Research Paper

Epigenetic mutation load is weakly correlated with epigenetic age acceleration

Qi Yan[1], Kimberly C. Paul[1], Ake T. Lu[2], Cynthia Kusters[1], Alexandra M. Binder[1,3], Steve Horvath[2,4], Beate Ritz[1,5]

[1]Department of Epidemiology, UCLA Fielding School of Public Health, Los Angeles, CA 90095, USA
[2]Department of Human Genetics, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA 90095, USA
[3]Population Sciences in the Pacific Program (Cancer Epidemiology), University of Hawaii Cancer Center, University of Hawaii at Manoa, Honolulu, HI 96813, USA
[4]Department of Biostatistics, Fielding School of Public Health, University of California Los Angeles, Los Angeles, CA 90095, USA
[5]Department of Neurology, UCLA School of Medicine, Los Angeles, CA 90095, USA

Correspondence to: Beate Ritz; **email:** britz@ucla.edu

ABSTRACT

DNA methylation (DNAm) age estimators are widely used to study aging-related conditions. It is not yet known whether DNAm age is associated with the accumulation of stochastic epigenetic mutations (SEMs), which reflect dysfunctions of the epigenetic maintenance system. Here, we defined epigenetic mutation load (EML) as the total number of SEMs per individual. We assessed associations between EML and DNAm age acceleration estimators using biweight midcorrelations in four population-based studies (total n = 6,388). EML was not only positively associated with chronological age (meta r = 0.171), but also with four measures of epigenetic age acceleration: the Horvath pan tissue clock, intrinsic epigenetic age acceleration, the Hannum clock, and the GrimAge clock (meta-analysis correlation ranging from r = 0.109 to 0.179). We further conducted pathway enrichment analyses for each participant's SEMs. The enrichment result demonstrated the stochasticity of epigenetic mutations, meanwhile implicated several pathways: signaling, neurogenesis, neurotransmitter, glucocorticoid, and circadian rhythm pathways may contribute to faster DNAm age acceleration. Finally, investigating genomic-region specific EML, we found that EMLs located within regions of transcriptional repression (TSS1500, TSS200, and 1stExon) were associated with faster age acceleration. Overall, our findings suggest a role for the accumulation of epigenetic mutations in the aging process.

INTRODUCTION

Epigenetic changes are an important hallmark of aging [1–3]. DNA methylation analysis provided promising molecular biomarkers of aging [4], with several epigenetic aging clocks having been introduced and used by aging researchers in recent years [5–12]. Age-adjusted epigenetic age estimates (referred to as

epigenetic age acceleration) have been linked to a large number of age-related conditions [6, 7, 13–20].

Here we set out to investigate whether DNAm clocks possibly capture any dysfunction of the epigenetic maintenance system (EMS) of a cell [5, 13, 21]. Age is known to greatly increase the variability of DNA methylation levels and the epigenetic profiles of

monozygotic twins diverge considerably with age [22, 23]. Gentilini et al [24] proposed that stochastic epigenetic mutations (SEMs) increase exponentially with chronological age. The association of SEMs and aging was for the first time longitudinally assessed in the Swedish twin cohort [25] which confirmed that epigenetic mutations accumulate with age in an individual. In addition, SEMs have recently been associated with hepatocellular carcinoma staging [26], exposure to endocrine-disrupting compounds [27], socioeconomic position, and lifestyle factors [28]. Despite extensive research in this field, to our knowledge, most previous studies focused on chronological age rather than epigenetic age and epigenetic age acceleration. One study found SEM counts to be positively associated with epigenetic age acceleration based on both the Horvath and Hannum clocks [27]. Another recent study focused on Hannum, GrimAge, and intrinsic epigenetic age estimators within the Generation Scotland and the Lothian Birth Cohort, and reported positive associations between SEM counts and all three epigenetic age measurements [29]. To address the complexity of the aging process and the biological mechanisms underlying different epigenetic clocks, it may be useful to systematically study multiple clocks at the same time. In addition, biologic pathway enrichment analysis may help us gain an understanding of the pathophysiology of accelerated aging.

We pooled four population-based studies (total n = 6,388) to systematically investigate whether SEM counts are associated with epigenetic age acceleration. We included four DNAm aging clocks that represent different manifestations of the epigenetic aging processes, including: the pan-tissue chronological age estimator by Horvath (2013, Horvath clock) [5]; an intrinsic epigenetic age measure derived from the Horvath clock by additionally regressing out cell compositions (intrinsic clock) [30]; the leukocyte-based chronological age estimator by Hannum et al. (2013, Hannum clock) [11]; and the epigenetic mortality risk predictor developed recently by Lu et al. (2019, GrimAge clock) [7]. Age-adjusted versions of these biomarkers are generally being referred to as measures of epigenetic age acceleration and denoted as AgeAccelHorvath, intrinsic epigenetic age acceleration (IEAA), AgeAccelHannum, and AgeAccelGrim, respectively. We also coined the new term "epigenetic mutation load (EML)" as representing the total number of SEMs observed for each individual. In this article, we will 1) relate EML to different epigenetic age acceleration measures; 2) functionally annotate mutated CpG sites; 3) conduct biological pathway enrichment analysis; 4) relate DNA region-specific EMLs to epigenetic measures of age acceleration; and 5) compare SEMs with the Shannon entropy measure as the latter

can be interpreted as alternative measure for the decline of epigenetic maintenance.

RESULTS

Study population demographics

Our study includes 6,388 individuals from 4 studies: the Framingham Heart Study (FHS) Offspring Cohort, the Women's Health Initiative (WHI), the Jackson Heart Study (JHS), and the Parkinson's Environment and Genes (wave 1) known as the PEG1 study.

The main characteristics of the study populations are shown in Table 1. Briefly, FHS provided data for 2,326 individuals, with nearly half of them male (n = 1077; 46%) and all are white. Of the 2,091 female participants from the WHI, 989 (47%) are non-Hispanic white, 431 (21%) Hispanic, and 671 (32%) African American. JHS investigated 1,734 African American individuals with a majority of female participants (n = 1086; 63%). The 237 PEG1 control study participants were mostly non-Hispanic white (n = 207; 87%), and half were male (n = 126; 53%). The age ranges varied with the JHS having the largest range (22-93; mean = 56.2), and WHI the smallest (50-80; mean = 65.4). Mean ages of all populations ranged between 56.2 and 67.4. Additional details on cohorts and participant characteristics can be found in the Methods.

Epigenetic mutation load is the number of SEMs

All DNA methylation data was extracted from blood samples with the Illumina Infinium platform (450K array for PEG1, FHS, and WHI studies; EPIC array for WHI). Following a published and validated approach [24, 26, 31], a SEM is observed for a given person at a specific CpG site if an individual's methylation level is more than three times the interquartile range (IQR) lower than the 25th percentile $(Q1 - 3 \times IQR)$, or more than three times the IQR higher than the 75th percentile $(Q3 + 3 \times IQR)$. The 25th and 75th percentile, and correspondingly the IQR, for each CpG locus was estimated across all samples. Furthermore, we defined the epigenetic mutation load (EML) of each study participant according to the total number of SEMs.

EML was highly variable across people (Supplementary Table 1), with a mean value ranging from 1647 to 3401 depending on the total number of CpGs measured on different arrays (FHS: 2433; WHI: 1647; JHS: 3401; PEG1: 2137). Since EMLs were not normally distributed, natural log-transformed EML values were used in all analyses.

EML was not associated with microarray slides (ANOVA $p = 0.135$) or position on the array (ANOVA

Table 1. Distribution of demographics and DNAm aging clocks.

	FHS (n = 2326)	WHI (n= 2091)	JHS (n= 1734)	PEG 1 (n = 237)
Age				
Min	40	50	22	35
Max	92	80	93	92
Mean (SD)	66.36 (8.94)	65.34 (7.10)	56.21 (12.30)	67.42 (12.82)
Sex				
Male (%)	1,077 (46)	0 (0)	648 (37)	126 (53)
Female (%)	1,249 (54)	2,091 (100)	1,086 (63)	111 (47)
Race/Ethnicity				
White (%)	2,326 (100)	989(47)	0 (0)	207 (87)
Hispanic (%)	0 (0)	431 (21)	0 (0)	19 (8)
African American (%)	0 (0)	671 (32)	1734 (100)	0 (0)
Native American (%)	0 (0)	0 (0)	0 (0)	11 (5)
AgeAccelHorvath				
Min	-16.03	-22.56	-16.57	-13.44
Median	-0.38	-0.07	-0.07	-0.13
Max	41.62	29.35	22.81	22.98
Mean (SD)	-0.08 (4.81)	0.10 (5.18)	0.04 (4.45)	0.00 (5.31)
IEAA				
Min	-21.83	-21.46	-15.67	-12.17
Median	-0.17	-0.05	0.07	-0.13
Max	26.93	24.89	22.40	20.28
Mean (SD)	-0.03 (4.59)	0.02 (4.88)	0.05 (4.34)	0.00 (4.92)
AgeAccelHannum				
Min	-19.25	-19.50	-11.59	-12.92
Median	-0.18	0.02	-0.15	-0.27
Max	27.97	18.19	19.35	12.53
Mean (SD)	-0.02 (4.83)	0.02 (4.80)	0.03 (3.49)	0.00 (4.42)
AgeAccelGrim				
Min	-10.92	-10.03	-13.66	-8.74
Median	-0.76	-0.47	-0.81	-0.64
Max	22.51	16.35	24.94	14.62
Mean (SD)	0.02 (4.86)	0.01 (3.80)	0.01 (4.81)	0.00 (4.50)

p = 0.458). Also, EML was not correlated with the average intensity of bisulfite conversion controls (Pearson r = -0.085, p = 0.194). Thus, we concluded that the EML was independent of batches or other technical aspects.

Correlations among DNAm aging clocks

We calculated all DNAm aging estimators including the Horvath clock, the Hannum clock, the GrimAge clock, the PhenoAge clock, the SkinBlood clock, as well as an epigenetic estimate of telomere length (DNAmTL) using the online DNA Methylation Age Calculator (https://dnamage.genetics.ucla.edu/).

As expected, chronological age was strongly positively correlated with all epigenetic age estimators (Pearson r ranging from 0.79 to 0.93, Supplementary Figure 1), and these aging clocks were also strongly correlated with each other (Pearson r ranging from 0.73 to 0.90,

Supplementary Figure 1). Meanwhile, the epigenetic estimate of telomere length, DNAmTL, was negatively correlated with chronological age and the epigenetic age estimates (Pearson r ranging from -0.63 to -0.72, Supplementary Figure 1).

For each clock, we calculated DNA methylation-based age acceleration based on the residuals of the regression of DNA methylation age on each participants' chronological age. Thus, due to this approach, none of the epigenetic measures of age accelerations are correlated with chronological age (Pearson r = 0) as can be seen from Figure 1 and Supplementary Figure 2. AgeAccelHorvath is highly correlated (Pearson r = 0.93) with IEAA because both are based on the Horvath pan tissue clock. AgeAccelHannum was moderately associated with both AgeAccelHorvath and IEAA (Pearson r = 0.48 and 0.4 respectively). AgeAccelGrim showed only weak correlations with the other epigenetic measures of age acceleration which reflects the fact that

GrimAge clock is a mortality risk predictor as opposed to an age estimator. Overall, the moderate pairwise correlations between the DNAm based biomarkers reflect different properties: some are highly confounded by blood cell composition and capture immunosenescence (Hannum, GrimAge, DNAmTL) while others are not (Horvath pan tissue, IEAA) [13, 32].

Association between EML and DNAm aging clocks

We estimated the association between EML, chronological age, cell composition, and DNAm age acceleration using biweight midcorrelation (bicor) for each dataset separately and calculated pooled statistics using Stouffer's method. Bicor is a median-based measurement of correlation that is robust to outliers [33]. We adjusted for potential confounders including age, sex,

race/ethnicity, and cell compositions (naïve CD8 cells, CD8+CD28-CD45RA- T cells, Plasma Blasts, CD4 T cells, and Granulocytes) by regressing out the effects of these factors and retaining the residuals only for analysis. Results for AgeAccelHorvath, IEAA, AgeAccelHannum, and AgeAccelGrim are shown in Table 2 and Figure 2, while other clocks can be found in Supplementary Table 2. These analyses show that EML per study participant was positively correlated with chronological age (meta r = 0.171, meta P-value = 1.64E-42). Furthermore, EML was negatively correlated with CD4+ T cells (meta r = -0.121, meta P-value = 4.24E-22), plasmablasts (meta r = -0.085, meta P-value = 1.14E-11), and granulocytes (meta r = -0.064, meta P-value = 3.70E-07), but positively with exhausted CD8+ (defined as CD8+CD28-CD45RA-) T cells. These results are consistent with known age-related changes in blood cell composition [34, 35].
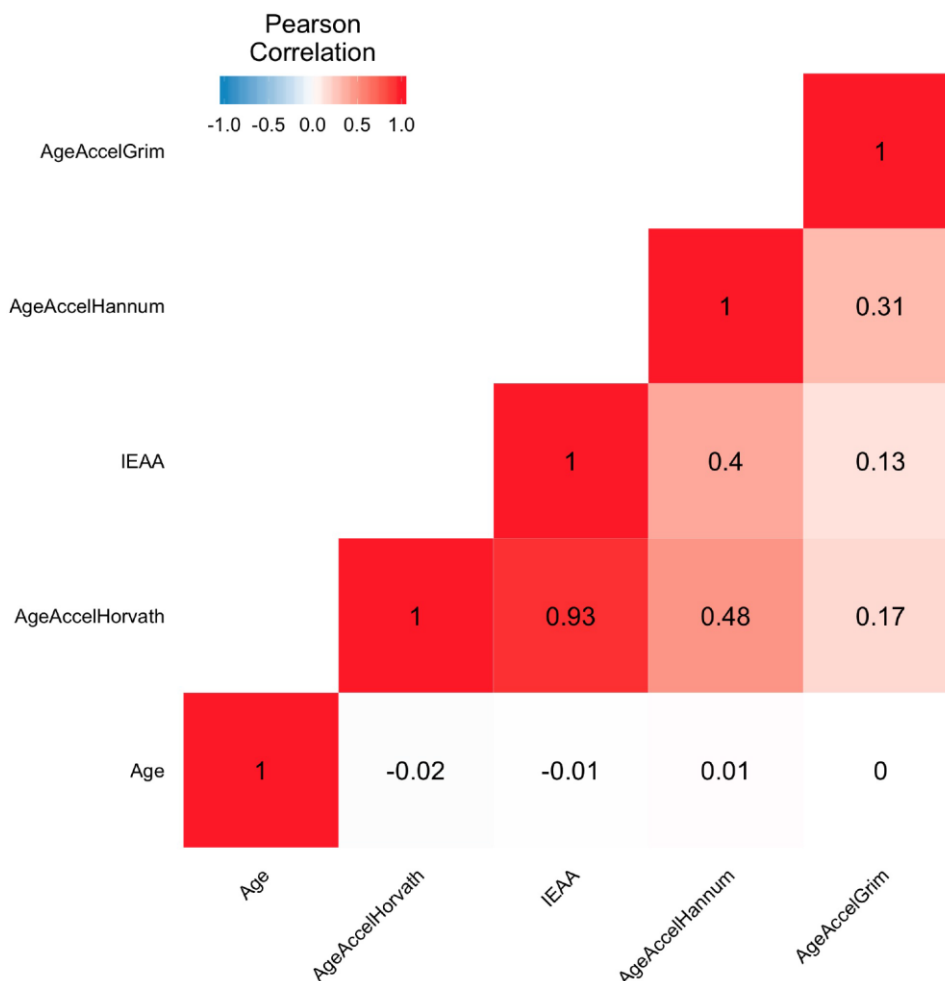
Figure 1. Heatmap of pairwise correlations of chronological age and epigenetic age accelerations. The heat map color-codes the pairwise Pearson correlations of chronological age and epigenetic age accelerations in the Framingham Heart Study (N=2326). Age represents the chronological age. AgeAccelHorvath, IEAA, AgeAccelHannum, and AgeAccelGrim represent measures of epigenetic age acceleration derived from the Horvath pan tissue clock, the intrinsic clock, the Hannum clock, and the GrimAge clock, respectively. The shades of color (blue, white, and red) visualize correlation values from -1 to 1. Each square reports a Pearson correlation coefficient.

Table 2. Biweight midcorrelation analysis of EML.

