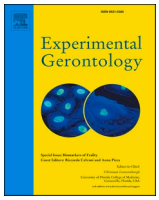
Experimental Gerontology 211 (2025) 112918

Contents lists available at ScienceDirect

Experimental Gerontology

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



EpiClock; biological age measurement from blood DNA methylation using a minimal CpG marker set for high-throughput iPlex mass spectrometry assay for screening in drug development and population health☆

Hyojung Kim [a], Ah-Hyun Park [a], Minjae Kwon [a], Kyoung Joo Lee [a], Min-Jeong Kim [a], Min-Seob Lee [a,b,*]

[a] Eone-Diagnomics Genome Center, 143 Gaetbeol-ro, Yeonsu-gu, Incheon, Republic of Korea
[b] Diagnomics, 5050 Murphy Canyon Rd, San Diego, CA, 92123, USA

ARTICLE INFO

Keywords:
Biological aging
CpGs
DNA methylation
epigenetic clocks
MassARRAY
iplex
EpiClock

ABSTRACT

Measuring biological aging—the underlying rate of physiological decline—has become increasingly important in predicting disease risk, guiding personalized health strategies, and advancing age-targeted therapeutics. Unlike chronological age, biological age provides a more accurate reflection of an individual's functional health and longevity potential.

DNA methylation-based epigenetic clocks are established tools for estimating biological age, but existing assays often require hundreds to thousands of CpG sites, resulting in high costs, complex analysis, and poor scalability. These limitations hinder their practical use in large-scale population screening and routine clinical applications.

To address this gap, we developed EpiClock, a streamlined and cost-effective biological age prediction model based on just eight age-associated CpG markers (ASPA, FHL2, MIR29B2CHG, Chr16q24.1, SLC12A5, SST, LDB2, and COL1A1) from blood-derived DNA. The model showed high accuracy ($R^2 = 0.9332$; mean absolute deviation (MAD) = 3.78 years), strong inter-operator reproducibility ($R^2 = 0.9667$), and robust intra-assay precision (CV < 6 %, SD ≤ 0.05), with optimal performance using buffy coat samples.

By combining a minimal marker set with high analytical reliability, EpiClock enables scalable, high-throughput biological age assessment—supporting its use in drug development, population-level screening, and accessible precision aging diagnostics.

1. Introduction

Aging is a multifactorial biological process characterized by a progressive decline in physiological function, increased vulnerability to chronic diseases, and eventual mortality. While chronological age provides a simple metric of time, it does not reflect the true biological heterogeneity in aging trajectories among individuals. In contrast, biological age—the molecular and functional state of an organism—has emerged as a more meaningful indicator of healthspan, disease risk, and lifespan potential (Bell et al., 2019), (Bocklandt et al., 2011).

Among biological aging biomarkers, epigenetic modifications, particularly DNA methylation, have shown exceptional promise. DNA methylation involves the addition of methyl groups to cytosine residues, primarily at CpG dinucleotides, and plays a critical role in regulating gene expression, genomic stability, and cellular identity. With age, the methylome undergoes predictable changes, including site-specific hypermethylation and global hypomethylation—patterns that correlate strongly with chronological age but also capture deviations reflective of stress, disease, and lifestyle exposures (Chen et al., 2016), (Dhingra et al., 2020).

These methylation changes form the basis for epigenetic clocks, which are computational models that estimate biological age by analyzing DNA methylation at specific CpG sites. Among these, Steve Horvath's pan-tissue clock was the first to demonstrate high-precision age prediction across diverse tissues using 353 CpG sites (median error < 4 years) (Rivier, 2025). Subsequent clocks—such as Hannum's

☆ This article is part of a Special issue entitled: 'Molecular and Epigenetic Aspects of Aging' published in Experimental Gerontology.

* Corresponding author at: Eone-Diagnomics Genome Center, 143 Gaetbeol-ro, Yeonsu-gu, Incheon, Republic of Korea.
E-mail address: mlee@edgc.com (M.-S. Lee).

https://doi.org/10.1016/j.exger.2025.112918
Received 24 April 2025; Received in revised form 29 September 2025; Accepted 30 September 2025
Available online 9 October 2025

whole-blood model (Marttila et al., 2015), DNAm PhenoAge (Jones et al., 2015), GrimAge (Lu et al., 2019), and the DunedinPoAm algorithm (Hong et al., 2017)—have expanded on this concept by associating methylation-based age acceleration with morbidity, cognitive decline, cardiovascular disease, cancer, and all-cause mortality. Importantly, these clocks do more than track time—they capture the biological consequences of stress, inflammation, and disease at the molecular level, thus offering a dynamic readout of health and aging.

In comparison, traditional aging biomarkers like telomere length, though historically well-studied, suffer from low resolution, high inter-individual variability, and inconsistent associations with functional outcomes (Müezzinler et al., 2013). In contrast, DNA methylation-based biomarkers provide quantitative, tissue-consistent, and highly reproducible readouts of aging and health status, positioning them as the gold standard for modern biological age assessment.

Despite their scientific strengths, current epigenetic clocks typically rely on genome-wide microarray or sequencing platforms that analyze hundreds to thousands of CpG sites. These methods are expensive, technically complex, and not easily scalable—factors that hinder their application in large-scale health screening or routine clinical practice (Field et al., 2018).

To overcome these barriers, we developed EpiClock, a highly efficient, cost-effective, and scalable DNA methylation-based biological age test. The EpiClock model uses only eight carefully selected CpG markers (ASPA, FHL2, MIR29B2CHG, Chr16q24.1, SLC12A5, SST, LDB2, and COL1A1), chosen through literature review and empirical validation for strong age correlation. These markers are implemented in a single multiplexed assay, enabling simultaneous amplification and extension reactions, drastically reducing hands-on time, cost, and sample requirements.

The assay is conducted on the Agena MassARRAY platform, a MALDI-TOF-based system that enables high-resolution, label-free quantification of methylation through mass-based detection. Compared to fluorescence-based or sequencing assays, MassARRAY offers exceptional precision, reproducibility, low sample input, and high-throughput compatibility, making it well-suited for clinical and population-scale applications.

In this study, we present the development and validation of EpiClock, and its robustness across sample types. By leveraging minimal methylation markers and scalable analysis technology, EpiClock delivers an accessible and powerful tool for biological age assessment—supporting early disease detection, risk stratification, therapeutic monitoring, and population-level health interventions.

