Human aging, DNA methylation, and telomere length: Investigating indices of biological aging

By Yunsung Lee, MSc.

Dissertation presented for the degree of Philosophical Doctor (Ph.D.)

Institute of Health and Society,
Faculty of Medicine,
University of Oslo

Department of Genetics and Bioinformatics,
Division of Health Data and Digitalization,
Norwegian Institute of Public Health

Series of dissertations submitted to the
Faculty of Medicine, University of Oslo

ISBN 978-82-8377-838-0

Cover: Hanne Baadsgaard Utigard.
Print production: Reprosentralen, University of Oslo.

Contents

Acknowledgments

First and foremost, I would like to show my gratitude to my supervisors, Astanand (Anil) Jugessur, Jennifer R. Harris, Håkon K. Gjessing, and Per M. Magnus. Especially, I want to thank Anil for sharing his scientific insight, knowledge, and wisdom. His limitless enthusiasm and perfectionism constituted a solid basis for writing this thesis and all the enclosed papers. I thank Jennifer for her challenging and inspiring questions, suggestions, and comments. Jennifer's input always made me ask even more questions to myself. I thank Håkon for sharing his statistical expertise. It was so exciting to discuss the statistical modeling/inference aspects of the project with him. I send special thanks to Per for recruiting me back in 2016 and getting me the opportunity for higher education. Thanks to his consistent support, I could overcome the nonlinearity of this Ph.D. program.

I also express many thanks to Steve Horvath and Ake Lu at the University of California Los Angeles. The six-month-long research stay at Horvath's lab was a perfect rescue plan when the original plan was delayed. During this period, Anil P.S. Ori and Abraham Aviv gladly shared their expertise in telomere biology. I much appreciate their help.

I would like to pay special regard to Siri E. Håberg. Without her START data, I would never have published the third paper or finalized this thesis. I also want to thank Jon Bohlin for being a senior author and providing insightful comments and suggestions on my thesis. Also, I express my gratitude to the START research group for all the creative suggestions and feedback.

Furthermore, I would like to thank my fiancé, Xiang Yong He, for her endless understanding and support. She has been my sunshine at all times. I always remember the love and support from my parents, sister, friends, other family members, and family-in-law. Lastly, I thank God for sending all the guidance and protection from above.

Oslo, Norway

Yunsung Lee

Abbreviations

5mC	5-methylcytosine
ABEC	Adult Blood-based EPIC clock
AGA	Appropriate for gestational age
BMI	Body mass index
BMIQ	Beta-mixture quantile dilation
cABEC	common ABEC
CCDS	Consensus coding sequence
CGI	CpG islands
CI	Confidence interval
CPC	Control placental clock
CpH	A cytosine followed by either adenine, thymine, or another cytosine
DIG	Digoxigenin
DNAm	DNA methylation
DNMT	DNA methyltransferase
EAA	Epigenetic age acceleration
eABEC	extended ABEC
EM algorithm	Expectation-maximization algorithm
EPIC	Illumina Infinium MethylationEPIC BeadChip
EWAS	Epigenome-wide association study
FDR	False discovery rate
FWER	Family-wise error rate
GEO	Gene Expression Omnibus
GWAS	Genome-wide association study
HM27	Illumina HumanMethylation27 BeadChip
HM450	Illumina HumanMethylation450 BeadChip
ITS	Interstitial telomeric sequence
LASSO	Least absolute shrinkage and selection operator
LBC1921	The Lothian Birth Cohorts of 1921
LBC1936	The Lothian Birth Cohorts of 1936
LGA	Low for gestational age
LTL	Leukocyte telomere length
MAD	Mean absolute difference
MoBa	The Norwegian Mother, Father, and Child Cohort Study
MW ladder	Molecular weight ladder
Noob	Normal-exponential out-of-band probes
OD	Optical density
OR	Odds ratio
PC	Principal component
PLR	Penalized linear regression
PNA	Peptide nucleic acid

Q-FISH	Quantitative fluorescent in situ hybridization
qPCR	Quantitative Polymerase Chain Reaction
r	Pearson correlation coefficient
RPC	Robust placental clock
SNP	Single-nucleotide polymorphism
START	STudy of Assisted Reproductive Technology
STELA	Single Telomere Length Analysis
SWAN	Subset-quantile Within Array Normalization
T/S ratio	Telomere to single-copy gene ratio
TeSLA	Telomere Shortest Length Assay
TL	Telomere length
TRF	Terminal restriction fragment
WGBS	Whole-genome bisulfite sequencing
WHO	World Health Organization

List of papers

Paper I

Lee Y, Haftorn KL, Denault WRP, Nustad HE, Page CM, Lyle R, Lee-Ødegård S, Moen GH, Prasad RB, Groop LC, Sletner L, Sommer C, Magnus MC, Gjessing HK, Harris JR, Magnus P, Håberg SE, Jugessur A[†], Bohlin J[†]. Blood-based epigenetic estimators of chronological age in human adults using DNA methylation data from the Illumina MethylationEPIC array. BMC Genomics. 2020 Oct 27;21(1):747. doi: 10.1186/s12864-020-07168-8. PMID: 33109080; PMCID: PMC7590728. (Published, †joint senior authors)

Paper II

Lee Y, Choufani S, Weksberg R, Wilson SL, Yuan V, Burt A, Marsit C, Lu AT, Ritz B, Bohlin J, Gjessing HK, Harris JR, Magnus P, Binder AM[†], Robinson WP[†], Jugessur A[†] and Horvath S[†]. Placental epigenetic clocks: estimating gestational age using placental DNA methylation levels. Aging (Albany NY). 2019 Jun 24;11(12):4238-4253. doi: 10.18632/aging.102049. PubMed PMID: 31235674; PubMed Central PMCID: PMC6628997. (Published, [†]joint senior authors)

Paper III

Lee Y[†], Sun D[†], Ori APS, Lu AT, Seeboth A, Harris SE, Deary IJ, Marioni RE, Soerensen M, Mengel-From J, Hjelmborg J, Christensen K, Wilson JG, Levy D, Reiner AP, Chen W, Li S, Harris JR, Magnus P, Aviv A[‡], Jugessur A[‡] and Horvath S[‡]. Epigenome-wide association study of leukocyte telomere length. Aging (Albany NY). 2019 Aug 26;11(16):5876-5894. doi: 10.18632/aging.102230 (Published, [†]joint first authors, [‡]joint senior authors)

1 Introduction

Human life expectancy has increased steadily over the last century due to significant improvements in public health. Although mortality is markedly lower among youth, an increasing proportion of the elderly suffers from age-related diseases such as cancer, cardiovascular disease, dementia, diabetes mellitus, hypertension, osteoarthritis, and osteoporosis [1]. These age-related diseases not only significantly undermine the quality of life at an individual level, but they also impose an enormous economic burden at a societal level [2]. Contemporary medicine has, therefore, focused primarily on managing and preventing diseases, including the promotion of well-being and longevity among the elderly [3].

Aging research is rooted in the science of extending the healthy human life span and has emerged as one of the fundamental solutions to the challenges facing contemporary medicine. The US National Institute on Aging has long recognized the importance of identifying reliable and modifiable biomarkers of aging [4]. At the time of writing, a myriad of candidate biomarkers of aging are currently being developed based on known biological hallmarks, including DNA methylation (DNAm) [5-8], telomere length [9, 10], CD4+ and CD8+ T cell ratio [11], metabolic rate [12], proteomic alterations [13], and gut microbiota [14].

The scope of aging research has been further broadened to include growth and development in the early stages of life. As Raiten *et al.* [15] stated in their report, there is a pressing need for new biomarkers of growth, given that anthropometric measurements such as height, weight, and body composition fall short of providing a complete picture of the biological mechanisms underlying growth. Accordingly, multiple candidate biomarkers of growth based on DNAm [16, 17], telomere length [18], renal function [19], serum alanine aminotransferase [20], and homocysteine concentration [21] have been proposed.

This thesis is centered on the development of epigenetic biomarkers of aging and growth in humans. **Paper I** proposes blood-based epigenetic estimators of chronological age in adults, based on the use of DNAm data from the Illumina MethylationEPIC (EPIC) array, which is, at the time of writing, the latest methylation platform developed by Illumina (San Diego, CA, USA). We used this array to measure DNAm in 2000 mother-father-newborn trios from the Norwegian Mother, Father, and Child Cohort Study (MoBa). **Paper II** proposes placental tissue-based epigenetic clocks to estimate the gestational age of fetuses, using DNAm data from an earlier Illumina platform, the Infinium HumanMethylation450 (HM450) BeadChip, containing nearly half the number of probes as the EPIC array. **Paper III** reports an association between two prominent age-related biomarkers, leukocyte telomere length (LTL) and blood-derived DNAm, using seven multi-ethnic cohorts. Leveraging prior evidence that LTL is highly heritable and shortens with chronological age, this study identified differentially methylated regions that are associated with LTL.

2 Background

2.1 Aging and diseases in adults

Aging in adults is characterized by the gradual deterioration of biological functions over time. The reduced functions indicate an increased vulnerability to environmental challenges that contribute to an elevated risk of disease and death [22]. A rich body of literature describes the many physical and physiological manifestations of aging in humans [4]. For example, young children gradually lose their ability to hear high-frequency sounds as they become teenagers [23]. Adults experience hair loss [24] and frailty [25] with advanced chronological age. An increased risk of miscarriage is associated with advanced maternal age [26].

Advanced chronological age is a long-recognized risk factor for various diseases, the most prominent of which are atherosclerosis [27], type 2 diabetes [28], hypertension [29], Alzheimer’s disease [30], and Parkinson’s disease [31]. The risk of cardiovascular disease in older males aged 85 to 94 years is 20-fold higher than in younger males aged 35 to 44 years [32]. The risk of chronic obstructive pulmonary disease increases by 94% for each 10-year increment in age [33]. According to a 2018-report from the World Health Organization (WHO), ischemic heart disease, stroke, and chronic obstructive pulmonary disease were the top three causes of death in 2016 [34].

2.2 Growth and development in early life

The process of human growth starts with fertilization. It subsequently spans embryonic and fetal development during gestation and continues through infancy into adulthood. A large body of research describes physical and developmental changes in the early life course of humans [35].

For example, the sense of hearing in fetuses forms around the 19th week of gestation [36], and the canalicular period of lung development ends around the 25th week of gestation [37]. International reference curves of fetal birth weight, length, and head circumference are widely used to assess gestational age [38, 39]. Postnatally, the child's weight, height, and body mass index (BMI) are used to estimate chronological age [40-42]. Later in the child's development, specific time windows are used to estimate motor and language skill development [43].

The growth status in early life has been extensively studied with regard to perinatal morbidity, mortality, and health outcomes in adulthood. To the extent that there is a measurably large variation in fetal development across individuals [44, 45], gestational age and birth outcomes such as birth weight, head circumference, and length have been hypothesized to be useful proxies for predicting health status in later life [46-49]. For example, preterm newborns with gestational age between 23 and 27 weeks showed a higher risk of infant mortality and autism than those born at term [50, 51]. Moreover, individuals whose birth weights are low for gestational age (LGA) are more likely to have a higher BMI than those whose birth weights are appropriate for gestational age (AGA) [52, 53].

The next section elaborates on why chronological (or gestational) age is insufficient in capturing individual variation in biological deterioration, growth, and development, and why there is a need to develop reliable biomarkers of aging and growth.

2.3 Rationale for biomarkers of aging and growth

A basic premise of the abovementioned studies in Section 2.1 and 2.2 is that chronological (or gestational) age is an effective surrogate for assessing an individual's aging and growth. To some extent, this is sensible because chronological (or gestational) age readily reflects functional differences among individuals with a large age gap, e.g., teenagers versus golden-agers or extreme preterm versus term newborns. However, chronological age is not as informative when assessing health outcomes in individuals of the same chronological age or groups of individuals with a narrow age gap. This is because individuals may exhibit divergent health outcomes, even though they have the same chronological age [54, 55].