Outcome = log(EML) [**]	Meta [*]		FHS (n = 2326)		WHI (n= 2091)		JHS (n=1734)		PEG 1 (n = 237)	
	Meta r	Meta P_value	Bicor r	P_value	Bicor r	P_value	Bicor r	P_value	Bicor r	P_value
Age	0.171	1.64E-42	0.244	7.15E-33	0.104	1.73E-06	0.145	1.50E-09	0.176	6.45E-03
DNAm Age Acceleration										
AgeAccelHorvath	0.109	3.25E-18	0.106	3.11E-07	0.140	1.34E-10	0.079	9.68E-04	0.071	2.75E-01
IEAA	0.112	4.04E-19	0.109	1.26E-07	0.144	3.85E-11	0.080	8.50E-04	0.073	2.63E-01
AgeAccelHannum	0.179	2.43E-46	0.225	4.12E-28	0.156	6.55E-13	0.148	6.33E-10	0.095	1.46E-01
AgeAccelGrim	0.162	2.25E-38	0.173	3.74E-17	0.180	9.91E-17	0.111	3.46E-06	0.224	5.23E-04
Cell types										
CD8.naive	-0.021	9.19E-02	-0.072	5.20E-04	0.020	3.66E-01	-0.011	6.58E-01	0.042	5.23E-01
CD8pCD28nCD45RAn	0.077	9.23E-10	0.085	3.90E-05	0.086	8.16E-05	0.052	2.88E-02	0.082	2.07E-01
PlasmaBlast	-0.085	1.14E-11	-0.054	8.94E-03	-0.070	1.39E-03	-0.140	4.89E-09	-0.110	9.23E-02
CD4T	-0.121	4.24E-22	-0.146	1.68E-12	-0.113	2.17E-07	-0.096	6.79E-05	-0.118	6.95E-02
Gran	-0.064	3.70E-07	-0.075	2.72E-04	-0.016	4.74E-01	-0.091	1.60E-04	-0.170	8.67E-03

[*]Meta-analysis using Stouffer's method with weights given by the square root of the number of (non-missing) samples in each data set.

[**]Adjusted for Age, Sex, Race/ethnicity, Cell types.

EML was also positively correlated with AgeAccelHorvath, IEAA, AgeAccelHannum, and AgeAccelGrim, with AgeAccelHannum exhibiting the strongest correlation (meta r = 0.179; meta P-value = 2.43E-46).

We further distinguished between epigenetic age acceleration and deceleration to determine correlations with EML. The correlation between EML and age acceleration was largely the same as what we presented originally. Interestingly, the correlation between EML and age deceleration was much smaller in size and less statistically significant (see Supplementary Table 3).

Sensitivity analyses

We evaluated associations between EML, chronological age, cell compositions, and age accelerations in males and females separately (Supplementary Table 4). For both sexes, EML remained positively correlated with chronological age, exhausted CD8+ T cells, and age acceleration suggesting that EML and age acceleration are independent of sex.

Several sensitivity analyses were conducted to ensure the reliability and reproducibility of the observed associations. To address a possibly non-linear relationship between epigenetic aging and chronological age, we additionally adjusted for a square term in age, (age^2, Supplementary Table 5). Also, to assess the potential for additional confounding, we adjusted for body mass index (Supplementary Table 6). This led to qualitatively similar results.

In order to explore whether the criteria used to define SEM will change results, we conducted another sensitivity analysis using two new SEM measures: 1) loose SEM: defined as a specific CpG site with its methylation level exceeding two times the interquartile range (IQR) of the first quartile ($Q1 - 2 \times IQR$) or the third quartile ($Q3 + 2 \times IQR$) across all subjects; and 2) stringent SEM: defined as a specific CpG site with its methylation level exceeding four times the interquartile range (IQR) of the first quartile ($Q1 - 4 \times IQR$) or the third quartile ($Q3 + 4 \times IQR$) across all subjects. We then calculated the total number of SEMs according to the loose and stringent definition for each person (loose or stringent EML, respectively). The biweight midcorrelations between loose or stringent EMLs and measures of epigenetic age accelerations were very similar to the original results (Supplementary Table 7).

We also explored the effect of different normalization methods for the methylation data (Illumina background correction, functional normalization, Noob, and quantile normalization). We found that the association between EML and age acceleration was not influenced by the normalization method (Supplementary Table 8).

Functional annotations

To test whether individual SEMs were randomly distributed across the genome or were more likely to be found in certain genomic regions or biological pathways, we conducted enrichment analyses to assess whether SEMs were enriched in clock-CpGs (Horvath clock, PhenoAge clock, Hannum clock), genomic regions (transcription start sites (TSS1500, TSS200), untranslated regions (5'UTR, 3'UTR), 1st Exon, and gene body), or regulatory features (i.e., enhancers, DNase hypersensitive sites, open chromatin regions, transcription factor binding site, promoters). For each participant, we first annotated the probes and each mutation based on the location related to genes (i.e.,

TSS1500, TSS200, 5'UTR, 1st EXON, gene body, 3'UTR), or regulatory features using the manifest provided by Illumina. Then we conducted hypergeometric tests for each region and each subject separately using a nominal significance threshold of 0.05. Last, for each region, we summarized the number of individuals for which the test was significant (Supplementary Tables 9–11).

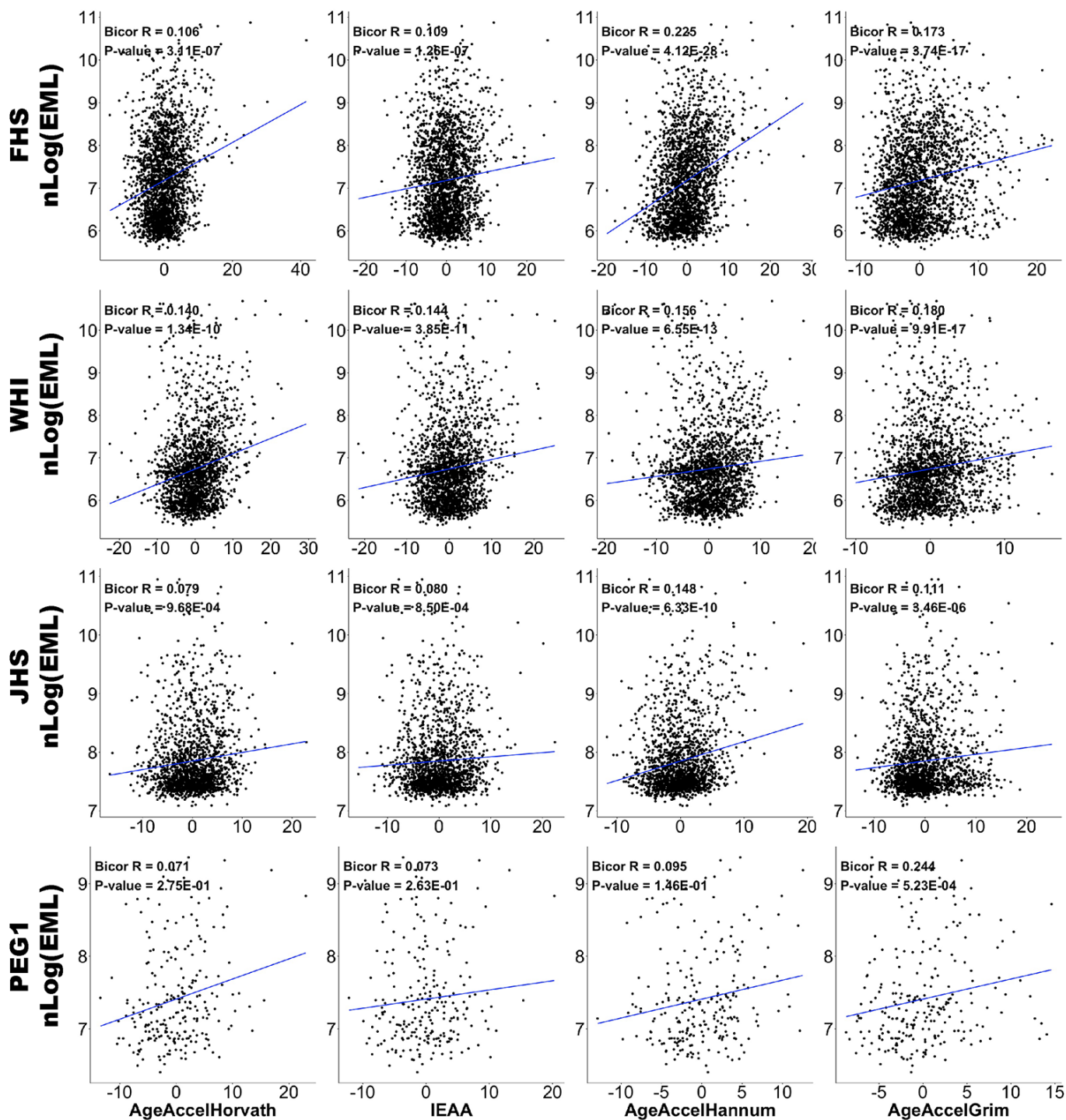**Figure 2. Correlations between EML and epigenetic age accelerations.** Scatter plots of DNAm age acceleration estimators (x-axis; AgeAccelHorvath, IEAA, AgeAccelHannum, and AgeAccelGrim in each column, respectively) versus natural log-transformed EMLs (y-axis). Data from FHS, WHI, JHS, and PEG1 are plotted in four rows respectively. Each panel reports a biweight midcorrelation coefficient and correlation test p-value.

The results show that for each clock-CpG set or region, only a small proportion of participants have their SEMs enriched which illustrating both the stochastic nature and inter-personal variation of SEMs.

Pathway enrichment analysis

Next, we examined whether, among study participants who exhibit faster age accelerations, SEMs are enriched in particular biological pathways. We first conducted KEGG pathway enrichment analyses for each study participant. Then for each KEGG pathway, we calculated the number of people with enriched SEMs for this particular pathway. Finally, we investigated the association between SEMs enrichment in a particular pathway and age accelerations using linear regression. Additional details can be found in the method section.

Supplementary Tables 12–16 showed the top 10 pathways that were commonly enriched in each study, and generally we found that SEMs enriched in these pathways were also statistically significantly associated with faster age accelerations (AgeAccelHorvath, IEAA, AgeAccelHannum, and AgeAccelGrim, respectively).

Briefly, in all four study populations, we identified similar enrichment patterns, and SEMs enriched in signaling pathways, axon guidance, glutamatergic synapse, morphine addiction, glucocorticoid pathway (Cushing syndrome), or circadian rhythm pathways were associated with faster AgeAccelHorvath, AgeAccelHannum, and IEAA. Whereas the associations between pathway enrichment and AgeAccelGrim or AgeAccelPheno were less strong and not necessarily statistically significant. We only observed SEMs enriched in neuroactive ligand-receptor interaction was associated with faster AceAccelGrim and AgeAccelPheno.

Region-specific EML

To address the functionality of SEMs on biological age acceleration, we calculated the number of SEMs co-located with clock CpGs for each study participant (i.e., clock-specific SEMs) and assessed whether there were any clock-specific EMLs corresponding to age acceleration (Supplementary Table 17), but we observed no statistically significant association with the three clocks tested (Horvath clock: 353 CpGs; Hannum clock: 71 CpGs; PhenoAge clocks: 513 CpGs).

Next, we divided CpGs into different genomics region/regulatory feature groups based on the annotations, and then calculated EMLs within each region for each study participant (i.e., genomic region-specific EML; regulatory region-specific EML) (Table 3 and Supplementary Table 18). EMLs in TSS1500, TSS200, and the 1stExon regions were related to faster age accelerations. Also, EML in DNase hypersensitive regions was positively correlated with faster age accelerations. In contrast, 3'UTR specific-EML was associated with younger chronologic age and slower age acceleration.

The direction of SEM

Based on the direction of the mutation, we separated SEMs into hypomethylated SEM (Q1 − 3 × IQR) and hypermethylated SEM (Q3 + 3 × IQR) and calculated hypomethylated and hypermethylated EMLs respectively within FHS. We assessed the correlations between the newly calculated directional EMLs and epigenetic age acceleration. The results remained largely the same (See Supplementary Table 19). Furthermore, consistent with previous studies [29], hypermethylated SEMs were mainly located in CpG islands, while hypomethylated SEMs were enriched in the open seas (see Supplementary Tables 20, 21).

Shannon entropy, EML, and DNAm age acceleration

As an alternate measure for a well-functioning EMS that maintains genomic stability, we calculated the Shannon entropy of the whole methylome based on the 450K or EPIC array. An increase in entropy means that the methylome becomes less predictable across the population of cells, i.e. when the methylation fractions (beta values) tend towards 50%.

We found the Shannon entropy to be positively correlated with chronologic age (meta r = 0.046, meta P-value = 2.16E-04), EML (meta r = 0.234, meta P-value = 7.71E-78), and all four measures of age accelerations (meta P-value: AgeAccelHorvath = 8.79E-09, IEAA = 6.56E-04, AgeAccelHannum = 1.80E-22, AgeAccelGrim = 1.67E-22) (Table 4 and Supplementary Table 22).

DISCUSSION

It has previously been proposed that aging-related decline in epigenetic maintenance increases the occurrence of SEMs in individuals [24, 25]. Our data suggest that the EML per study participant are weakly but statistically significantly associated with several widely used measures of epigenetic age acceleration based on epigenetic clocks.

It has been hypothesized that DNA methylation clocks may capture the imperfection of the EMS resulting in epigenetic instability [5, 13, 21]. Our study provides new evidence for this hypothesis showing that the accumulation of stochastic epigenetic mutations is

Table 3. Meta-analysis *: Biweight midcorrelation analysis of genomic region-specific EML.

Outcome = log(Region-EML) **	TSS1500 (50999 CpGs)		TSS200 (41175 CpGs)		5'UTR (23024 CpGs)		1stExon (7669 CpGs)		Gene Body (135960 CpGs)		3'UTR (14010 CpGs)	
	Meta r	Meta P_value	Meta r	Meta P_value	Meta r	Meta P_value	Meta r	Meta P_value	Meta r	Meta P_value	Meta r	Meta P_value
Age	0.103	1.42E-16	0.074	3.39E-09	-0.046	2.44E-04	0.126	7.86E-24	-0.122	1.81E-22	-0.085	1.38E-11
DNAm Age Acceleration												
AgeAccelHorvath	0.050	7.35E-05	-0.004	7.52E-01	-0.055	9.43E-06	0.074	4.13E-09	-0.060	1.96E-06	-0.064	3.35E-07
IEAA	0.058	3.16E-06	-0.006	6.52E-01	-0.052	2.94E-05	0.067	9.83E-08	-0.063	4.04E-07	-0.064	2.66E-07
AgeAccelHannum	0.098	5.06E-15	0.076	1.62E-09	-0.044	4.93E-04	0.154	1.00E-34	-0.140	4.87E-29	-0.111	6.74E-19
AgeAccelGrim	-0.001	9.40E-01	0.054	1.45E-05	-0.012	3.57E-01	0.082	6.30E-11	-0.053	1.94E-05	-0.073	5.52E-09
Cell types												
CD8.naive	-0.048	1.23E-04	-0.035	5.18E-03	-0.023	6.55E-02	-0.044	4.66E-04	0.053	1.96E-05	0.037	3.10E-03
CD8pCD28nCD45RAn	0.043	5.51E-04	-0.032	1.07E-02	0.015	2.25E-01	-0.020	1.08E-01	0.016	2.06E-01	-0.018	1.41E-01
PlasmaBlast	0.041	1.04E-03	-0.019	1.28E-01	-0.019	1.23E-01	-0.022	7.45E-02	-0.023	6.96E-02	-0.012	3.32E-01
CD4T	0.043	5.19E-04	-0.066	1.34E-07	-0.013	2.91E-01	-0.095	2.97E-14	0.002	8.74E-01	0.024	5.09E-02
Gran	-0.008	5.43E-01	-0.082	5.17E-11	-0.039	1.85E-03	-0.113	1.82E-19	0.067	9.65E-08	0.065	1.72E-07

* Meta-analysis using Stouffer's method with weights given by the square root of the number of (non-missing) samples in each data set.

** Adjusted for Age, Sex, Race/ethnicity, Cell types, Log(total EML).

Table 4. Association between Shannon entropy and age, AgeAccel, EML.

