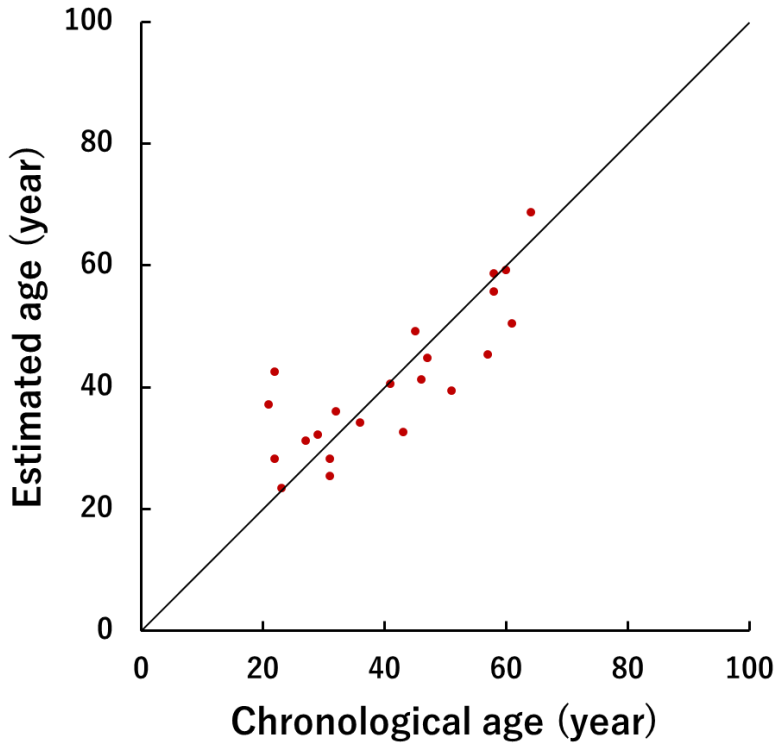

Fig. 7. Relationship between estimated and chronological age of the samples of dental pulp using multiple regression equations.

In the sample (n = 22), the mean absolute deviation (MAE) was 5.86 years.