WILEY-VCH

ELECTROPHORESIS

RESEARCH ARTICLE OPEN ACCESS

A Ge.F.I. Collaborative Study: Evaluating Reproducibility and Accuracy of a DNA-Methylation-Based Age-Predictive Assay for Routine Implementation in Forensic Casework

Martina Onofri[1] | Federica Alessandrini[2] | Serena Aneli[3] | Loredana Buscemi[2] | Elena Chierto[3] | Matteo Fabbri[4] | Paolo Fattorini[5] | Paolo Garofano[6] | Fabiano Gentile[7] | Silvano Presciuttini[8] | Carlo Previderè[9] | Carlo Robino[3] | Simona Severini[10] | Federica Tommolini[10] | Pamela Tozzo[11] | Andrea Verzeletti[12] | Eugenia Carnevali[10]

[1]Section of Legal Medicine, Department of Medicine and Surgery, University of Perugia, Terni, Italy | [2]Department of Biomedical Sciences and Public Health, Polytechnic University of Marche, Ancona, Italy | [3]Department of Public Health Sciences and Pediatrics, University of Turin, Turin, Italy | [4]Section of Legal Medicine, Department of Translational Medicine, University of Ferrara, Ferrara, Italy | [5]Department of Medicine, Surgery and Health, University of Trieste, Trieste, Italy | [6]Forensic Genetics Laboratory – Regional Antidoping Centre "A. Bertinaria", Orbassano, Italy | [7]Reparto Carabinieri Investigazioni Scientifiche di Parma, Biology Section, Parma, Italy | [8]Department of Translational Research and New Technologies in Medicine and Surgery, University of Pisa, Pisa, Italy | [9]Department of Public Health, Experimental and Forensic Medicine, University of Pavia, Pavia, Italy | [10]Forensic Sciences Laboratory, Section of Legal Medicine, Department of Medicine and Surgery, Santa Maria Hospital, University of Perugia, Terni, Italy | [11]Department of Cardiac, Thoracic, Vascular Sciences and Public Health, University of Padova, Padova, Italy | [12]Institute of Legal Medicine of Brescia, University of Brescia, Brescia, Italy

Correspondence: Martina Onofri (martina.onofri@live.it)

Received: 28 August 2024 | **Accepted:** 24 December 2024

Funding: The authors received no specific funding for this work.

Keywords: age prediction | CpG marker | DNA methylation | inter-laboratory validation | Single Base Extension

ABSTRACT

The increasing interest in DNA methylation (DNAm) analysis within the forensic scientific community prompted a collaborative project by Ge.F.I. (Genetisti Forensi Italiani). The study evaluated a standardized bisulfite conversion–based Single Base Extension (SBE) protocol for the analysis of the methylation levels at five age-predictive loci (ELOVL2, FHL2, KLF14, C1orf132/MIR29B2C, and TRIM59). The study encompassed three phases: (1) setting up and validating the protocol to ensure consistency and reproducibility; (2) comparing fresh peripheral blood with blood spots; and (3) evaluating sources of intra- and inter-laboratory variability. Samples from 22 Italian volunteers were analyzed by 6 laboratories in replicates for a total of 528 records. From phase I emerged that the choice of genetic sequencer significantly contributed to inter-laboratory data variation, resulting in separate regression analyses performed for each laboratory. In phase II, blood spots were found to be a reliable source for DNAm analysis, despite exhibiting increased experimental variation compared to fresh peripheral blood. In phase III, a strong correlation between the individual's predicted and true ages was observed across different laboratories. Analysis of variance (ANOVA) of the residuals indicated that one-third of the total variance could be attributed to laboratory-specific factors, whereas two-thirds could be attributed to inter-individual biological differences. The leave-one-out cross-validation (LOO-CV) method yielded an overall mean absolute deviation (MAD) value of 4.41 years, with an average 95% confidence interval of 5.24 years. Stepwise regression analysis proved that a restricted model (ELOVL2, C1orf132/MIR29B2C, and TRIM59) produced results virtually indistinguishable from the five-loci model. Additionally, the analysis of samples in replicates greatly improved the fit of the regression model, balancing the slight effects of intra-laboratory variability. In conclusion, the bisulfite conversion–based SBE protocol, combined with replicate

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2025 The Author(s). *Electrophoresis* published by Wiley-VCH GmbH.

analysis and in-lab calibration of a regression-prediction model, proves to be a reliable and easily implementable method for age prediction in forensic laboratories.

1 | Introduction

In the last decade, epigenetic modifications, such as DNA methylation (DNAm) patterns, have become the object of a rapidly growing interest from the forensic scientific community. In particular, DNAm consists of the transfer of a methyl group to the 5′ carbon of a cytosine (C), producing a 5-methyl-cytosine (5mC). For the most part, DNAm occurs at CpG sites, which are dinucleotide sequences consisting of a cytosine followed by a guanine [1]. For forensic applications, bisulfite-dependent sequencing techniques are mainly employed for DNAm analysis [2, 3]. Bisulfite conversion is, as of today, the most widely used method for "transferring" the epigenetic information into a nucleotide sequence modification [4, 5]. Treating the DNA with sodium bisulfite converts unmethylated cytosines into uracils (U), whereas methylated cytosines remain unchanged. A subsequent polymerase chain reaction (PCR) step substitutes uracils into thymines (T) [4, 5], and the region of interest is then sequenced through a variety of methods, such as Single Base Extension (SBE). SBE, also known as minisequencing, consists of a primer extension reaction in which primers that bind 1 bp upstream to the CpG site of interest and dideoxynucleotides (ddNTP) are used, thus typing exclusively the target nucleotide. Despite the availability of next-generation sequencing (NGS) assays [6–9], SBE sequencing is still frequently employed, thanks to its multiplexing ability, single-base resolution, and ease of implementation in most forensic genetics laboratories [10–14].

Given its hereditability, reversible nature, and modulation due to intrinsic and extrinsic factors, DNAm is currently being studied for implementation in forensic genetics analysis for tissue and body fluid identification, monozygotic twins discrimination, lifestyle prediction, and biological and chronological age prediction [15–21]. For its relevance in generating investigative leads, chronological age prediction is a particularly active area of research aimed at developing accurate and robust age-predictive models for forensic purposes [22–24]. However, this proves to be a challenging task due to both biological and analytical factors (refer to the reviews by Naue [25], Refn et al. [26], and Castagnola et al. [27] for an in-depth analysis). Biologically, epigenetic modifications are not exclusively linked to the aging process. An individual's biological background (population of origin, disease status, etc.) [28, 29], environmental conditions, and lifestyle choices [17, 20, 30] are some of the factors found to have a compound effect on DNAm levels, leading to a substantial inter-individual variation. The observed multi-factorial nature of DNAm directly correlates to technical aspects, because strictly age-related markers have to be selected for an accurate DNAm-based chronological age-predictive model. Additionally, routine forensic casework samples may be multi-source, scarce, degraded, and oftentimes not suitable for DNAm level quantification [7, 26, 31–35]. As of today, a variety of multi-tissue and tissue-specific loci have been identified for age prediction from forensically relevant substrates [11, 34, 36–39].

The present study consisted of an Italian collaborative project aimed at evaluating the robustness, reproducibility, and accuracy of a standardized bisulfite conversion–based SBE protocol for DNAm analysis of the five most researched loci for age prediction from blood (ELOVL2, FHL2, KLF14, C1orf132/MIR29B2C, and TRIM59) [11–13]. Beyond the age-predictive application of DNAm analysis, the purpose of this study was the validation of the proposed DNAm analysis protocol for its future implementation in the routine forensic workflow. A cohort of 10 Italian laboratories, part of the Genetisti Forensi Italiani (Ge.F.I.), the Italian-speaking group of the International Society for Forensic Genetics (ISFG), collaborated in analyzing bisulfite-converted DNA extracted from fresh blood samples and mock-casework blood spots.

2 | Materials and Methods

2.1 | Outline of the Study

The study was organized in three phases, with 10 participating Ge.F.I. laboratories. Due to the extended time frame in which the analytical part of this study was conducted (approximately 2 years), not all laboratories took part in all phases. To minimize variability caused by different laboratory processing conditions and possible time-dependent alterations to DNAm status, laboratories were instructed to perform the analysis within 10 days of receiving the samples. Laboratories that could not guarantee immediate processing were excluded from participation in specific phases of the study. Moreover, during the experiment's extended course, some laboratories' genetic analyzers were replaced. The participants are listed in Table 1, along with the genetic analyzers used in each phase.

Peripheral blood samples were collected from a total of 22 healthy volunteers after they had provided their informed consent, according to each laboratory's Institutional guidelines and the approval by the Ethics Committee of Perugia University. All samples were anonymized, and only the information considered relevant to the study was collected, namely, chronological age and healthy status, without any additional specifications. The protocol from Onofri et al. [13] was employed throughout the study, and all participants received detailed instructions on the entire workflow as well as an electropherogram representative of the expected results (Figure S1). The organizing laboratory provided all participating laboratories with PCR and SBE primer mixes for the entire duration of the study.

The phases were organized as follows:

A. Phase I: The aim of this phase was the protocol set-up and validation in all participating laboratories. To assess inter-laboratory variability, the organizing lab sent each participating lab a sample of extracted DNA and a sample

TABLE 1 | List of the participating laboratories and the genetic analyzers used in each phase.

Laboratory	Code	Phase I	Phase II	Phase III
Università Politecnica delle Marche, Ancona	Lab1	—	3130	3130
Università degli Studi di Brescia	Lab2	3500	3500	3500
Centro Regionale Antidoping "A Bertinaria"	Lab3	3500	3500	3500
Università degli Studi di Ferrara	Lab4	310	310	—
Università degli Studi di Padova	Lab5	—	3130	3130
Università di Pavia	Lab6	SeqStudio	SeqStudio	SeqStudio
Reparto Carabinieri Investigazioni Scientifiche di Parma	Lab7	—	3500	3500
Azienda Ospedaliera "S.Maria" di Terni[a]	Lab8	310	SeqStudio	SeqStudio
Università degli Studi di Torino	Lab9	SeqStudio	SeqStudio	SeqStudio
Università degli Studi di Trieste	Lab10	—	—	—

Note: Lab10 did not participate in the analytical process but provided samples in Phase III.
[a]Organizing Lab.

of bisulfite-converted DNA from the same individual (a 44-year-old female). As shown in Figure 1, labs agreed to bisulfite-convert the extracted DNA in duplicate. Both the in-house converted and the pre-converted DNA samples were to be amplified by PCR and subjected to SBE analysis in triplicate.