2. Methods

2.1. Overview of quantitative methylation detection by MALDI-TOF mass spectrometry

The process of methylation detection using MALDI-TOF mass spectrometry begins with the isolation of genomic DNA from blood cells. The extracted DNA is then subjected to bisulfite conversion, a chemical

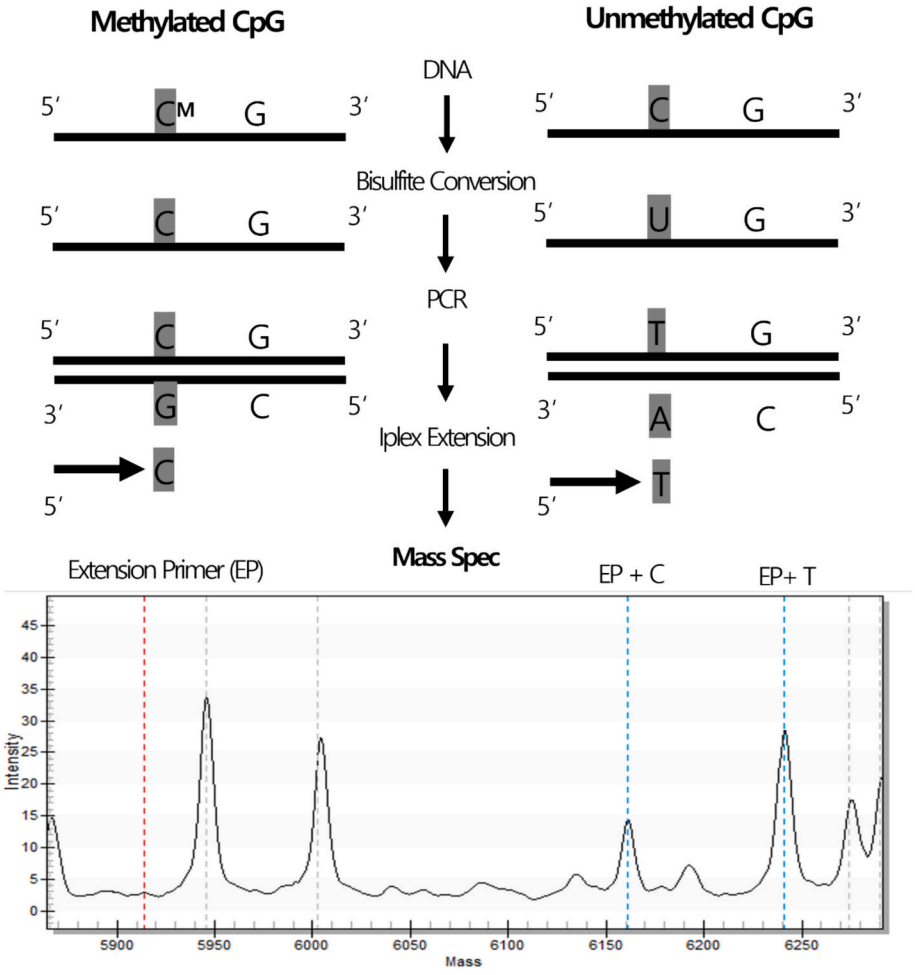


Fig. 1. Workflow of DNA methylation analysis using bisulfite conversion and MALDI-TOF MS.
Genomic DNA is bisulfite-converted to distinguish methylated and unmethylated cytosines. After PCR and single-base extension, the products are analyzed by MALDI-TOF mass spectrometry to determine site-specific methylation based on mass differences.

treatment that distinguishes methylated from unmethylated cytosines (Fig. 1). During this process, unmethylated cytosines are deaminated and converted into uracils, whereas methylated cytosines remain unchanged.

Following bisulfite conversion, the DNA undergoes PCR amplification, during which uracil residues (from unmethylated cytosines) are amplified as thymine, ultimately pairing with adenine, while methylated cytosines are preserved as cytosines and amplified as guanine. This sequence-dependent change provides the basis for detecting methylation status.

Before the primer extension step, a Shrimp Alkaline Phosphatase (SAP) reaction is performed to eliminate any residual dNTPs, preventing nonspecific incorporation during the downstream iPLEX reaction.

In the iPLEX single-base extension step, an extension primer anneals just upstream of the CpG site of interest and incorporates a single terminating nucleotide. The identity of the incorporated base depends on whether the original cytosine at that site was methylated or not. Because the extended products differ in mass depending on the incorporated nucleotide, they can be accurately resolved using MALDI-TOF mass spectrometry.

This method enables high-throughput, base-specific, and quantitative measurement of DNA methylation at targeted CpG sites.

2.2. Target CpG sites selection for epigenetic clock

Our CpG site selection followed a systematic two-step approach combining literature-based screening with empirical validation. We initially compiled a list of over 100 CpG markers previously reported to show significant age-associated methylation changes, based on comprehensive reviews of the literature and available datasets. From this pool, we shortlisted 37 CpG sites located in 31 genes that demonstrated consistent correlation with chronological age across independent studies (Table 1).

We subsequently evaluated these 37 candidate CpGs in our dataset of 113 individuals. The selection of final markers was based on statistical strength of association with chronological age. Specifically, we selected the eight CpG sites that exhibited the most robust and statistically significant correlations, ensuring both biological relevance and predictive performance.

Notably, several of the CpGs included in our model are also present in established epigenetic clock models. FHL2, SST, and COL1A1 are included in Horvath's 353-CpG clock, underscoring their established relevance in aging biology. ASPA and MIR29B2CHG are present in the Hannum clock and other independent studies of age-related methylation patterns. Chr16q24.1 was identified during our literature and dataset mining as a CpG region with strong age-associated methylation dynamics, although it is not widely represented in existing clock models. SLC12A5 and LDB2, while not part of prior clocks, showed highly significant age correlations in our data and have known roles in neurodevelopmental and metabolic pathways linked to aging, justifying their inclusion.

In summary, our final CpG panel represents a refined selection of markers that combine prior biological evidence with new empirical validation, ensuring strong performance while maintaining clinical feasibility.

2.3. Sample collection and preparation

Peripheral blood samples were collected in K2-EDTA tubes with a minimum volume of 3 mL and stored at 4 °C prior to processing. Buffy coat was isolated within 5 days of collection to minimize degradation. Genomic DNA (gDNA) was extracted using the QIAamp DNA Mini Kit (QIAGEN), following the manufacturer's protocol. DNA quality was assessed using Nanodrop One and Qubit 2.0 fluorometry, with inclusion criteria of A260/A280 ratio between 1.6 and 2.1 and a concentration ≥ 10 ng/μL (total yield ≥500 ng).

Table 1
List of CpG sites associated with age from previous studies.