Outcome =Entropy **	Meta *		FHS (n = 2326)		WHI (n= 2091)		JHS (n= 1734)		PEG 1 (n = 237)	
	Meta r	Meta P_value	Bicor r	P_value	Bicor r	P_value	Bicor r	P_value	Bicor r	P_value
Age	0.046	2.16E-04	0.001	9.55E-01	0.068	2.01E-03	0.071	2.92E-03	0.117	7.30E-02
DNAm Age Acceleration										
AgeAccelHorvath	0.072	8.79E-09	0.081	9.11E-05	0.160	1.76E-13	-0.039	1.02E-01	0.006	9.22E-01
IEAA	0.043	6.56E-04	0.035	9.02E-02	0.131	1.70E-09	-0.052	3.16E-02	0.018	7.83E-01
AgeAccelHannum	0.122	1.80E-22	0.155	6.23E-14	0.136	4.60E-10	0.063	9.10E-03	0.096	1.41E-01
AgeAccelGrim	0.122	1.67E-22	0.077	1.89E-04	0.228	3.86E-26	0.043	7.17E-02	0.164	1.14E-02
EML	0.234	7.71E-78	0.089	1.63E-05	0.294	6.87E-43	0.325	7.22E-44	0.281	1.10E-05

* Meta-analysis using Stouffer's method with weights given by the square root of the number of (non-missing) samples in each data set.

** Adjusted for Age, Sex, Race/ethnicity, Cell types.

associated with epigenetic age acceleration according to four clocks: the Horvath, the intrinsic, the Hannum, and the GrimAge clock. The first three clocks have been built to predict chronological age while the GrimAge clock was designed as a mortality risk predictor that explicitly uses chronological age as one of its predictors. We observed statistically significant associations between EML and age acceleration measured by all four clocks, with the AgeAccelHannum and AgeAccelGrim being most strongly associated with EML. One explanation might be that these clocks have different relationships to blood cell composition. While measures of epigenetic age acceleration based on Horvath's pan tissue clock, AgeAccelHorvath and IEAA, are at best weakly related to changes in blood cell composition, AgeAccelHannum and AgeAccelGrim correlate more strongly with blood cell counts and markers of immunosenescence. Therefore, similar to the Hannum and GrimAge clocks, EML also reflects changes in blood cell composition i.e. the immune system. Previously, studies showed that DNAm

biomarkers of aging that capture altered immune cell composition are better predictors of mortality [7, 13]. Thus, not only the intracellular accumulation of epigenetic mutations we investigated here, but also changes in cell composition contribute to EML as part of the biological aging processes that diverge from chronological age. This finding is also consistent with several previous studies [25, 27, 29].

It is worth noting that only the intracellular accumulation of epigenetic mutations suggests that an insufficient EMS may be involved in increasing the EML, thus we ascribe greater weight and importance to the correlations with the Horvath clock and IEAA. The relatively weak correlations between EML and AgeAccelHorvath or IEAA indicate that these DNAm age estimators also capture other hallmarks of the aging process apart from a dysfunctional EMS [3]. The size of the correlations for these age acceleration measures and EML are comparable to those reported for many other known risk factors of aging. For example, in the

WHI cohort, the correlation between BMI and AgeAccelGrim is 0.14; the correlation between exercise and AgeAccelGrim is -0.1; and the correlations for many other risk factors are below +/- 0.3 (See Lu et al) [7]. Correlations between risk factors and IEAA are even smaller (r between 0.08 and -0.06, see Quach et al. Figure 1) [36]. Nevertheless, the fact that associations between EML and AgeAccelHorvath or IEAA are, at best, weak reminds us that other mechanisms and factors apart from EMS also play important roles in the ageing process. Indeed, future in vivo or in vitro studies are needed to better understand the causal relationship between EML and epigenetic aging.

EML exhibited much stronger correlations with age acceleration than deceleration. This result suggests that epigenetic age acceleration and deceleration may have different biological mechanisms, and that the maintenance of epigenetic stability plays more of a role in the acceleration of epigenetic age than the deceleration.

Our finding that EML is statistically significantly albeit weakly correlated with various measures of epigenetic age acceleration is consistent with several previous studies [27, 29]. In order to understand the biological foundation for EML contributions to the epigenetic aging process, we conducted several functional and pathway enrichment analyses. Functional annotation and pathway enrichment analysis showed no predominant regions or biological pathways as being enriched with SEMs. This is in line with previous observations that a majority of SEMs are randomly distributed across the genome and that the locations necessarily differ between individuals as the name suggests [27]. Despite this inherent inter-individual variation, we found that individuals with SEMs enriched in signaling pathways, neurogenesis, neurotransmission, glucocorticoid, or circadian rhythm pathways were more likely to show age acceleration as measured by AgeAccelHorvath, IEAA, and AgeAccelHannum. These non-random patterns – if confirmed – may very well reflect the accumulation of SEMs in pathways related to biological mechanisms that are involved in aging. For example, some signaling pathways such as oxytocin signaling and MAPK/ERK signaling pathway have been associated with age-related muscle maintenance and regeneration [37], while excess glucocorticoid levels may reflect a lifelong accumulation of stressors and this pathway plays a key role in frailty [38] and the aging process [21]. Furthermore, some clock-CpGs are located in glucocorticoid response elements [39]. We also found SEMs enriched in several neurogenesis or neurotransmission-related pathways that may be

contributing to the ticking rate of clocks. This is consistent with the previous finding that DNAm age acceleration is linked to neuropathology [18, 40], especially Parkinson's disease [16] and Down syndrome [15]. Moreover, our finding of SEM enrichment in the circadian entrainment pathway supports the hypothesis that the DNAm age estimators are related to the oscillation of the circadian rhythm [13]. Interestingly, although patterns were similar, we found less evidence for pathway enrichment with SEMs and age acceleration based on the GrimAge and PhenoAge clocks. This may again underscore that different clocks indeed capture different aspects of the aging process.

EMLs within TSS1500, TSS200, and especially the 1stExon regions were found to be associated with faster age accelerations, and for these regions methylation levels have been shown to be related to gene expression [41, 42]. Therefore, our result may suggest that the accumulation of random epigenetic mutations in these regions may influence biological aging processes through gene expression regulation. Interestingly, even though we would also have expected this, we did not observe such associations for promoter regions. Further studies are needed to investigate the biologic consequences of region-specific effects of epigenetic mutations on aging.

The Shannon Entropy measure reflects higher levels of entropy such that the methylome becomes less predictable across the population of cells due to the failure of DNAm maintenance [11]. Epigenetic Shannon entropy as well as this measure's variability increase with age [10, 11, 43, 44]. In our study, EML and Shannon entropy was strongly correlated, confirming that both measure aspects of the EMS, even though EML and entropy capture different aspects of epigenetic stability. SEM represents rare methylation value extremes at a site due to the accumulation of maintenance failures whereas entropy reflects an ongoing 'smoothing' of the epigenetic landscape such that beta values tend towards 50% [45].

There are limitations of our study. First, it is possible that some unmeasured confounders biased our results. Sensitivity analyses, however, showed that the SEM measure was not affected by potential technical artifacts or poor sample quality, and the association between EML and age acceleration was independent of potential confounders including chronological age, sex, race/ethnicity, and BMI. Hence, although technical effects and confounding are hard to avoid, the observed associations between EML and age accelerations were robust to adjustments for a number of covariates. Second, from all four studies, we only had cross-sectional data available. Therefore, we were unable to

investigate the accumulation of epigenetic mutations over time within individuals. Finally, DNA methylation was measured in blood samples only. Therefore, pathway results need to be interpreted with caution as many of the identified pathways listed above have no direct relevance to the function of the blood tissue. While this seems to support the inherent randomness of SEMs, stochastic epigenetic mutations may still accumulate in a non-random pattern within certain biological pathways if repair mechanisms fail systematically due to properties related to these pathways across different tissues. Also, it has been shown that epigenetic changes in blood may indeed reflect epigenetic fingerprints of other target tissues [46, 47]. Nevertheless, tissue- and cell-specific analyses are needed to better understand the relationship between stochastic epigenetic mutations and aging processes in different tissues.

In summary, using large datasets from multiple population-based studies, we were able to show that EML per study participant is associated with different epigenetic aging markers (aging clocks) and importantly with epigenetic age acceleration. Moreover, epigenetic mutations enriched in particular biological pathways or genomic regions related to gene expression were associated with accelerated aging and these may contribute to the ticking of the epigenetic clock. Our findings from pathway enrichment analyses also suggest some interesting biological mechanisms that may influence the ticking of the epigenetic aging clocks and drive the acceleration of the biological aging process.

MATERIALS AND METHODS

Study population

Our study is based on data from four studies: the Framingham Heart Study (FHS) Offspring Cohort, the Women's Health Initiative (WHI), the Jackson Heart Study (JHS), and the Parkinson's Environment and Genes (wave 1) known as the PEG1 study.

We used 2,326 individuals from the FHS Offspring cohort [48]. The FHS cohort is a large-scale longitudinal study started in 1948, initially investigating the common factors of characteristics that contribute to cardiovascular disease (CVD) (https://www.framinghamheartstudy.org/index.php). The study at first enrolled participants living in the town of Framingham, Massachusetts, who were free of overt symptoms of CVD, heart attack, or stroke at enrollment. In 1971, the study started FHS Offspring Cohort to enroll a second generation of the original participants' adult children and their spouses (n= 5124) for conducting similar examinations. The FHS Offspring Cohort collected medical history and measurement data,

immunoassays at exam 7, and blood DNA methylation profiling at exam 8. Participants from the FHS Offspring Cohort were eligible for our study if they attended both the seventh and eighth examination cycles and consented to have their molecular data used for the study. We used the 2,326 participants from the group of Health/ Medical/Biomedical (IRB, MDS) consent and available for both Immunoassay array DNA methylation array data. The FHS data are available in dbGaP (accession number: phs000363.v16.p10 and phs000724.v2.p9).

The WHI is a national study that enrolled postmenopausal women aged 50-79 years into the clinical trials (CT) or observational study (OS) cohorts between 1993 and 1998 [49, 50]. We included 2,091 WHI participants with available phenotype and DNA methylation array data from "*Broad Agency Award 23*" (WHI BA23). WHI BA23 focuses on identifying miRNA and genomic biomarkers of coronary heart disease (CHD), integrating the biomarkers into diagnostic and prognostic predictors of CHD and other related phenotypes, and other objectives can be found in https://www.whi.org/researchers/data/WHIStudies/StudySites/BA23/Pages/home.aspx.

The JHS is a large, population-based observational study evaluating the etiology of cardiovascular, renal, and respiratory diseases among African Americans residing in the three counties (Hinds, Madison, and Rankin) that make up the Jackson, Mississippi metropolitan area [51]. The age at enrollment for the unrelated cohort was 35-84 years; the family cohort included related individuals >21 years old. Participants provided extensive medical and social history, had an array of physical and biochemical measurements and diagnostic procedures, and provided genomic DNA during a baseline examination (2000-2004) and two follow-up examinations (2005-2008 and 2009-2012). Annual follow-up interviews and cohort surveillance are ongoing. In our analysis, we used the visits at baseline from 1,734 individuals as part of project JHS ancillary study ASN0104, available with both phenotype and DNA methylation array data.

The PEG1 study was conducted during 2000-2007 to investigate the causes of Parkinson's disease (PD) in agricultural regions of the California central valley. We analyzed blood samples from 238 healthy controls enrolled from Kern, Tulare, or Fresno counties. Controls were required to be over the age of 35, having lived within one of the counties for at least 5 years prior to enrollment, and do not have a diagnosis of Parkinsonism. Demographic information, lifestyle factors, and medication use were collected in standardized interviews, including lifetime information of cigarette smoking and coffee/ tea consumption.

For PEG1, FHS, and WHI, peripheral blood samples were collected, and bisulfite conversion using the Zymo EZ DNA Methylation Kit (Zymo Research, Orange, CA, USA) as well as subsequent hybridization of the HumanMethylation450k Bead Chip (Illumina, San Diego, CA), and scanning (iScan, Illumina) were performed according to the manufacturer's protocols by applying standard settings. For JHS, DNA methylation quantification was conducted using HumanMethylation EPIC Bead Chip (Illumina, San Diego, CA).

Preprocess

For PEG1 samples, raw signal intensities were retrieved using the function *read.metharray.exp* of the R package *minfi* from the Bioconductor open-source software (http://www.bioconductor.org/), followed by linear dye bias correction, noob background correction, and functional normalization using the same R package [52–55]; β-value was used for all the analyses. One sample was identified as low quality due to low median methylated and unmethylated signal intensities across the entire array and thus removed from the study population. Detection p-values were derived using the function *detectionP* as the probability of the total signal (methylation + unmethylated) being detected above the background signal level, as estimated from negative-control probes. All in all, 845 probes with a detection p-value above 0.05 in at least 5% of samples were removed. Also, 645 probes with a bead count <3 in at least 5% of samples; 11,334 probes on the X or Y chromosome; 7,306 probes containing a SNP at the CpG interrogation site and/or at the single nucleotide extension for 5% maf; and 27,332 cross-reactive probes were also removed. In total, 438,050 probes were included for downstream analyses.

For FHS and WHI samples, 11,334 probes on the X or Y chromosome; 7,306 probes containing a SNP at the CpG interrogation; and 27,332 cross-reactive probes were also removed. In total, 439,540 probes were included for downstream analyses.

For JHS samples, 19,532 probes on the X or Y chromosome; 53,435 probes containing a SNP at the CpG interrogation and cross-reactive were also removed. In total, 793,869 probes were included for downstream analyses.

SEM calculation

The calculation of SEM was consistent with a previously published and validated approach [24, 26, 31]. CpG with methylation levels three times the

interquartile range above the third quartile or below the first quartile was identified as a SEM. Toward this end, we calculated the IQR for each of the 438,050 probes in each dataset (for PEG1, FHS, and WHI) or the 793,969 probes (for JHS). Then, SEMs were identified based on extreme methylation levels. Finally, we summed across the count of all SEMs per sample and defined the total number of SEMs of each study participant as epigenetic mutation load (EML). EML was not normally distributed; therefore, we used the natural log of the number of SEMs for all regression analyses.

In FHS, we separated SEMs into hypermethylated and hypomethylated SEMs based on the direction of the mutation. We also defined consistently hypermethylated or hypomethylated SEMs as a CpG mutated in the same direction in more than 10 participants.

In order to assess whether the criteria used for SEMs will change the results, we defined loose SEM and stringent SEM as described above. We then calculated the total numbers of loose and stringent SEMs for each person.

DNA methylation age

We included eight different DNAm aging biomarkers in this study. Utilizing our online DNA Methylation Age Calculator (https://dnamage.genetics.ucla.edu/), we calculated DNA methylation-based ages and the age accelerations based on the residuals of the regression of DNA methylation age on each participants' chronological age for each clock.

Four types of DNA methylation-based biomarkers were included in the main analyses. Briefly, Horvath clock was calculated using a linear combination of 353 CpGs that have previously been shown to predict chronological age in multiple tissues [5]; and the intrinsic clock was derived from the Horvath clock by additionally regressing out cell compositions [30]; Hannum clock was calculated using a linear combination of 71 CpGs to predict chronological age in blood [11]; and GrimAge clock) was calculated from a linear combination of 7 DNAm plasma protein surrogates and a DNAm-based estimator of smoking pack-years designed to predict mortality [7].

Other DNAm aging biomarkers were included in the Supplementary analyses, including: the extrinsic clock [30], PhenoAge clock [6], SkinBlood clock [56], DNAm based estimator of telomere length [57], each of the 7 DNAm protein surrogates underlying the definition of the GrimAge clock [7], as well as DNAm based estimate of smoking pack-years.

Cell composition

White blood cell composition was imputed for each study participant using our online published DNA Methylation Age Calculator, https://dnamage.genetics.ucla.edu/. The following imputed blood cell counts were included in downstream analyses: CD4+ T, naïve CD8+ T, exhausted cytotoxic CD8+ T cells (defined as CD8 positive CD28 negative CD45R negative, CD8+CD28-CD45RA-), plasmablasts, and granulocytes. Naïve CD8+ T, exhausted cytotoxic CD8+ T cells, and plasmablasts were calculated based on the Horvath method [58]. The remaining cell types were imputed using the Houseman method [59].

Shannon entropy

The formula of Shannon entropy is:

$$Entropy = \frac{1}{N \times log_2(\frac{1}{2})}$$
$$\times \sum_{i=1}^{N}[\beta_i \times log_2(\beta_i) + (1 - \beta_i)$$
$$\times log_2(1 - \beta_i)]$$

where β_i represents the methylation beta value for the ith probe (CpG site) in the array, N represents the total number of probes included in the formula [11].

Statistical analysis

All analyses were conducted using R v3.6.1. We used Pearson correlations to assess the relations between different DNAm ages and age accelerations. We evaluated potential batch effects by assessing the difference of EMLs in microarray slides or position on the array with the ANOVA test. To eliminate the possibility that SEMs are driven by incomplete bisulfite conversion, Pearson correlations between EMLs and the average intensity of bisulfite conversion controls were also calculated. Average intensity of bisulfite conversion controls was derived using the *ENmix* R package [60].