B. Phase II: Given the relevance of analytical accuracy for casework samples, the aim of this phase was the comparison of the results obtained from whole blood and bloodstains. Additionally, the intra- and inter-laboratory repeatability was tested, and an analysis of possible variability factors (e.g., different extraction protocols) was performed. From the peripheral blood of 8 volunteers, aged 25–60 (Table S1), the organizing laboratory set up dried blood spots and performed DNA extraction. Laboratories were provided with 8 DNA extracts and 8 dried blood spots, 1 per each individual. Laboratories were tasked with extracting the blood spots with their own preferred method and with performing downstream bisulfite conversion and PCR/SBE analysis. Both the in-lab extracted blood spots and the DNA extracts provided by the organizing laboratory were bisulfite-converted in duplicates. Each bisulfite-converted DNA was PCR/SBE amplified in duplicates (Figure 1).

C. Phase III: This phase aimed to compare the true and the predicted age of 22 individuals across labs and to determine components of the deviations due to experimental error and/or true biological differences. Seven participating laboratories collected peripheral blood samples from 2 individuals each and distributed the DNA extracts to all other laboratories. Thus, each laboratory analyzed the same pool of 14 volunteers aged 25–83, selected to represent the full spectrum of the human adult lifespan (Table S1). All laboratories were asked to convert the DNA extracts twice. Each bisulfite-converted DNA was PCR/SBE amplified in duplicates (Figure 1). The data gathered from the analysis of the 8 DNA extracts from Phase II was merged with Phase III results, for a total cohort of 22 individuals.

2.2 | Sample Preparation, DNA Extraction, and Conversion

Sample preparation is described in the Supporting Information section ("Sample preparation") for each phase. The extraction method used by each laboratory is reported in Table S2. The organizing laboratory performed genomic DNA extraction with the QIAamp DNA Mini kit (QIAGEN, Hilden, Germany) and quantitation with the PowerQuant System (Promega Corporation, Madison, WI, USA) according to the manufacturer's protocols [40, 41]. Bisulfite conversion was performed in all laboratories with the EZ DNA Methylation-Direct kit (Zymo Research, Irvine, CA, USA) [42]. As per the manufacturer's instructions, around 400 ng of DNA per sample were bisulfite-converted.

2.3 | PCR Amplification of Bisulfite-Converted DNA, SBE, and Enzymatic Clean-Ups

The preliminary results [13], the bisulfite conversion reaction yield was estimated to be around 60% of the initial input. Bisulfite-converted DNA samples were diluted accordingly to a concentration of 2.5 ng/µL. Bisulfite-converted DNA of 10 ng was amplified in a PCR pentaplex by using the QIAGEN Multiplex PCR Master Mix 2x (QIAGEN, Hilden, Germany) [43]. The PCR primers' sequences for the five investigated loci (ELOVL2, FHL2, KLF14, C1orf132/MIR29B2C, and TRIM59) were reported in the studies of Cho et al. and Jung et al. [11, 12], whereas primer mix concentrations and amplification conditions were the ones used by Onofri et al. [13]. PCR amplification was followed by an enzymatic clean-up with the ExoSAP Express PCR Product Cleanup reagent (Applied Biosystems, Waltham, MA, USA) [44].

The multiplex SBE reaction was performed using the SNaPshot Multiplex Kit (Applied Biosystems, Waltham, MA, USA). The SBE primers' sequences were reported in the studies of Cho et al. and Jung et al. [11, 12], whereas SBE primer mix concentrations were as in the studies of Onofri et al. [13]. Temperature conditions for SBE were as per the SNaPshot manufacturer's instructions

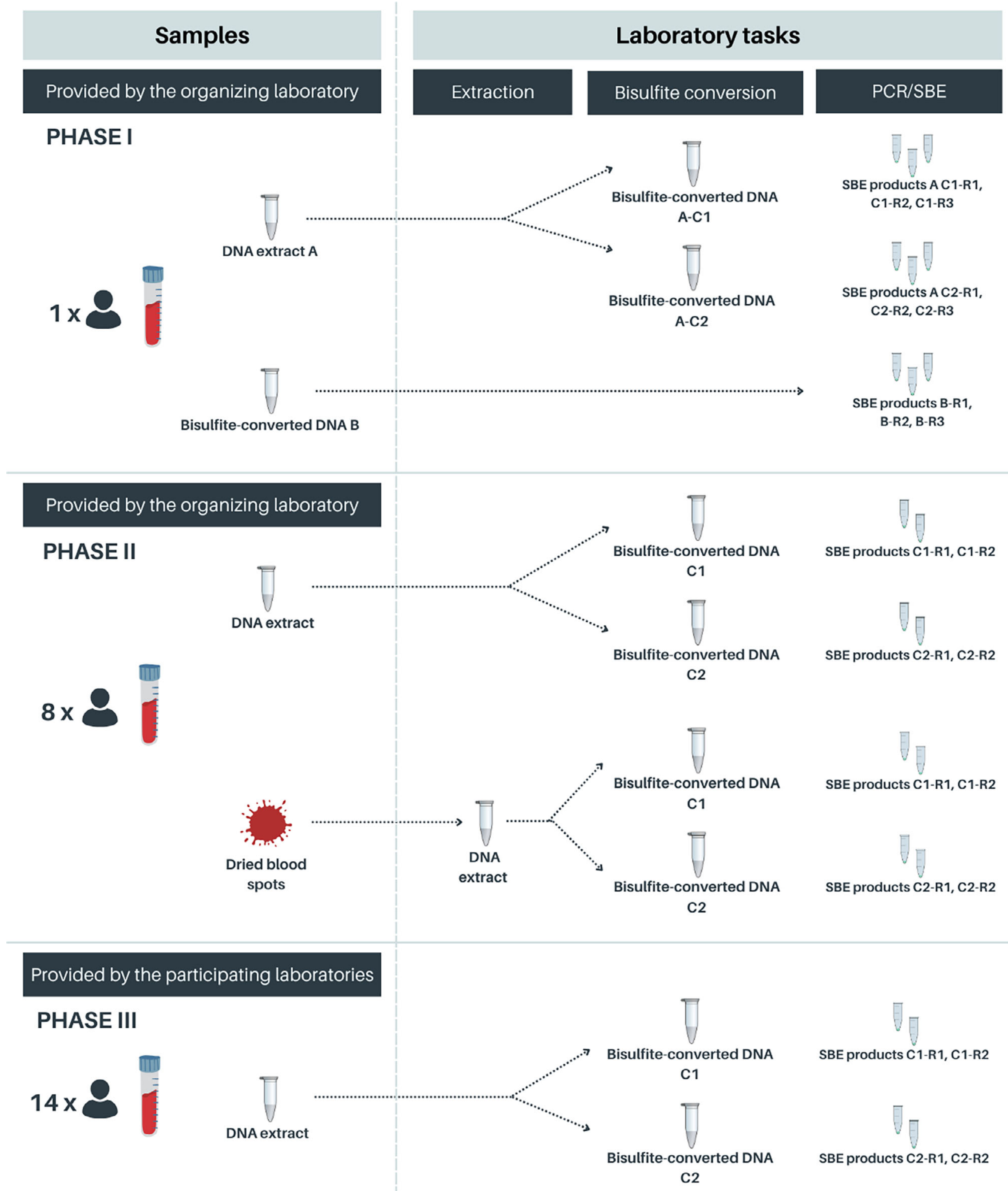

FIGURE 1 | Graphical representation of the samples and laboratory tasks of each phase. "C" represents the bisulfite conversion replicates, whereas "R" represents the PCR/SBE replicates. For sample information, refer to Table S1. PCR, polymerase chain reaction; SBE, Single Base Extension.

[45]. SBE products were purified by using the USB Shrimp Alkaline Phosphatase (SAP) (Applied Biosystems, Waltham, MA, USA) as per the manufacturer's instructions [46].

PCR/SBE primer mix aliquots were provided by the organizing laboratory for the entire duration of the study.

2.4 | Capillary Electrophoresis (CE)

All genetic analyzers are listed in Table 1. CE running conditions and analytical threshold (AT) routinely used by each laboratory for SNP analysis were applied. As above mentioned, laboratories were provided an electropherogram (EPG) representative of

the expected peak placement to aid in results interpretation (Figure S1) along with the information from the studies of Cho et al. and Jung et al. [11, 12] on the SBE fragments length. For the loci with forward SBE primers, namely, ELOV2 and FHL2, methylated cytosines were indicated by peaks in the yellow fluorescence channel and labeled C, whereas unmethylated cytosines appeared as red peaks and were labeled thymine (T). For the loci with reverse SBE primers (KLF14, C1orf132/MIR29B2C, and TRIM59), methylated cytosines appeared as blue peaks and were labeled guanines (G), whereas unmethylated cytosines were indicated by green peaks and labeled adenines (A). For each sample replicate, laboratories were asked to report the size, height, and area of the methylated and unmethylated peaks observed at each of the five loci by filling a table provided by the organizing laboratory (Table S3) and to accompany this information with the samples' EPGs.

2.5 | Statistical Analysis

Methylation levels at each locus were calculated by dividing the intensity in RFU (I) of the methylation signal by the sum of the intensity of the methylated (I_C or I_G) and the unmethylated peaks (I_T or I_A). For the loci with forward SBE primers, namely, ELOVL2 and FHL2, the methylation ratio corresponded to $I_C/I_C + I_T$, whereas for the loci with reverse SBE primers (KLF14, C1orf132/MIR29B2C, and TRIM59), the ratio was calculated as $I_G/I_G + I_A$. In the present study, these fractions are designated as peak ratios.

Mean absolute deviation (MAD) was reported to summarize data variability around a central tendency, whether an average or a predicted value from regression analysis. Two-tailed Student's *t*-tests (for two samples) and one-way analysis of variance (ANOVA) (for multiple samples) were performed to assess whether observed differences among variables were meaningful or likely due to random variation. For multiple regression results, we reported *R*-squared values, representing the percentage of variance explained by a model, the root mean squared error (RMSE, equivalent to the standard deviation of residuals), and MAD. Regression analyses were cross-validated using the leave-one-out cross-validation (LOO-CV) method, which provides unbiased performance estimates and is particularly suitable for small datasets. This method systematically excludes one data point at a time, fitting the regression model with the remaining data, in the end producing *N* regression analyses, *N* being the number of data points used in a regression analysis. Stepwise regression modeling was performed with a *p* value inclusion threshold of 0.2 to account for potential collinearity among locus data and to evaluate whether reducing the number of predictors could maintain the performance of a more parsimonious model.