Given that the rate of biological aging varies widely across individuals [56, 57], it is crucial to develop a marker of biological aging that captures the variation in functional capacity across different age groups and within same-aged peers. For an aging biomarker to be precise and valid, two conditions must be met [58-60]. First, it must be highly correlated with chronological age, e.g., the Pearson correlation coefficient should be higher than 0.8. Second, it must be highly predictive of age-related conditions, including the onset of cancer/cardiovascular diseases, all-cause morbidity, mortality, preeclampsia, and neonatal/postnatal death.

The current consensus views DNAm-based epigenetic clocks and telomere length as the best biomarker of aging [61]. The next section explains what these two biomarkers are, how they are developed, and to the extent to which they predict chronological age and age-related conditions.

2.4 Candidate biomarkers of aging: DNAm and telomere length

2.4.1 DNAm

DNAm refers to the process by which a methyl moiety ($-CH_3$) is added to specific nucleotides in DNA. Of the four DNA nucleotides present in human DNA (cytosine (C), adenine (A), guanine

(G), and thymine (T)), methylation occurs predominantly at cytosine and only occasionally at adenine. The methylated form of cytosine, commonly referred to as 5-methylcytosine (5mC), has a methyl group attached to the fifth carbon of its cytosine 6-atom ring (**Figure 1A**). The vast majority of cytosine methylations are observed in ‘cytosine-phosphate-guanine’ dinucleotide motifs, which are commonly referred to as CpG sites [62]. Cytosine methylation has also been observed at a cytosine that is followed by either adenine, thymine, or another cytosine (abbreviated as CpH, where H is either A, T, or C). This non-CpG methylation is prevalent in human embryonic stem [63, 64] and brain [65, 66] cells. The methylated form of adenine, N6-methyladenine (6mA) (**Figure 1B**), is less common than 5mC (~0.05% of the total adenines [67]).

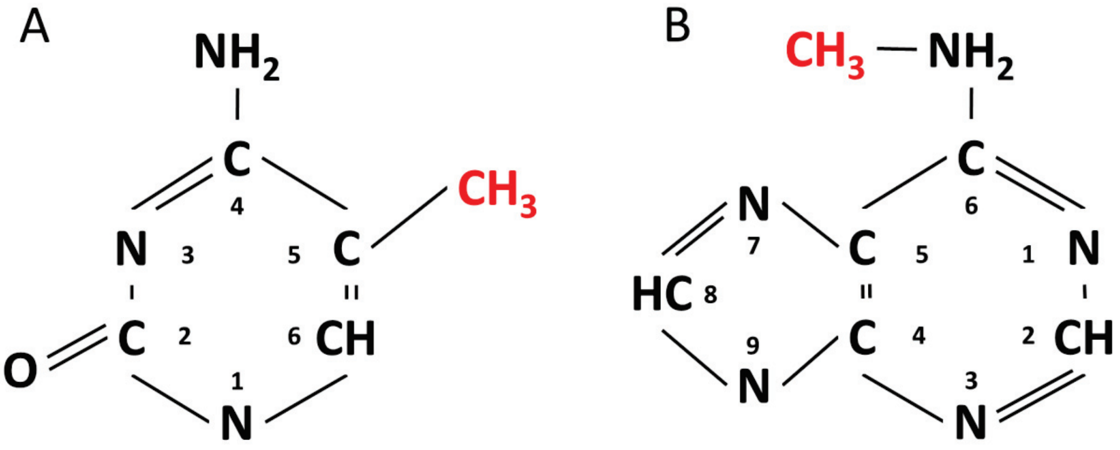


Figure 1. Chemical structure of (A) 5mC and (B) 6mA.

The human genome contains an estimated total of 28 million CpG sites [68]. The observed frequency of CpG sites is 25% of the expected frequency if one assumes random nucleotide composition, which means that CpG sites are underrepresented across the genome. The CpG sites tend to cluster in specific regions of the genome, such as promoters, with 60-70% of the CpGs located near transcription start sites [69]. These CpG-dense regions are termed CpG islands (CGIs) if they are 200-2,000 base pair long, have a high GC content (>50%), and show a high ratio (>0.6) of observed to expected number of CpG sites [70].

Nearly two decades ago, Lander *et al.* [71] reported the presence of 28,890 CGIs across the human genome. The CGIs are typically found in proximal promoters near transcription start sites [72] and in distal promoters that regulate transcription factor binding to gene bodies [73, 74]. Most CpG sites outside of CGIs are typically methylated, whereas most CpG sites within CGIs tend to be unmethylated [74].

DNA methyltransferases (DNMTs) are members of an important family of enzymes that catalyze the transfer of a methyl group from S-adenosyl-L-methionine to cytosines or adenines [75, 76]. These DNMTs belong to two families: DNMT1 and DNMT3. The primary function of DNMT1 is to maintain existing DNAm patterns [77, 78]. During cell division, for example, the parent strand maintains the methylated sites, whereas the daughter strand does not. DNMT1 binds to the daughter strand to maintain the established DNAm of the parent strand [79]. DNMT3 does not only preserve existing DNAm, similar to DNMT1, but it also creates *de novo* changes at non-methylated CpG sites [78]. The activities of DNMT3 are often found in early embryonic development [77, 80].

2.4.2 Biological mechanisms associated with DNAm

DNAm regulates gene expression by activating or repressing transcription in differentiated cells. Early studies reported correlations between cytosine methylation and gene expression in mammals and other vertebrates [81-83]. However, these studies only focused on a limited number of CpG sites in a small number of genes. The advent of high-throughput microarrays has subsequently enabled screening for associations between cytosine methylation and gene expression at a genome-wide level. The current consensus is that methylation in promoter regions suppresses gene expression, whereas methylation in gene bodies activates gene expression [84-87]. However, this pattern is not completely consistent across the genome because the direction of the association between methylation and gene expression may depend on the genomic location of the CpG sites [75, 88, 89].

Mammalian DNAm patterns show spatiotemporal variation in early development [90]. Investigations into the methylation mechanisms underlying genomic imprinting and X-chromosome inactivation are particularly active areas of research. Genomic imprinting refers to the expression or repression of a gene in a parent-of-origin specific manner [91, 92]. Imprinting marks are erased and reprogrammed during germline cell development [93]. Differentially methylated imprinting control regions are present in both maternal and paternal germline cells [91, 94]. X-chromosome inactivation is a process by which one of the two X chromosomes in females is silenced to maintain a similar gene dosage in males and females [95]. The majority of CGIs show higher methylation levels on the inactive X chromosome than the active X chromosome [95, 96].

Alterations in DNAm have also been associated with pathological processes in mammals [97]. For example, different types of cancer show promoter hypermethylation, which is also associated with reduced expression of tumor-suppressor genes [98]. In immune effector cells, differentially methylated regions have been reported to be associated with type 1 diabetes [99] and rheumatoid arthritis [100]. Differential DNAm, i.e., hypomethylation in CGIs but hypermethylation in open seas, has also been reported in type 2 diabetes [101, 102]. Moreover, associations between DNAm and other metabolic traits, including lipoprotein cholesterol [103], triglycerides [104], and coronary artery disease [105], have also been scrutinized [97]. For example, Guay *et al.* [105] reported that lower DNAm levels at the gene coding for 'troponin T1, slow skeletal type' (*TNNT1*) were associated with lower high-density lipoprotein cholesterol levels and a higher risk of coronary artery disease in men with familial hypercholesterolemia.

2.4.3 Age-related change in DNAm

A broad range of studies has assessed age-related changes in methylation across the human genome. Drinkwater *et al.* [106] reported that older subjects (mean age 75 years) had reduced 5-methylated cytosine levels in total peripheral blood DNA compared to younger subjects (mean age 25 years). Kwabi-Addo *et al.* [107] found a strong linear relationship between age and DNAm levels at the promoter regions of several genes in normal prostate tissue samples.

Further evidence of an association between DNAm and age has emerged from more recent microarray-based studies. For example, Christensen *et al.* [108] reported an association between DNAm and age in several tissues and organs, including the brain, lung, blood, head, and neck. Boks *et al.* [109] examined whole-blood samples of twins and healthy controls and reported similar associations. Alisch *et al.* [110] found 2,078 age-related loci using DNAm data from peripheral blood samples of boys aged 3-17 years. These findings add support to the early hypothesis of Cooney [111] that somatic cells inherit incomplete DNAm after each cell division, which eventually leads to genetic instability and senescence.

Microarray-based studies in newborns have revealed gestational age-related changes in DNAm from cord blood [16, 112-115] and placental tissues [116, 117]. In one of the first studies examining gestational age in newborns, Bohlin *et al.* [16] identified 5,474 CpGs associated with gestational age using DNAm data generated from cord blood samples of 1,753 newborns. In a recent meta-analysis of 3,648 newborns from 17 cohorts, Merid *et al.* [112] found 8,899 CpGs associated with gestational age. Novakovic *et al.* [116] reported differentially methylated regions between placental tissues from the 1st and 3rd trimester. Further, using DNAm data from placental samples of 170 newborns, Mayne *et al.* [117] identified 62 CpG sites that were predictive of gestational age

2.4.4 Microarrays for measuring DNAm

Whole-genome bisulfite sequencing (WGBS) is currently the gold standard for distinguishing methylated from unmethylated cytosines at a genome-wide level [118]. This method uses sodium bisulfite, which converts unmethylated cytosines to uracils but leaves methylated cytosines unchanged. Methylation levels are then quantified by contrasting bisulfite-converted DNA reads against non-converted reads. Although WGBS has been successful in measuring DNAm in diverse cells and tissues at a genome-wide level, it is still relatively expensive. It also requires advanced technical expertise to process the resulting sequence data.

An alternative to WGBS is high-throughput microarrays, such as the Infinium BeadChip® array produced by Illumina (San Diego, CA, USA). Microarrays have garnered substantial interest in recent years because they are user-friendly and can readily generate a comprehensive DNAm dataset for downstream analyses [118]. The Infinium technology is based on sodium bisulfite conversion of DNA, similar to WGBS, but instead of whole-genome sequences, this method targets CpG sites using specific probes on a microarray. The Infinium BeadChip, Illumina HumanMethylation27 BeadChip (HM27), was introduced in 2008 and contained 27,578 probes targeting CpG sites in proximal promoter regions of consensus coding sequence (CCDS) genes derived from the latest reference mouse and human genomes [119, 120]. The advent of HM27 heralded a new era of epigenetic studies by enabling epigenome-wide association studies (EWASs) of a wide range of phenotypes, including aging [121], type I diabetes [122], hearing ability [123], breast cancer [124], cigarette smoking [125], schizophrenia [126], and Kawasaki disease [127], among many others.

Since HM27 was launched, Illumina has introduced new microarrays by extending the genomic coverage of its Infinium BeadChips. HM450 was first introduced in 2011 and contained 485,577 probes that targeted 482,421 CpGs, 3,091 CpHs, and 65 single-nucleotide polymorphisms (SNPs). Subsequently, Illumina introduced EPIC in 2016, raising the number of probes to 865,918. These probes targeted 862,927 CpGs, 2,932 CpHs, and 59 SNPs (please refer to the Manifest v1.0 B5[1] released by Illumina for further details [128]).

[1] The annotations included in the manifest file correspond to the human genome assembly GRCh37 (hg19), unless stated otherwise.