To assess the association between EML and age/age accelerations/cell compositions, we applied biweight midcorrelation (bicor) implemented in the WGCNA R package. We adjusted for potential confounders including age, sex, race/ethnicity, and cell compositions (naïve CD8 cells, CD8+CD28-CD45RA- T cells, Plasma Blasts, CD4 T cells, and Granulocytes) by regressing out the effects of these factors and retaining the residuals of log(EML)only for analysis.

We also conducted stratified analyses to evaluate the associations between EML, chronological age, epigenetic estimates of cell composition, and epigenetic age acceleration in males and females separately for the PEG1, FHS, and JHS studies. To ensure the reliability and reproducibility of the associations, several sensitivity analyses were conducted. In addition to the potential confounders mentioned above, we adjusted for a quadratic term in age (age^2) to account for non-linear relationships and body mass index.

To evaluate the effect of data preprocessing steps, we repeated the analysis using four different normalization methods (Illumina background correction, noob, functional normalization, and quantile normalization implemented in the *minfi* package) in the PEG dataset.

We analyzed each dataset separately, therefore in order to obtain an overall p-value across all four studies, we conducted meta-analyses using Stouffer's method for meta-analysis estimates of the correlation coefficient (meta r), and the corresponding two-sided p-values (meta p-value).

Functional annotation, region-specific SEMs, and pathway enrichment analysis

For each participant, we annotated the probes and SEMs based on the location in relation to genes (TSS1500, TSS200, 5'UTR, 1st Exon, gene body, and 3'UTR), or regulatory features (enhancer region, DNase hypersensitive region, open chromatin region, transcription factor binding site, and promoter region) using the manifest provided by Illumina. To test for region specific enrichment, we conducted hypergeometric tests for each region and each subject separately. A p-value less than 0.05 was considered statistically significant. For biological pathway enrichment analysis, the *enrichKEGG* function implemented in the *ClusterProfiler* package [61] was used to assess whether study participant's SEMs were enriched in particular KEGG pathways (P-value threshold = 0.05). For each genomic region or KEGG pathway, we then summarized how many study participants had their SEMs enriched. We also investigated the association between SEMs enrichment in each KEGG pathway and age accelerations using linear regression, adjusted for the total number of SEMs per study participant:

$$For\ pathway\ j: AgeAccel_i$$
$$= \beta_0 + \beta_1\ Enrich_{ij}$$
$$+ \beta_2 \log(Total\ EML_i) + \varepsilon_i$$

where *AgeAccel* stands for age accelerations (AgeAccelHorvath, IEAA, AgeAccelHannum, AgeAccelGrim); *Enrich* stands for enrichment for pathway j (significant: 1, non-significant: 0); *Total EML* stands for log transformed total EML for participant i.

AUTHOR CONTRIBUTIONS

QY, SH and BR conceived of the study. QY carried out the statistical analyses. KP, AL, CK, and AB contributed to the data analysis plan, participated in the critical interpretation of the results, and contributed to the writing of the manuscript.

CONFLICTS OF INTEREST

These authors declare no conflicts of interest.

REFERENCES

1. Kennedy BK, Berger SL, Brunet A, Campisi J, Cuervo AM, Epel ES, Franceschi C, Lithgow GJ, Morimoto RI, Pessin JE, Rando TA, Richardson A, Schadt EE, et al. Geroscience: linking aging to chronic disease. Cell. 2014; 159:709–13.
https://doi.org/10.1016/j.cell.2014.10.039
PMID:25417146

2. Oberdoerffer P, Sinclair DA. The role of nuclear architecture in genomic instability and ageing. Nat Rev Mol Cell Biol. 2007; 8:692–702.
https://doi.org/10.1038/nrm2238 PMID:17700626

3. López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. The hallmarks of aging. Cell. 2013; 153:1194–217.
https://doi.org/10.1016/j.cell.2013.05.039
PMID:23746838

4. Jylhävä J, Pedersen NL, Hägg S. Biological age predictors. EBioMedicine. 2017; 21:29–36.
https://doi.org/10.1016/j.ebiom.2017.03.046
PMID:28396265

5. Horvath S. DNA methylation age of human tissues and cell types. Genome Biol. 2013; 14:R115.
https://doi.org/10.1186/gb-2013-14-10-r115
PMID:24138928

6. Levine ME, Lu AT, Quach A, Chen BH, Assimes TL, Bandinelli S, Hou L, Baccarelli AA, Stewart JD, Li Y, Whitsel EA, Wilson JG, Reiner AP, et al. An epigenetic biomarker of aging for lifespan and healthspan. Aging (Albany NY). 2018; 10:573–91.
https://doi.org/10.18632/aging.101414
PMID:29676998

7. Lu AT, Quach A, Wilson JG, Reiner AP, Aviv A, Raj K, Hou L, Baccarelli AA, Li Y, Stewart JD, Whitsel EA, Assimes TL, Ferrucci L, Horvath S. DNA methylation GrimAge strongly predicts lifespan and healthspan. Aging (Albany NY). 2019; 11:303–27.
https://doi.org/10.18632/aging.101684
PMID:30669119

8. Vidal-Bralo L, Lopez-Golan Y, Gonzalez A. Simplified assay for epigenetic age estimation in whole blood of

adults. Front Genet. 2016; 7:126.
https://doi.org/10.3389/fgene.2016.00126
PMID:27471517

9. Zhang Y, Wilson R, Heiss J, Breitling LP, Saum KU, Schöttker B, Holleczek B, Waldenberger M, Peters A, Brenner H. DNA methylation signatures in peripheral blood strongly predict all-cause mortality. Nat Commun. 2017; 8:14617.
https://doi.org/10.1038/ncomms14617
PMID:28303888

10. Weidner CI, Lin Q, Koch CM, Eisele L, Beier F, Ziegler P, Bauerschlag DO, Jöckel KH, Erbel R, Mühleisen TW, Zenke M, Brümmendorf TH, Wagner W. Aging of blood can be tracked by DNA methylation changes at just three CpG sites. Genome Biol. 2014; 15:R24.
https://doi.org/10.1186/gb-2014-15-2-r24
PMID:24490752

11. Hannum G, Guinney J, Zhao L, Zhang L, Hughes G, Sadda S, Klotzle B, Bibikova M, Fan JB, Gao Y, Deconde R, Chen M, Rajapakse I, et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. Mol Cell. 2013; 49:359–67.
https://doi.org/10.1016/j.molcel.2012.10.016
PMID:23177740

12. Garagnani P, Bacalini MG, Pirazzini C, Gori D, Giuliani C, Mari D, Di Blasio AM, Gentilini D, Vitale G, Collino S, Rezzi S, Castellani G, Capri M, et al. Methylation of ELOVL2 gene as a new epigenetic marker of age. Aging Cell. 2012; 11:1132–34.
https://doi.org/10.1111/acel.12005
PMID:23061750

13. Horvath S, Raj K. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. Nat Rev Genet. 2018; 19:371–84.
https://doi.org/10.1038/s41576-018-0004-3
PMID:29643443

14. Zheng SC, Widschwendter M, Teschendorff AE. Epigenetic drift, epigenetic clocks and cancer risk. Epigenomics. 2016; 8:705–19.
https://doi.org/10.2217/epi-2015-0017
PMID:27104983

15. Horvath S, Garagnani P, Bacalini MG, Pirazzini C, Salvioli S, Gentilini D, Di Blasio AM, Giuliani C, Tung S, Vinters HV, Franceschi C. Accelerated epigenetic aging in down syndrome. Aging Cell. 2015; 14:491–95.
https://doi.org/10.1111/acel.12325
PMID:25678027

16. Horvath S, Ritz BR. Increased epigenetic age and granulocyte counts in the blood of Parkinson's disease patients. Aging (Albany NY). 2015; 7:1130–42.
https://doi.org/10.18632/aging.100859
PMID:26655927

17. Maierhofer A, Flunkert J, Oshima J, Martin GM, Haaf T, Horvath S. Accelerated epigenetic aging in werner syndrome. Aging (Albany NY). 2017; 9:1143–52.
https://doi.org/10.18632/aging.101217
PMID:28377537

18. Levine ME, Lu AT, Bennett DA, Horvath S. Epigenetic age of the pre-frontal cortex is associated with neuritic plaques, amyloid load, and Alzheimer's disease related cognitive functioning. Aging (Albany NY). 2015; 7:1198–211.
https://doi.org/10.18632/aging.100864
PMID:26684672

19. Levine ME, Hosgood HD, Chen B, Absher D, Assimes T, Horvath S. DNA methylation age of blood predicts future onset of lung cancer in the women's health initiative. Aging (Albany NY). 2015; 7:690–700.
https://doi.org/10.18632/aging.100809
PMID:26411804

20. Ghosh J, Mainigi M, Coutifaris C, Sapienza C. Outlier DNA methylation levels as an indicator of environmental exposure and risk of undesirable birth outcome. Hum Mol Genet. 2016; 25:123–29.
https://doi.org/10.1093/hmg/ddv458 PMID:26566672

21. Epigenetics of Aging and Longevity: Translational Epigenetics, volume 4. Moskalev A, Vaiserman AM, (Editors). Academic Press, 2017.

22. Fraga MF, Ballestar E, Paz MF, Ropero S, Setien F, Ballestar ML, Heine-Suñer D, Cigudosa JC, Urioste M, Benitez J, Boix-Chornet M, Sanchez-Aguilera A, Ling C, et al. Epigenetic differences arise during the lifetime of monozygotic twins. Proc Natl Acad Sci USA. 2005; 102:10604–09.
https://doi.org/10.1073/pnas.0500398102
PMID:16009939

23. Talens RP, Christensen K, Putter H, Willemsen G, Christiansen L, Kremer D, Suchiman HE, Slagboom PE, Boomsma DI, Heijmans BT. Epigenetic variation during the adult lifespan: cross-sectional and longitudinal data on monozygotic twin pairs. Aging Cell. 2012; 11:694–703.
https://doi.org/10.1111/j.1474-9726.2012.00835.x
PMID:22621408

24. Gentilini D, Garagnani P, Pisoni S, Bacalini MG, Calzari L, Mari D, Vitale G, Franceschi C, Di Blasio AM. Stochastic epigenetic mutations (DNA methylation) increase exponentially in human aging and correlate with X chromosome inactivation skewing in females. Aging (Albany NY). 2015; 7:568–78.
https://doi.org/10.18632/aging.100792 PMID:26342808

25. Wang Y, Karlsson R, Jylhävä J, Hedman ÅK, Almqvist C, Karlsson IK, Pedersen NL, Almgren M, Hägg S. Comprehensive longitudinal study of epigenetic

mutations in aging. Clin Epigenetics. 2019; 11:187.
https://doi.org/10.1186/s13148-019-0788-9
PMID:31818313

26. Gentilini D, Scala S, Gaudenzi G, Garagnani P, Capri M, Cescon M, Grazi GL, Bacalini MG, Pisoni S, Dicitore A, Circelli L, Santagata S, Izzo F, et al. Epigenome-wide association study in hepatocellular carcinoma: identification of stochastic epigenetic mutations through an innovative statistical approach. Oncotarget. 2017; 8:41890–902.
https://doi.org/10.18632/oncotarget.17462
PMID:28514750

27. Curtis SW, Cobb DO, Kilaru V, Terrell ML, Marder ME, Barr DB, Marsit CJ, Marcus M, Conneely KN, Smith AK. Exposure to polybrominated biphenyl and stochastic epigenetic mutations: application of a novel epigenetic approach to environmental exposure in the michigan polybrominated biphenyl registry. Epigenetics. 2019; 14:1003–18.
https://doi.org/10.1080/15592294.2019.1629232
PMID:31200609

28. Fiorito G, McCrory C, Robinson O, Carmeli C, Rosales CO, Zhang Y, Colicino E, Dugué PA, Artaud F, McKay GJ, Jeong A, Mishra PP, Nøst TH, et al, and BIOS Consortium, and Lifepath consortium. Socioeconomic position, lifestyle habits and biomarkers of epigenetic aging: a multi-cohort analysis. Aging (Albany NY). 2019; 11:2045–70.
https://doi.org/10.18632/aging.101900
PMID:31009935

29. Seeboth A, McCartney DL, Wang Y, Hillary RF, Stevenson AJ, Walker RM, Campbell A, Evans KL, McIntosh AM, Hägg S, Deary IJ, Marioni RE. DNA methylation outlier burden, health, and ageing in generation Scotland and the lothian birth cohorts of 1921 and 1936. Clin Epigenetics. 2020; 12:49.
https://doi.org/10.1186/s13148-020-00838-0
PMID:32216821

30. Chen BH, Marioni RE, Colicino E, Peters MJ, Ward-Caviness CK, Tsai PC, Roetker NS, Just AC, Demerath EW, Guan W, Bressler J, Fornage M, Studenski S, et al. DNA methylation-based measures of biological age: meta-analysis predicting time to death. Aging (Albany NY). 2016; 8:1844–65.
https://doi.org/10.18632/aging.101020
PMID:27690265

31. Gentilini D, Somigliana E, Pagliardini L, Rabellotti E, Garagnani P, Bernardinelli L, Papaleo E, Candiani M, Di Blasio AM, Viganò P. Multifactorial analysis of the stochastic epigenetic variability in cord blood confirmed an impact of common behavioral and environmental factors but not of in vitro conception.

Clin Epigenetics. 2018; 10:77.
https://doi.org/10.1186/s13148-018-0510-3
PMID:29930742

32. Horvath S, Gurven M, Levine ME, Trumble BC, Kaplan H, Allayee H, Ritz BR, Chen B, Lu AT, Rickabaugh TM, Jamieson BD, Sun D, Li S, et al. An epigenetic clock analysis of race/ethnicity, sex, and coronary heart disease. Genome Biol. 2016; 17:171.
https://doi.org/10.1186/s13059-016-1030-0
PMID:27511193

33. Langfelder P, Horvath S. Fast R functions for robust correlations and hierarchical clustering. J Stat Softw. 2012; 46:i11.
PMID:23050260

34. Wang Y, Karlsson R, Lampa E, Zhang Q, Hedman ÅK, Almgren M, Almqvist C, McRae AF, Marioni RE, Ingelsson E, Visscher PM, Deary IJ, Lind L, et al. Epigenetic influences on aging: a longitudinal genome-wide methylation study in old swedish twins. Epigenetics. 2018; 13:975–87.
https://doi.org/10.1080/15592294.2018.1526028
PMID:30264654

35. Franceschi C, Bonafè M, Valensin S, Olivieri F, De Luca M, Ottaviani E, De Benedictis G. Inflamm-aging. An evolutionary perspective on immunosenescence. Ann N Y Acad Sci. 2000; 908:244–54.
https://doi.org/10.1111/j.1749-6632.2000.tb06651.x
PMID:10911963

36. Quach A, Levine ME, Tanaka T, Lu AT, Chen BH, Ferrucci L, Ritz B, Bandinelli S, Neuhouser ML, Beasley JM, Snetselaar L, Wallace RB, Tsao PS, et al. Epigenetic clock analysis of diet, exercise, education, and lifestyle factors. Aging (Albany NY). 2017; 9:419–46.
https://doi.org/10.18632/aging.101168
PMID:28198702

37. Elabd C, Cousin W, Upadhyayula P, Chen RY, Chooljian MS, Li J, Kung S, Jiang KP, Conboy IM. Oxytocin is an age-specific circulating hormone that is necessary for muscle maintenance and regeneration. Nat Commun. 2014; 5:4082.
https://doi.org/10.1038/ncomms5082 PMID:24915299

38. Clegg A, Hassan-Smith Z. Frailty and the endocrine system. Lancet Diabetes Endocrinol. 2018; 6:743–52.
https://doi.org/10.1016/S2213-8587(18)30110-4
PMID:30017798

39. Zannas AS, Arloth J, Carrillo-Roa T, Iurato S, Röh S, Ressler KJ, Nemeroff CB, Smith AK, Bradley B, Heim C, Menke A, Lange JF, Brückl T, et al. Lifetime stress accelerates epigenetic aging in an urban, African American cohort: relevance of glucocorticoid signaling. Genome Biol. 2015; 16:266.
https://doi.org/10.1186/s13059-015-0828-5