CpG ID	GRCh38 chromosome position	Associated gene	Age-associated methylation pattern	Reference
cg07211259	chr9:5510497	PDCD1LG2	hypo	Marttila et al., 2015
cg00674365	chr19:56507700	ZNF471	hyper	
cg06352730	chr7:47582094	TNS3	hyper	
cg16867657	chr6:11044644	ELOVL2	hyper	
cg16762684	chr18:77108537	MBP	hypo	
cg26063719	chr10:17231188	VIM	hypo	
cg08548498	chr20:45254905	SLPI	hyper	
cg04875128	chr15:31483692	OTUD7A	hyper	
cg08262002	chr4:16573700	LDB2	hyper	
cg18618815	chr17:50197963	COL1A1	hypo	
cg21801378	chr15:72319784	CELF6	hyper	Jones et al., 2015
cg22736354	chr6:18122488	NHLRC1	hyper	
cg00059225	chr5:151924796	GLRA1	hyper	
cg01820374	chr12:6772917	LAG3	hypo	
cg06291867	chr10:90857405	HTR7	hyper	
cg06493994	chr6:25652374	SCGN	hyper	
cg09809672	chr1:236394382	EDARADD	hypo	
cg17861230	chr19:18233091	PDE4C	hyper	
cg19722847	chr12:30696180	IPO8	hypo	
cg27320127	chr2:47571257	KCNK12	hyper	
cg21572722	chr6:11044661	ELOVL2	hyper	Freire-Aradas et al., 2016
cg02228185	chr17:3476273	ASPA	hypo	
rs1456897271	chr19:18233027	PDE4C	hypo	
cg06639320	chr2:105399282	FHL2	hyper	
cg19283806	chr18:68722183	CCDC102B	hypo	
rs1257890264	chr1:207823715	not associated	hypo	
cg07082267	chr16:85395429	not associated	hypo	
cg07547549	chr20:46029586	SLC12A5	hyper	Hong et al., 2017
cg17885226	chr6:104940856	LIN28B-AS1	hyper	
cg00481951	chr3:187669862	SST	hyper	
cg08097417	chr7:130734372	KLF14	hyper	
cg21296230	chr15:32718335	GREM1	hyper	
cg17110586	chr19:35963721	not associated	hyper	
cg15480367	chr14:92923140	not associated	hyper	
cg19802138	chr13:112068405	SOX1	hyper	
cg19671120	chr2:98346511	CNGA3	hyper	
cg05009601	chr7:122304446	FEZF1	hyper	

2.4. DNA extraction and bisulfite conversion

Genomic DNA (gDNA) was extracted from buffy coat samples using the QIAamp DNA Mini Kit (QIAGEN) in accordance with the manufacturer's protocol. The concentration and purity of the extracted DNA were assessed using the Nanodrop One spectrophotometer (Thermo Fisher Scientific) and the Qubit 2.0 Fluorometer (Invitrogen), with inclusion criteria of ≥10 ng/μL in concentration (≥500 ng total yield) and an A260/A280 ratio between 1.6 and 2.1.

For methylation analysis, 500 ng of high-quality gDNA was subjected to bisulfite conversion using the Epi-geno Kit (EDGC, Cat. No. EPGN50R). In this procedure, 20 μL of the DNA sample was combined with 10 μL of denaturation buffer in a PCR tube. After thorough mixing, 150 μL of freshly prepared CT conversion reagent was added to the mixture. The sample was then placed in a thermal cycler and subjected to the following program to facilitate denaturation and bisulfite conversion: 98 °C for 5 min, 50 °C for 1 h, and 4 °C for 10 min, followed by a hold at 4 °C in the dark.

Following the conversion reaction, 400 μL of binding buffer was added to a purification column, and the bisulfite-treated DNA mixture was transferred to the column and gently mixed by inversion. The column was centrifuged at ≥10,000 ×*g* for 30 s to bind the DNA.

To remove sulfonated cytosine residues, the column was first washed

with 100 μL of washing buffer and centrifuged, followed by the addition of 50 μL of desulphonation buffer and a 15-min incubation at room temperature. The column was then washed twice with 200 μL of washing buffer, each time followed by centrifugation, and a final centrifugation step of 3 min was performed to eliminate any residual buffer.

Finally, the purified, bisulfite-converted DNA was eluted by adding ≥10 μL of elution buffer to the column, incubating for 5 min, and centrifuging for 1 min. The resulting DNA was stored at −20 °C until use in downstream methylation analysis.

2.5. Primer design for iPLEX MassARRAY

For each of the eight selected CpG sites used in the EpiClock model, flanking genomic sequences were obtained from the UCSC Genome Browser (GRCh37/hg19 assembly). A target region comprising 100 base pairs upstream and downstream of the CpG site of interest was extracted to ensure optimal primer placement.

To simulate the effects of bisulfite conversion, all cytosines (C) in the non-CpG context were replaced with thymines (T). The CpG sites of interest were annotated as putative C/T polymorphisms (i.e., G) to enable methylation quantification via mass spectrometry. Any known SNPs within the amplicon or extension region were labeled using IUPAC nucleotide codes to account for potential allelic variation.

PCR and single-base extension primers were designed using Assay Design Suite v2.0 (Agena Bioscience, San Diego, CA, USA). Primer sets were optimized to avoid overlapping mass spectra and to ensure compatibility within a single multiplex reaction. The design parameters included amplicon size (~150–250 bp), melting temperature (Tm), GC content, and absence of secondary structures or primer-dimer formation. All primers were synthesized by a commercial provider and validated prior to analytical use.

2.6. iPlex PCR and primer extension

Bisulfite-converted DNA (2 μL) was used as the template for multiplex PCR amplification targeting eight age-associated CpG markers: ASPA, FHL2, MIR29B2CHG, Chr16q24.1, SLC12A5, SST, LDB2, and COL1A1. A single multiplexed primer pool was designed to enable the simultaneous amplification of all eight loci in one reaction. PCR was performed under optimized cycling conditions (95 °C for 2 min; 45 cycles of 95 °C for 30 s, 56 °C for 30 s, and 72 °C for 1 min; final extension at 72 °C for 5 min).

PCR products were treated with Shrimp Alkaline Phosphatase (SAP) to remove unincorporated dNTPs, followed by a single-base primer extension reaction using the iPLEX Pro kit (Agena Bioscience). Extension primers were designed to target CpG sites of interest and differentiate methylated vs. unmethylated cytosines via mass shift. Extension cycling followed the manufacturer's protocol.

2.7. MassARRAY mass spectrometry analysis

Quantitative methylation analysis was performed using the MassARRAY® system (Agena Bioscience), which employs matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry for high-resolution detection of base-specific methylation at selected CpG sites. This technology enables precise differentiation between methylated and unmethylated cytosines based on their mass signatures following bisulfite conversion and single-base primer extension.