Calculations and visualizations were performed using Microsoft Excel with the Real Statistics Resource Pack [47] and custom macros, as well as R (version 4.2.1) and Python (version 3.10.12). Analyses were cross-checked by multiple authors (especially S.A. and S.P.).

3 | Results

3.1 | Results of Phase I

3.1.1 | Repeatability of the Conversion Reactions

As the main aim of Phase I was to establish a consistent bisulfite conversion protocol across all labs, the participating laboratories received an aliquot of DNA pre-converted by the organizing lab, alongside a sample of extracted DNA to be converted. Both samples were subjected to triplicate PCR/SBE analysis, resulting in nine data points per locus per laboratory. Figure 2 depicts a comparison between the pre-converted DNA (blue points, left) and the DNA locally converted in duplicates (yellow points, right) by the five laboratories that completed the protocol. To assess the repeatability of the conversion reaction performed by each lab, data from the 2 in-lab conversion experiments were analyzed separately by locus using Student's *t*-test, totaling 25 comparisons. Four labs showed no significant heterogeneity among experiments for any locus. However, one lab (Lab2) did not detect any peak at the KLF14 locus for the replicates of one bisulfite-converted sample. Therefore, a comparison for the KLF14 locus for Lab2 could not be performed, and, additionally, Lab2 displayed *p* values <0.02 at two other loci, suggesting some inconsistency. Additional evidence of difficulty was reflected in the smaller overall MAD from the means of each locus in the experiments conducted on provided bisulfite-converted DNA (blue points in Figure 2, MAD = 0.010) compared to the MAD for in-lab bisulfite-converted DNA (yellow points in Figure 2, MAD 0.018). Overall, these results demonstrate the effectiveness of the standardized protocol in achieving consistent conversion, despite minor challenges that were addressed as the study progressed.

3.1.2 | Heterogeneity of Peak Ratios Among Labs

The previous section demonstrated fair consistency in the data produced by each lab across different DNA sources (the DNA converted by the organizing lab and two independently in-lab-converted DNAs). However, Figure 2 shows that calculated peak ratios vary widely across labs. For example, Lab2 and Lab3 produced, for ELOVL2, average peak ratios exceeding 0.4, Lab4 reported values between 0.3 and 0.4, and Labs 6 and 9 reported values around 0.3. To further investigate, we focused on the data from DNA pre-converted by the organizing lab (blue points in Figure 2), which was a single-DNA batch shared among labs. Because the organizing lab also provided its own data (four PCR reactions performed on the same converted DNA used for the exercise), we were able to compare results from six labs on the exact same DNA source. Figure 3 presents side-by-side data from each of the six labs in the same order, separately by locus. A pattern emerges in the data; for simplicity, if we label high values as "*H*" and low values as "*L*," we observe the sequences HHLLHL for the first two loci (ELOVL2 and FHL2) and LLHLHL for the last two loci.

A plausible explanation for this pattern is the use of different electrophoresis instruments across labs. Specifically, two labs used the ABI PRISM 310 Genetic Analyzer (Applied Biosystems,

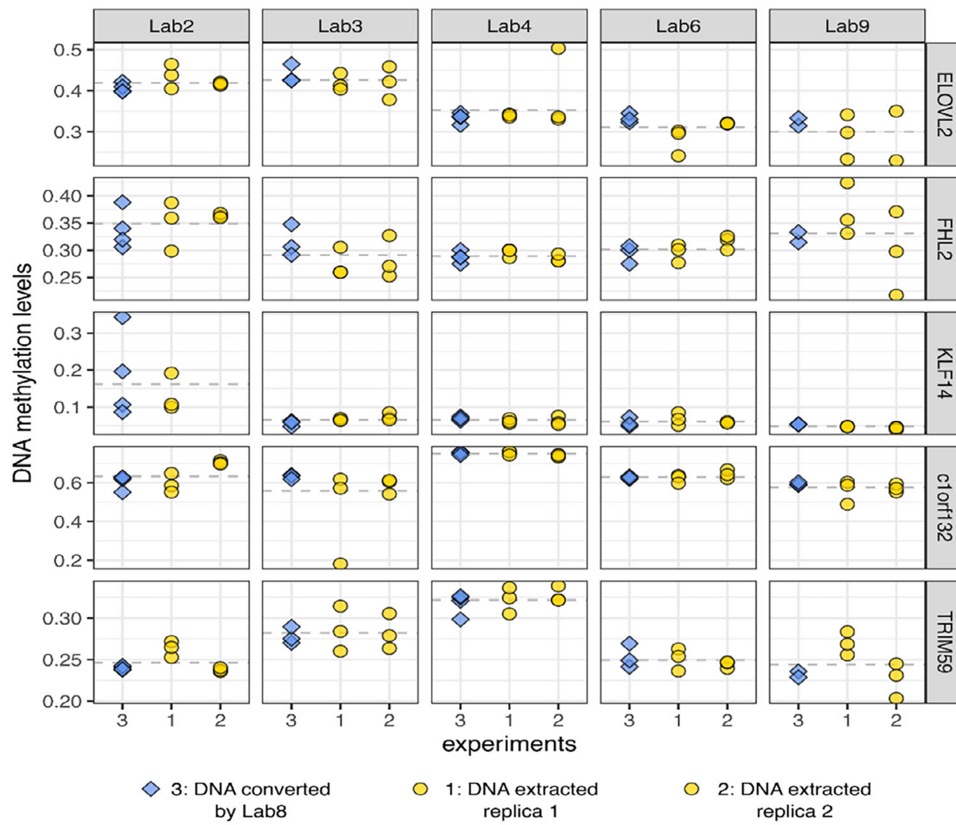

FIGURE 2 | DNA methylation levels of all experiments performed by each laboratory in Phase I. Yellow dots: three PCRs from the two conversion reactions performed by the participating labs. Blue dots: three PCRs from the converted DNA provided by the organizing laboratory. The dashed line represents the mean of the results obtained on the DNA converted by Lab8.

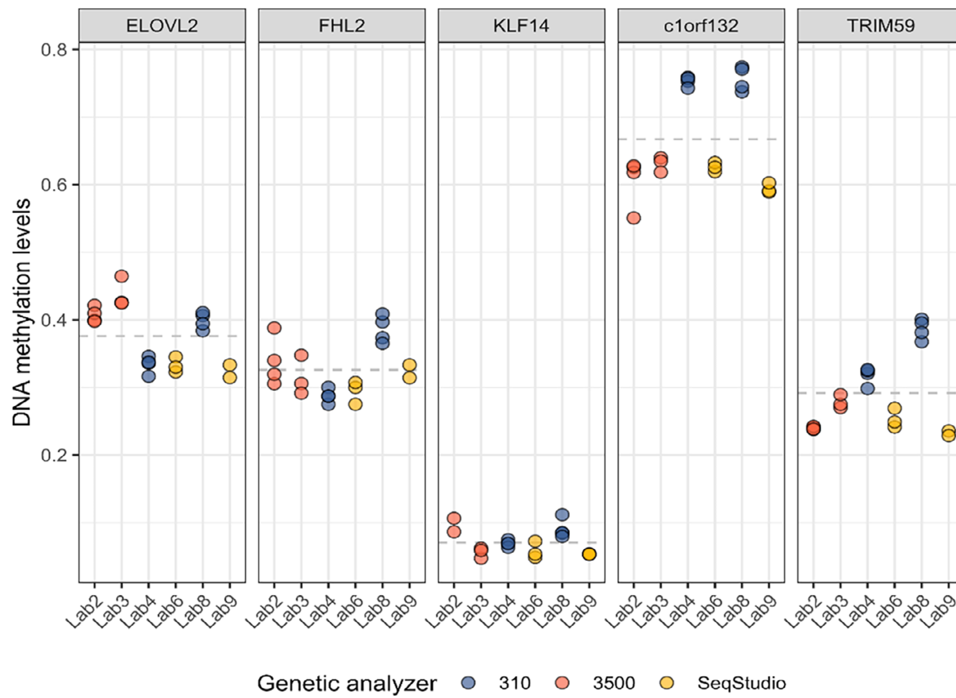

FIGURE 3 | Variation of DNA methylation levels among labs in repeated PCR/SBE/capillary electrophoresis. For each locus, the three or four data points obtained by six labs on the same bisulfite-converted DNA are arranged vertically. The dashed lines refer to the mean DNA methylation levels among all laboratories.

Waltham, MA, USA), two labs used the SeqStudio Genetic Analyzer (Applied Biosystems, Waltham, MA, USA), and two labs used the 3500 Genetic Analyzer (Applied Biosystems, Waltham, MA, USA). Analyzing deviations of lab means from the grand mean for each locus (dashed line in Figure 3), we found that labs using the SeqStudio consistently reported lower means in 10 out of 10 cases, whereas labs using the 310 reported higher means in 8 out of 10 cases. A chi-square test for independence on the contingency table (not shown) yielded a chi-square value of 13.71 with two degrees of freedom ($p = 0.001$), indicating a significant association. Furthermore, ANOVA confirmed that the type of instrument used had a significant impact. We examined the deviations of lab peak ratios from each locus mean and grouped them by the three instrument types. A substantial portion of the total variance (32%) was attributed to differences among instruments ($p = 7.6 \times 10\hat{}-9$), highlighting their notable contribution to the observed variability.

3.2 | Results of Phase II

3.2.1 | Outline of Phase II

In the second phase, the organizing lab provided the participating labs with DNA extracted from eight subjects and the corresponding blood spots (see the first eight samples in Table S1). The participating laboratories were responsible for two tasks: (1) converting the DNA in duplicate and performing PCR/SBE on both bisulfite-converted DNA products in duplicate, generating four peak ratios for each of the five loci, and (2) extracting DNA from the blood spots and applying the same protocol, finally providing eight data points for each of the eight individuals. The mean and standard deviation of methylation ratios are reported in Figure S2.

3.2.2 | Comparing Blood With Bloodstains

To compare DNAm levels from whole blood and blood spots, we performed a locus-by-locus regression analysis on the methylation data from each laboratory. In this analysis, methylation values from blood spots were regressed against those from whole blood. This approach allows us to assess the agreement between the two sample types by examining regression parameters for paired observations. Perfect consistency between the two would yield a correlation (r) of 1, a slope of 1, and an intercept of 0. Any deviations from these values provide insight into the reliability of blood spots, compared to whole blood, for predicting chronological age and for other methylation-based epigenetic studies.