PMID:26673150

40. Marioni RE, Shah S, McRae AF, Ritchie SJ, Muniz-Terrera G, Harris SE, Gibson J, Redmond P, Cox SR, Pattie A, Corley J, Taylor A, Murphy L, et al. The epigenetic clock is correlated with physical and cognitive fitness in the lothian birth cohort 1936. Int J Epidemiol. 2015; 44:1388–96. https://doi.org/10.1093/ije/dyu277 PMID:25617346

41. Jiao Y, Widschwendter M, Teschendorff AE. A systems-level integrative framework for genome-wide DNA methylation and gene expression data identifies differential gene expression modules under epigenetic control. Bioinformatics. 2014; 30:2360–66. https://doi.org/10.1093/bioinformatics/btu316 PMID:24794928

42. Brenet F, Moh M, Funk P, Feierstein E, Viale AJ, Socci ND, Scandura JM. DNA methylation of the first exon is tightly linked to transcriptional silencing. PLoS One. 2011; 6:e14524. https://doi.org/10.1371/journal.pone.0014524 PMID:21267076

43. Bell CG, Lowe R, Adams PD, Baccarelli AA, Beck S, Bell JT, Christensen BC, Gladyshev VN, Heijmans BT, Horvath S, Ideker T, Issa JJ, Kelsey KT, et al. DNA methylation aging clocks: challenges and recommendations. Genome Biol. 2019; 20:249. https://doi.org/10.1186/s13059-019-1824-y PMID:31767039

44. Slieker RC, van Iterson M, Luijk R, Beekman M, Zhernakova DV, Moed MH, Mei H, van Galen M, Deelen P, Bonder MJ, Zhernakova A, Uitterlinden AG, Tigchelaar EF, et al, and BIOS consortium. Age-related accrual of methylomic variability is linked to fundamental ageing mechanisms. Genome Biol. 2016; 17:191. https://doi.org/10.1186/s13059-016-1053-6 PMID:27654999

45. Field AE, Robertson NA, Wang T, Havas A, Ideker T, Adams PD. DNA methylation clocks in aging: categories, causes, and consequences. Mol Cell. 2018; 71:882–95. https://doi.org/10.1016/j.molcel.2018.08.008 PMID:30241605

46. Lowe R, Rakyan VK. Correcting for cell-type composition bias in epigenome-wide association studies. Genome Med. 2014; 6:23. https://doi.org/10.1186/gm540 PMID:25031617

47. Breeze CE, Paul DS, van Dongen J, Butcher LM, Ambrose JC, Barrett JE, Lowe R, Rakyan VK, Iotchkova V, Frontini M, Downes K, Ouwehand WH, Laperle J, et al. eFORGE: a tool for identifying cell type-specific signal in epigenomic data. Cell Rep. 2016; 17:2137–50.

https://doi.org/10.1016/j.celrep.2016.10.059 PMID:27851974

48. Dawber TR, Meadors GF, Moore FE Jr. Epidemiological approaches to heart disease: the framingham study. Am J Public Health Nations Health. 1951; 41:279–81. https://doi.org/10.2105/ajph.41.3.279 PMID:14819398

49. Design of the women's health initiative clinical trial and observational study. The women's health initiative study group. Control Clin Trials. 1998; 19:61–109. https://doi.org/10.1016/s0197-2456(97)00078-0 PMID:9492970

50. Anderson GL, Manson J, Wallace R, Lund B, Hall D, Davis S, Shumaker S, Wang CY, Stein E, Prentice RL. Implementation of the women's health initiative study design. Ann Epidemiol. 2003; 13:S5–17. https://doi.org/10.1016/s1047-2797(03)00043-7 PMID:14575938

51. Taylor HA Jr, Wilson JG, Jones DW, Sarpong DF, Srinivasan A, Garrison RJ, Nelson C, Wyatt SB. Toward resolution of cardiovascular health disparities in African Americans: design and methods of the jackson heart study. Ethn Dis. 2005; 15:S6–4. PMID:16320381

52. Triche TJ Jr, Weisenberger DJ, Van Den Berg D, Laird PW, Siegmund KD. Low-level processing of illumina infinium DNA methylation BeadArrays. Nucleic Acids Res. 2013; 41:e90. https://doi.org/10.1093/nar/gkt090 PMID:23476028

53. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, Irizarry RA. Minfi: a flexible and comprehensive bioconductor package for the analysis of infinium DNA methylation microarrays. Bioinformatics. 2014; 30:1363–69. https://doi.org/10.1093/bioinformatics/btu049 PMID:24478339

54. Fortin JP, Labbe A, Lemire M, Zanke BW, Hudson TJ, Fertig EJ, Greenwood CM, Hansen KD. Functional normalization of 450k methylation array data improves replication in large cancer studies. Genome Biol. 2014; 15:503. https://doi.org/10.1186/s13059-014-0503-2 PMID:25599564

55. Fortin JP, Triche TJ Jr, Hansen KD. Preprocessing, normalization and integration of the illumina HumanMethylationEPIC array with minfi. Bioinformatics. 2017; 33:558–60. https://doi.org/10.1093/bioinformatics/btw691 PMID:28035024

56. Horvath S, Oshima J, Martin GM, Lu AT, Quach A, Cohen H, Felton S, Matsuyama M, Lowe D, Kabacik S, Wilson JG, Reiner AP, Maierhofer A, et al. Epigenetic

clock for skin and blood cells applied to hutchinson gilford progeria syndrome and ex vivo studies. Aging (Albany NY). 2018; 10:1758–75.
https://doi.org/10.18632/aging.101508
PMID:30048243

57. Lu AT, Seeboth A, Tsai PC, Sun D, Quach A, Reiner AP, Kooperberg C, Ferrucci L, Hou L, Baccarelli AA, Li Y, Harris SE, Corley J, et al. DNA methylation-based estimator of telomere length. Aging (Albany NY). 2019; 11:5895–923.
https://doi.org/10.18632/aging.102173
PMID:31422385

58. Horvath S, Levine AJ. HIV-1 infection accelerates age according to the epigenetic clock. J Infect Dis. 2015; 212:1563–73.
https://doi.org/10.1093/infdis/jiv277 PMID:25969563

59. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, Wiencke JK, Kelsey KT. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinformatics. 2012; 13:86.
https://doi.org/10.1186/1471-2105-13-86
PMID:22568884

60. Xu Z, Niu L, Li L, Taylor JA. ENmix: a novel background correction method for illumina HumanMethylation450 BeadChip. Nucleic Acids Res. 2016; 44:e20.
https://doi.org/10.1093/nar/gkv907
PMID:26384415

61. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012; 16:284–87.
https://doi.org/10.1089/omi.2011.0118
PMID:22455463
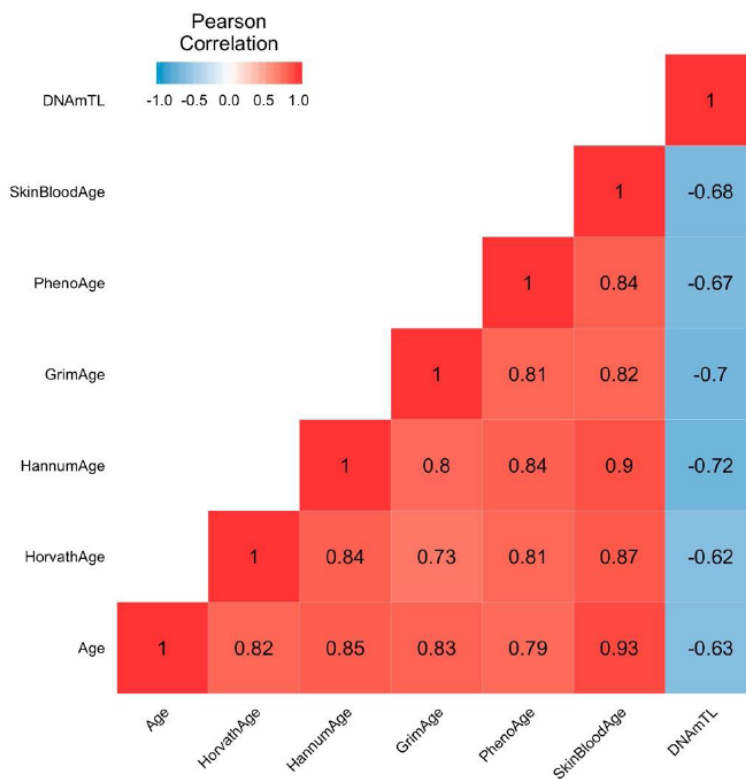
Supplementary Figure 1. Heatmap of pairwise correlations of chronological age and epigenetic ages. The heatmap color-codes the pairwise Pearson correlations of chronological age and epigenetic age in the Framingham Heart Study (N=2326). Age represents the chronological age. HorvathAge, HannumAge, GrimAge, PhenoAge, SkinBloodAge represent measures of epigenetic age derived from the Horvath pan tissue clock, the Hannum clock, the GrimAge clock, the PhenoAge clock, and the SkinBloodAge clock, respectively. DNAmTL represent DNAm-based surrogate markers of telomere length. The shades of color (blue, white, and red) visualize correlation values from -1 to 1. Each square reports a Pearson correlation coefficient.
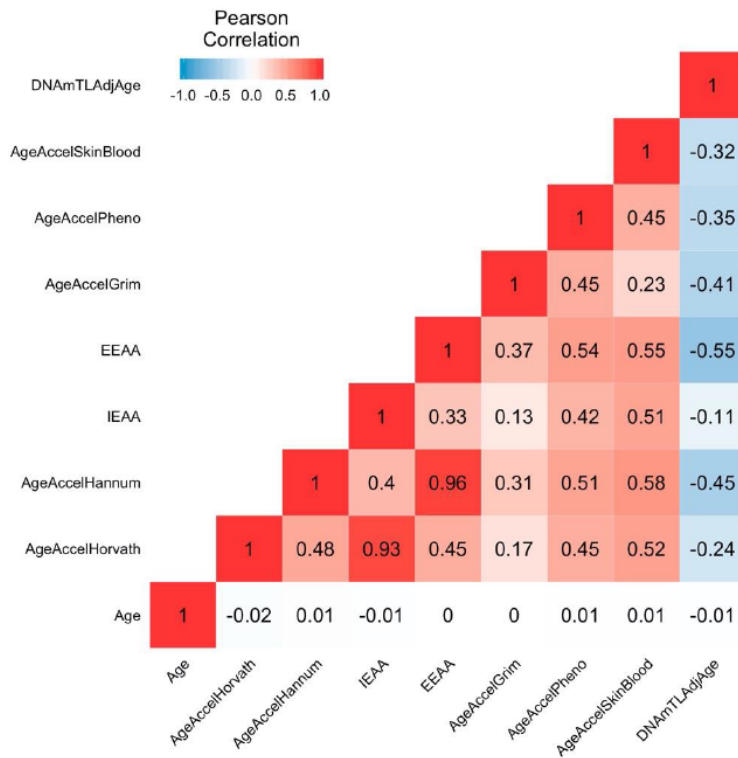
Supplementary Figure 2. Heatmap of pairwise correlations of chronological age and epigenetic age accelerations. The heatmap color-codes the pairwise Pearson correlations of chronological age and epigenetic age accelerations in the Framingham Heart Study (N=2326). Age represents the chronological age. AgeAccelHorvath, AgeAccelHannum, IEAA, EEAA, AgeAccelGrim, AgeAccelPheno, and AgeAccelSkinBlood represent measures of epigenetic age acceleration derived from the Horvath pan tissue clock, the Hannum clock, the intrinsic clock, the extrinsic clock, the GrimAge clock, the PhenoAge clock, the SkinBloodAge clock, respectively. DNAmTLAdjAge represents age adjusted DNAm-based surrogate markers of telomere length. The shades of color (blue, white, and red) visualize correlation values from -1 to 1. Each square reports a Pearson correlation coefficient.

Supplementary Tables

Supplementary Table 1. Distribution of EML for each dataset.

	FHS (450K)	WHI (450K)	JHS (EPIC)	PEG1 (450K)
min(EML)	264	213	1194	603
max(EML)	52783	43768	56691	11645
mean(EML)	2433	1647	3401	2137
sd(EML)	4183	3937	4512	1953
% of probes as SEM	86%	85%	77%	45%

Supplementary Table 2. Biweight midcorrelation analysis of EML.

Outcome = log(EML)**	Meta*		FHS (n = 2326)		WHI (n= 2091)		JHS (n= 1734)		PEG 1 (n = 237)	
	Meta r	Meta P_value	Bicor r	P_value	Bicor r	P_value	Bicor r	P_value	Bicor r	P_value
Age	0.171	1.64E-42	0.244	7.15E-33	0.104	1.73E-06	0.145	1.50E-09	0.176	6.45E-03
BMI	0.005	7.94E-01	0.017	4.23E-01	-0.009	6.91E-01				
DNAm Age Acceleration										
AgeAccelHorvath	0.109	3.25E-18	0.106	3.11E-07	0.140	1.34E-10	0.079	9.68E-04	0.071	2.75E-01
IEAA	0.112	4.04E-19	0.109	1.26E-07	0.144	3.85E-11	0.080	8.50E-04	0.073	2.63E-01
EEAA (Unadjusted for cell types)	0.236	5.82E-79	0.297	1.12E-48	0.212	1.31E-22	0.166	3.09E-12	0.211	1.05E-03
AgeAccelHannum	0.179	2.43E-46	0.225	4.12E-28	0.156	6.55E-13	0.148	6.33E-10	0.095	1.46E-01
AgeAccelGrim	0.162	2.25E-38	0.173	3.74E-17	0.180	9.91E-17	0.111	3.46E-06	0.224	5.23E-04
DNAmADMAdjAge	0.068	4.75E-08	0.121	4.58E-09	0.070	1.41E-03	-0.017	4.71E-01	0.157	1.58E-02
DNAmB2MAdjAge	0.143	3.47E-30	0.189	4.69E-20	0.129	3.44E-09	0.074	2.05E-03	0.287	7.20E-06
DNAmCystatinCAdjAge	0.153	3.02E-34	0.119	8.71E-09	0.241	6.48E-29	0.081	7.30E-04	0.183	4.64E-03
DNAmGDF15AdjAge	0.148	3.64E-32	0.167	6.09E-16	0.127	5.80E-09	0.142	3.01E-09	0.162	1.27E-02
DNAmLeptinAdjAge	-0.025	4.65E-02	-0.021	3.19E-01	-0.035	1.11E-01	-0.031	2.03E-01	0.063	3.34E-01
DNAmPACKYRSAdjAge	0.161	8.96E-38	0.166	7.99E-16	0.172	2.23E-15	0.137	1.10E-08	0.150	2.13E-02
DNAmPAI1AdjAge	0.031	1.47E-02	0.024	2.49E-01	-0.004	8.59E-01	0.061	1.07E-02	0.173	7.56E-03
DNAmTIMP1AdjAge	0.028	2.35E-02	0.070	7.04E-04	0.010	6.41E-01	-0.001	9.72E-01	-0.010	8.80E-01
AgeAccelPheno	0.123	7.28E-23	0.178	4.45E-18	0.106	1.13E-06	0.078	1.20E-03	0.042	5.22E-01
DNAmTLAdjAge	-0.065	1.91E-07	-0.080	1.17E-04	-0.033	1.36E-01	-0.082	6.05E-04	-0.082	2.11E-01
DNAmAgeSkinBloodClockAdjAge	0.047	1.93E-04	0.102	8.32E-07	0.022	3.17E-01	0.019	4.27E-01	-0.080	2.22E-01

Biweight midcorrelation analyses of EML with chronological age, BMI, AgeAccelHorvath, IEAA, extrinsic epigenetic age acceleration (EEAA) derived from the Hannum clock by up-weighting the contribution of age-related blood cell counts, AgeAccelHannum, AgeAccelGrim, age adjusted DNAm-based surrogate markers of adrenomedullin (DNAmADMAdjAge), beta-2 microglobulin (DNAmB2MAdjAge), cystatin C (DNAmCystatinCAdjAge), growth differentiation factor 15 (DNAmGDF15AdjAge), leptin (DNAmLeptinAdjAge), plasminogen activation inhibitor 1 (DNAmPAI1AdjAge), tissue inhibitor metalloproteinase 1 (DNAmTIMP1AdjAge), smoking pack-years (DNAmPACKYRSAdjAge), AgeAccelPheno, age adjusted DNAm-based surrogate markers of telomere length (DNAmTLAdjAge), and age adjusted SkinBlood clock.
*Meta-analysis using Stouffer's method with weights given by the square root of the number of (non-missing) samples in each data set.
**All analyses except EEAA were adjusted for Age, Sex, Race/ethnicity, Cell types; EEAA was not adjusted for cell types.