After bisulfite conversion and multiplex PCR amplification, residual dNTPs were enzymatically removed using Shrimp Alkaline Phosphatase (SAP). This was followed by an iPLEX® single-base extension (SBE) reaction, in which extension primers were specifically designed to anneal immediately adjacent to the target CpG site. During the iPLEX reaction, either a dideoxycytidine (ddC, indicating a methylated site) or a dideoxythymidine (ddT, indicating an unmethylated site) was incorporated into the primer depending on the original methylation state of the cytosine.

The resulting extension products were desalted and loaded onto a SpectroCHIP® Array using the NanoDispenser RS1000 (Agena Bioscience). Each chip was then analyzed using the MassARRAY Analyzer 4, a MALDI-TOF mass spectrometer that detects the distinct mass differences between the extension products. Each methylation call was represented by a specific mass peak, with methylated and unmethylated forms producing clearly separated signals.

Methylation ratios were quantified using the MassARRAY Typer Analyzer software, which calculates the relative abundance of methylated versus unmethylated cytosines at each site based on peak intensity. The DNA methylation percentage (%) is calculated by dividing the signal intensity of the methylated cytosine (C) by the total signal intensity of both methylated (C) and unmethylated (T) cytosines, then multiplying by 100.

Methylation (%) = (Signal intensity of C / (Signal intensity of C + Signal intensity of T)) × 100.

This formula provides a quantitative measure of the methylation level at a specific CpG site.

This method provides high analytical sensitivity and specificity, with single-nucleotide resolution and the ability to multiplex multiple CpG targets in a single reaction. In the current study, eight CpG markers were analyzed per sample in a single well, providing robust and reproducible methylation profiles suitable for downstream biological age estimation.

2.8. Data analysis

Statistical analysis was performed to evaluate the accuracy, precision, and reproducibility of the EpiClock model. The relationship between predicted biological age and actual chronological age was assessed using linear regression, with the coefficient of determination (R^2) indicating model accuracy. The mean absolute deviation (MAD) was calculated to measure the average prediction error in years.

To assess reproducibility, inter-operator consistency was evaluated by analyzing the same samples processed independently by different operators, and intra-assay precision was determined using replicate measurements within a single run. For each CpG site, the methylation percentage was calculated as the ratio of methylated signal intensity (C) to the sum of methylated and unmethylated signals (C + T), expressed as a percentage.

3. Results

3.1. Selection of aging-related DNA methylation markers and EpiClock model development

To identify CpG sites associated with biological aging, we conducted both an extensive literature review and a data-driven analysis using available DNA methylation databases. Specifically, CpG sites that showed consistent differences in methylation levels across age groups were selected through in silico screening of large-scale epigenomic datasets, including age-stratified cohorts.

These sites had previously been reported in the literature to correlate with chronological age and were further supported by database evidence showing age-associated methylation shifts (Fraga et al., 2005), (Hannum et al., 2013).

To experimentally validate their association with aging, we analyzed blood samples from 113 individuals aged 19 to 92 years (Table 2). Among the 37 candidates, eight CpG sites showed statistically significant correlations with chronological age in our dataset. These eight validated markers were subsequently used to construct the EpiClock model for estimating biological age.

Table 2
Demographic characteristics of participants.

Age (years):	
Mean	51.7
Median	51.0
1st Quartile	34.0
3rd Quartile	69.0
Gender:	
Male	38 (33.6 %)
Female	75 (66.4 %)
Geographic/ethnic background:	
East Asian/Korean	113 (100 %)

3.2. Correlation between individual biomarkers and biological age

The association between DNA methylation levels at each of the eight selected CpG sites and biological age was assessed using linear regression analysis. As shown in Fig. 2, all eight biomarkers demonstrated statistically significant correlations with biological age (p-value <0.0001), though the strength of the correlation varied among markers.

Among the eight markers, FHL2 showed the strongest individual association with biological age ($R^2 = 0.8451$), followed by MIR29B2CHG ($R^2 = 0.7739$), SST ($R^2 = 0.6547$), and chr16q24.1 ($R^2 = 0.6107$). Markers such as ASPA ($R^2 = 0.5309$), COL1A1 ($R^2 = 0.5024$), and LDB2 ($R^2 = 0.5658$) also demonstrated moderate correlations. The lowest, though still significant, correlation was observed in SLC12A5, with an R^2 of 0.2953.

These results indicate that each of the selected CpG markers individually reflects age-related methylation changes. However, the combination of these markers in a multivariate model—as described in the following section—enhances overall predictive accuracy and robustness for estimating biological age.

3.3. Comprehensive predictive model utilizing eight biomarkers

To construct a robust and accurate biological age predictor, we developed the EpiClock model by integrating methylation data from eight validated CpG biomarkers. We constructed the EpiClock model using a weighted linear combination of methylation score (beta values) for the eight CpG sites. The weights were derived from a multivariate linear regression using a training subset of the dataset, with chronological age as the dependent variable. Coefficients were selected to minimize the mean squared error, and the final model is expressed as: 83.961 + (cg02228185 score x − 22.7087) + (rs1257890264 score x − 29.9418) + (cg07082267 score x − 47.3211) + (cg18618815 score x − 47.3102) + (cg06639320 score x 93.23488) + (cg07547549 score x − 12.2298) + (cg00481951 score x 72.33255) + (cg08262002 score x − 25.812). The model's performance was evaluated by comparing predicted biological age with chronological age in blood samples, as illustrated in Fig. 3A. The model achieved a high coefficient of determination ($R^2 = 0.9392$), indicating that 93.92 % of the variation in biological age was explained by the EpiClock. The model also demonstrated strong accuracy, with a mean absolute deviation (MAD) of just 3.78 years, and a highly significant p-value (<0.001).

We also calculated 95 % confidence intervals for our performance metrics using bootstrap resampling methods. The R^2 confidence interval [0.914, 0.954] demonstrates that even at the lower bound, our model explains over 91 % of age variance. The MAD confidence interval [3.2, 4.6 years] shows that even in the worst-case scenario, our prediction accuracy remains clinically excellent with errors under 5 years.