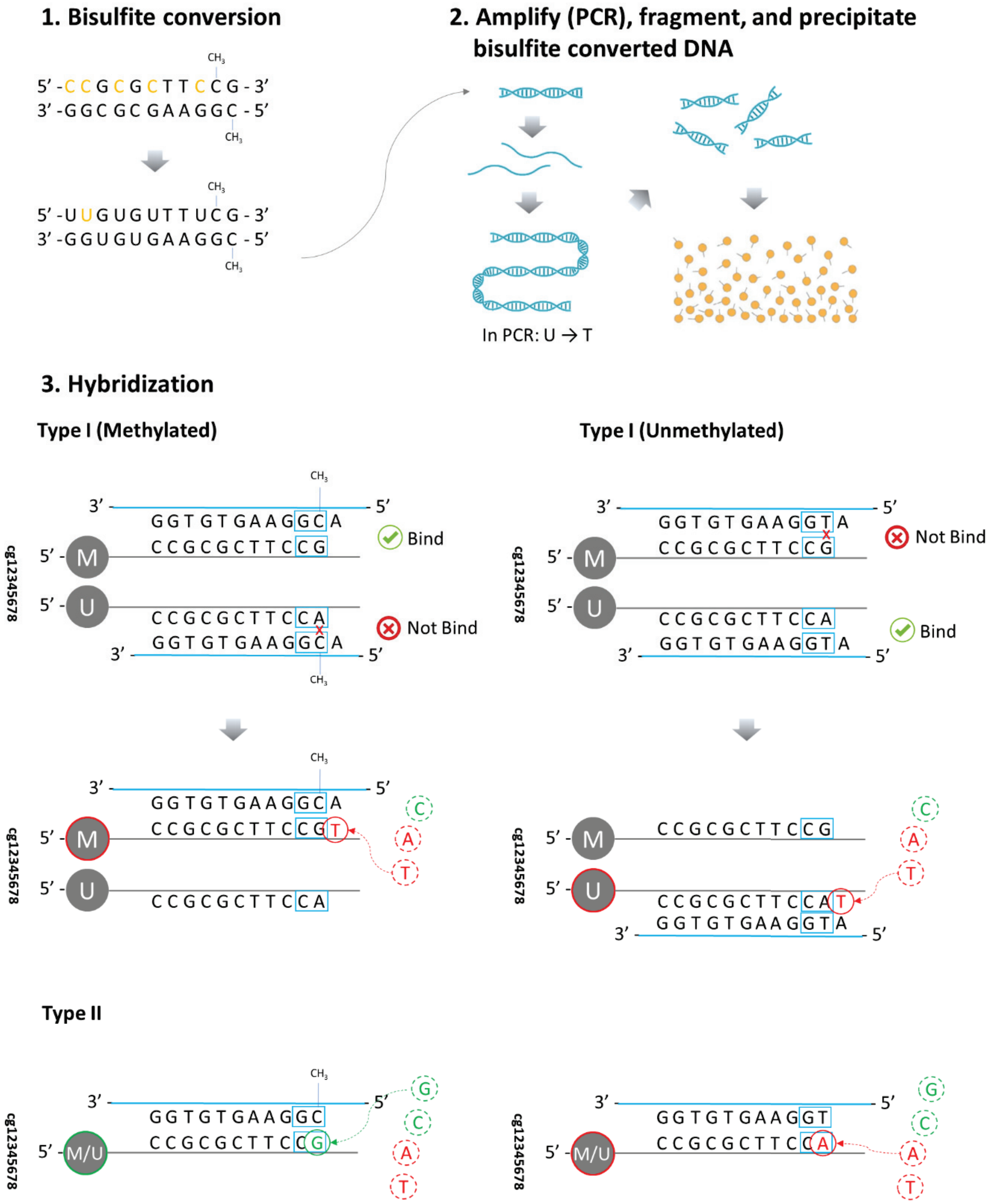


Figure 2. Illustrative workflow of the Illumina Infinium BeadChip.
Source: Infinium® HD Assay Methylation Protocol Guide, and Pidsley *et al.* [118].
This figure outlines the laboratory workflow and hybridization of DNA fragments to the probes on the BeadChip. The pattern of hybridization, i.e., the binding to the beads and single-base extension for each type of probe (Type I and Type II) is outlined in step 3. M=methylated, U=unmethylated.

Laboratory workflow of the Illumina Infinium BeadChips

The laboratory workflow of all the three Illumina Infinium BeadChips (HM27, HM450, and EPIC) starts with sodium bisulfite conversion of DNA samples (**Figure 2**). Sodium bisulfite converts unmethylated cytosine to uracil but leaves methylated cytosines unaltered. The uracil from the bisulfite conversion is then converted to thymine after DNA amplification.

For HM27, DNA samples are hybridized to the two types of Infinium Type I probes, each of which is designed to measure the methylation level at one of the targeted CpGs. Each of the Infinium Type I probes has methylated (M) and unmethylated (U) beads (**Figure 2**). The methylated bead has a 5' to 3' sequence that ends with the target CpG. By contrast, the unmethylated bead has a 5' to 3' sequence that ends with a dinucleotide of C followed by A. A 3' to 5' fragment that includes the target GC site (reversed CpG site) with methylation can only bind to the methylated bead. By contrast, a 3' to 5' fragment that includes the target GC site without methylation (GT after bisulfite conversion and amplification) can only bind to the unmethylated bead.

If the binding is successful, beads carry out single-base extension by adding one complementary nucleotide downstream of the target GC site. This complementary nucleotide is labeled with either a green or a red fluorophore. Cytosine and guanine are labeled green, whereas adenine and thymine are labeled red. Therefore, the intensity of methylation is quantified either as green or red, depending on the nucleotide downstream of each target GC site. HM450 and EPIC also include the Infinium Type I probes, the principle of which is identical to that of HM27.

The remarkable improvement in HM450 and EPIC stems from their use of Infinium Type II probes. The Infinium Type II probes carry only one bead consisting of a 5' to 3' sequence that ends with a cytosine from a target CpG. Therefore, if a 3' to 5' fragment including the target GC site with methylation binds to the bead, a complementary guanine labeled green is added to the 3' end of the bead. Conversely, if a 3' to 5' fragment including the target GC site without methylation (GT after bisulfite conversion and amplification) binds to the bead, a complementary adenine labeled red is added to the 3' end of the bead. Thus, the intensity of methylation is measured by two colors (green and red) in Type II probes.

Bead chips with single-base extension are scanned using the Illumina HiScan or iScan System. The scanner generates two high-resolution raw image files (iDAT) for each sample, one with the red label and the other with the green label, based on the fluorophores emitted from the bead chips.

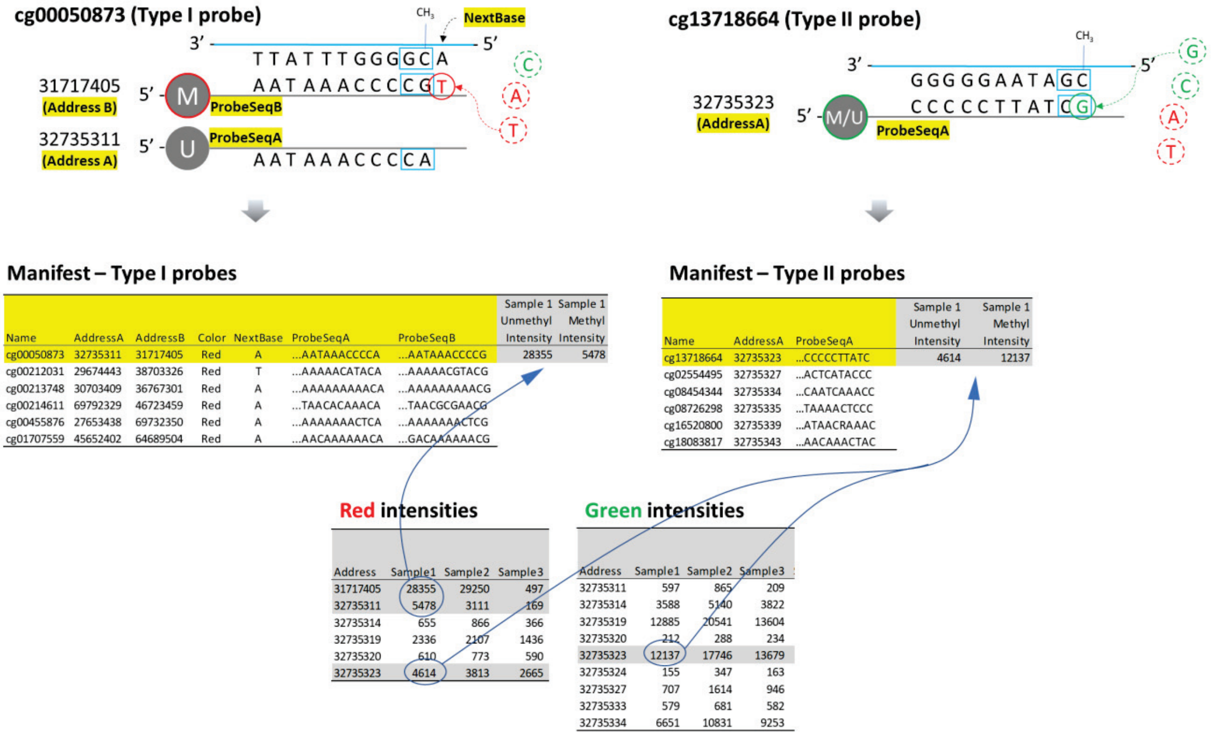
Development of a bioinformatics pipeline

The minfi R package [129] is widely used to process iDAT files because it enables building a fine-tuned pipeline with automated quality control steps. This includes probe/sample exclusion, background noise correction, and normalization. In addition to the minfi package [129], the RnBeads package [130] is another popular software for downstream analyses because it provides an all-in-one automated quality control procedure. Although the user-friendly interface of these bioinformatics packages increases efficiency, it is still important to understand the mechanisms behind each function for a fair assessment of the quality of DNAm data.

The first step in the quality-control pipeline is to read iDAT files and quantify the fluorescence intensities at each probe using the `read.methylarray.exp` function, which produces an `RGChannelSet` object. The `RGChannelSet` includes 'manifest' information in the form of two datasets: one for red fluorescence intensities and the other for green fluorescence intensities (please refer to **Figure 3** for further details regarding each of these datasets).

The next step is to determine the methylated and unmethylated intensity at each locus using different background correction and normalization functions (introduced in Section 4.1.2). Given these functions are based on different normalization methods, the function `proprocessRaw` assigns the red and green fluorescence intensities to corresponding loci using the probe addresses (denoted as AddressA and AddressB in the manifest file in **Figure 3**). For example, the unmethylated intensity of sample 1 at cg00050873 must be obtained from the red intensity of sample 1 at nucleotide position 31717405. In contrast, the methylated intensity must come from the red intensity at nucleotide position 32735311. The important point here is that the addresses of unmethylated beads can be found in AddressA, whereas those of methylated beads can be found in AddressB.

To illustrate, in the case of cg13718664, the methylated intensity of sample 1 must come from the red intensity at 32735323, and the unmethylated intensity must come from the green intensity at the same address (327353223). Normalization methods are then applied to the methylated and unmethylated intensities (please see Section 4.1.2 for further details).