The correlation coefficients (Pearson r values) for ELOVL2 and FLH2 exhibited a high level of agreement, averaging 0.93 across all six laboratories for both loci (Table 2). Slightly lower r values were observed for C1orf132/MIR29B2C and TRIM59, both averaging 0.87. In contrast, the mean correlation for KLF14 was low, with an average of 0.35 and remarkable variation among labs. Considering the slopes of the regression lines, the averages for all loci were below 1, indicating a slight increase in variability in the spots compared to blood across all laboratories. Additionally, there was significant variation among labs, with the KLF14 locus again standing out as an outlier compared to the others. Examining the intercepts, the average values were close to zero, suggesting an overall fair consistency in the data. However, the variation among laboratories was substantial.

In conclusion, the KLF14 locus appears as a problematic one, whereas the other loci, despite showing increased variability in the bloodstain dataset, appear to be useful for methylation analysis in the context of age prediction from spotted blood.

3.3 | Results of Phase III

3.3.1 | Outline of Phase III

In the third phase of the study, seven of the participant labs agreed to extract and quantify DNA from the blood of two individuals each and to deliver the extracts to all other labs (Figure S3).

As the protocol applied in Phase III was the same as in Phase II, the results from fresh blood extract analysis in Phase II and the data from Phase III were merged. Thus, the final database of the study included 22 individuals (Table S1), 12 females and 10 males, with mean ages of 42.8 and 50.9, respectively. Women were more represented in the younger age class, though the sex mean difference was not significant. Boks et al. [48] reported a significant association of DNAm with sex, albeit of a modest level (more comprehensive review in [25]). We tested the sex effect in a regression model in our sample, and it was not significant (data not shown). Therefore, the data were not stratified for sex in the present work.

Each laboratory conducted two DNA conversions for each sample. Both bisulfite reaction products were subjected to duplicate PCR/SBE analysis, resulting in four data points per individual per locus. Six laboratories submitted complete or nearly complete data, contributing to a total of 528 records (22 subjects × 6 labs × 4 replicates). Of the 2640 total data points (528 records × 5 loci), only 31 (1.2%) were missing from the final dataset.

3.3.2 | Single-Locus Regressions

We began by visually inspecting the methylation data from each lab, stratified by study phases (II and III), duplicate conversion reactions, and loci, resulting in a total of 120 regression lines. This initial analysis confirmed no systematic differences between the two conversion reactions within each lab, allowing us to average the four measurements for each subject. However, clear differences emerged in some cases between Phases II and III, particularly in the intercepts of the regression lines (data not shown). To account for these differences, we aligned the Phase II data to Phase III data. The data of phase II $y_{i,0}$, (where the index i refers to the ith subject, $i = 1,2,\ldots,8$, and the index 0 refers to Phase II), were aligned to the data of Phase III $y_{j,1}$, ($j = 1, 2,\ldots, 14$) by computing $y_{i,1} = \hat{y}_{i,1} + e_{i,0}$, where caret refers to predicted values and $e_{i,0}$ are the residuals from the regression line of Phase II (i.e., $e_{i,0} = y_{i,0} - \hat{y}_{i,0}$). Figure 4 illustrates the aligned methylation data for all 22 subjects across both phases, with regression lines showing methylation ratios by chronological age at each locus.

TABLE 2 | Regression analysis of DNA methylation levels in spotted samples versus blood samples.

Locus	Regression values	Lab1	Lab2	Lab3	Lab6	Lab8	Lab9	Mean
ELOVL2	*R*	0.97	0.88	0.95	0.96	0.82	0.98	0.93
	Slope	1.09	0.78	0.91	1.06	0.96	1.05	0.98
	Intercept	−0.07	0.08	−0.03	−0.03	0.03	−0.03	−0.01
FHL2	*R*	0.96	0.87	0.98	0.98	0.84	0.96	0.93
	Slope	0.98	0.72	0.91	1.13	0.68	0.87	0.88
	Intercept	−0.02	0.1	−0.02	−0.02	0.13	0.02	0.03
KLF14	*R*	N.D.	−0.25	−0.11	0.99	0.44	0.68	0.35
	Slope	N.D.	−0.28	−0.04	1	0.35	0.55	0.32
	Intercept	N.D.	0.13	0.02	0	0.02	0.01	0.04
C1orf132/MIR29B2C	R	N.D.	0.89	0.65	0.99	0.9	0.89	0.87
	Slope	N.D.	1	0.93	1.06	0.74	0.84	0.91
	Intercept	N.D.	0.04	0.19	−0.05	0.18	0.12	0.1
TRIM59	*R*	0.86	0.65	0.92	0.99	0.85	0.96	0.87
	Slope	1.08	0.34	1.58	1.06	0.95	0.84	0.98
	Intercept	0.08	0.14	−0.17	0	0.01	0.03	0.02

Note: R, Pearson correlation coefficient.
Abbreviation: N.D, no data.

3.3.3 | Multiple Regressions

3.3.3.1 | Full Model. On the basis of current literature [10–13, 29, 32, 49], methylation levels at the five studied markers are commonly used as independent variables in age prediction. Accordingly, a multiple linear regression model was developed using the adjusted data from Phases II and III (see previous section). This five-loci model will thereafter be referred to as the "full model." The full model was applied both to each laboratory dataset individually (22 records per lab) and to the combined dataset (132 records). Model fit summary statistics (Table 3) reveal some heterogeneity across laboratories, as indicated by variations in *R*-squared and MAD values. Notably, the fit for the combined dataset is substantially lower, suggesting that attempting to create a generalizable model across different data sources without specific adjustments may be ineffective. Even with lab-based adjustments (Table 4, last column), the model fit statistics remain poorer than those of all but one individual lab.

Figure 5 presents the results of the multiple regression analysis of the full model across the six laboratories. Panel (a) shows the ages predicted by each lab for each individual, plotted against their chronological age. The high correlation of predicted ages across laboratories is evident; a one-way ANOVA on the residuals revealed that 33.8% of the variance could be attributed to intrinsic laboratory variability. The remaining 66.2% of the variation could be attributed to biological factors, indicating that such deviations between chronological and biological age may be due to true individual differences in methylation levels rather than experimental error. Further, no lab consistently over- or underestimated age relative to the others. Instead, predicted ages for each individual varied randomly around the overall mean, indicating an absence of systematic bias across labs. Interestingly, two individuals, a 38-year-old and a 46-year-old, showed predicted ages consistently much higher and much lower, respectively, than their chronological ages, suggesting a real biological explanation for their deviation.

Figure 5b illustrates each individual's mean predicted age and standard deviations, with six color-coded regression lines for each lab and an overall regression line (black), shaded with a 95% confidence interval (CI). Notably, this CI includes the line where chronological and predicted ages are equal (dashed line in both panels). We also computed the 95% CI for each individual's mean age prediction, which ranged from 2.43 to 8.82 years, with an average width of 5.24 years across the 22 subjects. This interval indicates that, with 95% confidence, the true deviation of a new sample's predicted age from its chronological age will fall within approximately $X \pm 2.62$ years, where X represents the mean of the six lab predictions. The relatively narrow CI range can be attributed to the independent measurements provided by the six labs. Notably, eight of the 22 CIs do not contain the actual chronological age of the individual, whereas only one would be expected to deviate by chance. This finding supports the idea that much of the observed variation among individuals likely reflects true differences in methylation levels rather than purely technical variability.

3.3.3.2 | Restricted Model. Redundancy of the CpG loci concerning age prediction, due to their generally high level of collinearity, has been previously reported [36, 50–52]. In our analysis, 15 out of the 30 p values calculated for the variables in the full model exceeded 0.2; this is a commonly used threshold for retaining predictors in a stepwise selection of the backward elimination method [53]. This threshold allows potentially relevant predictors to remain in the model while accounting for collinearity. Consequently, a stepwise regression analysis applied to the datasets from each of the six laboratories was conducted.

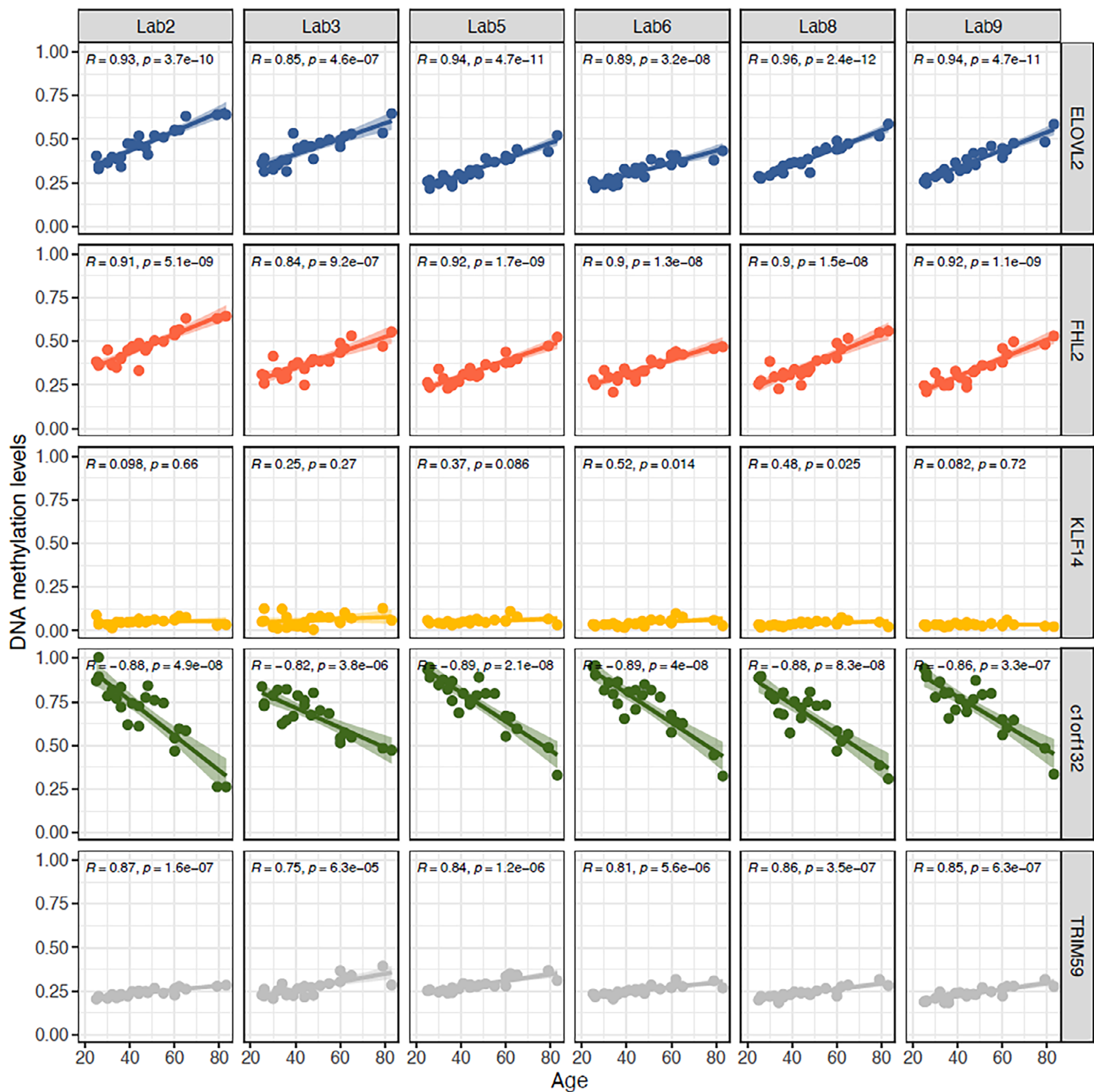

FIGURE 4 | Pearson correlation between DNA methylation levels at each locus and chronological age of the 22 individuals analyzed in Phases II and III by each laboratory.