To assess the contribution of individual CpG markers within the model, we analyzed their individual correlations with chronological age. As shown in Fig. 3B, all eight markers exhibited statistically significant associations with age ($p < 0.01$). Among them, FHL2 ($R^2 = 0.86$), COL1A1 ($R^2 = 0.80$), MIR29B2CHG ($R^2 = 0.77$), and chr16:85395429 ($R^2 = 0.74$) showed particularly strong age-related

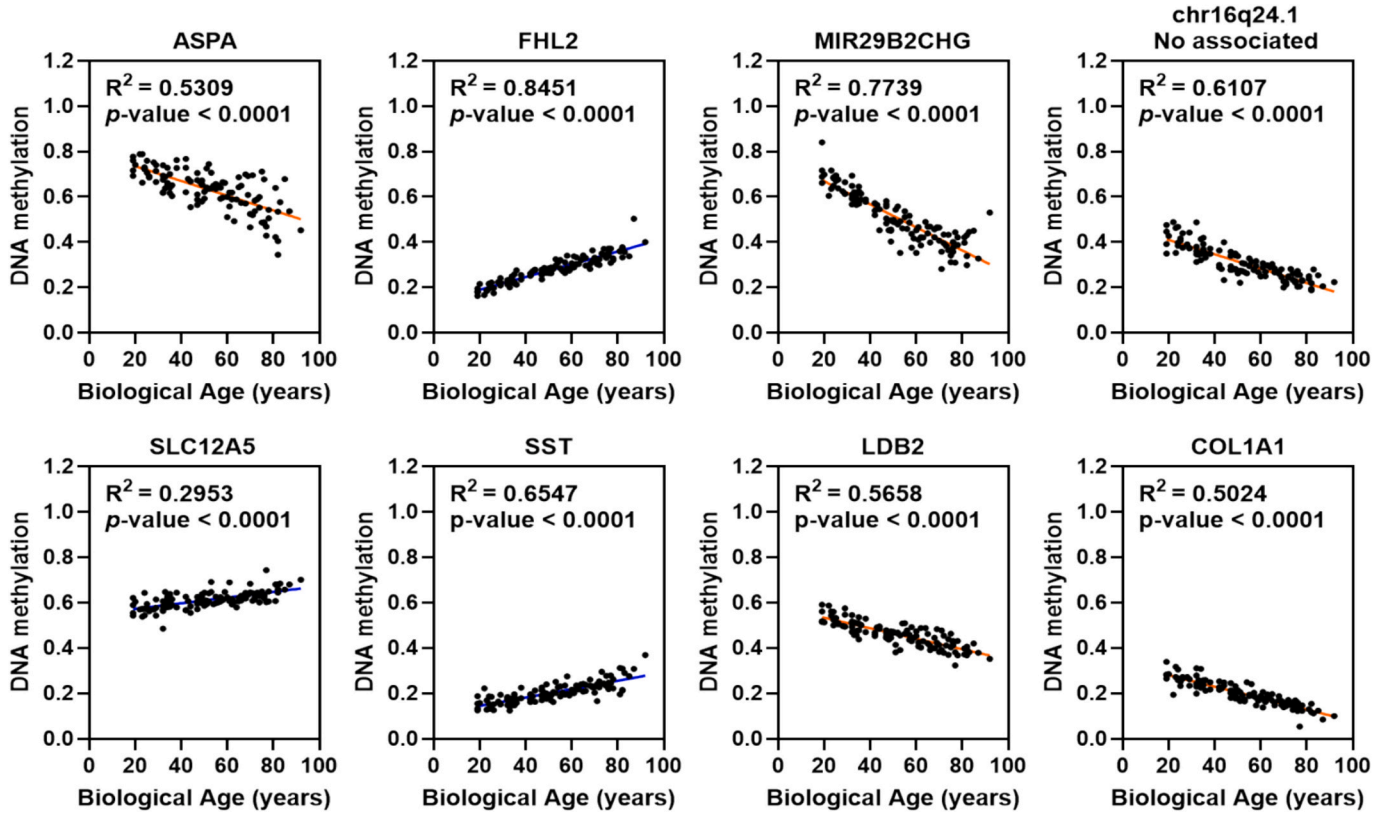
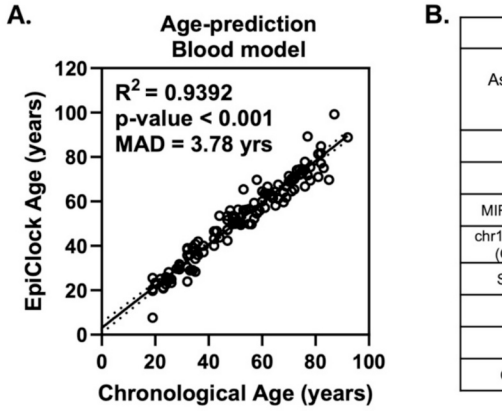


Fig. 2. Epigenetic age prediction model for blood biomarkers. Linear regression plots show the correlation between DNA methylation levels and biological age for each of the eight selected CpG sites. Each plot includes R^2 and p-value to indicate the strength and significance of the association.



B.

Individual Marker-Age Correlations			
Associated Gene	DNA methylation	Correlation with actual age	
		R^2	P-value
ASPA	hypo	0.50	3.42E-03
FHL2	hyper	0.86	3.94E-05
MIR29B2CHG	hypo	0.77	1.64E-03
chr16:85395429 (GRCh38)	hypo	0.74	2.30E-03
SLC12A5	hyper	0.41	5.24E-01
SST	hyper	0.64	3.89E-04
LDB2	hypo	0.68	1.55E-01
COL1A1	hypo	0.80	3.97E-02

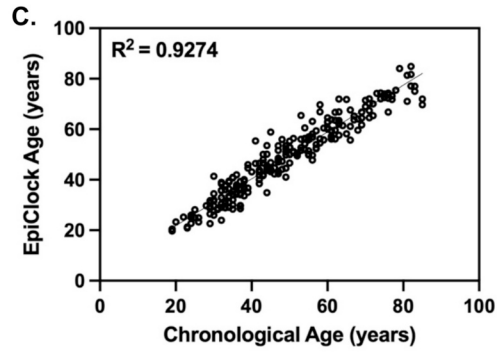


Fig. 3. EpiClock model performance and marker validation. A. Scatter plot comparing EpiClock-predicted biological age with chronological age, showing overall model performance (R^2 = 0.9392). B. Individual CpG marker correlations with chronological age, demonstrating the age-associated methylation changes that justified marker selection. C. Scatter plot of 236 validation samples.

methylation changes. The remaining markers—ASPA, SLC12A5, SST, and LDB2—also contributed meaningfully to the model, demonstrating consistent hypo- or hypermethylation trends across the age spectrum (Supplementary Data).

We also validated EpiClock model using a set of 236 test samples, which is showing a strong relationship between the estimated biological EpiClock and their corresponding chronological ages (Fig. 3C). These findings support the EpiClock model as a reliable, statistically validated, and scalable approach for estimating biological age based on targeted DNA methylation analysis.

3.4. Limit of detection (LOD)

To evaluate the minimum input DNA amount required for reliable methylation measurement, a limit of detection (LOD) study was conducted using four normal samples analyzed at five different DNA input concentrations: 1000 ng, 750 ng, 500 ng, 200 ng, and 100 ng. All inputs were applied prior to bisulfite conversion, and the resulting methylation scores were compared to those obtained under the reference condition of 1000 ng input DNA.