Manifest – Type I probes

Name	AddressA	AddressB	Color	NextBase	ProbeSeqA	ProbeSeqB	Sample 1 Unmethyl Intensity	Sample 1 Methyl Intensity
cg00050873	32735311	31717405	Red	A	...AATAAACCCCA	...AATAAACCCCG	28355	5478
cg00212031	29674443	38703326	Red	T	...AAAAACATACA	...AAAAACGTACG		
cg00213748	30703409	36767301	Red	A	...AAAAAAAAACA	...AAAAAAAAACG		
cg00214611	69792329	46723459	Red	A	...TAACACAAACA	...TAACGCGAACG		
cg00455876	27653438	69732350	Red	A	...AAAAAAACTCA	...AAAAAAACTCG		
cg01707559	45652402	64689504	Red	A	...AACAAAAAACA	...GACAAAAAACG		

Manifest – Type II probes

Name	AddressA	ProbeSeqA	Sample 1 Unmethyl Intensity	Sample 1 Methyl Intensity
cg13718664	32735323	...CCCCCTTATC	4614	12137
cg02554495	32735327	...ACTCATACCC		
cg08454344	32735334	...CAATCAAACC		
cg08726298	32735335	...TAAAACTCCC		
cg16520800	32735339	...ATAACRAAAC		
cg18083817	32735343	...AACAAACTAC		

Red intensities

Address	Sample1	Sample2	Sample3
31717405	28355	29250	497
32735311	5478	3111	169
32735314	655	866	366
32735319	2336	2107	1436
32735320	610	773	590
32735323	4614	3813	2665

Green intensities

Address	Sample1	Sample2	Sample3
32735311	597	865	209
32735314	3588	5140	3822
32735319	12885	20541	13604
32735320	212	288	234
32735323	12137	17746	13679
32735324	155	347	163
32735327	707	1614	946
32735333	579	681	582
32735334	6651	10831	9253

Figure 3. Underlying functionality for generating methylated and unmethylated intensities. A manifest file contains technical and genomic information on all targeted loci (mostly CpG sites). This file includes the type and address of probes, the color of fluorescence, the name of neighboring genes, the type of genomic region, and much more. For illustrative purposes, the main manifest file is displayed here as two separate excerpts, one for Type I probes and the other for Type II probes.

The final step is to compute beta and M values based on the levels of methylation and unmethylation at each locus. The beta values are defined as $\frac{M}{M+U+100}$, where M and U are the levels of methylation and unmethylation, respectively. The M values are defined as $\log(\frac{M}{U})$.

The next section describes the existing biomarkers of aging trained on microarray DNAm data.

2.4.5 DNA methylation-based aging biomarkers

The development of aging biomarkers begins with a search for associations between human chronological age and DNAm levels at each CpG site [108-110, 122, 131-133]. Although several individual CpG sites show modest correlations with chronological age [134, 135], these individual CpG sites are insufficient to obtain a high and robust age correlation across independent cohorts.

The approach of using a linear combination of DNAm at multiple CpG sites has been proposed to increase the prediction accuracy of aging biomarkers [5, 132]. Mathematically, the epigenetic age of the ith individual can be defined as $DNAmAge_i = \hat{\beta}_0 + \sum_{j=1}^{k} X_{i,j}\hat{\beta}_j$, where $X_{i,j}$ is the ith individual's DNAm level at the jth CpG site, and $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k$ are the estimated coefficients. Penalized regression has been widely used to estimate these coefficients owing to its efficacy in selecting variables that are most predictive of an outcome of interest from high dimensional data (i.e., data in which the number of predictor variables is much larger than the sample size). The next sections will elaborate on the modeling procedures and performance of various existing epigenetic clocks.

Table 1. Summary of previously published epigenetic clocks.

Target	Name of epigenetic clock	Training set	Source	Method	Number of selected CpGs
Adults	**First-generation clock**				
	Bocklandt Saliva clock (2011)	34 identical male twin pairs (HM27)	Saliva	LASSO[1]	
	Hannum Blood-based clock (2013)	656 individuals (HM450)	Peripheral blood	Elastic net regression (Chronological age ← CpGs)	71
	Horvath Pan-tissue clock (2013)	3,931 individuals (HM27 & HM450)	51 different tissues	Elastic net regression (Chronological age ← CpGs)	353
	Horvath Skin & Blood clock (2018)	896 individuals (HM450 & EPIC)	Whole blood, buccal, epithelium, fibroblast, skin, and cord blood	Elastic net regression (Chronological age ← CpGs)	391
	Second-generation clock				
	Levine PhenoAge clock (2018)	Step 1: 9,926 individuals Step 2: 456 individuals at two time points (HM27, HM450 & EPIC)	Step 1: NA[2] Step 2: peripheral blood	Step 1: Penalized Cox regression. (Time-to-death ← 42 clinical biomarkers + chronological age + sex) * selected 10 clinical biomarkers and calculated PhenoAge through transformation. Step 2: Elastic net regression. (PhenoAge ← CpGs) * Selected 513 CpGs.	513
	Lu GrimAge clock (2019)	Step 1 & 2: 1,731 individuals (450K & EPIC)	Step 1: peripheral blood Step 2: NA[1]	Step 1: Elastic net regression. (each of 88 Plasma proteins ← CpGs) * Selected 12 plasma proteins. Step 2: Penalized Cox regression. (Time-to-death ← chronological age + sex + DNAm imputed smoking pack-years + 12 DNAm imputed plasma proteins) * Selected 7 DNAm-imputed plasma proteins.	NA
Newborns	Bohlin Cord Blood clock (2016) - Ultrasound[3]	1,068 newborns (HM450)	Cord blood	LASSO[2]	96
	Knight Cord Blood clock (2016)	207 newborns (HM27 & HM450)	Cord blood	Elastic net regression	148

[1] No epigenetic data were used in this step.

[2] LASSO: Least absolute shrinkage and selection operator.

[3] Bohlin and colleagues reported two epigenetic clocks, one for ultrasound-estimated gestational age and the other for last menstrual-estimated gestaional age.

Single-tissue epigenetic clocks: Bocklandt et al. (2011) Saliva Clock and Hannum et al. (2013) Blood-based Clock

Bocklandt *et al.* [132] developed an epigenetic clock using saliva-based HM27-derived DNAm data from 34 identical male twin pairs aged 21-55 years. They used penalized linear regression, i.e., Least absolute shrinkage and selection operator (LASSO), together with a regularization parameter chosen through a 'leave-one-out strategy,' which selected the three strongest predictors of chronological age. Their results showed a correlation coefficient of r=0.83 between predicted and observed age, and a mean absolute difference (MAD) of 5.2 years between predicted and observed age in the training set. However, the Bocklandt *et al.* [132] Saliva clock was not validated in an independent cohort.

Subsequently, Hannum *et al.* [5] proposed another epigenetic clock that was trained on blood-based HM450-derived DNAm data from 656 individuals aged 19-101 years. Like the Bocklandt *et al.* [132] Saliva clock, Hannum and colleagues applied penalized linear regression (i.e., elastic net regression – a mixture of LASSO and ridge regression) of chronological age on DNAm levels at more than 450,000 CpG sites. The regularization parameter was chosen through 10-fold cross-validation, resulting in the selection of 71 CpG sites. The Hannum *et al.* [5] Blood-based clock was validated in 174 independent samples (r=0.90, Root Mean Square Error=4.89 years). However, later applications of this method showed that this clock was suboptimal for age prediction in children [136, 137].

The two epigenetic clocks described above were trained on single-tissue DNAm data (saliva or blood) that can be obtained from human subjects non-invasively. These single-tissue clocks are thus easily applicable to other studies with DNAm from saliva or blood samples. As the Hannum *et al.* [5] Blood-based clock has shown highly accurate age predictions, it has widely been used to estimate the epigenetic age of human blood tissues from healthy controls and subjects with a condition [138-140].

However, a criticism of these single-tissue clocks was that they were insufficient in reflecting an innate aging profile functioning across all human tissues and cells [60]. In an attempt to address this limitation, Hannum *et al.* [5] applied their epigenetic clock to DNAm generated from breast, kidney, lung, and skin tissue (Figure 4A in Hannum *et al.* [5]). As there were substantial linear offsets in age predictions, they had to calibrate the offsets using linear regressions.

Later in 2013, the criticism against single-tissue epigenetic clocks was addressed by the Horvath [6] Pan-tissue clock, which was applicable to DNAm generated from various human tissues.

Multi-tissue epigenetic clock: Horvath [6] Pan-tissue Clock

Horvath [6] developed an epigenetic clock using 82 DNAm datasets (a mixture of HM27 and HM450) from 51 healthy tissues[2] taken from 8,000 individuals aged 0-100 years. Elastic net regression with a mixture parameter of 0.5 resulted in the selection of 353 CpG sites that were present on both HM27 and HM450. In a test set comprising DNAm from multiple tissues, the Horvath [6] Pan-tissue clock showed a correlation coefficient of 0.96 and a MAD of 3.6 years, underscoring its high accuracy in predicting epigenetic age across multiple tissues. As the Horvath [6] Pan-tissue clock enables the measurement of epigenetic age across a broad spectrum of tissues, it has been used to investigate epigenetic age in a wide variety of age-related conditions (see Table 1 in the review article by Horvath and Raj [60]).

[2] White blood cell, peripheral blood mononuclear cell, cord blood, cerebellum, frontal cortex, prefrontal cortex, temporal cortex, breast, buccal, cartilage knee, colon, dermal fibroblast, epidermis, gastric, head, neck, heart, kidney, liver, lung, bone marrow, placenta, prostate, saliva, stomach and thyroid.

An underlying hypothesis for these association studies was that 'epigenetic age acceleration' (EAA) might be associated with the risk of age-related conditions. In other words, individuals with a given condition might be epigenetically older or younger than same-aged subjects without the condition. Note that EAA is obtained from the residuals of a regression of epigenetic age on chronological age. As Horvath and Raj [60] highlighted in their review, several studies found modest to strong associations between EAA and a variety of health conditions, including cognitive function, frailty, Down syndrome, Huntington disease, and insulin level, indicating that EAA might explain why people of the same chronological age show different risks of age-related diseases. However, EAA derived from the Horvath [6] Pan-tissue clock showed only weak associations with BMI, time to death, and markers of immunosenescence, underscoring the need for additional clocks that target other traits and tissues.

Levine *et al.* [8] PhenoAge clock

Levine *et al.* [8] proposed a new epigenetic clock by penalized-regressing 'phenotypic age,' instead of chronological age, on blood-based DNAm. They hypothesized that the phenotypic age estimated from clinical biomarkers[3] would be better at capturing individual variations in the onset of age-related disease, functional deterioration, and death than chronological age. The development of the Levine *et al.* [8] PhenoAge clock comprised two steps. First, phenotypic age was estimated using a penalized Cox regression of the time to aging-related mortality on 42 clinical biomarkers in 9,926 adult samples. This procedure selected ten clinical biomarkers[3]. Second, the phenotypic age was penalized-regressed on DNAm levels at 20,169 CpG sites that were common

[3] Albumin, creatinine, serum glucose, C-reactive protein, lymphocyte percent, mean (red) cell volume, red cell distribution width, alkaline phosphatase, white blood cell count, and chronological age.

between HM27, HM450, and EPIC. This regression resulted in the selection of a total of 581 CpG sites.

Although the Levine *et al.* [8] PhenoAge clock was suboptimal at predicting chronological age (r=0.62 to 0.89), it substantially improved prediction of age-related conditions compared to the earlier Horvath [6] Pan-tissue clock and Hannum *et al.* [5] Blood-based clock. Based on analyses in independent cohorts, the Levine *et al.* [8] PhenoAge clock showed strong associations with age-related morbidity (P=1.95E-20), all-cause mortality (P=7.9E-47), smoking (P=0.0033), ethnicity (P =5.1E-5), higher education (P=6E-9), higher income (P=9E-5), and blood cell counts (naïve CD8+ T cells: P=9.2E-65, naïve CD4+ T cells: P=4.2E-42, and CD4+ helper T cells: P=3.6E-58), after adjusting for chronological age.