Supplementary Table 3. Biweight midcorrelation analysis of EML in FHS, stratified by the direction of acceleration.

Outcome = log(EML) [*]	Acceleration		Deceleration	
	Bicor r	P_value	Bicor r	P_value
AgeAccelHorvath	0.178	3.81E-09	-0.039	1.71E-01
IEAA	0.150	4.09E-07	-0.064	2.66E-02
EEAA (Unadjusted for cell types)	0.299	4.74E-25	0.089	2.20E-03
AgeAccelHannum	0.239	3.96E-16	0.085	3.12E-03
AgeAccelGrim	0.136	1.64E-05	0.060	2.77E-02
DNAmADMAdjAge	0.149	5.88E-07	0.029	3.14E-01
DNAmB2MAdjAge	0.130	1.90E-05	0.070	1.34E-02
DNAmCystatinCAdjAge	0.135	8.79E-06	-0.026	3.54E-01
DNAmGDF15AdjAge	0.149	1.61E-06	0.078	4.90E-03
DNAmLeptinAdjAge	0.015	6.04E-01	0.068	2.26E-02
DNAmPACKYRSAdjAge	0.077	2.43E-02	0.049	6.15E-02
DNAmPAI1AdjAge	0.075	1.33E-02	-0.035	2.25E-01
DNAmTIMP1AdjAge	0.092	2.17E-03	-0.058	4.30E-02
AgeAccelPheno	0.206	4.49E-12	0.036	2.08E-01
DNAmTLAdjAge	0.064	2.51E-02	-0.127	2.27E-05
DNAmAgeSkinBloodClockAdjAge	0.168	1.40E-08	-0.023	4.17E-01

* Adjusted for Age, Sex, Cell types.

Supplementary Table 4. Stratification analysis: biweight midcorrelation analysis of EML, stratified by sex.

	PEG 1				FHS				WHI		JHS			
Outcome = log(EML) [*]	Male (n = 126)		Female (n = 111)		Male (n = 1077)		Female (n = 1249)		Femal (n= 2091)		Male (n = 648)		Female (n = 1086)	
	Bicor r	P_value	Bicor r	P_value	Bicor r	P_value	Bicor r	P_value	Bicor r	P_value	Bicor r	P_value	Bicor r	P_value
Age	0.201	2.40E-02	0.200	3.53E-02	0.225	7.53E-14	0.259	1.54E-20	0.104	1.73E-06	0.134	5.98E-04	0.150	7.24E-07
BMI					0.026	3.90E-01	0.025	3.75E-01	-0.009	6.91E-01				
DNAm Age Acceleration														
AgeAccelHorvath	0.105	2.43E-01	0.041	6.71E-01	0.086	4.69E-03	0.126	7.67E-06	0.140	1.34E-10	0.068	8.27E-02	0.082	6.87E-03
IEAA	0.110	2.21E-01	0.033	7.32E-01	0.108	4.02E-04	0.116	4.06E-05	0.144	3.85E-11	0.076	5.19E-02	0.086	4.63E-03
EEAA (Unadjusted for cell types)	0.205	2.13E-02	0.245	9.45E-03	0.305	1.27E-24	0.315	3.35E-30	0.212	1.31E-22	0.180	4.20E-06	0.155	2.90E-07
AgeAccelHannum	0.080	3.73E-01	0.106	2.66E-01	0.215	9.77E-13	0.257	3.02E-20	0.156	6.55E-13	0.158	5.16E-05	0.144	2.05E-06
AgeAccelGrim	0.192	3.13E-02	0.299	1.43E-03	0.218	4.20E-13	0.167	2.71E-09	0.180	9.91E-17	0.082	3.79E-02	0.134	9.43E-06
DNAmADMAdjAge	0.100	2.67E-01	0.223	1.88E-02	0.207	6.85E-12	0.091	1.35E-03	0.070	1.41E-03	-0.024	5.37E-01	0.011	7.06E-01
DNAmB2MAdjAge	0.300	6.40E-04	0.230	1.50E-02	0.238	2.16E-15	0.147	1.94E-07	0.129	3.44E-09	0.063	1.08E-01	0.077	1.13E-02
DNAmCystatinCAdjAge	0.195	2.86E-02	0.187	4.89E-02	0.181	2.42E-09	0.075	7.81E-03	0.241	6.48E-29	0.039	3.25E-01	0.105	5.43E-04
DNAmGDF15AdjAge	0.129	1.49E-01	0.172	7.03E-02	0.211	2.85E-12	0.134	2.06E-06	0.127	5.80E-09	0.097	1.37E-02	0.167	2.91E-08
DNAmLeptinAdjAge	-0.073	4.19E-01	0.138	1.49E-01	0.071	1.94E-02	-0.037	1.89E-01	-0.035	1.11E-01	0.043	2.75E-01	0.046	1.27E-01
DNAmPACKYRSAdjAge	0.184	3.89E-02	0.120	2.10E-01	0.161	1.06E-07	0.179	2.09E-10	0.172	2.23E-15	0.113	4.05E-03	0.147	1.19E-06
DNAmPAI1AdjAge	0.026	7.70E-01	0.362	9.25E-05	0.028	3.66E-01	0.039	1.65E-01	-0.004	8.59E-01	0.053	1.76E-01	0.065	3.21E-02
DNAmTIMP1AdjAge	-0.038	6.72E-01	0.031	7.50E-01	0.100	1.07E-03	0.060	3.41E-02	0.010	6.41E-01	-0.031	4.27E-01	0.016	5.97E-01
AgeAccelPheno	0.123	1.68E-01	-0.080	4.02E-01	0.211	2.66E-12	0.152	6.90E-08	0.106	1.13E-06	0.079	4.47E-02	0.080	8.40E-03
DNAmTLAdjAge	-0.052	5.64E-01	-0.166	8.26E-02	-0.061	4.61E-02	-0.101	3.63E-04	-0.033	1.36E-01	-0.088	2.56E-02	-0.069	2.21E-02
DNAmAgeSkinBloodClockAdjAge	-0.029	7.45E-01	-0.141	1.40E-01	0.105	5.37E-04	0.102	2.90E-04	0.022	3.17E-01	-0.010	7.99E-01	0.032	2.88E-01
Cell types														

	Bicor r	P_value	Bicor r	P_value	Bicor r	P_value	Bicor r	P_value	Bicor r	P_value	Bicor r	P_value	Bicor r	P_value
CD8.naive	0.043	6.32E-01	0.044	6.47E-01	-0.076	1.27E-02	-0.068	1.62E-02	0.020	3.66E-01	-0.008	8.30E-01	-0.010	7.32E-01
CD8pCD28nCD45RAn	0.018	8.44E-01	0.179	6.00E-02	0.076	1.26E-02	0.096	7.00E-04	0.086	8.16E-05	0.068	8.43E-02	0.030	3.28E-01
PlasmaBlast	-0.089	3.19E-01	-0.125	1.92E-01	-0.012	6.90E-01	-0.088	1.81E-03	-0.070	1.39E-03	-0.148	1.60E-04	-0.139	4.57E-06
CD4T	-0.134	1.35E-01	-0.133	1.65E-01	-0.150	7.97E-07	-0.143	4.03E-07	-0.113	2.17E-07	-0.049	2.17E-01	-0.120	7.55E-05
Gran	-0.182	4.18E-02	-0.128	1.80E-01	-0.065	3.37E-02	-0.084	2.81E-03	-0.016	4.74E-01	-0.065	9.80E-02	-0.111	2.57E-04

[*] All analyses except EEAA were adjusted for Age, Sex, Race/ethnicity, Cell types; EEAA was not adjusted for cell types.

Supplementary Table 5. Sensitivity analysis: biweight midcorrelation analysis of EML, additionally adjusted for age[2].

	PEG 1 (n = 237)				FHS (n = 2326)				WHI (n= 2091)				JHS (n= 1734)			
	Model 1[*]		Model 2[**]		Model 1[*]		Model 2[**]		Model 1[*]		Model 2[**]		Model 1[*]		Model 2[**]	
Outcome = log(EML)	Bicor r	P_value	Bicor r	P_value	Bicor r	P_value	Bicor r	P_value	Bicor r	P_value	Bicor r	P_value	Bicor r	P_value	Bicor r	P_value
AgeAccelHorvath	0.071	2.75E-01	0.089	1.72E-01	0.106	3.11E-07	0.106	3.08E-07	0.140	1.34E-10	0.141	9.63E-11	0.079	9.68E-04	0.100	2.84E-05
IEAA	0.073	2.63E-01	0.091	1.61E-01	0.109	1.26E-07	0.109	1.23E-07	0.144	3.85E-11	0.144	3.19E-11	0.080	8.50E-04	0.102	2.18E-05
EEAA (Unadjusted for cell types)	0.211	1.05E-03	0.225	4.80E-04	0.297	1.12E-48	0.298	7.40E-49	0.212	1.31E-22	0.210	2.35E-22	0.166	3.09E-12	0.168	2.13E-12
AgeAccelHannum	0.095	1.46E-01	0.111	8.81E-02	0.225	4.12E-28	0.225	3.79E-28	0.156	6.55E-13	0.156	8.30E-13	0.148	6.33E-10	0.158	3.69E-11
AgeAccelGrim	0.224	5.23E-04	0.219	7.01E-04	0.173	3.74E-17	0.173	3.87E-17	0.180	9.91E-17	0.180	1.09E-16	0.111	3.46E-06	0.118	8.87E-07
DNAmADMAdjAge	0.157	1.58E-02	0.156	1.63E-02	0.121	4.58E-09	0.121	4.65E-09	0.070	1.41E-03	0.069	1.56E-03	-0.017	4.71E-01	-0.019	4.34E-01
DNAmB2MAdjAge	0.287	7.20E-06	0.288	6.61E-06	0.189	4.69E-20	0.188	4.84E-20	0.129	3.44E-09	0.128	4.33E-09	0.074	2.05E-03	0.072	2.75E-03
DNAmCystatinCAdjAge	0.183	4.64E-03	0.182	4.96E-03	0.119	8.71E-09	0.119	8.84E-09	0.241	6.48E-29	0.240	8.64E-29	0.081	7.30E-04	0.075	1.69E-03
DNAmGDF15AdjAge	0.162	1.27E-02	0.162	1.26E-02	0.167	6.09E-16	0.167	6.12E-16	0.127	5.80E-09	0.126	7.40E-09	0.142	3.01E-09	0.143	2.11E-09
DNAmLeptinAdjAge	0.063	3.34E-01	0.061	3.46E-01	-0.021	3.19E-01	-0.021	3.23E-01	-0.035	1.11E-01	-0.034	1.20E-01	-0.031	2.03E-01	-0.031	2.02E-01
DNAmPACKYRSAdjAge	0.150	2.13E-02	0.144	2.64E-02	0.166	7.99E-16	0.166	8.60E-16	0.172	2.23E-15	0.173	1.67E-15	0.137	1.10E-08	0.148	6.41E-10
DNAmPAI1AdjAge	0.173	7.56E-03	0.170	8.84E-03	0.024	2.49E-01	0.024	2.48E-01	-0.004	8.59E-01	-0.001	9.57E-01	0.061	1.07E-02	0.067	5.03E-03
DNAmTIMP1AdjAge	-0.010	8.80E-01	-0.008	9.08E-01	0.070	7.04E-04	0.070	7.11E-04	0.010	6.41E-01	0.008	7.01E-01	-0.001	9.72E-01	-0.002	9.42E-01
AgeAccelPheno	0.042	5.22E-01	0.051	4.36E-01	0.178	4.45E-18	0.178	4.22E-18	0.106	1.13E-06	0.107	9.47E-07	0.078	1.20E-03	0.083	5.33E-04
DNAmTLAdjAge	-0.082	2.11E-01	-0.087	1.84E-01	-0.080	1.17E-04	-0.080	1.16E-04	-0.033	1.36E-01	-0.033	1.30E-01	-0.082	6.05E-04	-0.093	1.08E-04
DNAmAgeSkinBloodClockAdjAge	-0.080	2.22E-01	-0.045	4.91E-01	0.102	8.32E-07	0.102	7.97E-07	0.022	3.17E-01	0.024	2.79E-01	0.019	4.27E-01	0.047	5.25E-02

[*] Adjusted (Age, Sex, Race/ethnicity, Cell types)
[**] Adjusted (Age, Age2, Sex, Race/ethnicity, Cell types)

Supplementary Table 6. Sensitivity analysis: biweight midcorrelation analysis of EML, additionally adjusted for BMI.

	FHS (n = 2326)				WHI (n= 2091)			
	Model 1[*]		Model 2[**]		Model 1[*]		Model 2[**]	
Outcome = log(EML)	Bicor r	P_value	Bicor r	P_value	Bicor r	P_value	Bicor r	P_value
AgeAccelHorvath	0.106	3.11E-07	0.110	1.16E-07	0.140	1.34E-10	0.144	4.13E-11
IEAA	0.109	1.26E-07	0.113	4.37E-08	0.144	3.85E-11	0.148	1.36E-11
EEAA (Unadjusted for cell types)	0.297	1.12E-48	0.298	1.29E-48	0.212	1.31E-22	0.216	2.92E-23
AgeAccelHannum	0.225	4.12E-28	0.226	3.19E-28	0.156	6.55E-13	0.157	5.85E-13
AgeAccelGrim	0.173	3.74E-17	0.172	8.56E-17	0.180	9.91E-17	0.181	1.02E-16
DNAmADMAdjAge	0.121	4.58E-09	0.123	2.93E-09	0.070	1.41E-03	0.074	7.79E-04
DNAmB2MAdjAge	0.189	4.69E-20	0.190	2.84E-20	0.129	3.44E-09	0.130	2.44E-09
DNAmCystatinCAdjAge	0.119	8.71E-09	0.117	1.51E-08	0.241	6.48E-29	0.244	1.31E-29
DNAmGDF15AdjAge	0.167	6.09E-16	0.165	1.37E-15	0.127	5.80E-09	0.127	7.21E-09
DNAmLeptinAdjAge	-0.021	3.19E-01	-0.021	3.21E-01	-0.035	1.11E-01	-0.029	1.93E-01
DNAmPACKYRSAdjAge	0.166	7.99E-16	0.164	1.88E-15	0.172	2.23E-15	0.169	1.09E-14
DNAmPAI1AdjAge	0.024	2.49E-01	0.021	3.16E-01	-0.004	8.59E-01	0.007	7.34E-01
DNAmTIMP1AdjAge	0.070	7.04E-04	0.068	1.03E-03	0.010	6.41E-01	0.016	4.71E-01

AgeAccelPheno	0.178	4.45E-18	0.180	2.10E-18	0.106	1.13E-06	0.111	4.26E-07
DNAmTLAdjAge	-0.080	1.17E-04	-0.079	1.47E-04	-0.033	1.36E-01	-0.033	1.35E-01
DNAmAgeSkinBloodClockAdjAge	0.102	8.32E-07	0.104	5.72E-07	0.022	3.17E-01	0.024	2.71E-01

[*] Adjusted (Age, Sex, Race/ethnicity, Cell types)
[**] Adjusted (Age, Sex, Race/ethnicity, Cell types, BMI)

Supplementary Table 7. Sensitivity analysis: Biweight midcorrelation analysis of EML in FHS using different SEM cutoffs in FHS.