As shown in Fig. 4, methylation measurements remained highly consistent at higher input levels. Specifically, the average correlation coefficient (R^2) between methylation scores at 750 ng and 1000 ng was 0.986, while 500 ng samples showed an R^2 of 0.965, indicating minimal variability. However, further reductions in DNA input led to noticeable declines in correlation. At 200 ng, the average R^2 dropped to 0.949, and at 100 ng, it further declined to 0.861, reflecting increased signal variability and reduced assay precision.

Based on these findings, an input threshold of 500 ng or greater was established as the minimum recommended DNA amount for accurate and reproducible methylation quantification using the EpiClock assay.

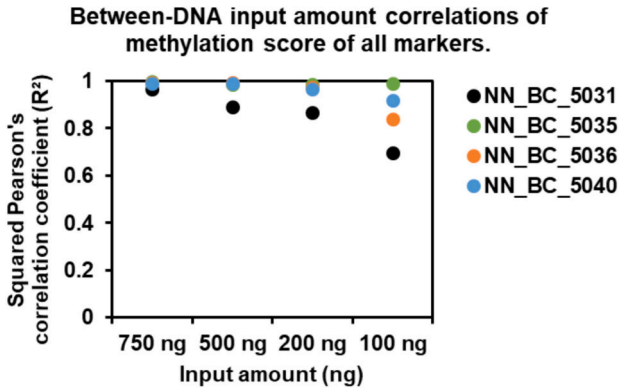


Fig. 4. Correlation between methylation scores derived from 1000 ng input DNA and those from lower input concentrations (750 ng, 500 ng, 200 ng, 100 ng), showing decreasing correlation with reduced DNA quantity.

3.5. Verification of precision and repeatability

3.5.1. Inter-operator consistency

To evaluate the analytical precision and reproducibility of the EpiClock assay, an inter-operator consistency test was performed. Two independent operators processed and analyzed the same set of 24 samples, and methylation scores (%) for the eight CpG markers were compared.

As illustrated in Fig. 5A, the methylation scores obtained by the two operators demonstrated a strong linear relationship, with a regression coefficient (R^2) of 0.9667, indicating excellent agreement between independent analyses. This result confirms that the assay yields highly consistent measurements regardless of the operator.

To further assess assay precision, the coefficient of variation (CV%)

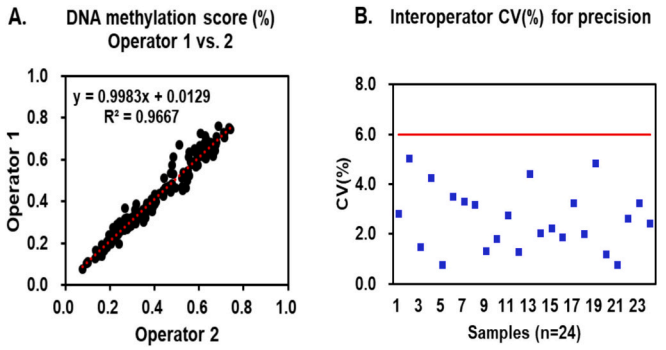


Fig. 5. A. Correlation of methylation scores (%) between two operators across 24 samples. B. Coefficient of variation (percent error) between the two operators.

was calculated for the final biological age estimates between the two operators across all samples. As shown in Fig. 4B, the CV values for all samples remained below 6 %, with the majority clustering well under this threshold. These findings confirm the high repeatability and robustness of the assay for routine biological age estimation.

To assess the inter-operator consistency, two independent operators analyzed the same set of 24 samples, and the methylation scores (%) for the eight markers were recorded (Fig. 5A). The regression analysis between the two independently conducted runs yielded an $R^2 = 0.9667$, indicating no significant differences between the results obtained by the two operators.

Furthermore, the coefficient of variation (CV%) for the final biological age calculations across the two operators was found to be less than 6 %, confirming high consistency (Fig. 5B).

3.5.2. *Intra-assay consistency*

To evaluate the intra-assay consistency of the EpiClock assay, six representative samples were tested in triplicate within the same experimental run. Methylation scores (%) were obtained for all eight CpG biomarkers, and the standard deviation (SD) of each marker across the three replicates was calculated.

As shown in Table 3, the standard deviation for all methylation measurements remained consistently low across all markers and samples. Notably, all SD values were ≤ 0.05, indicating high reproducibility and minimal technical variation within the same assay run. These results confirm that the assay maintains strong intra-assay precision, supporting its suitability for routine and high-throughput biological age testing.

3.6. *Robustness assessment based on sample type*

To evaluate the robustness of the EpiClock assay across different biological sample types, DNA was extracted from both buffy coat and whole blood collected from the same 23 participants (200 μL each). The goal was to assess whether the assay maintains accuracy and consistency despite differences in sample origin.

As shown in Fig. 6A, methylation scores (%) across the eight CpG biomarkers exhibited a high degree of correlation between DNA from buffy coat and whole blood, with an R^2 of 0.9533, confirming strong technical consistency in methylation measurement across sample types.

However, when applying the buffy coat–trained EpiClock model to predict biological age, notable differences in performance were observed. In buffy coat samples, the correlation between predicted biological age and actual age was $R^2 = 0.6451$ (Fig. 6B), whereas in whole blood, the correlation dropped slightly to $R^2 = 0.572$ (Fig. 6C), indicating reduced predictive strength in whole blood samples.

Further comparison of model performance, shown in Fig. 6D, revealed that the mean absolute error (MAE) was 3.30 years for buffy coat samples and 4.64 years for whole blood. Moreover, 5 out of 23 whole blood samples (21.7 %) showed an absolute error greater than 6 years, compared to none in the buffy coat group. This suggests that direct application of a buffy coat–trained model to whole blood samples may lead to increased variability and reduced accuracy.

These findings highlight the importance of sample-type specificity in DNA methylation–based age prediction. While the assay exhibits strong robustness within a given sample type, application to other sample matrices such as whole blood may require independent calibration and validation to maintain performance standards.

Table 3
Standard deviation of triplicate determinations for six samples across eight EpiClock biomarkers. All SD values remained ≤0.05, confirming high intra-assay reproducibility.