There are two important considerations regarding the Levine *et al.* [8] PhenoAge clock. First, the authors could have used a penalized Cox regression of time-to-mortality on DNAm, instead of the two-step procedure. Indeed, Zhang *et al.* [141] published an epigenetic clock where the mortality score was estimated from DNAm levels at ten CpG sites, using a penalized Cox regression on the 58 preselected CpGs from an EWAS of mortality (False discovery rate (FDR)<0.05). However, according to the supplementary analyses by Levine *et al.* [8], the Zhang *et al.* [141] clock presented a weaker association with all-cause mortality than the Levine *et al.* [8] PhenoAge clock. The Zhang *et al.* [141] clock would have performed better if the authors had included a larger number of CpG sites. Second, one can simply use the phenotypic age estimated from the ten clinical biomarkers rather than the epigenetic estimator of the estimated phenotypic age. The latter has fundamental relevance for the utility of epigenetic biomarkers of aging. Whether epigenetic modification is directly responsible for aging is not known, but if this were the case, epigenetic biomarkers of aging would be able to contribute important insights to therapeutic interventions, either by attenuating or even reversing biological aging [142].

Horvath *et al.* [7] Skin & Blood clock and Lu *et al.* [143] GrimAge clock

Horvath *et al.* [7] recently proposed a novel and accurate epigenetic estimator of chronological age, commonly referred to as the Horvath *et al.* [7] Skin & Blood clock, trained on DNAm data from skin, blood, and saliva samples. The authors focused on the CpG sites that were shared by HM450 and EPIC. Among these CpG sites, they preselected CpGs that were significantly associated with chronological age and those that were only weakly associated with chronological age. They then penalized-regressed chronological age on these preselected CpGs. The mixture parameter was set to 0.5, and the regularization parameter was selected through cross-validation. This clock outperformed both the Horvath [6] Pan-tissue clock and the Hannum *et al.* [5] Blood-based clock in predicting chronological age in skin and blood samples. Moreover, it showed high age correlations in neurons, glia, brain, liver, and bone tissues. Not surprisingly, therefore, this clock was deemed to be highly useful in forensics.

More recently, Lu *et al.* [143] published a compelling epigenetic clock, referred to as GrimAge, using a DNAm-based surrogate for smoking pack-years and seven DNAm-based surrogates for plasma protein levels[4]. Similar to the Levine *et al.* [8] PhenoAge clock, a two-step strategy was used as follows: First, Lu and colleagues penalized-regressed each of 88 plasma protein levels and smoking pack-years on chronological age, sex, and the CpGs in common between 450K and EPIC,

[4] **Adrenomedullin**, **beta-2-microglobulin**, CD56, ceruloplasmin, **cystatin-C**, EGF fibulin-like ECM protein1, **growth differentiation factor 15**, **leptin**, myoglobin, **plasminogen activator inhibitor 1**, serum paraoxonase/arylesterase 1, and **tissue Inhibitor Metalloproteinases 1 (**those highlighted in bold here in this list were selected for the GrimAge clock).

and then selected 12 plasma proteins[4] that showed a moderate correlation coefficient (r>0.35) between observed and imputed values in the test sets. Second, the authors penalized-regressed time-to-death (all-cause mortality) on chronological age, sex, the DNAm-imputed smoking pack-years, and 12 DNAm-imputed plasma proteins. The elastic net Cox regression selected chronological age, sex, DNAm pack-years, and the seven DNAm-imputed plasma proteins[4].

The strength of GrimAge, as a second-generation epigenetic clock along with the Levine *et al.* [8] PhenoAge clock, lies in its strong association with age-related conditions. The EAA stemming from GrimAge shows a very strong association with time-to-death (P=2.0E-75), time-to-coronary heart disease (P=6.2E-24), time-to-cancer (P=1.3E-12), age-at-menopause (P=1.6E-12), and comorbidity count (P=3.45E-17). A recent study showed that having a higher GrimAge is also associated with cognitive decline [144].

Despite its strong associations with mortality and age-related diseases, GrimAge has not yet been able to replace existing clinical biomarkers, such as blood glucose and blood pressure, because measuring DNAm levels is still not simple or cost-effective. The processing time depends on the turnaround of a given core facility, which may range from several weeks to months. Furthermore, the price of the EPIC BeadChip Kit for 16 samples is currently at €4,295 (updated quote from September 8th, 2020, according to Illumina's website [128]).

Bohlin *et al.* [16] Cord Blood clock and Knight *et al.* [17] Cord Blood clock.

Bohlin *et al.* [16] and Knight *et al.* [17] each developed an epigenetic estimator of gestational age trained on DNAm data derived from newborns' cord blood. Both used penalized regression to select the CpG sites that were most predictive of gestational age. Specifically, the mixture parameters in Bohlin *et al.* [16] and Knight *et al.* [17] were 1 and 0.5, respectively. Simpkin *et* *al.* [145] validated these two clocks in a publicly accessible population-based resource of DNA methylation data known as the Accessible Resource for Integrated Epigenomic Studies (ARIES) [146]. The Bohlin *et al.* [16] Cord Blood clock showed a higher correlation with gestational age (r=0.65) than the Knight *et al.* [17] Cord Blood clock (r=0.37).

The above epigenetic estimators of gestational age are useful as proxies for assessing developmental maturity and gestational age in newborns. According to Khouja *et al.* [147], greater gestational age acceleration by the Bohlin *et al.* [16] clock was associated with higher maternal BMI ($P<0.001$), birth weight ($P<0.001$), birth length ($P<0.001$), and head circumference ($P<0.001$). In addition, Bright *et al.* [148] reported that newborns with one-week greater gestational age acceleration were 0.14 kilograms heavier and 0.55 centimeters taller, and this effect persisted until nine months of age but attenuated thereafter.

2.4.6 Telomere length

Telomeres are the repetitive hexanucleotide sequences $(TTAGGG)_n$ that tag the end of each chromosome. Their broad function is to prevent genomic instability [149]. Telomeres in somatic cells shorten after each cell division due to the repressed activities of the enzyme telomerase whose role is to maintain the ends of chromosomes [150-153]. Analyses in large cohorts have shown that telomere length (TL) shortening in leukocytes is correlated with advanced chronological age (r=-0.29 to -0.45 in the blood samples of individuals aged 20-90 years, Lee *et al.* [154]). TL ranges from nine to 11 kilobases at birth [155] but gradually shortens to approximately four kilobases around 70-80 years of age [156]. The age-dependent attrition and considerable individual variation in TL make it a strong candidate biomarker of aging. To examine the validity of TL as an aging biomarker, there have been many attempts to search for associations between TL, mostly LTL, and age-related conditions (e.g., mortality, cardiovascular disease, and cancer) after adjustment for chronological age [157].

Strikingly, shorter LTL appears to be associated with increased mortality and risk of cardiovascular disease. Deelen *et al.* [158] reported a significant LTL-mortality association using data from 870 siblings aged between 90 and 99 (nonagenarians), 1,580 of their offspring, and 725 spouses from the Leiden Longevity Study [159]. Analyses in twin cohorts corroborated these findings (Bakaysa *et al.* [160] and Kimura *et al.* [161]). For example, Fitzpatrick *et al.* [162] reported a three-fold elevated risk of myocardial infarction and stroke per one unit decrease in LTL (in kilobase of the terminal restriction fragment (TRF)) in 419 subjects from the Cardiovascular Health Study [163]. Mwasongwe *et al.* [164] analyzed 2,518 individuals from the Jackson Heart Study [165] and reported that the a higher risk of subclinical atherosclerosis and peripheral arterial disease was associated with having shorter LTL. Furthermore, findings from an analysis of 3,259 adults showed that the risk of all-cause mortality was three-fold higher per one unit decrease in LTL [166].

By contrast, the risk of certain types of cancer appears to be lower with shorter LTL [157]. Nan *et al.* [167] found a protective effect of shorter LTL against melanoma in 557 melanoma cases and 579 age-matched controls. Anic *et al.* [168] replicated the association between shorter LTL and reduced risk of melanoma using the data from 198 melanoma cases and 372 controls. The protective effect of shorter LTL was examined in other cancer types, including cancer of the lung [169, 170], breast [171, 172], pancreas [173], and prostate [174]. Further, Telomeres Mendelian Randomization Collaboration [175] also showed that longer LTL was associated with an increased risk of several types of cancers, including glioma (odds ratio (OR) 5.27, 95% confidence interval

(CI): 3.15, 8.81), ovarian cancer (OR 4.35, 95% CI: 2.39, 7.94), lung adenocarcinoma (OR 3.19, 95% CI: 2.40,4.22), neuroblastoma (OR 2.98, 95% CI: 1.92, 4.62), and bladder cancer (OR 2.19, 95% CI: 1.32, 3.66).

To explain this apparent cancer-cardiovascular disease trade-off, i.e., shorter TL is associated with an elevated risk of cardiovascular disease, whereas longer TL is associated with an increased risk of cancer, Stone *et al.* [157] and Aviv and Shay [176] invoked a hypothesis that involves the rate of cellular replications in the association between TL and cancer. Notably, short TL induced by repressed telomerase limits the replicative capacity of cells, which lowers the odds of malignant transformation in the cell cycle [157]. By contrast, the limited cell replications from having a shorter TL also point to a higher risk of degenerative diseases such as myocardial infarction, stroke, and atherosclerosis.

These findings have led to the analysis of several determinants of TL. First, higher oxidative stress predicts shorter LTL. The guanines in the TTAGGG tandem repeats in telomeres are particularly vulnerable to the hydroxyl radicals produced by oxidative stress. The guanines are oxidized to 8-oxoguanines [177-179], which interfere with telomerase activity and contribute to telomere attrition. The association between oxidative stress and cardiovascular disease [180] supports the link between oxidative stress and TL. Secondly, several genome-wide analyses have shown that TL is associated with several genetic variants [181-184]. Although 11-14 loci have thus far been recognized as genetic determinants of LTL, the exact role of the involved genes remains obscure, except for the associations with *TERT* and *TERC* [157]. TERT is the catalytic subunit of telomerase, while TERC provides the template to synthesize the telomere repeats [149].

3 Aims of the thesis

The overarching goal of this thesis was to develop biomarkers of aging and growth in humans. Accordingly, I developed several novel DNA methylation-based estimators of chronological age in adults and gestational age in fetuses. Further, I conducted a genome-wide investigation of LTL in relation to DNAm.

The specific aims of this thesis were as follows:

1. Develop EPIC-derived blood-based epigenetic clocks that predict chronological age in adults.

2. Develop placental epigenetic clocks that predict fetal gestational age.

3. Conduct an EWAS of LTL in seven large adult cohorts.

4 Methods

This section outlines quality control procedures for microarray-based DNAm data, including the calculation of detection p-values, background correction, and normalization. As mentioned in Section 2.4.4, both background correction and normalization are crucial steps for processing iDAT files because they minimize the impact of non-biological signals and systematic discrepancies in the intensities between the Type I and Type II probes. Although a plethora of approaches for background correction and normalization have been proposed in the literature, this section only covers those that were used in **Paper I-III**. For illustrative purposes, mathematical reasoning and schematic flowcharts will be actively employed to reflect the core idea behind each approach.