Outcome = log(EML) [*]	EML (3IQR)		EML (2IQR) [**]		EML (4IQR) [***]	
	Bicor r	P_value	Bicor r	P_value	Bicor r	P_value
Age	0.244	7.15E-33	0.223	1.02E-27	0.247	1.13E-33
BMI	0.017	4.23E-01	0.006	7.60E-01	0.019	3.49E-01
DNAm Age Acceleration						
AgeAccelHorvath	0.106	3.11E-07	0.073	4.38E-04	0.112	5.44E-08
IEAA	0.109	1.26E-07	0.077	2.00E-04	0.135	5.28E-11
EEAA (Unadjusted for cell types)	0.297	1.12E-48	0.281	1.89E-43	0.293	2.46E-47
AgeAccelHannum	0.225	4.12E-28	0.224	5.97E-28	0.204	1.98E-23
AgeAccelGrim	0.173	3.74E-17	0.147	1.03E-12	0.180	1.75E-18
DNAmADMAdjAge	0.121	4.58E-09	0.084	4.41E-05	0.133	1.31E-10
DNAmB2MAdjAge	0.189	4.69E-20	0.178	3.79E-18	0.180	2.28E-18
DNAmCystatinCAdjAge	0.119	8.71E-09	0.116	2.03E-08	0.111	6.85E-08
DNAmGDF15AdjAge	0.167	6.09E-16	0.157	2.52E-14	0.161	4.38E-15
DNAmLeptinAdjAge	-0.021	3.19E-01	-0.019	3.62E-01	-0.027	1.97E-01
DNAmPACKYRSAdjAge	0.166	7.99E-16	0.142	5.64E-12	0.176	1.05E-17
DNAmPAI1AdjAge	0.024	2.49E-01	0.005	8.17E-01	0.037	7.75E-02
DNAmTIMP1AdjAge	0.070	7.04E-04	0.068	9.46E-04	0.071	6.02E-04
AgeAccelPheno	0.178	4.45E-18	0.173	3.96E-17	0.164	1.37E-15
DNAmTLAdjAge	-0.080	1.17E-04	-0.068	1.04E-03	-0.073	4.37E-04
DNAmAgeSkinBloodClockAdjAge	0.102	8.32E-07	0.072	5.20E-04	0.116	1.77E-08
Cell types						
CD8.naive	-0.072	5.20E-04	-0.038	6.97E-02	-0.084	4.77E-05
CD8pCD28nCD45RAn	0.085	3.90E-05	0.112	6.64E-08	0.038	6.37E-02
PlasmaBlast	-0.054	8.94E-03	-0.096	3.13E-06	-0.021	3.10E-01
CD4T	-0.146	1.68E-12	-0.176	9.14E-18	-0.149	5.34E-13
Gran	-0.075	2.72E-04	-0.074	3.37E-04	-0.066	1.52E-03

[*] Adjusted for age, sex, cell types.
[**] SEMs (2IQR) were defined as DNA methylation mutations that were greater than 2 times the IQR above the upper quartile of a given CpG or less than 2 times the IQR below the lower quartile of a given CpG.
[***] SEMs (4IQR) were defined as DNA methylation mutations that were greater than 4 times the IQR above the upper quartile of a given CpG or less than 4 times the IQR below the lower quartile of a given CpG.

Supplementary Table 8. Sensitivity analysis: biweight midcorrelation analysis of EML derived from methylation data with various normalization methods (Illumina, Noob and Quantile normalization).

	PEG 1 (n = 237)							
	Illumina		Functional		Noob		Quantile	
Outcome = log(EML)	Bicor r	P_value	Bicor r	P_value	Bicor r	P_value	Bicor r	P_value
Adjusted (Age, Sex, Race/ethnicity, Cell types)								
AgeAccelHorvath	0.075	2.49E-01	0.070	2.82E-01	0.077	2.40E-01	0.021	7.50E-01
IEAA	0.080	2.17E-01	0.067	3.07E-01	0.074	2.59E-01	0.029	6.55E-01
AgeAccelHannum	0.109	9.29E-02	0.093	1.52E-01	0.104	1.10E-01	0.036	5.86E-01
AgeAccelGrim	0.203	1.71E-03	0.213	9.69E-04	0.241	1.82E-04	0.134	3.93E-02
Cell types								
CD8.naive	0.070	2.84E-01	0.042	5.23E-01	0.029	6.52E-01	0.066	3.10E-01
CD8pCD28nCD45RAn	0.063	3.34E-01	0.082	2.07E-01	0.097	1.36E-01	0.074	2.58E-01
PlasmaBlast	-0.104	1.11E-01	-0.110	9.23E-02	-0.111	8.91E-02	-0.068	2.98E-01
CD4T	-0.080	2.20E-01	-0.118	6.95E-02	-0.115	7.66E-02	0.018	7.85E-01
Gran	-0.152	1.90E-02	-0.170	8.67E-03	-0.163	1.21E-02	-0.013	8.48E-01

Supplementary Table 9. SEMs enriched within clock CpGs.

	FHS (n = 2326)	WHI (n= 2091)	JHS (n= 1734)	PEG 1 (n = 237)
Clock	N	N	N	N
HorvathClock (353 CpGs)				
Significant	46	34	22	4
Non-Significant	2280	2057	1712	233
PhenoClock (513 CpGs)				
Significant	151	97	121	19
Non-Significant	2175	1994	1613	218
HannumClock (71 CpGs)				
Significant	26	10	19	5
Non-Significant	2300	2081	1715	232

Supplementary Table 10. SEMs enriched within gene regions

	FHS (n = 2326)	WHI (n= 2091)	JHS (n= 1734)	PEG 1 (n = 237)
Genomic region	N	N	N	N
TSS1500				
Significant	545	110	229	23
Non-Significant	1781	1981	1505	214
TSS200				
Significant	908	385	372	142
Non-Significant	1418	1706	1362	95
5'UTR				
Significant	213	71	64	8
Non-Significant	2113	2020	1670	229
1stExon				
Significant	595	346	322	84
Non-Significant	1731	1745	1412	153
Gene Body				
Significant	296	693	357	24
Non-Significant	2030	1398	1377	213
3'UTR				
Significant	110	276	179	6
Non-Significant	2216	1815	1555	231

Supplementary Table 11. SEMs enriched within regulatory regions.

	JHS (n= 1734)
Genomic region	**N**
Enhancer from FANTOM5	
Significant	54
Non-Significant	1680
DNase hypersensitive from ENCODE	
Significant	284
Non-Significant	1450
Open chromatin from ENCODE	
Significant	436
Non-Significant	1298
Transcription factor binding site from ENCODE	
Significant	383
Non-Significant	1351
Promoter from Methylation Consortium	
Significant	165
Non-Significant	1578

Supplementary Table 12. Top SEMs enriched KEGG pathways significantly associated with faster AgeAccelHorvath.

	KEGG Pathway	Description	Number of people with SEMs enriched	AgeAccel _Coef[*]	AgeAccel_Pvalue
FHS (n = 2326)	hsa05165	Human papillomavirus infection	355	1.664	7.62E-07
	hsa04360	Axon guidance	347	1.726	8.53E-07
	hsa05032	Morphine addiction	346	2.138	1.13E-11
	hsa04713	Circadian entrainment	341	1.796	4.97E-08
	hsa05033	Nicotine addiction	339	2.239	4.33E-13
	hsa04510	Focal adhesion	332	1.684	1.25E-06
	hsa04080	Neuroactive ligand-receptor interaction	326	2.487	2.84E-15
	hsa04724	Glutamatergic synapse	300	2.412	7.38E-13
	hsa04934	Cushing syndrome	297	2.041	1.15E-08
	hsa04727	GABAergic synapse	294	2.419	7.00E-13
WHI (n= 2091)	hsa00053	Ascorbate and aldarate metabolism	312	0.800	1.88E-02
	hsa04514	Cell adhesion molecules (CAMs)	237	1.255	8.19E-04
	hsa04724	Glutamatergic synapse	231	1.220	5.95E-03
	hsa04360	Axon guidance	229	2.071	8.38E-06
	hsa04713	Circadian entrainment	208	2.083	4.77E-06
	hsa05166	Human T-cell leukemia virus 1 infection	193	-1.660	5.10E-05
	hsa05032	Morphine addiction	189	2.541	1.45E-08
	hsa04934	Cushing syndrome	171	1.504	3.23E-03
	hsa04921	Oxytocin signaling pathway	167	1.445	6.29E-03
	hsa04020	Calcium signaling pathway	165	2.235	8.79E-06
JHS (n= 1734)	hsa05032	Morphine addiction	402	1.167	9.84E-04
	hsa04713	Circadian entrainment	395	0.830	2.73E-02
	hsa04921	Oxytocin signaling pathway	372	0.798	3.33E-02
	hsa04020	Calcium signaling pathway	327	1.258	8.01E-04
	hsa04512	ECM-receptor interaction	317	1.032	1.44E-02
	hsa04727	GABAergic synapse	282	1.093	9.28E-03
	hsa04934	Cushing syndrome	249	1.370	1.00E-03
	hsa04014	Ras signaling pathway	245	0.957	3.90E-02
	hsa05412	Arrhythmogenic right ventricular cardiomyopathy (ARVC)	227	1.690	1.44E-04
	hsa04261	Adrenergic signaling in cardiomyocytes	226	1.006	2.95E-02

	hsa04360	Axon guidance	59	0.319	9.06E-01
	hsa04713	Circadian entrainment	52	0.226	9.06E-01
	hsa05224	Breast cancer	46	0.586	8.45E-01
	hsa05165	Human papillomavirus infection	45	2.202	3.04E-01
PEG 1	hsa05032	Morphine addiction	44	0.182	9.21E-01
(n = 237)	hsa05226	Gastric cancer	44	0.868	7.78E-01
	hsa04934	Cushing syndrome	43	0.907	7.78E-01
	hsa04390	Hippo signaling pathway	39	0.916	7.78E-01
	hsa04724	Glutamatergic synapse	39	1.630	5.97E-01
	hsa04015	Rap1 signaling pathway	37	-0.390	8.94E-01

[*] Age acceleration residual as dependent variable, significant enrichment as independent variable, adjusted for nlog(EML)

Supplementary Table 13. Top SEMs enriched KEGG pathways significantly associated with faster AgeAccelHannum

	KEGG Pathway	Description	Number of people with SEMs enriched	AgeAccel_Coef [*]	AgeAccel_Pvalue
	hsa05165	Human papillomavirus infection	355	1.143	1.14E-03
	hsa04360	Axon guidance	347	0.792	3.98E-02
	hsa05032	Morphine addiction	346	2.222	1.67E-12
	hsa04713	Circadian entrainment	341	1.959	2.24E-09
FHS	hsa05033	Nicotine addiction	339	2.083	1.69E-11
(n = 2326)	hsa04510	Focal adhesion	332	1.074	3.35E-03
	hsa04080	Neuroactive ligand-receptor interaction	326	2.262	1.13E-12
	hsa04724	Glutamatergic synapse	300	2.029	1.84E-09
	hsa04934	Cushing syndrome	297	2.283	1.57E-10
	hsa00053	Ascorbate and aldarate metabolism	296	0.917	5.61E-03
	hsa04724	Glutamatergic synapse	231	1.461	5.59E-04
	hsa04360	Axon guidance	229	1.477	1.19E-03
	hsa04713	Circadian entrainment	208	1.987	7.57E-06
	hsa05166	Human T-cell leukemia virus 1 infection	193	-1.133	5.13E-03
WHI	hsa05032	Morphine addiction	189	2.373	5.72E-08
(n= 2091)	hsa04934	Cushing syndrome	171	2.160	9.38E-06
	hsa04921	Oxytocin signaling pathway	167	1.330	9.42E-03
	hsa04020	Calcium signaling pathway	165	1.844	1.77E-04
	hsa04510	Focal adhesion	145	-1.175	4.02E-02
	hsa04024	cAMP signaling pathway	137	2.147	3.52E-05
	hsa04360	Axon guidance	451	0.630	3.43E-02
	hsa05032	Morphine addiction	402	0.927	5.24E-04
	hsa04713	Circadian entrainment	395	0.582	3.85E-02
	hsa04020	Calcium signaling pathway	327	0.898	1.62E-03
JHS	hsa04512	ECM-receptor interaction	317	0.841	6.33E-03
(n= 1734)	hsa04727	GABAergic synapse	282	1.503	4.07E-08
	hsa05165	Human papillomavirus infection	277	0.685	3.43E-02
	hsa04934	Cushing syndrome	249	1.445	9.84E-07
	hsa04024	cAMP signaling pathway	239	0.906	7.88E-03
	hsa05412	Arrhythmogenic right ventricular cardiomyopathy (ARVC)	227	0.865	1.91E-02
	hsa04360	Axon guidance	59	0.418	8.86E-01
	hsa04713	Circadian entrainment	52	1.182	7.74E-01
	hsa05224	Breast cancer	46	-0.554	8.86E-01
PEG 1	hsa05165	Human papillomavirus infection	45	0.654	8.68E-01
(n = 237)	hsa05032	Morphine addiction	44	0.060	9.81E-01
	hsa05226	Gastric cancer	44	0.345	8.86E-01
	hsa04934	Cushing syndrome	43	0.820	8.66E-01

	hsa04390	Hippo signaling pathway	39	0.524	8.86E-01
	hsa04724	Glutamatergic synapse	39	1.877	6.22E-01
	hsa04015	Rap1 signaling pathway	37	-0.117	9.61E-01

[*] Age acceleration residual as dependent variable, significant enrichment as independent variable, adjusted for nlog(EML)

Supplementary Table 14. Top SEMs enriched KEGG pathways significantly associated with faster IEAA.

	KEGG Pathway	Description	Number of people with SEMs enriched	AgeAccel_Coef [*]	AgeAccel_Pvalue
FHS (n = 2326)	hsa05165	Human papillomavirus infection	355	0.887	2.56E-02
	hsa05032	Morphine addiction	346	1.055	3.76E-03
	hsa04713	Circadian entrainment	341	0.939	1.50E-02
	hsa05033	Nicotine addiction	339	1.560	3.24E-06
	hsa04080	Neuroactive ligand-receptor interaction	326	1.678	1.16E-06
	hsa04724	Glutamatergic synapse	300	1.392	2.08E-04
	hsa04934	Cushing syndrome	297	1.055	1.22E-02
	hsa04727	GABAergic synapse	294	1.400	2.06E-04
	hsa04020	Calcium signaling pathway	292	1.850	7.65E-07
	hsa05224	Breast cancer	291	1.021	1.53E-02
WHI (n= 2091)	hsa04724	Glutamatergic synapse	231	1.048	1.74E-02
	hsa04360	Axon guidance	229	1.439	2.09E-03
	hsa04713	Circadian entrainment	208	1.645	2.88E-04
	hsa05166	Human T-cell leukemia virus 1 infection	193	-1.187	4.08E-03
	hsa05032	Morphine addiction	189	2.108	3.21E-06
	hsa04921	Oxytocin signaling pathway	167	1.072	4.54E-02
	hsa04020	Calcium signaling pathway	165	1.789	3.44E-04
	hsa04015	Rap1 signaling pathway	163	1.306	1.63E-02
	hsa04024	cAMP signaling pathway	137	2.296	1.60E-05
	hsa05033	Nicotine addiction	135	2.201	1.11E-05
JHS (n= 1734)	hsa05032	Morphine addiction	402	0.884	3.66E-02
	hsa04934	Cushing syndrome	249	1.005	4.40E-02
	hsa05412	Arrhythmogenic right ventricular cardiomyopathy (ARVC)	227	1.322	8.85E-03
	hsa05033	Nicotine addiction	222	1.367	3.57E-03
	hsa05226	Gastric cancer	166	1.293	2.91E-02
	hsa04080	Neuroactive ligand-receptor interaction	164	2.127	2.41E-05
	hsa05414	Dilated cardiomyopathy (DCM)	155	1.893	2.23E-04
	hsa04721	Synaptic vesicle cycle	115	2.110	2.54E-04
	hsa05410	Hypertrophic cardiomyopathy (HCM)	108	2.627	1.26E-05
	hsa04950	Maturity onset diabetes of the young	81	1.958	8.00E-03
PEG 1 (n = 237)	hsa04360	Axon guidance	59	-0.324	9.32E-01
	hsa04713	Circadian entrainment	52	-0.461	9.10E-01
	hsa05224	Breast cancer	46	-0.076	9.81E-01
	hsa05165	Human papillomavirus infection	45	2.479	1.66E-01
	hsa05032	Morphine addiction	44	-0.302	9.32E-01
	hsa05226	Gastric cancer	44	0.063	9.81E-01
	hsa04934	Cushing syndrome	43	0.413	9.32E-01
	hsa04390	Hippo signaling pathway	39	0.472	9.10E-01
	hsa04724	Glutamatergic synapse	39	1.412	7.33E-01
	hsa04015	Rap1 signaling pathway	37	-1.111	8.06E-01

[*] Age acceleration residual as dependent variable, significant enrichment as independent variable, adjusted for nlog(EML)

Supplementary Table 15. Top SEMs enriched KEGG pathways significantly associated with faster AgeAccelGrim.