Sample	Replicate	Methylated score (%)							
		ASPA	MIR29B2CHG	Chr.16q24.1	COL1A1	FHL2	SLC12A5	SST	LDB2
NN_BC_0565	1	0.623	0.638	0.312	0.177	0.277	0.631	0.217	0.441
	2	0.612	0.625	0.307	0.180	0.286	0.610	0.229	0.427
	3	0.613	0.650	0.308	0.183	0.287	0.604	0.244	0.453
STDEV		0.006	0.012	0.002	0.003	0.005	0.014	0.014	0.013
NN_BC_0717	1	0.556	0.442	0.224	0.077	0.372	0.679	0.324	0.350
	2	0.560	0.505	0.234	0.062	0.362	0.687	0.321	0.322
	3	0.523	0.513	0.233	0.078	0.371	0.709	0.329	0.301
STDEV		0.020	0.039	0.005	0.009	0.006	0.016	0.004	0.025
NN_BC_0808	1	0.666	0.503	0.340	0.177	0.254	0.572	0.226	0.478
	2	0.652	0.466	0.317	0.198	0.266	0.604	0.244	0.448
	3	0.659	0.553	0.327	0.220	0.259	0.575	0.253	0.486
STDEV		0.007	0.043	0.012	0.021	0.006	0.018	0.014	0.020
NN_BC_0842	1	0.678	0.549	0.337	0.162	0.317	0.574	0.269	0.437
	2	0.650	0.588	0.313	0.127	0.284	0.623	0.299	0.423
	3	0.682	0.548	0.330	0.139	0.319	0.584	0.285	0.445
STDEV		0.018	0.023	0.012	0.018	0.020	0.026	0.015	0.011
NN_BC_0847	1	0.714	0.528	0.264	0.133	0.346	0.675	0.327	0.468
	2	0.734	0.457	0.253	0.130	0.363	0.665	0.300	0.426
	3	0.724	0.538	0.272	0.129	0.353	0.670	0.331	0.461
STDEV		0.010	0.045	0.010	0.002	0.008	0.005	0.017	0.023
NN_BC_5041	1	0.713	0.596	0.294	0.168	0.264	0.625	0.225	0.402
	2	0.693	0.588	0.305	0.173	0.275	0.643	0.221	0.389
	3	0.707	0.574	0.282	0.168	0.263	0.615	0.233	0.388
STDEV		0.010	0.011	0.011	0.003	0.007	0.015	0.006	0.008

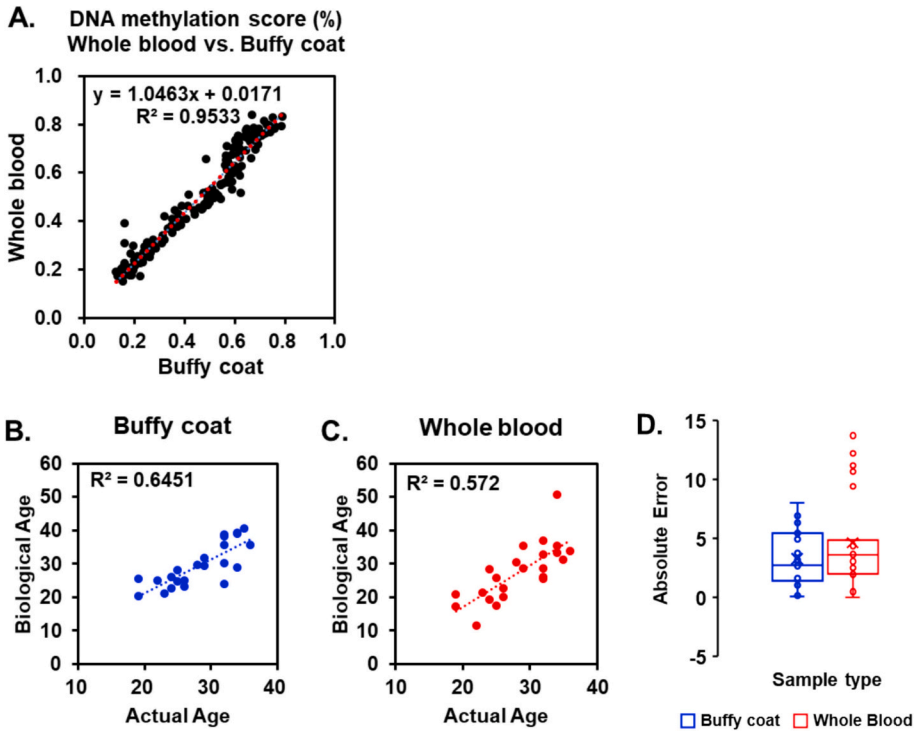


Fig. 6. Comparative analysis of biological age estimation by EpiClock across different sample types. A. Correlation of methylation scores between buffy coat and whole blood ($R^2 = 0.9533$). B. Biological age prediction using buffy coat samples ($R^2 = 0.6451$).
C. Biological age prediction using whole blood samples ($R^2 = 0.572$). D. Comparison of absolute error between predicted and actual age by sample type. Whole blood samples show increased variability and larger errors.

4. Discussion

In this study, we developed and validated EpiClock, a streamlined and highly accurate epigenetic age prediction model based on DNA methylation profiling from blood-derived samples. The model incorporates eight carefully selected age-associated CpG sites—ASPA, FHL2, MIR29B2CHG, Chr16q24.1, SLC12A5, SST, LDB2, and COL1A1-—and was implemented using the MassARRAY iPLEX platform. The EpiClock model demonstrated excellent performance, achieving a coefficient of determination (R^2) of 0.9392 and a mean absolute deviation (MAD) of 3.78 years, indicating that over 93 % of biological age variability could be explained with high precision and minimal error.

Unlike traditional aging biomarkers such as telomere length, which often show high inter-individual variability and low predictive resolution, DNA methylation–based biomarkers offer superior sensitivity and stability. Methylation changes occur in a predictable, tissue-specific, and time-dependent manner, making them more suitable for tracking subtle biological aging processes. In particular, methylation clocks capture cumulative environmental and lifestyle impacts on the epigenome, offering a more comprehensive reflection of biological state than telomere erosion alone.

Among the eight markers in the model, FHL2 ($R^2 = 0.86$), COL1A1 ($R^2 = 0.80$), MIR29B2CHG ($R^2 = 0.77$), and chr16q24.1 ($R^2 = 0.74$) showed particularly strong age correlations. While some markers, such as SLC12A5 ($R^2 = 0.41$), demonstrated weaker individual associations, their inclusion contributed to the multivariate model's improved robustness and overall accuracy. Future iterations of the model may benefit from re-weighting marker contributions or exploring additional age-sensitive loci for further refinement.

The assay also exhibited excellent technical reproducibility and operational robustness. Inter-operator testing yielded a high R^2 of 0.9667, and intra-assay triplicate measurements across six samples produced standard deviations $\leq$0.05, confirming minimal experimental variability. The limit of detection (LOD) study confirmed that at least 500 ng of DNA input is required to maintain analytical reliability, as lower DNA amounts (e.g., 200 ng or less) led to increased variability in methylation scores and reduced correlation with the reference condition.