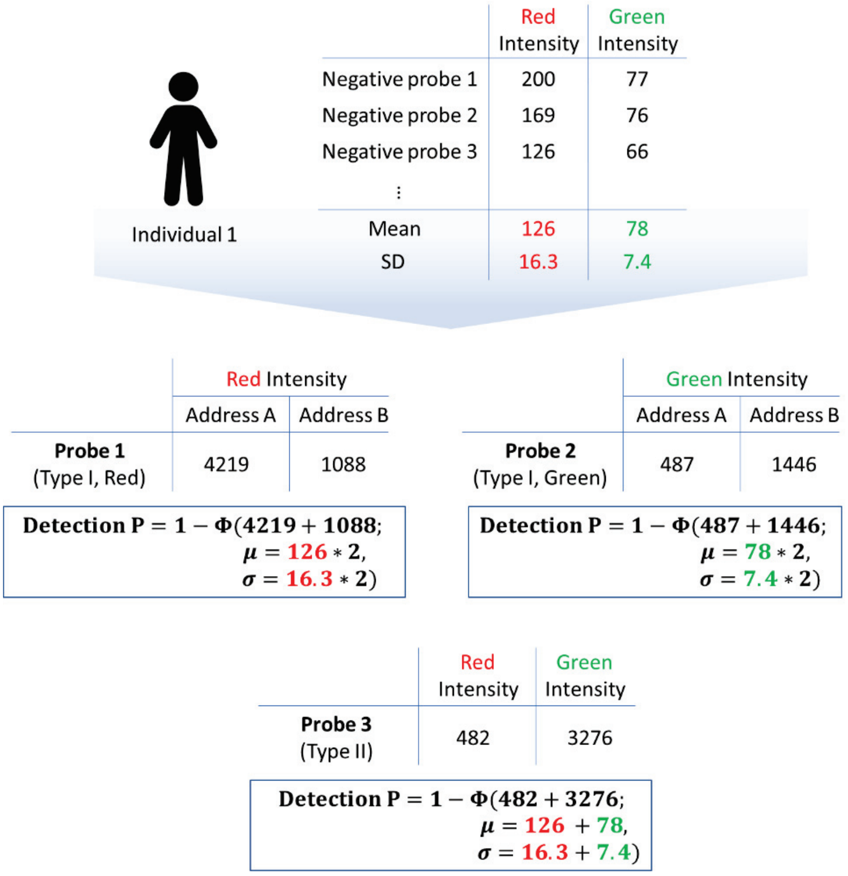
Furthermore, this section describes the methodological underpinnings of telomere length measurement (Southern blot analysis of terminal restriction fragment and quantitative polymerase chain reaction). Lastly, I elaborate on penalized linear regression and multiple testing that were mainly used in **Paper I-III**.

This section does not repeat the descriptions of the study populations used in **Paper I-III**. Relevant information can be found in the Methods section of each paper.

4.1 Quality control for microarray DNAm data

4.1.1 Detection p-values and exclusion criteria for samples and probes

Quality control in microarray-derived DNAm data is necessary to assess the reliability of each data point [185]. It starts with using a detection p-value to evaluate whether the total fluorescence intensity ($T = M + U$, Section 2.4.4) at each probe in each sample is strong enough, i.e., whether it deviates enough from the Gaussian distribution of noise signals. **Figure 4** displays how the `detectionP` function from the minfi R package computes the detection p-values for the Type I and Type II probes in an individual. First, `detectionP` estimates the two parameters, mean and standard error, for the Gaussian distribution of noise signals emitted from negative control probes. Second, `detectionP` calculates $\Pr(X > t)$, where $X \sim N(\hat{\mu}, \hat{\sigma}^2)$ and t is a realized value of the total intensity (T).



	Red Intensity	Green Intensity
Negative probe 1	200	77
Negative probe 2	169	76
Negative probe 3	126	66
⋮		
Mean	126	78
SD	16.3	7.4

	Red Intensity	
	Address A	Address B
Probe 1 (Type I, Red)	4219	1088

$$\textbf{Detection P} = 1 - \Phi(4219 + 1088;\ \mu = 126 * 2,\ \sigma = 16.3 * 2)$$

	Green Intensity	
	Address A	Address B
Probe 2 (Type I, Green)	487	1446

$$\textbf{Detection P} = 1 - \Phi(487 + 1446;\ \mu = 78 * 2,\ \sigma = 7.4 * 2)$$

	Red Intensity	Green Intensity
Probe 3 (Type II)	482	3276

$$\textbf{Detection P} = 1 - \Phi(482 + 3276;\ \mu = 126 + 78,\ \sigma = 16.3 + 7.4)$$

Figure 4. Calculation of detection p-values for Type I and Type II probes in an individual.

The `DetectionP` function from the minfi R package repeats this procedure across all probes and all individuals. The number used in this figure has been selected arbitrarily for illustrative purposes.

Data points with detection p-values>0.01, i.e., the total intensities are as weak as noise signals, are left out in the resulting DNAm data. Moreover, samples or probes with a high proportion of large detection p-values are also excluded from the data. Here, it is worth noting that laboratory settings, e.g., the kit used for bisulfite conversion and the type of scanning instrument used for measurement, can influence the detection p-values and probe/sample exclusions.

A quality control (QC) process also filters out probes near SNPs and those with low bead counts or cross-hybridization [186]. This is because the fluorescence intensities measured at these probes are unlikely to reflect methylation levels but rather genotypic variant callings or technical noises. Additionally, samples with sex mismatch, which are determined using the methylation signals from the sex chromosomes, are excluded. Relevant information about these probes can be obtained through the minfi-compatible packages[5] derived from the released manifest file from Illumina (e.g., the RnBeads.hg19 annotation in the case of the RnBeads package [130]).

4.1.2 Background correction and normalization of DNAm data

The next step after the exclusion of probes and samples is background correction and normalization. Although these two terms are often used interchangeably in the literature, they are different procedures. Background correction is for minimizing background noise in intensities by calculating conditional expectations, whereas normalization is for reshaping the distributions of intensities across samples. However, the conceptual difference between these two terms does not necessarily mean that they are mutually exclusive. For example, **in Paper I**, we removed background noise using normal-exponential out-of-band probes (Noob) and then applied Beta-

[5] IlluminaHumanMethylation450kanno.ilmn12.hg19 or IlluminaHumanMethylationEPICanno.ilm10b4.hg19

mixture quantile dilation (BMIQ) to the background corrected beta values in the RnBeads package [130].

4.1.2.1 Normal-exponential out-of-band probes (Noob)

The Noob method is to correct for background noise by deriving conditional expectations of true biological signals given the observed signals. Although Triche *et al.* [187] is widely cited for this Noob method, an earlier publication by Xie *et al.* [188] had already provided all the details of the underlying mathematical framework, including the derivation of conditional expectations and estimation of parameters. Hence, the mathematical elaborations in the following paragraphs will employ the notations of Xie *et al.* [188].

First, the observed intensity at the ith locus is assumed to be the sum of the true signal and background noise.

$$X_i = S_i + B_i,$$

$$where\ S_i \sim Exp(\alpha)\ and\ B_i{\sim}N(\mu, \sigma^2)$$

Then, the observed signal at the jth negative control probe is assumed to be identical to the background noise.

$$X_{0j} = B_{0j}\ ,$$

$$where\ B_{0j}{\sim}N(\mu, \sigma^2)$$

Again, as mentioned above, the purpose of the Noob normalization is to find the conditional expectation of the true intensity given the observed intensity, $E(S_i|X_i)$. To do this, the conditional probability density function of S_i given X_i, $f(s_i|x_i)$, is defined as follows:

$$f(s_i|x_i;\mu,\sigma^2,\alpha) = \frac{f(s_i, x_i;\ \mu,\sigma^2,\alpha)}{f(x_i;\ \mu,\sigma^2,\alpha)}$$

To derive the numerator, the joint probability density function of S_i and B_i is defined as follows:

$$f(s_i, b_i;\ \mu,\sigma^2,\alpha) = \frac{1}{\alpha}\exp\left(-\frac{s_i}{\alpha}\right) * \phi(b_i;\ \mu,\sigma^2)\ \because S_i \perp B_i$$

Then, the variable transformation of $B_i = X_i - S_i$ is applied. Here, the Jacobian is $\frac{dB_i}{dX_i} = 1$

$$\begin{aligned} f(s_i, x_i;\mu,\sigma^2,\alpha) &= \frac{1}{\alpha}\exp\left(-\frac{s_i}{\alpha}\right) * \frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{(x_i - s_i - \mu)^2}{2\sigma^2}\right\} \\ &= \frac{1}{\alpha}\exp\left(\frac{\sigma^2}{2\alpha^2} - \frac{x-\mu}{\alpha}\right) * \phi\left\{s_i;\ x_i - \mu - \frac{\sigma^2}{\alpha}, \sigma^2\right\} \end{aligned}$$

Next, we integrate $f(s_i, x_i;\mu,\sigma^2,\alpha)$ over s_i to obtain $f(x_i;\mu,\sigma^2,\alpha)$.

$$\begin{aligned} f(x_i;\mu,\sigma^2,\alpha) &= \int_0^{+\infty} f(s_i, x_i;\ \mu,\sigma^2,\alpha)ds_i \\ &= \frac{1}{\alpha}\exp\left(\frac{\sigma^2}{2\alpha^2} - \frac{x_i-\mu}{\alpha}\right) * \left\{1 - \Phi\left(0;\ x_i - \mu - \frac{\sigma^2}{\alpha}, \sigma^2\right)\right\} \end{aligned}$$

Therefore, the conditional probability function of S_i given X_i is as follows:

$$\begin{aligned} f(s_i|x_i;\ \mu,\sigma^2,\alpha) &= \frac{\phi\left(s_i;\ x_i - \mu - \frac{\sigma^2}{\alpha}, \sigma^2\right)}{1 - \Phi\left(0;\ x_i - \mu - \frac{\sigma^2}{\alpha}, \sigma^2\right)} \\ &= \frac{\phi(s_i;\ \mu_{sx}, \sigma^2)}{1 - \Phi(0; \mu_{sx}, \sigma^2)}\ , where\ \mu_{sx} = x_i - \mu - \frac{\sigma^2}{\alpha} \end{aligned}$$

Based on the conditional probability function of S_i given X_i, the corresponding conditional expectation is as follows:

$$E(S_i|X_i) = \int_0^{\infty} s_i\, f(s_i|x_i;\ \mu, \sigma^2, \alpha) ds_i$$

$$= \frac{1}{1 - \Phi(0;\ \mu_{sx}, \sigma^2)} * \int_0^{+\infty} s_i \phi(s_i;\ \mu_{sx}, \sigma^2) ds_i$$

$$= \frac{1}{1 - \Phi(0;\ \mu_{sx}, \sigma^2)}$$

$$* \left\{ \int_0^{+\infty} (s_i - \mu_{sx})\ \phi(s_i;\ \mu_{sx}, \sigma^2) ds_i + \int_0^{+\infty} \mu_{sx}\ \phi(s_i;\ \mu_{sx}, \sigma^2) ds_i \right\}$$

$$= \frac{-\sigma^2}{1 - \Phi(0;\ \mu_{sx}, \sigma^2)} * \int_0^{+\infty} \frac{-(s_i - \mu_{sx})}{\sigma^2}\ \phi(s_i;\ \mu_{sx}, \sigma^2) ds_i\ + \mu_{sx}$$

$$= \frac{-\sigma^2}{1 - \Phi(0;\ \mu_{sx}, \sigma^2)} * \left\{ \phi\left(+\infty;\ \mu_{sx}, \sigma^2\right) - \phi(0;\ \mu_{sx}, \sigma^2) \right\} + \mu_{sx}$$

$$= \frac{\sigma^2 \phi(0;\ \mu_{sx}, \sigma^2)}{1 - \Phi(0;\ \mu_{sx}, \sigma^2)} + \mu_{sx},$$

where $\mu_{sf} = x_i - \mu - \frac{\sigma^2}{\alpha}$, and $\phi(\cdot)$ and $\Phi(\cdot)$ are the probability density functions and the cumulative density functions of a normal distribution with a mean and variance, respectively.