TABLE 3 | Summary of performance metrics for the full regression model applied separately to each lab, to all data without any correction, and to all data adjusted for labs.

Metric	Lab2	Lab3	Lab5	Lab6	Lab8	Lab9	All data	Adjusted
R square	0.940	0.901	0.939	0.926	0.956	0.960	0.850	0.924
RMSE	4.600	5.950	4.670	5.150	3.940	3.760	6.380	4.625
MAD	3.310	4.040	3.280	3.020	2.200	2.570	5.000	3.415

Abbreviations: MAD, mean absolute deviation; RMSE, root mean square error.

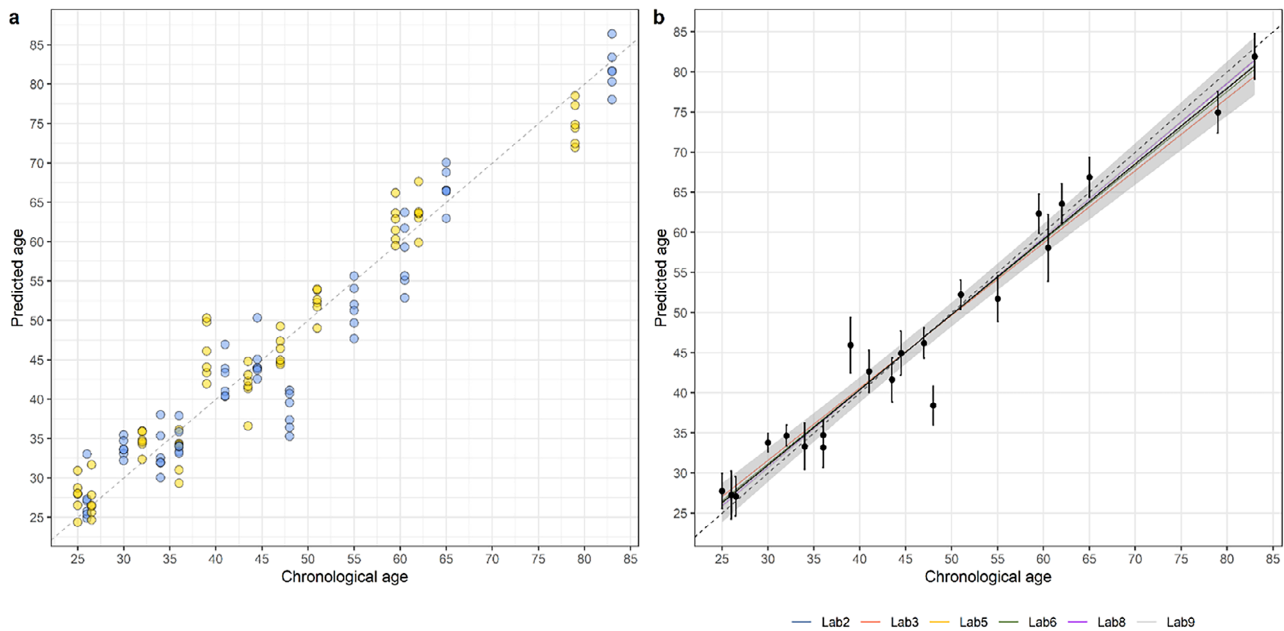

FIGURE 5 | Predicted and chronological age of the 22 subjects analyzed by 6 independent laboratories. (a) The six points for each individual represent the predictions made by each laboratory (the colors are alternated among the individuals); (b) mean and standard deviation across all laboratory predictions. A regression line between the mean values of the predicted age and the chronological age (in black with a grey shade being the 95% confidence interval) as well as colored regression lines for each laboratory are reported. The dashed lines, both in (a) and in (b), represent the line of identity.

TABLE 4 | Results of the stepwise regression applied to the data by each lab.

Locus	Lab2	Lab3	Lab5	Lab6	Lab8	Lab9
ELOVL2		Y	Y	Y	Y	Y
FHL2	Y	Y				
KLF14						
C1orf132/MIR29B2C	Y		Y	Y	Y	Y
TRIM59	Y	Y	Y		Y	Y

Note: "Y" indicates that the model retained the locus. Five out of the six laboratories retain three loci.

The goal of this approach is to establish a parsimonious model that retains only the most relevant predictors. The process initiates with all variables included and systematically removes one variable at a time, eliminating the least contributing variable to the model's fit until no more variables meet the removal criteria. On the basis of the results (Table 4), we retained the three loci ELOVL2, C1orf132/MIR29B2C, and TRIM59 in a "restricted model."

The performance metrics of the full and restricted models were comparable. The mean *R*-squared values were 0.930 and 0.937, respectively. The mean RMSE values were 4.771 versus 4.687, and the mean MAD values were 3.125 versus 3.173. Additionally, the MAD values were even lower in three out of the six labs.

At the individual level, the SE of the residuals was similar between the restricted and full model (averages across all subjects: 1.15 ± 0.33 years and 1.02 ± 0.27 years, respectively), and the individual mean MAD values were highly correlated ($r = 0.93$, Figure 6). For 14 subjects, the absolute difference between the deviations of the two models was <1 year; for 6 other subjects, it was <2 years, and for only 2 subjects out of 22 it was slightly higher than 3 years (namely, 3.1 years).

Comparable results were obtained by applying another approach that deals with multicollinearity in the predictor variables, namely, the least absolute shrinkage and selection operator (LASSO) regression. In this method, all the original variables are retained in the analysis, but the model penalizes the coefficient of each variable and simultaneously performs variable selection. Notably, the KLF14 coefficient was set to zero in five of the six laboratories. Table S4 shows the final coefficients obtained with this approach and the basic performance metrics.

3.3.4 | Data Cross-Validation

Although the primary goal of this study was not to create a fully predictive model for chronological age based on methylation levels—an endeavor that would ideally require a larger sample size—the accuracy and robustness of the regression model remain important to assess. Evaluating the generalizability of a model developed from a small dataset is especially critical, as adding or removing even a single subject can significantly affect regression parameters and predictions. In the preceding sections, we evaluated the model fit across the entire dataset, which is effective for comparing repeated measurements but does not estimate generalizability. To better understand how this model might perform with new data, a cross-validation approach is required. LOO-CV is a method well suited for small datasets, as it systematically excludes each data point in turn, refitting the model with the remaining data to provide an unbiased performance estimate. We applied LOO-CV to each laboratory's

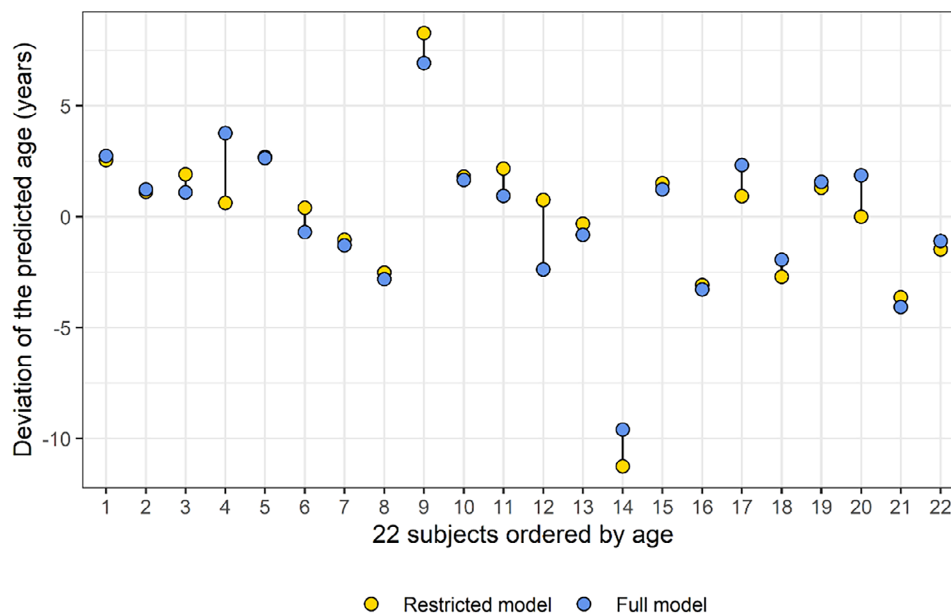

FIGURE 6 | Mean absolute deviation (MAD) values of both the full and restricted models for the 22 subjects analyzed in Phases II and III.

data, evaluating the MAD for the full model (including all five genetic markers) across the six labs. The cross-validated MAD values, shown in Table 4 of Section 3.3.2, increased as expected across all laboratories, with the mean rising from 3.07 to 4.41 years. This increase illustrates the potential over-optimism that can arise when regression analysis is performed without cross-validation.

4 | Discussion

This collaborative effort among Ge.F.I laboratories evaluated the robustness, reproducibility, and accuracy of an SBE-based assay for age-predictive DNAm analysis and aimed at easing its implementation in forensic laboratories. A significant strength of the study design is the concurrent analysis of the same panel of test individuals by multiple forensic laboratories, with each lab working independently and blind to the others' results throughout all phases.