Importantly, our sample-type comparison revealed that while methylation measurements were highly correlated between buffy coat and whole blood DNA ($R^2 = 0.9533$), the biological age predictions showed notable discrepancies. The EpiClock model performed better on buffy coat samples (MAD = 3.30 years) compared to whole blood (MAD = 4.64 years), with 20 % of whole blood samples exceeding an absolute error of 6 years. These results emphasize that the model is currently optimized for buffy coat DNA, and applying it to other sample types will require further calibration and validation.

One of the most powerful aspects of EpiClock is its minimal-marker design, utilizing only eight CpG sites. This contrasts with many conventional epigenetic clocks that require the analysis of hundreds or thousands of loci via microarray or sequencing-based methods, which are cost-prohibitive and technically complex. Our simplified model significantly reduces assay time, cost, and sample requirements while maintaining comparable accuracy. This streamlined approach is particularly advantageous for clinical implementation, population-level screening, and routine monitoring. While our sample size was statistically justified for initial model development, larger validation studies ($n \geq 500$) are needed to assess generalizability across diverse populations and enable robust subgroup analyses.

Moreover, the EpiClock model holds substantial potential for use in high-throughput drug discovery and therapeutic screening, especially in the field of anti-aging and rejuvenation research. Because the assay quantitatively reflects biological aging with fine resolution, it can serve as an endpoint biomarker in intervention studies, allowing researchers to rapidly assess the epigenetic impact of candidate compounds, lifestyle changes, or emerging therapies targeting aging-related pathways. Its scalability makes it ideal for screening large cohorts or compound libraries in pharmaceutical pipelines.

From a technological perspective, we selected the iPLEX MassARRAY system due to its unique advantages in precision, scalability, and cost-efficiency. Compared to other methylation analysis platforms—such as pyrosequencing, SNaPshot, qPCR, or even EpiTYPER—the iPLEX system offers reliable performance with lower DNA input requirements and efficient multiplexing. While platforms like pyrosequencing offer high single-site resolution, they lack throughput and scalability. In contrast, our MassARRAY-based model balances accuracy, cost, and practicality, making it an attractive solution for translational and applied aging research.

However, this study also has several important limitations that should be considered when interpreting the results. Because predictor CpGs were both selected and evaluated within the same cohort, the possibility of overfitting cannot be excluded, and the relatively high R^2 observed here may represent an optimistic estimate of performance. The model was developed exclusively in a Korean population, which reduces confounding within the study but also limits generalizability, as methylation patterns and age-associated changes can vary across populations depending on genetic background, lifestyle, and environmental exposures. In addition, the predictor was optimized on a targeted MALDI-TOF assay, while most publicly available reference datasets are generated using array-based platforms. Such technical differences can influence methylation measurements and complicate direct replication efforts. Future research should aim to validate the model in independent, multi-ethnic cohorts and, where possible, conduct cross-platform benchmarking to ensure robustness and reproducibility.

In summary, EpiClock is a robust, reproducible, and clinically adaptable biological age prediction tool. Its minimal design, high predictive accuracy, and compatibility with scalable, cost-effective platforms position it as a powerful asset for both personalized medicine and therapeutic screening in the emerging field of epigenetic longevity science. Future work will focus on expanding its utility across different tissues and validating its responsiveness to anti-aging interventions in longitudinal human studies.

5. Conclusion

EpiClock is a next-generation epigenetic clock model that delivers accurate, reproducible, and cost-efficient biological age prediction using a minimal set of eight CpG methylation markers. Achieving a strong correlation with chronological age ($R^2 = 0.9392$) and low error (MAD = 3.78 years), the model is optimized for buffy coat DNA with a minimum input of 500 ng and demonstrates excellent intra- and inter-operator consistency.

By combining precision with scalability, EpiClock is well-suited for clinical use, population screening, and high-throughput therapeutic evaluation. Its minimal-marker design reduces complexity and cost while maintaining high performance, making it a practical tool for personalized aging assessment and early intervention strategies.

With further validation, EpiClock has the potential to become a foundational platform in aging research, supporting anti-aging drug discovery, monitoring of healthspan interventions, and broader applications in precision healthcare.

CRediT authorship contribution statement

Hyojung Kim: Visualization, Validation, Methodology, Data curation, Writing – original draft. **Ah-Hyun Park:** Resources, Methodology. **Minjae Kwon:** Methodology. **Kyoung Joo Lee:** Methodology. **Min-Jeong Kim:** Supervision, Methodology, Data curation, Writing – review & editing. **Min-Seob Lee:** Supervision, Project administration, Methodology, Conceptualization, Writing – review & editing, Writing – original draft.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.exger.2025.112918.

Data availability

Data will be made available on request.

References

Bell, C.G., et al., 2019. DNA methylation aging clocks: challenges and recommendations. Genome Biol. 20 (1), 249.
Bocklandt, S., et al., 2011. Epigenetic predictor of age. PLoS One 6 (6), e14821.
Chen, B.H., et al., 2016. DNA methylation-based measures of biological age: meta-analysis predicting time to death. Aging 8 (9), 1844–1865.
Dhingra, R., et al., 2020. DNA methylation age acceleration is associated with cardiovascular disease risk in a multiethnic study. Clin. Epigenetics 12, 123.
Field, A.E., et al., 2018. DNA methylation clocks in aging: categories, causes, and consequences. Mol. Cell 71 (6), 882–895.
Fraga, M.F., et al., 2005. Epigenetic differences arise during the lifetime of monozygotic twins. Proc. Natl. Acad. Sci. 102 (30), 10604–10609.
Freire-Aradas, A., et al., 2016. Development of a methylation marker set for forensic age estimation using analysis of public methylation data and the Agena Bioscience EpiTYPER system. Forensic Sci. Int. Genet. 24, 65–74.
Hannum, G., et al., 2013. Genome-wide methylation profiles reveal quantitative views of human aging rates. Mol. Cell 49 (2), 359–367.
Hong, S.R., et al., 2017. DNA methylation-based age prediction from saliva: high age predictability by combination of 7 CpG markers. Forensic Sci. Int. Genet. 29, 118–125.
Jones, M.J., Goodman, S.J., Kobor, M.S., 2015. DNA methylation and healthy human aging. Aging Cell 14 (6), 924–932.
Lu, A.T., et al., 2019. DNA methylation GrimAge strongly predicts lifespan and healthspan. Aging (Albany NY) 11 (2), 303–327.
Marttila, S., et al., 2015. Ageing-associated changes in the human DNA methylome: genomic locations and effects on gene expression. BMC Genomics 16, 179.
Müezzinler, A., et al., 2013. Critical evaluation of telomere length measurement techniques and their contribution to aging research. Obes. Rev. 15 (3), 192–201.
Rivier, C.A., 2025. Bidirectional relationship between epigenetic age and stroke, dementia, and late-life depression. Nat. Commun. 16, 1261.