The next step is to estimate the distribution parameters (μ, σ^2, and α) used in the $E(S_i|X_i)$. Two components are required for this: 1) negative control data and 2) an estimation method. The Noob normalization method employs the out-of-band intensities for negative control. The out-of-band intensities, a type of negative control from Type I probes, are the intensities of the fluorescence opposite to that which each bead pair is supposed to exhibit. For example, the bead pair for cg00050873, introduced in Section 2.4.4, is supposed to exhibit a red fluorescence because the next base is adenine, labeled red. Nevertheless, we can still retrieve the intensities of green fluorescence from this bead pair, and these out-of-band intensities can be used to estimate the distribution parameters (μ and σ^2) for the background noise.

The `preprocessNoob` function in the minfi package employs the non-parametric approach introduced by Xie *et al.* [188]. First, it estimates the parameters as follows:

$$\hat{\mu} = \bar{X}_0 = \sum_{j=1}^{J} X_{0j}/J \, ,$$

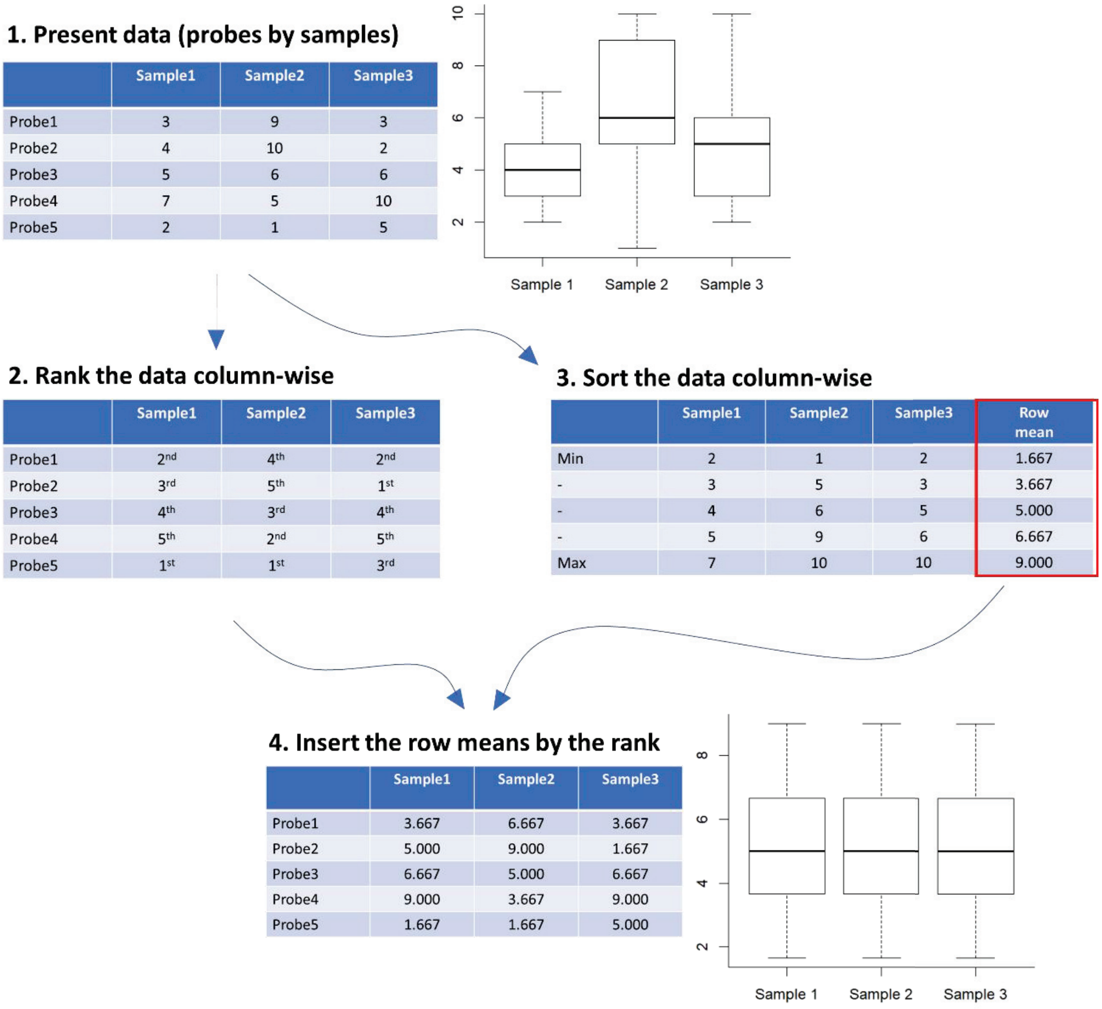
$$\hat{\sigma}^2 = \frac{\sum_{j=1}^{J} \left(X_{0j} - \bar{X}_0\right)^2}{J-1},$$

$$\hat{\alpha} = \max(\bar{X} - \bar{X}_0, 10), \text{where } \bar{X} = \sum_{i=1}^{I} X_i/I$$

Based on these estimated parameters, the conditional expectation of the true intensity, given the observed intensity, is calculated separately for the red and green fluorescence.

4.1.2.2 Quantile normalization

The quantile normalization [189] starts with the simple idea of forcing the same distribution of intensities across samples. **Figure 5** details the procedure of the quantile normalization. Step 1 is a restructuring of the data by placing samples in columns and probes in rows. Step 2 is determining a rank for each column. Step 3 is sorting the intensities from lowest to highest for each column and calculating the mean for each row (as red-boxed in **Figure 5**). Step 4 is assigning the row means according to the ranks calculated in Step 2 above.



1. Present data (probes by samples)

	Sample1	Sample2	Sample3
Probe1	3	9	3
Probe2	4	10	2
Probe3	5	6	6
Probe4	7	5	10
Probe5	2	1	5

2. Rank the data column-wise

	Sample1	Sample2	Sample3
Probe1	2nd	4th	2nd
Probe2	3rd	5th	1st
Probe3	4th	3rd	4th
Probe4	5th	2nd	5th
Probe5	1st	1st	3rd

3. Sort the data column-wise

	Sample1	Sample2	Sample3	Row mean
Min	2	1	2	1.667
-	3	5	3	3.667
-	4	6	5	5.000
-	5	9	6	6.667
Max	7	10	10	9.000

4. Insert the row means by the rank

	Sample1	Sample2	Sample3
Probe1	3.667	6.667	3.667
Probe2	5.000	9.000	1.667
Probe3	6.667	5.000	6.667
Probe4	9.000	3.667	9.000
Probe5	1.667	1.667	5.000

Figure 5. Procedure of the quantile normalization.

The `preprocessQuantile` function [190] from the minfi package performs the quantile normalization in each of multiple subsets of methylated and unmethylated intensities data, separately. The subgroup criteria are the type of chromosome (autosomal or sex), the sex of the sample, and probe type (**Figure 6**). The yellow boxes in **Figure 6** indicate all the subgroups where the `preprocessQuantile` function performs the quantile normalization. The same subsetting rule is applied to the data of unmethylated intensities. Further details can be found in Hansen and colleagues' GitHub repository for `preprocessQuantile` [191].

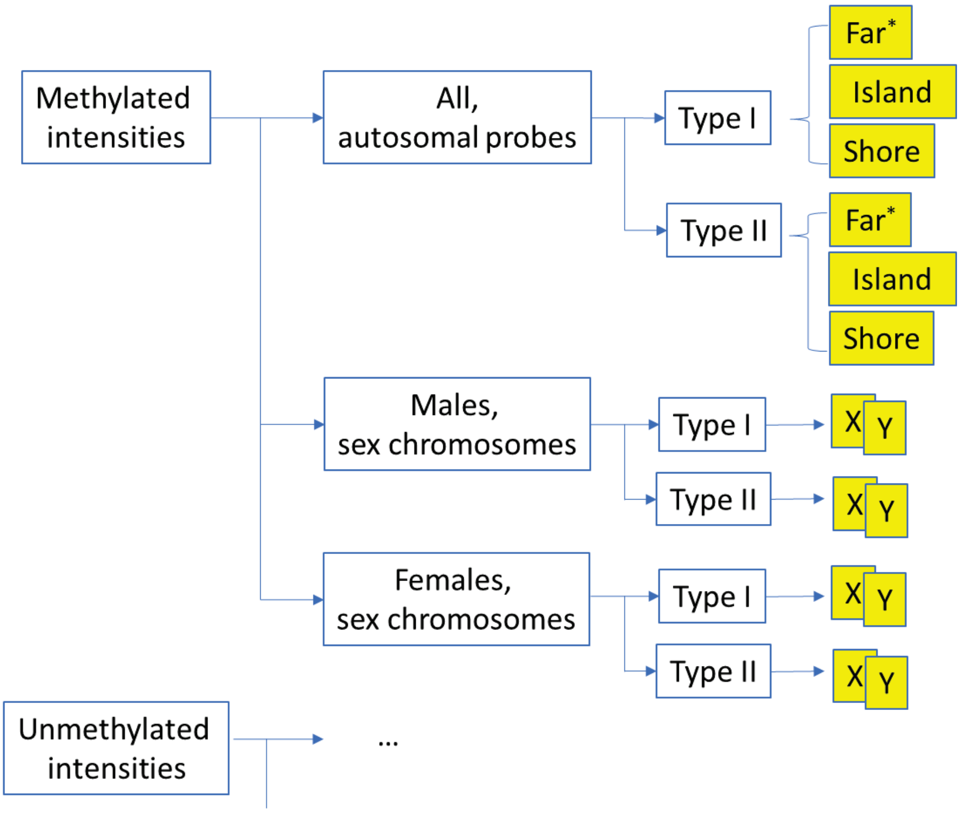


Figure 6. Subgroups for the quantile normalization.
[*] These subsets include probes in the shelf and open sea regions. The subgrouping procedure for the methylated intensities is identical to that of the unmethylated intensities.

4.1.2.3 Subset-quantile Within Array Normalization (SWAN)

Maksimovic *et al.* [192] scrutinized the different proportions of the two types of probes found in CGIs. They found 57% of Type I probes and 21% of Type II probes in the CGIs, which indicates that the distribution of the intensities from the two types of probes differs substantially (please refer to Figure 1 in Maksimovic *et al.* [192]). To address this point, the Subset-quantile Within Array Normalization (SWAN) method – the `preprocessSWAN` from the minfi package – has been developed to allow both types of probes on a single array to be normalized together. The procedure starts with classifying probes into several subgroups according to the type of probes and the number of CpG sites underlying the 50 base pair probe body. **Table 2** shows the number of probes in each subgroup. From each subgroup, SWAN randomly selects the same number of probes (which is the minimum number of probes in the red-colored subgroups in **Table 2**; 11,303 in case of HM450) so that the intensities from the selected Type I and Type II probes have similar distributions (please refer to Figure 2 in Maksimovic *et al.* [192]).

Table 2. The number of probes according to the type of probe (Type I and Type II) and the number of CpG sites in the probe body on the Illumina HumanMethylation450 Beadchip.

	The number of CpG sites in the probe body							
Probe Type	0	1	2	3	4	5	…	Total
Type I	0	11,303*	26,201*	36,401*	33,398	17,519		135,476
Type II	151,164	111,590*	61,313*	23,362*	2,607	0		350,036

* SWAN randomly selects 11,303 probes out of each of the red-colored cells.