4.1 | The Study Protocol

A crucial step of the exercise was the establishment of a common protocol. In Phase I, we began by assessing the replicability of the bisulfite conversion reaction across different laboratories. The organizing laboratory provided all participating laboratories with one DNA extract and one bisulfite-converted DNA, both derived from the peripheral blood of a single individual. The participating laboratories were tasked with bisulfite-converting the DNA extract in duplicate. Both the in-house bisulfite-converted DNA samples and the bisulfite-converted sample provided by the organizing laboratory were to be PCR/SBE amplified in triplicates, for a total of 9 replicates. To assess whether blood spots could be a reliable source of material for DNAm analysis, in Phase II, the protocol was tested on the DNA extracts and dried blood spots derived from the peripheral blood of eight individuals. The organizing laboratory provided one DNA extract and one blood spot per individual. All laboratories were tasked with extracting the dried blood spots according to their preferred method. Both the DNA extracts provided by the organizing laboratory and the in-house-extracted DNA extracts were to be bisulfite converted in duplicate. Each bisulfite-converted DNA was PCR/SBE amplified twice for a total of 8 replicates per individual (4 PCR/SBE replicates for the DNA extract and 4 PCR/SBE replicates for the dried blood spots), resulting in a total of 192 replicas for the provided DNA extracts and 192 replicas for the dried blood spots. In Phase III, seven participating laboratories provided DNA extracts derived from the peripheral blood of two individuals. The samples were distributed to all laboratories; thus, all laboratories analyzed the same pool of 14 samples using the same protocol established for Phase II, that is, 2 bisulfite conversions per sample and 2 PCR/SBE replicates for each bisulfite-converted DNA. Given that the same experimental protocol was employed, Phase III results were merged with data obtained from the analysis of the DNA extracts from Phase II for a final cohort of 22 individuals, resulting in a database of 528 records (22 subjects times 6 labs times 4 replicates).

Although some challenges arose during initial implementation (Phase I), the results obtained from Phases II and III demonstrate that the protocol was easily and successfully adopted by all participating laboratories and produced promising results in terms of the accuracy of the predicted age (Section 3).

4.2 | Intra- and Inter-Lab Variations of Methylation Levels

Overall, a slight intra-laboratory variation was observed for DNAm levels estimation, whereas results were substantially more heterogeneous among the laboratories (Figures 3 and 5). This variation can be attributed to differences in the equipment and technical procedures among the laboratories, which may

result in a cumulative effect when considering the number of analytical steps of the protocol [19, 32, 54, 55]. In fact, after DNA extraction, samples are bisulfite converted and must then undergo multiplex PCR, multiplex SBE, in-between purification steps, and CE. Given its crucial role in the adopted protocol, bisulfite conversion was first investigated as a possible source of variability, and its replicability was assessed (Phase I). Some minor initial challenges were encountered, as shown by the incomplete and more heterogeneous results provided by Lab2. In subsequent Phases II and III, a negligible intra-laboratory variation was observed, demonstrating the consistency of the bisulfite conversion reaction.

Downstream processes following bisulfite conversion and the laboratory equipment seem to impact the inter- and intra-lab variability in a greater fashion. Differences between Phases II and III were observed for some participants, especially in the intercepts of their regression lines. Such discrepancies could be attributed to several factors, including the aliquots of PCR and SBE primer mixes sent at different times for the duration of the study. Nonetheless, it was possible to adjust the data from both phases per each laboratory by applying a correction factor (as described in Section 3.3.2). After adjustment, it was revealed that the PCR and SBE steps contribute more to intra-laboratory variation than the DNA conversion step. On the basis of the overall results, we may suggest that future implementations could reduce costs by performing a single-bisulfite conversion per sample and at least three independent amplifications/SBE steps. Such an approach has been previously proposed [13, 54]. As far as laboratory equipment is concerned, the CE instruments appeared to substantially contribute to inter-laboratory variations observed in DNAm levels measurement. In fact, in Phase I, the greatest source of inter-lab variation was attributed to the genetic analyzer used by each laboratory. This finding was confirmed by Phases II and III of the study and is consistent with previously reported results [7, 19, 31].

Lastly, unknown factors and DNA extraction methods can have an impact on the analysis of DNAm levels [56–58]. Phase II results proved how DNA extraction, among other factors, could contribute to inter-laboratory variation and the discrepancy in the DNAm levels measurements between fresh samples and dried blood spots. Such consideration is more thoroughly discussed in Section 4.6.

4.3 | Regression Analysis of the Full Model

The CpG sites examined in this study, namely, ELOVL2, FHL2, KLF14, C1orf132/MIR29B2C, and TRIM59, have previously been used for age prediction from blood samples [10–13, 29, 32, 49]. We refer to the multiple linear regression based on these loci as the full model. As mentioned above, the CE instruments greatly contribute to inter-laboratory variation. To adjust the data produced by different laboratories based on the genetic analyzer used, a z-score correction could be applied to the results, as demonstrated in the studies of Freire-Aradas et al. and Lee et al. [7, 19]. However, our primary goal was to compare the different labs' results using their original data, not to develop a comprehensive model. Considering the number of factors responsible for inter-laboratory variability (see previous section), it does not seem useful nor scientifically sound to produce a universal regression model applicable to all participating laboratories or to conduct data correction to transpose the results solely based on the genetic analyzer. As described in Section 3.3.3 and reported in Table 3, the model constructed using the entirety of the collected data (not adjusted) performed worse than any of the laboratory-tailored models. Similarly, the model built on the adjusted data presents poorer summary statistics than those of all laboratories except one. Precisely, the ages predicted by the laboratory-specific models were highly correlated, with a prediction error (MAD) ranging from 2.22 to 4.04 years (mean of 3.07 years). One of the most relevant outcomes of this study was the finding that 66.2% of the differences between predicted and chronological ages are due to individual-dependent factors, indicating true biological variations in the DNAm levels. This suggests that only 33.8% of the variation can be attributed to lab-dependent factors related to technical procedures. Large deviations of the "biological age" from the chronological age were apparent for two of the 22 subjects (Figure 5a); otherwise, the analysis of the CIs, whose overall mean was 5.24 years, showed an excess of individuals with smaller but yet significant deviations. Quite interestingly, the evidence of consistent over- and under-estimation in the predicted age for the two above-cited individuals, a 38-year-old and a 46-year-old, confirmed that factors unrelated to chronological age contribute to substantial inter-individual variation in DNAm patterns. Therefore, additional studies are necessary to identify significant DNAm changes that occur during the lifespan of an individual, to understand and quantify the effect of additional influences on DNAm status, and to further improve the true age prediction accuracy and robustness of DNAm-based tools for forensic purposes.

4.4 | Cross-Validation and Sample Size

The full model was cross-validated by applying to each of the six datasets the LOOCV method; this revealed a marked increase in the mean MAD value, from 3.07 years to 4.41 years (a rise of 44%). While a value of 4.5 years aligns with the literature [10–13, 29, 32, 49], the large increase suggests that prediction errors are sensitive to outliers in a sample of 22 subjects. In other words, including a few subjects with large deviations from the predicted age could significantly alter the model's performance metrics. However, this sensitivity is precisely the issue that cross-validation approaches like LOOCV address by providing a more robust estimate of model performance.

It is important to acknowledge that the sample size used in this study (n = 22) is relatively small. However, the focus of this analysis was to assess the feasibility of using DNAm for age prediction by forensic labs with limited budgets. The high degree of correlation between the predicted ages from the six independent labs shows that even with such a sample size, the model can effectively capture the core relationship between DNAm and age.

4.5 | Regression Analysis of a Reduced Model

Given the redundancy observed for some commonly used CpG loci in age prediction [36, 50–52] and the high p values observed

for many variables in the full model regression analysis, we explored a model with a lower number of loci. Stepwise regression analysis, using a p value threshold of <0.2 [53], identified ELOV2, C1orf132/MIR29B2C, and TRIM59 as the three loci to be kept in a "reduced model." Although p value thresholds for variable inclusion are not fixed rules, they are useful tools for model selection, aiming to balance model complexity reduction with the preservation of predictive accuracy. The fact that roughly half of the p values exceeded 0.2 in our case suggests that further lowering the threshold would not be advisable. Notably, the exclusion of the KLF14 locus, which was problematic in the analysis of bloodstains, was also supported by LASSO regression analysis (Table S4). As the performance metrics of the reduced and the full models were found to be similar, we suggest that these three loci may be sufficient and may even be more accurate for establishing an efficient and cost-effective age prediction test suitable for forensic applications (Section 3.3.3.2, Figure 6). However, further research is warranted to definitively prove the efficacy of this reduced panel for robust age prediction. A direction for future research could be to replicate the study on additional datasets of comparable size to ensure that the observed results aren't sample-specific and that the simplified protocol is broadly applicable.

4.6 | Blood Spots vs. Fresh Blood Extracts

Dried blood traces are a much more common substrate than fresh blood for forensic genetics investigations. For potential casework applications, it was essential to assess the reliability and accuracy of DNAm analysis protocol on dried blood spots. As a proof of concept, in Phase II, the organizing laboratory provided dried blood spots and fresh blood extract from eight individuals. As noted in Section 3.2.2, blood spots can serve as a reliable source for DNAm analysis based on the five loci investigated here; however, a degree of inter- and inter-laboratory variability greater than that registered for fresh blood extracts was observed (Table 2). Overall, DNAm level measurements at all loci appeared to be less consistent in blood spots than in DNA extracts, with KLF14 standing out as a problematic locus. It is worth noting that KLF14 performed poorly in both whole blood extracts and blood spots, albeit more so in the latter. Considering that both extracts and spots originated from the same biological material (blood), it would be reasonable to retain the loci that performed well in both declinations of the same biological source material rather than choosing a different set of age-predictive CpG loci. In addition to the locus selection carried out by stepwise regression analysis and the performance metrics of the reduced model, the exclusion of the KLF14 locus is further supported by its performance in dried blood stains.

From a technical standpoint, because all laboratories followed the same bisulfite conversion and PCR/SBE protocol, the increased variation observed for blood spots could be reasonably attributed to the different extraction methods employed by the laboratories [58]. It is important to note that the blood traces analyzed in this study were of relatively fresh deposition because they were processed within 10–15 days from peripheral blood collection by the organizing laboratory (Supporting Information section, Sample Preparation). In fact, despite the data reported in the literature on global methylation levels stability for dried blood spots [59–63], to minimize variability dependent on possible time-dependent effects on DNAm status, blood spots were immediately analyzed upon being received by the participating laboratories. However, real casework evidence is rarely of these pristine conditions. A considerable amount of time may pass between the deposition of a trace at the crime scene and its collection, all the while exposing the biological material to external agents that affect its quality and quantity. After collection, time and storage conditions can contribute to the trace's alteration. Therefore, it is not possible to exclude issues arising from scarce amounts of template DNA and degradation, which commonly hinder forensic DNA profiling [64–68] and could also affect the correct detection and quantification of DNAm patterns.