	KEGG Pathway	Description	Number of people with SEMs enriched	AgeAccel_Coef[*]	AgeAccel_Pvalue
FHS (n = 2326)	hsa04360	Axon guidance	347	-0.798	4.97E-02
	hsa04510	Focal adhesion	332	-0.855	3.24E-02
	hsa04080	Neuroactive ligand-receptor interaction	326	-0.742	4.31E-02
	hsa04390	Hippo signaling pathway	284	-1.431	5.72E-04
	hsa04151	PI3K-Akt signaling pathway	265	-0.956	2.56E-02
	hsa05205	Proteoglycans in cancer	263	-1.454	8.02E-04
	hsa05226	Gastric cancer	256	-1.399	1.04E-03
	hsa04512	ECM-receptor interaction	252	-0.865	3.81E-02
	hsa04550	Signaling pathways regulating pluripotency of stem cells	249	-1.598	1.57E-04
	hsa04010	MAPK signaling pathway	240	-1.168	9.80E-03
WHI (n= 2091)	hsa04510	Focal adhesion	145	-1.069	2.69E-02
	hsa04010	MAPK signaling pathway	135	-1.251	1.37E-02
	hsa05033	Nicotine addiction	135	0.969	2.45E-02
	hsa04080	Neuroactive ligand-receptor interaction	130	1.157	8.32E-03
	hsa04072	Phospholipase D signaling pathway	129	-1.556	1.20E-03
	hsa04310	Wnt signaling pathway	117	-1.251	1.72E-02
	hsa04152	AMPK signaling pathway	111	-1.792	4.40E-04
	hsa04911	Insulin secretion	102	1.178	1.86E-02
	hsa04014	Ras signaling pathway	95	-1.463	1.03E-02
	hsa04144	Endocytosis	91	-1.763	1.12E-03
JHS (n= 1734)	hsa04724	Glutamatergic synapse	573	-0.148	9.05E-01
	hsa04360	Axon guidance	451	0.254	8.92E-01
	hsa04725	Cholinergic synapse	437	0.164	9.05E-01
	hsa04510	Focal adhesion	413	-0.163	9.07E-01
	hsa05032	Morphine addiction	402	0.352	7.97E-01
	hsa00053	Ascorbate and aldarate metabolism	400	0.093	9.38E-01
	hsa04713	Circadian entrainment	395	0.240	8.92E-01
	hsa04015	Rap1 signaling pathway	391	0.228	8.95E-01
	hsa00040	Pentose and glucuronate interconversions	374	-0.038	9.62E-01
	hsa04080	Neuroactive ligand-receptor interaction	164	1.793	9.56E-03
PEG 1 (n = 237)	hsa04360	Axon guidance	59	-0.468	9.20E-01
	hsa04713	Circadian entrainment	52	-0.521	9.20E-01
	hsa05224	Breast cancer	46	-2.263	4.84E-01
	hsa05165	Human papillomavirus infection	45	-1.114	8.36E-01
	hsa05032	Morphine addiction	44	-0.895	9.20E-01
	hsa05226	Gastric cancer	44	-1.497	6.67E-01
	hsa04934	Cushing syndrome	43	0.378	9.20E-01
	hsa04390	Hippo signaling pathway	39	-0.786	9.20E-01
	hsa04724	Glutamatergic synapse	39	0.526	9.20E-01
	hsa04015	Rap1 signaling pathway	37	-1.882	6.67E-01

[*] Age acceleration residual as dependent variable, significant enrichment as independent variable, adjusted for nlog(EML)

Supplementary Table 16. Top SEMs enriched KEGG pathways significantly associated with faster AgeAccelPheno.

	KEGG Pathway	Description	Number of people with SEMs enriched	AgeAccel_Coef[*]	AgeAccel_Pvalue
FHS (n = 2326)	hsa05165	Human papillomavirus infection	355	0.003	1.00E+00
	hsa04360	Axon guidance	347	-1.074	2.14E-01
	hsa05032	Morphine addiction	346	0.409	8.14E-01
	hsa04713	Circadian entrainment	341	0.078	9.62E-01
	hsa05033	Nicotine addiction	339	0.300	8.43E-01
	hsa04510	Focal adhesion	332	0.049	9.74E-01
	hsa04080	Neuroactive ligand-receptor interaction	326	0.120	9.13E-01
	hsa04724	Glutamatergic synapse	300	-0.022	9.84E-01
	hsa04934	Cushing syndrome	297	-0.255	8.99E-01
	hsa00053	Ascorbate and aldarate metabolism	296	0.168	9.02E-01
WHI (n= 2091)	hsa04940	Type I diabetes mellitus	493	-0.525	4.15E-01
	hsa05330	Allograft rejection	337	-0.167	8.74E-01
	hsa05332	Graft-versus-host disease	313	-0.344	6.60E-01
	hsa00053	Ascorbate and aldarate metabolism	312	0.285	7.38E-01
	hsa00040	Pentose and glucuronate interconversions	305	0.127	9.19E-01
	hsa05416	Viral myocarditis	299	-0.086	9.67E-01
	hsa00860	Porphyrin and chlorophyll metabolism	272	0.048	9.81E-01
	hsa05320	Autoimmune thyroid disease	272	-0.228	8.32E-01
	hsa04612	Antigen processing and presentation	257	-0.228	8.32E-01
	hsa00980	Metabolism of xenobiotics by cytochrome P450	253	-0.017	9.88E-01
JHS (n= 1734)	hsa04724	Glutamatergic synapse	573	0.029	9.58E-01
	hsa04360	Axon guidance	451	0.290	8.08E-01
	hsa04725	Cholinergic synapse	437	0.150	8.98E-01
	hsa04510	Focal adhesion	413	-0.604	5.25E-01
	hsa05032	Morphine addiction	402	0.785	3.30E-01
	hsa00053	Ascorbate and aldarate metabolism	400	0.464	5.82E-01
	hsa04713	Circadian entrainment	395	1.052	1.57E-01
	hsa04080	Neuroactive ligand-receptor interaction	164	2.014	3.77E-02
	hsa05414	Dilated cardiomyopathy (DCM)	155	2.054	3.77E-02
	hsa05410	Hypertrophic cardiomyopathy (HCM)	108	2.306	4.22E-02
PEG 1 (n = 237)	hsa04360	Axon guidance	59	-0.606	9.29E-01
	hsa04713	Circadian entrainment	52	0.351	9.53E-01
	hsa05224	Breast cancer	46	-1.177	9.29E-01
	hsa05165	Human papillomavirus infection	45	0.292	9.62E-01
	hsa05032	Morphine addiction	44	-0.684	9.29E-01
	hsa05226	Gastric cancer	44	0.589	9.29E-01
	hsa04934	Cushing syndrome	43	0.075	9.86E-01
	hsa04390	Hippo signaling pathway	39	-2.425	7.03E-01
	hsa04724	Glutamatergic synapse	39	1.064	9.29E-01
	hsa04015	Rap1 signaling pathway	37	-0.464	9.29E-01

[*] Age acceleration residual as dependent variable, significant enrichment as independent variable, adjusted for nlog(SEM)

Supplementary Table 17. Association between Clock region-specific EML and corresponding DNAm AgeAccel.

Outcome = log(clock EML) Model [*]	PEG 1 (n = 237)		FHS (n = 2326)		WHI (n= 2091)		JHS (n= 1734)	
	Bicor r	P_value	Bicor r	P_value	Bicor r	P_value	Bicor r	P_value
HorvathClock (353 CpGs) v. AgeAccel	0.008	9.07E-01	0.020	8.35E-01	0.034	1.23E-01	0.180	4.16E-14
PhenoClock (513 CpGs) v. AgeAccelPheno	0.008	9.03E-01	0.013	5.29E-01	0.058	8.26E-03	0.064	7.57E-03
HannumClock (71CpGs) v. AgeAccelHannum	0.071	2.77E-01	-0.030	1.51E-01	0.017	4.28E-01	0.044	6.84E-02

[*] Adjusted for Age, Sex, Cell types, Race/ethnicity, Log(total EML)

Supplementary Table 18. Biweight midcorrelation analysis of regulatory region-specific EML in JHS.

Outcome = log(Region EML) [*]	Enhancer (CpGs = 26395)		DNAse (CpGs = 466862)		Open Chromatin (CpGs = 108758)		TFBS (CpGs = 122647)		Promoter (CpGs = 110008)	
	Bicor r	P_value	Bicor r	P_value	Bicor r	P_value	Bicor r	P_value	Bicor r	P_value
Age	-0.045	6.01E-02	0.169	1.24E-12	0.019	4.34E-01	-0.065	7.11E-03	-0.027	2.58E-01
DNAm Age Acceleration										
AgeAccelHorvath	-0.027	2.66E-01	0.129	6.75E-08	0.026	2.73E-01	-0.012	6.14E-01	-0.003	8.94E-01
IEAA	-0.027	2.64E-01	0.132	3.26E-08	0.025	3.04E-01	-0.023	3.34E-01	-0.002	9.32E-01
AgeAccelHannum	-0.059	1.36E-02	0.236	2.08E-23	0.002	9.26E-01	-0.089	1.94E-04	-0.029	2.35E-01
AgeAccelGrim	0.022	3.66E-01	0.098	4.01E-05	0.018	4.56E-01	-0.038	1.17E-01	-0.013	5.99E-01
Cell types										
CD8.naive	-0.004	8.69E-01	-0.101	2.40E-05	0.029	2.20E-01	0.027	2.67E-01	-0.029	2.23E-01
CD8pCD28nCD45RAn	0.016	5.10E-01	0.073	2.33E-03	-0.004	8.57E-01	-0.024	3.21E-01	0.038	1.12E-01
PlasmaBlast	-0.020	3.98E-01	-0.011	6.41E-01	0.038	1.11E-01	0.016	4.99E-01	-0.040	9.44E-02
CD4T	-0.012	6.05E-01	-0.090	1.79E-04	0.086	3.33E-04	0.109	5.20E-06	0.016	5.03E-01
Gran	-0.067	5.15E-03	-0.181	3.39E-14	0.106	8.95E-06	0.151	2.47E-10	-0.049	4.21E-02

[*] Adjusted for Age, Sex, Cell types, Log(total EML).

Supplementary Table 19. Biweight midcorrelation analysis of EML in FHS, stratified by the direction of SEM.

Outcome = log(EML) [*]	All EML		Hypermethylated EML		Hypomethylated EML	
	Bicor r	P_value	Bicor r	P_value	Bicor r	P_value
Age	0.244	7.15E-33	0.271	1.81E-40	0.139	1.53E-11
BMI	0.017	4.23E-01	0.016	4.49E-01	0.012	5.53E-01
DNAm Age Acceleration						
AgeAccelHorvath	0.106	3.11E-07	0.101	1.19E-06	0.092	9.46E-06
IEAA	0.109	1.26E-07	0.100	1.34E-06	0.107	2.22E-07
EEAA (Unadjusted for cell types)	0.297	1.12E-48	0.268	1.20E-39	0.286	7.07E-45
AgeAccelHannum	0.225	4.12E-28	0.225	3.85E-28	0.192	1.05E-20
AgeAccelGrim	0.173	3.74E-17	0.167	5.80E-16	0.125	1.29E-09
DNAmADMAdjAge	0.121	4.58E-09	0.121	4.26E-09	0.076	2.62E-04
DNAmB2MAdjAge	0.189	4.69E-20	0.197	7.10E-22	0.088	2.26E-05
DNAmCystatinCAdjAge	0.119	8.71E-09	0.125	1.39E-09	0.031	1.29E-01
DNAmGDF15AdjAge	0.167	6.09E-16	0.181	1.58E-18	0.093	7.75E-06
DNAmLeptinAdjAge	-0.021	3.19E-01	-0.003	8.74E-01	-0.047	2.23E-02
DNAmPACKYRSAdjAge	0.166	7.99E-16	0.142	5.10E-12	0.172	7.31E-17
DNAmPAI1AdjAge	0.024	2.49E-01	0.056	6.70E-03	-0.053	1.12E-02
DNAmTIMP1AdjAge	0.070	7.04E-04	0.083	5.69E-05	0.021	3.06E-01
AgeAccelPheno	0.178	4.45E-18	0.168	3.52E-16	0.133	1.14E-10
DNAmTLAdjAge	-0.080	1.17E-04	-0.088	1.99E-05	-0.057	5.93E-03

DNAmAgeSkinBloodClockAdjAge	0.102	8.32E-07	0.127	7.26E-10	0.044	3.36E-02
Cell types						
CD8.naive	-0.072	5.20E-04	-0.071	5.91E-04	-0.040	5.38E-02
CD8pCD28nCD45RAn	0.085	3.90E-05	0.045	2.83E-02	0.104	5.07E-07
PlasmaBlast	-0.054	8.94E-03	-0.043	3.92E-02	-0.060	3.98E-03
CD4T	-0.146	1.68E-12	-0.138	2.08E-11	-0.096	3.89E-06
Gran	-0.075	2.72E-04	-0.106	3.10E-07	-0.017	4.15E-01

[*]Adjusted for age, sex, cell types.

Supplementary Table 20. FHS: Distribution of hype- and hypomethylated SEMs in relation to CpG island.

	Constantly Hypermethylated SEM [*]	Constantly Hypomethylated SEM [**]
Open Sea	8403	41533
Island	45807	5366
N_Shelf	711	5101
N_Shore	11205	6769
S_Shelf	619	4913
S_Shore	8483	5233

[*] Constantly hypermethylated SEMs were defined as DNA methylation mutations identified in more than 10 participants that were greater than three times the IQR above the upper quartile of a given CpG.
[**] Constantly hypomethylated SEMs were defined as DNA methylation mutations identified in more than 10 participants that were less than three times the IQR above the lower quartile of a given CpG.

Supplementary Table 21. FHS: Distribution of hype- and hypomethylated SEMs in relation to genomic region.

	Constantly Hypermethylated SEM [*]	Constantly Hypomethylated SEM [**]
TSS1500	11769	5841
TSS200	13399	1639
5'UTR	4536	3350
1st Exon	2160	628
Gene body	12757	26197
3' UTR	559	3442

[*] Constantly hypermethylated SEMs were defined as DNA methylation mutations identified in more than 10 participants that were greater than three times the IQR above the upper quartile of a given CpG.
[**] Constantly hypomethylated SEMs were defined as DNA methylation mutations identified in more than 10 participants that were less than three times the IQR above the lower quartile of a given CpG.

Supplementary Table 22. Association between Shannon entropy and age, AgeAccel, EML.

Outcome =Entropy [*]	FHS (n = 2326)		WHI (n= 2091)		JHS (n= 1734)		PEG 1 (n = 237)	
	Bicor r	P_value	Bicor r	P_value	Bicor r	P_value	Bicor r	P_value
Age	0.001	9.55E-01	0.068	2.01E-03	0.071	2.92E-03	0.117	7.30E-02
DNAm Age Acceleration								
AgeAccelHorvath	0.081	9.11E-05	0.160	1.76E-13	-0.039	1.02E-01	0.006	9.22E-01
IEAA	0.035	9.02E-02	0.131	1.70E-09	-0.052	3.16E-02	0.018	7.83E-01
EEAA (Unadjusted for cell types)	0.038	6.39E-02	0.090	3.60E-05	0.017	4.73E-01	0.123	5.93E-02
AgeAccelHannum	0.155	6.23E-14	0.136	4.60E-10	0.063	9.10E-03	0.096	1.41E-01
AgeAccelGrim	0.077	1.89E-04	0.228	3.86E-26	0.043	7.17E-02	0.164	1.14E-02
DNAmADMAdjAge	0.001	9.62E-01	0.205	2.80E-21	-0.022	3.49E-01	0.075	2.53E-01
DNAmB2MAdjAge	0.033	1.12E-01	0.057	9.76E-03	-0.019	4.19E-01	0.157	1.59E-02
DNAmCystatinCAdjAge	0.110	9.87E-08	0.422	5.62E-91	0.075	1.67E-03	0.200	1.93E-03
DNAmGDF15AdjAge	0.095	4.70E-06	0.226	1.53E-25	0.230	2.54E-22	0.076	2.45E-01
DNAmLeptinAdjAge	-0.023	2.68E-01	-0.060	6.11E-03	0.031	2.01E-01	-0.034	6.05E-01
DNAmPACKYRSAdjAge	0.091	1.17E-05	0.125	9.64E-09	0.054	2.46E-02	0.112	8.44E-02
DNAmPAI1AdjAge	-0.036	8.02E-02	0.009	6.68E-01	-0.007	7.78E-01	0.099	1.30E-01
DNAmTIMP1AdjAge	0.015	4.73E-01	-0.041	5.96E-02	-0.085	4.17E-04	-0.113	8.17E-02
AgeAccelPheno	0.079	1.35E-04	0.069	1.68E-03	-0.008	7.47E-01	-0.018	7.88E-01
DNAmTLAdjAge	0.002	9.36E-01	0.042	5.41E-02	0.017	4.78E-01	0.010	8.80E-01
DNAmAgeSkinBloodClockAdjAge	-0.025	2.35E-01	-0.063	4.18E-03	-0.001	9.81E-01	-0.072	2.67E-01
SEM	0.089	1.63E-05	0.294	6.87E-43	0.325	7.22E-44	0.281	1.10E-05

[*] Adjusted for Age, Sex, Race/ethnicity, Cell types.