As already proposed in the studies of Peng et al. [69], additional research is necessary to more comprehensively assess the reliability and accuracy of the bisulfite conversion–dependent SBE analysis for age prediction when performed on real forensic casework samples. DNAm stability in trace blood samples will need to be examined for varying storage times and conditions as well as for external factors known for affecting DNA integrity (e.g., humidity, temperature, UV radiation, chemical and microbial agents) [64–68]. On the basis of Phase II results, different forensically used DNA extraction methods will have to be more thoroughly investigated in relation to DNAm measurement variability. Additionally, the performance of a reduced model comprising the most informative loci in both fresh DNA extracts and dried blood spots will have to be evaluated to develop a single age-predictive tool that can be applied for both declinations of the same source material.

5 | Conclusions

This study successfully validated an efficient protocol for DNAm-based age prediction, which can be easily replicated by any forensic genetics laboratory. The results show that, despite the availability and accuracy of NGS methodologies, laboratories that do not yet dispose of this technology may easily implement the bisulfite conversion–based SBE protocol herein employed. In fact, the protocol was highly reproducible with minimal intra-laboratory variations, which could be mitigated by performing PCR/SBE replicates, and the apparatus and methodologies employed by different laboratories appeared to have minimal impact on age-prediction accuracy. Laboratories wishing to implement the method can analyze a reference sample of 20–25 individuals to calibrate their in-house regression model and then proceed to predict the age of unknown samples. Both fresh and spotted peripheral blood proved to be reliable sources of DNA for age methylation studies, though blood spot analysis has a wider statistical error, highlighting the need for additional analysis from blood stains. Our findings also suggest the possibility of further simplifying the protocol by reducing the analyzed loci to three, although this simplification requires validation through dedicated collaborative studies in the future. The evidence collected on inter-individual variability suggests that the true age prediction accuracy and robustness of DNAm-based tools for forensic purposes would benefit from analyzing multiple individuals in the same age range to account for true inter-individual differences and from further research on the

impact of factors unrelated to chronological aging (e.g., lifestyle habits) on DNAm levels.

Author Contributions

Martina Onofri: investigation, data curation, writing–original draft, writing–review and editing, project administration. **Federica Alessandrini**: investigation, data curation, writing–review and editing. **Serena Aneli**: formal analysis, investigation, writing–original draft, data curation. **Loredana Buscemi**: supervision. **Elena Chierto**: formal analysis, investigation, data curation. **Matteo Fabbri**: investigation, data curation, resources. **Paolo Fattorini**: data curation, resources, validation. **Paolo Garofano**: investigation, data curation. **Fabiano Gentile**: investigation, data curation, resources. **Silvano Presciuttini**: formal analysis, data curation, validation, writing–original draft, writing–review and editing. **Carlo Previderè**: investigation, data curation, resources. **Carlo Robino**: data curation, resources, validation. **Simona Severini**: investigation, data curation, project administration. **Federica Tommolini**: investigation, data curation, project administration. **Pamela Tozzo**: investigation, data curation, resources. **Andrea Verzeletti**: investigation, data curation, resources. **Eugenia Carnevali**: conceptualization, methodology, validation, writing–review and editing, supervision, project administration, resources.

Ethics Statement

This study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of the University of Perugia, Italy (protocol code 43615, approved on the 23rd of February 2021).

Consent

Informed consent was obtained from all subjects involved in the study.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. L. D. Moore, T. Le, and G. Fan, "DNA Methylation and Its Basic Function," *Neuropsychopharmacology* 38, no. 1 (January 2013): 23–38.

2. F. Kader and M. Ghai, "DNA Methylation and Application in Forensic Sciences," *Forensic Science International* 249 (April 2015): 255–265.

3. E. Khodadadi, L. Fahmideh, E. Khodadadi, et al., "Current Advances in DNA Methylation Analysis Methods," *BioMed Research International* 2021, no. 1 (January 2021): 8827516.

4. S. J. Clark, A. Statham, C. Stirzaker, P. L. Molloy, and M. Frommer, "DNA Methylation: Bisulphite Modification and Analysis," *Nature Protocols* 1, no. 5 (December 2006): 2353–2364.

5. S. Kundu, R. Das, S. Laskar, Y. Choudhury, and S. K. Ghosh, "Chapter Two—Principles of Bi-Sulfite Conversion of DNA and Methylation-Specific PCR (MSP) in Biological Research," in *Epigenetics Methods*, ed. T Tollefsbol (Cambridge: Academic Press, 2020), 17–36.

6. A. Vidaki, D. Ballard, A. Aliferi, T. H. Miller, and L. P. Barron, "Syndercombe Court D. DNA Methylation-Based Forensic Age Prediction Using Artificial Neural Networks and Next Generation Sequencing," *Forensic Science International: Genetics* 28 (May 2017): 225–236.

7. A. Freire-Aradas, E. Pośpiech, A. Aliferi, et al., "A Comparison of Forensic Age Prediction Models Using Data From Four DNA Methylation Technologies," *Frontiers in Genetics* 11 (2020): 932.

8. M. Pfeifer, T. Bajanowski, J. Helmus, and M. Poetsch, "Inter-Laboratory Adaption of Age Estimation Models by DNA Methylation Analysis—Problems and Solutions," *International Journal of Legal Medicine* 134, no. 3 (May 2020): 953–961.

9. A. Woźniak, A. Heidegger, D. Piniewska-Róg, et al., "Development of the VISAGE Enhanced Tool and Statistical Models for Epigenetic Age Estimation in Blood, Buccal Cells and Bones," *Aging (Albany NY)* 13, no. 5 (March 2021): 6459–6484.

10. R. Zbieć-Piekarska, M. Spólnicka, T. Kupiec, et al., "Development of a Forensically Useful Age Prediction Method Based on DNA Methylation Analysis," *Forensic Science International: Genetics* 17 (July 2015): 173–179.

11. S. Cho, S. E. Jung, S. R. Hong, et al., "Independent Validation of DNA-Based Approaches for Age Prediction in Blood," *Forensic Science International: Genetics* 29 (July 2017): 250–256.

12. S. E. Jung, S. M. Lim, S. R. Hong, E. H. Lee, K. J. Shin, and 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 Science International: Genetics* 38 (January 2019): 1–8.

13. M. Onofri, A. Delicati, B. Marcante, et al., "Forensic Age Estimation Through a DNA Methylation-Based Age Prediction Model in the Italian Population: A Pilot Study," *International Journal of Molecular Sciences* 24, no. 6 (March 2023): 5381.

14. Y. Takahashi, M. Asari, S. Isozaki, et al., "Age Prediction by Methylation Analysis of Small Amounts of DNA Using Locked Nucleic Acids," *Journal of Forensic Science* 68, no. 1 (January 2023): 267–274.

15. Å. Johansson, S. Enroth, and U. Gyllensten, "Continuous Aging of the Human DNA Methylome Throughout the Human Lifespan," *PLoS ONE* 8, no. 6 (June 2013): e67378.

16. A. Vidaki and M. Kayser, "Recent Progress, Methods and Perspectives in Forensic Epigenetics," *Forensic Science International: Genetics* 37 (November 2018): 180–195.

17. C. Liu, R. E. Marioni, Å. K. Hedman, et al., "A DNA Methylation Biomarker of Alcohol Consumption," *Molecular Psychiatry* 23, no. 2 (February 2018): 422–433.

18. H. S. E. Sabeeha, "Forensic Epigenetic Analysis: The Path Ahead," *Medical Principles and Practice* 28, no. 4 (July 2019): 301–308.

19. J. E. Lee, J. M. Lee, J. Naue, et al., "A Collaborative Exercise on DNA Methylation-Based Age Prediction and Body Fluid Typing," *Forensic Science International: Genetics* 57 (March 2022): 102656.

20. A. Vidaki, B. P. Jiménez, B. Poggiali, et al., "Targeted DNA Methylation Analysis and Prediction of Smoking Habits in Blood Based on Massively Parallel Sequencing," *Forensic Science International: Genetics* 65 (July 2023): 102878.

21. G. Hannum, J. Guinney, L. Zhao, et al., "Genome-Wide Methylation Profiles Reveal Quantitative Views of Human Aging Rates," *Molecular Cell* 49, no. 2 (January 2013): 359–367.

22. C. G. Bell, R. Lowe, P. D. Adams, et al., "DNA Methylation Aging Clocks: Challenges and Recommendations," *Genome Biology* 20, no. 1 (November 2019): 249.

23. C. Maulani and E. I. Auerkari, "Age Estimation Using DNA Methylation Technique in Forensics: A Systematic Review," *Egyptian Journal of Forensic Sciences* 10, no. 1 (December 2020): 38.

24. M. Kayser, W. Branicki, W. Parson, and C. Phillips, "Recent Advances in Forensic DNA Phenotyping of Appearance, Ancestry and Age," *Forensic Science International: Genetics* 65 (July 2023): 102870.

25. J. Naue, "Getting the Chronological Age Out of DNA: Using Insights of Age-Dependent DNA Methylation for Forensic DNA Applications," *Genes & Genomics* 45, no. 10 (October 2023): 1239–1261.

26. M. R. Refn, M. L. Kampmann, N. Morling, J. Tfelt-Hansen, C. Børsting, and V. Pereira, "Prediction of Chronological Age and Its Applications in Forensic Casework: Methods, Current Practices, and

Future Perspectives," *Forensic Sciences Research* 8, no. 2 (August 2023): 85–97.

27. M. J. Castagnola, F. Medina-Paz, and S. C. Zapico, "Uncovering Forensic Evidence: A Path to Age Estimation Through DNA Methylation," *International Journal of Molecular Sciences* 25, no. 9 (January 2024): 4917.

28. M. Spólnicka, R. Zbieć-Piekarska, M. Karp, et al., "DNA Methylation Signature in Blood Does Not Predict Calendar Age in Patients With Chronic Lymphocytic Leukemia But May Alert to the Presence of Disease," *Forensic Science International: Genetics* 34 (May 2018): e15–e17.

29. J. Fleckhaus, P. Bugert, N. A. M. Al-Rashedi, and M. A. Rothschild, "Investigation of the Impact of Biogeographic Ancestry on DNA Methylation Based Age Predictions Comparing a Middle East and a Central European Population," *Forensic Science International: Genetics* 67 (November 2023): 102923.

30. U. Lim and M. A. Song, "Dietary and Lifestyle Factors of DNA Methylation," *Methods in Molecular Biology* 863 (2012): 359–376.

31. M. H. So and H. Y. Lee, "Genetic Analyzer-Dependent DNA Methylation Detection and Its Application to Existing Age Prediction Models," *Electrophoresis* 42, no. 14–15 (2021): 1497–1506.

32. S. R. Hong and K. J. Shin, "Can We Integrate Method-Specific Age-Predictive Models?: Analysis Method-Induced Differences in Detected DNA Methylation Status," *Forensic Science International: Genetics* 62 (January 2023): 102805.

33. Y. Hamano, S. Manabe, C. Morimoto, S. Fujimoto, and K. Tamaki, "Forensic Age Prediction for Saliva Samples Using Methylation-Sensitive High Resolution Melting: Exploratory Application for Cigarette Butts," *Scientific Reports* 7, no. 1 (September 2017): 10444.

34. J. W. Lee, C. M. Choung, J. Y. Jung, H. Y. Lee, and S. K. Lim, "A Validation Study of DNA Methylation-Based Age Prediction Using Semen in Forensic Casework Samples," *Legal Medicine* 31 (March 2018): 74–77.

35. X. Han, C. Xiao, S. Yi, Y. Li, M. Chen, and D. Huang, "Accurate Age Estimation From Blood Samples of Han Chinese Individuals Using Eight High-Performance Age-Related CpG Sites," *International Journal of Legal Medicine* 136, no. 6 (November 2022): 1655–1665.

36. H. Y. Lee, S. E. Jung, Y. N. Oh, A. Choi, W. I. Yang, and K. J. Shin, "Epigenetic Age Signatures in the Forensically Relevant Body Fluid of Semen: A Preliminary Study," *Forensic Science International: Genetics* 19 (November 2015): 28–34.

37. B. Bekaert, A. Kamalandua, S. C. Zapico, W. Van de Voorde, and R. Decorte, "Improved Age Determination of Blood and Teeth Samples Using a Selected Set of DNA Methylation Markers," *Epigenetics* 10, no. 10 (October 2015): 922–930.

38. S. Gopalan, J. Gaige, and B. M. Henn. *DNA Methylation-Based Forensic Age Estimation in Human Bone*. bioRxiv (2019): 801647.

39. K. Fokias, L. Dierckx, V. W. V. de, and B. Bekaert, "Age Determination Through DNA Methylation Patterns in Fingernails and Toenails," *Forensic Science International: Genetics* 64 (May 2023): 102846.

40. QIAGEN. *QIAamp DNA Mini Kit—Protocol* (2018), https://whttps://www.qiagen.com/us/resources/resourcedetail?id=566f1cb1-4ffe-4225-a6de-6bd3261dc920&lang=enww.qiagen.comhttps://www.qiagen.com/us/resources/resourcedetail?id=566f1cb1-4ffe-4225-a6de-6bd3261dc920&lang=enhttps://www.qiagen.com/us/resources/resourcedetail?id=566f1cb1-4ffe-4225-a6de-6bd3261dc920&lang=en/us/resources/resourcedetail?id=566f1cb1-4ffe-4225-a6de-6bd3261dc920&lang=en. Last Accessed November 26, 2024.

41. Promega. *PowerQuant System—Technical Manual* (2022), https://ita.promega.com/-/media/files/resources/protocols/technical-manuals/tmd/powerquant-system-technical-manual.pdf?rev=eed77410b69d43a2b0ea03fc78c7b69c&sc_lang=en. Last Accessed November 26, 2024.

42. Zymo Research. *EZ DNA Methylation-DirectTM Kit—Protocol* (2022), https://files.zymoresearch.com/protocols/_d5020_d5021_ez_dna_methylation-direct_kit.pdf. Last Accessed November 26, 2024.

43. QIAGEN. *QIAGEN Multiplex PCR Handbook* (2010), https://www.qiagen.com/us/resources/resourcedetail?id=a541a49c-cd06-40ca-b1d2-563d0324ad6c&lang=en. Last Accessed November 11, 2024.

44. Applied BiosystemsTM. *ExoSAP-ITTM Express PCR Product Cleanup—Brief Protocol* (2017), https://assets.thermofisher.com/TFS-Assets/LSG/manuals/75001b.pdf. Last Accessed November 26, 2024.

45. Applied BiosystemsTM. *ABI PRISM SNaPshotTM Multiplex Kit—Quick Reference Card* (2015), https://assets.thermofisher.com/TFS-Assets/LSG/manuals/cms_041205.pdf. Last Accessed November 26, 2024.

46. Applied BiosystemsTM. *Shrimp Alkaline Phosphatase (SAP)—Brief Protocol* (2016), https://assets.thermofisher.com/TFS-Assets/LSG/manuals/78390B.pdf. Last Accessed November 26, 2024.

47. C. Zaiontz, *Real Statistics Using Excel*, https://real-statistics.com/. Last Accessed November 26, 2024.

48. M. P. Boks, E. M. Derks, D. J. Weisenberger, et al., "The Relationship of DNA Methylation With Age, Gender and Genotype in Twins and Healthy Controls," *PLoS ONE* 4, no. 8 (August 2009): e6767.

49. H. C. Dias, C. Cordeiro, J. Pereira, et al., "DNA Methylation Age Estimation in Blood Samples of Living and Deceased Individuals Using a Multiplex SNaPshot Assay," *Forensic Science International* 311 (June 2020): 110267.

50. H. S. G. Al-Ghanmy, N. A. M. Al-Rashedi, and A. Y. Ayied, "Age Estimation by DNA Methylation Levels in Iraqi Subjects," *Gene Reports* 23 (June 2021): 101022.

51. E. Paparazzo, V. Lagani, S. Geracitano, et al., "An ELOVL2-Based Epigenetic Clock for Forensic Age Prediction: A Systematic Review," *International Journal of Molecular Sciences* 24, no. 3 (January 2023): 2254.

52. L. Feng, F. Peng, S. Li, et al., "Systematic Feature Selection Improves Accuracy of Methylation-Based Forensic Age Estimation in Han Chinese Males," *Forensic Science International: Genetics* 35 (July 2018): 38–45.

53. B. R. Kirkwood and J. A. C. Sterne, *Essential Medical Statistics* (Hoboken: John Wiley & Sons, 2010): 514.

54. I. Garali, M. Sahbatou, A. Daunay, et al., "Improvements and Inter-Laboratory Implementation and Optimization of Blood-Based Single-Locus Age Prediction Models Using DNA Methylation of the ELOVL2 Promoter," *Scientific Reports* 10, no. 1 (September 2020): 15652.

55. A. Aliferi, S. Sundaram, D. Ballard, et al., "Combining Current Knowledge on DNA Methylation-Based Age Estimation Towards the Development of a Superior Forensic DNA Intelligence Tool," *Forensic Science International: Genetics* 57 (March 2022): 102637.

56. P. M. Warnecke, C. Stirzaker, J. R. Melki, D. S. Millar, C. L. Paul, and S. J. Clark, "Detection and Measurement of PCR Bias in Quantitative Methylation Analysis of Bisulphite-Treated DNA," *Nucleic Acids Research* 25, no. 21 (November 1997): 4422–4426.

57. E. A. Moskalev, M. G. Zavgorodnij, S. P. Majorova, et al., "Correction of PCR-Bias in Quantitative DNA Methylation Studies by Means of Cubic Polynomial Regression," *Nucleic Acids Research* 39, no. 11 (June 2011): e77.

58. C. Soriano-Tárraga, J. Jiménez-Conde, E. Giralt-Steinhauer, et al., "DNA Isolation Method Is a Source of Global DNA Methylation Variability Measured With LUMA. Experimental Analysis and a Systematic Review," *PLoS ONE* 8, no. 4 (April 2013): e60750.

59. N. H. Staunstrup, A. Starnawska, M. Nyegaard, et al., "Genome-Wide DNA Methylation Profiling With MeDIP-Seq Using Archived Dried Blood Spots," *Clinical Epigenetics* 8 (2016): 81.

60. K. Groen, R. A. Lea, V. E. Maltby, R. J. Scott, and J. Lechner-Scott, "Letter to the Editor: Blood Processing and Sample Storage Have Negligible Effects on Methylation," *Clinical Epigenetics* 10, no. 1 (February 2018): 22.

61. R. M. Walker, L. MacGillivray, S. McCafferty, et al., “Assessment of Dried Blood Spots for DNA Methylation Profiling,” *Wellcome Open Research* 4 (March 2019): 44.

62. L. Montrose, J. M. Goodrich, M. Morishita, et al., “Neonatal Lead (Pb) Exposure and DNA Methylation Profiles in Dried Bloodspots,” *International Journal of Environmental Research and Public Health* 17, no. 18 (September 2020): 6775.

63. H. R. Gosselt, P. H. Griffioen, Z. B. D. van, N. Oosterom, J. R. de, and S. G. Heil, “Global DNA (Hydroxy)Methylation Is Stable Over Time Under Several Storage Conditions and Temperatures,” *Epigenetics* 16, no. 1 (July 2020): 45–53.

64. H. R. Dash and S. Das, “Microbial Degradation of Forensic Samples of Biological Origin: Potential Threat to Human DNA Typing,” *Molecular Biotechnology* 60, no. 2 (February 2018): 141–153.

65. R. H. A. Hady, H. Z. Thabet, N. E. Ebrahem, and H. A. Yassa, “Thermal Effects on DNA Degradation in Blood and Seminal Stains: Forensic View,” *Academic Forensic Pathology* 11, no. 1 (March 2021): 7–23.

66. M. Poetsch, P. Markwerth, H. Konrad, T. Bajanowski, and J. Helmus, “About the Influence of Environmental Factors on the Persistence of DNA—A Long-Term Study,” *International Journal of Legal Medicine* 136, no. 3 (May 2022): 687–693.

67. L. Sliskovic, I. Milos, A. Zecic, S. Kuret, and D. Sutlovic, “Does Sunlight Affect the Quality for Purposes of DNA Analysis of Blood Stain Evidence Collected From Different Surfaces?,” *Genes (Basel)* 15, no. 7 (July 2024): 888.

68. L. Bhoyar, P. Mehar, and K. Chavali, “An Overview of DNA Degradation and Its Implications in Forensic Caseworks,” *Egyptian Journal of Forensic Sciences* 14, no. 1 (March 2024): 15.

69. F. Peng, L. Feng, J. Chen, et al., “Validation of Methylation-Based Forensic Age Estimation in Time-Series Bloodstains on FTA Cards and Gauze at Room Temperature Conditions,” *Forensic Science International: Genetics* 40 (May 2019): 168–174.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.