1. Present data

		Rank in a scale of 0-1		Sample 1 Methylated Intensity
Type I	probe1	0.11		5
	probe2	0.00	Subset!	1
	probe3	0.67		13
	probe4	0.22	Subset!	7
	probe5	0.89	Subset!	18
	probe6	0.33		8
	probe7	0.56	Subset!	11
	probe8	0.78		16
	probe9	1.00	Subset!	20
	probe10	0.44		9
Type II	probe1	0.93		19
	probe2	0.29		5
	probe3	0.57	Subset!	12
	probe4	0.14		3
	probe5	1.00	Subset!	20
	probe6	0.36		6
	probe7	0.00	Subset!	1
	probe8	0.79		16
	probe9	0.50		11
	probe10	0.21		4
	probe11	0.64		13
	probe12	0.71	Subset!	15
	probe13	0.86		18
	probe14	0.07	Subset!	2
	probe15	0.43		10

2. Subset

Sample 1 Type I Intensity	Sample 1 Type II Intensity
1	12
7	20
18	1
11	15
20	2

3. Sort in each column and obtain row means and rank in a scale of 0-1.

Sample 1 Type I Intensity	Sample 1 Type II Intensity	Row mean	Rank in a scale of 0-1
1	2	1.5	0
7	12	9.5	0.25
11	15	13	0.5
18	16	17	0.75
20	20	20	1

4. Plot the row means by the rank and interpolate

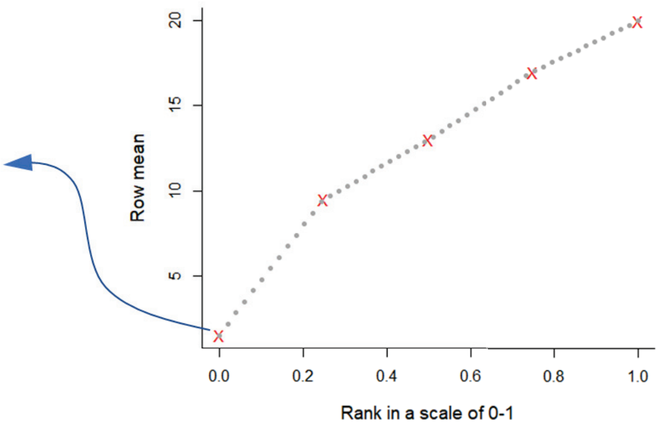


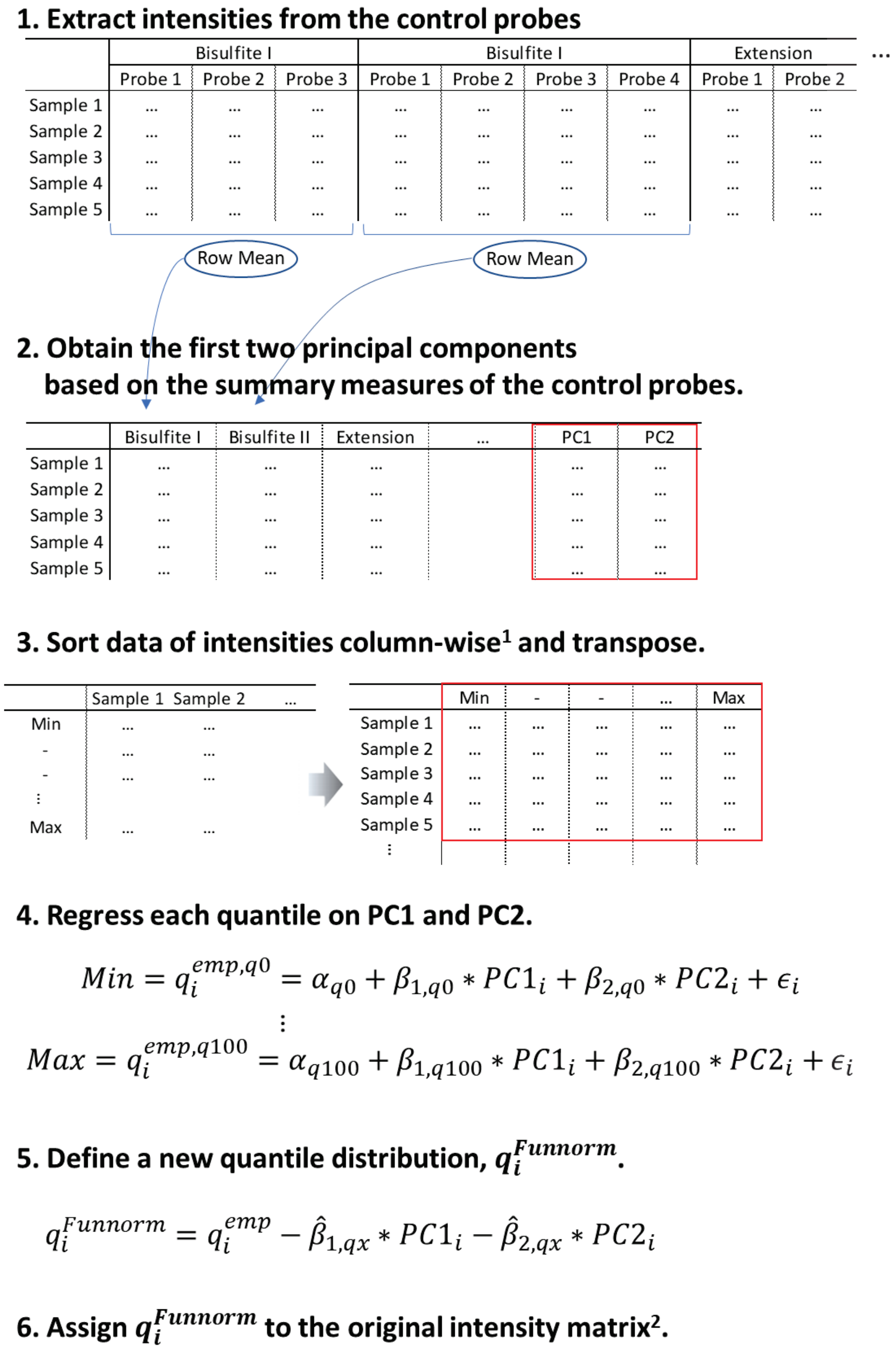
Figure 7. Details of the normalization procedure of SWAN.
This figure showcases a scale-down procedure of SWAN. In practice, the random selection in step 1 occurs on a much larger scale (e.g., 11,303*3 for each probe type).

The next step is to sort the intensities from the selected Type I and Type II probes and create a dataset, as illustrated by Step 3 in **Figure 7**. Based on this dataset, SWAN calculates the row means and ranks on a scale of 0-1 (Step 4 in **Figure 7**) and interpolates between the row means (Step 4 in **Figure 7**). The final step is to assign the row means resulting from the interpolation to the original data according to the pre-calculated ranks on a scale of 0-1. This procedure is applied to the methylated and unmethylated intensities for each sample separately.

4.1.2.4 Functional normalization

The functional normalization procedure proposed by Fortin *et al.* [193] aims to minimize unwanted technical variation across samples. This technical variation is commonly referred to as 'batch effects' that may confound the variables of interest and increase the chance of false-positive findings [194, 195]. To address these issues, the functional normalization extends the quantile normalization (Step 3 in **Figure 5**) by adjusting the empirical quantile distribution for the covariates that may explain the technical variation.

Details of the functional normalization are provided in **Figure 8**. First, the functional normalization extracts background-corrected intensities (by Noob) from the control probes (details are provided in the Supplementary material in Fortin *et al.* [193]) and out-of-band probes (also described in Section 4.1.2.1), and computes summary measures, e.g., row means. Second, it applies principal component analysis to the computed summary measures and selects the first two principal components (PCs). Third, it obtains an empirical quantile distribution by sorting methylated (or unmethylated) intensities within each sample, in the same way as the quantile normalization described in Section 4.1.2.2. Fourth, it regresses each quantile (from the minimum to the maximum) on the first two PCs, i.e., PC1 and PC2. Fifth, it computes the adjusted quantile distribution by subtracting the effect of PCs from the original quantile distribution. Finally, it assigns the derived adjusted (functional-normalized) quantile distribution to the initial intensity matrix, in the same way as the quantile normalization.



1. Extract intensities from the control probes

	Bisulfite I			Bisulfite I				Extension		…
	Probe 1	Probe 2	Probe 3	Probe 1	Probe 2	Probe 3	Probe 4	Probe 1	Probe 2	
Sample 1	…	…	…	…	…	…	…	…	…	
Sample 2	…	…	…	…	…	…	…	…	…	
Sample 3	…	…	…	…	…	…	…	…	…	
Sample 4	…	…	…	…	…	…	…	…	…	
Sample 5	…	…	…	…	…	…	…	…	…	

2. Obtain the first two principal components based on the summary measures of the control probes.

	Bisulfite I	Bisulfite II	Extension	…	PC1	PC2
Sample 1	…	…	…		…	…
Sample 2	…	…	…		…	…
Sample 3	…	…	…		…	…
Sample 4	…	…	…		…	…
Sample 5	…	…	…		…	…

3. Sort data of intensities column-wise[1] and transpose.

	Sample 1	Sample 2	…
Min	…	…	
-	…	…	
-	…	…	
⋮			
Max	…	…	

	Min	-	-	…	Max
Sample 1	…	…	…	…	…
Sample 2	…	…	…	…	…
Sample 3	…	…	…	…	…
Sample 4	…	…	…	…	…
Sample 5	…	…	…	…	…
⋮					

4. Regress each quantile on PC1 and PC2.

$$Min = q_i^{emp,q0} = \alpha_{q0} + \beta_{1,q0} * PC1_i + \beta_{2,q0} * PC2_i + \epsilon_i$$

$$\vdots$$

$$Max = q_i^{emp,q100} = \alpha_{q100} + \beta_{1,q100} * PC1_i + \beta_{2,q100} * PC2_i + \epsilon_i$$

5. Define a new quantile distribution, $q_i^{Funnorm}$.

$$q_i^{Funnorm} = q_i^{emp} - \hat{\beta}_{1,qx} * PC1_i - \hat{\beta}_{2,qx} * PC2_i$$

6. Assign $q_i^{Funnorm}$ to the original intensity matrix[2].

Figure 8. The workflow of the functional normalization procedure.

[1] This process is identical to the quantile normalization but targets 500 randomly selected probes here. [2] This step employs a linear interpolation to fill the gap between the 500 probes selected in the earlier step.

4.1.2.5 Beta-mixture quantile dilation (BMIQ)

BMIQ proposed by Teschendorff *et al.* [196] focuses on the different CpG densities between the Type I and Type II probe, similar to SWAN (Section 4.1.2.3). While SWAN attempts to find a common quantile distribution between the two types of probes, BMIQ modifies the distribution of beta values (ranges 0-1) from the Type II probes based on the beta values from the Type I probes. Each step of the modification procedure is explained in great detail in the Methods section in Teschendorff *et al.* [196]. Hence, this section only outlines the fundamental ideas of BMIQ.

The distributional modification of BMIQ starts with modeling a pair of three-state, unmethylation (U), hemimethylation (H) and methylation (M), beta mixtures: one for the beta values from the Type I probes and the other for those from the Type II probes.

$$p(\beta^{Type}) = \pi_U^{Type} Beta\left(\beta; a_U^{Type}, b_U^{Type}\right) + \pi_H^{Type} Beta\left(\beta; a_H^{Type}, b_H^{Type}\right) \\ + \pi_M^{Type} Beta\left(\beta; a_M^{Type}, b_M^{Type}\right),$$

where $Beta$ is the probability density function of a beta distribution with the two scale parameters a and b, and $Type$ denotes the type of probes, $Type \in (I, II)$. $\pi_U^{Type}, \pi_H^{Type}$ and π_M^{Type} are the probabilities that a beta value belongs to the respective beta distributions. When we define $s \in (U, H, M)$, the parameters, π_s^{Type}, a_s^{Type} and b_s^{Type} are estimated using the expectation-maximization (EM) algorithm [197].

The next step is to transform the beta values from the Type II probes in the U or M states to corresponding quantiles using the cumulative density function of the beta mixtures. First, BMIQ calculates a cumulative density of a beta value, denoted as β_0. If the beta value belongs to the U state, the cumulative density will be $p = \mathrm{F}(\beta_0; \alpha_U^{II}, \beta_U^{II}) = \int_0^{\beta_0} Beta(\beta; \alpha_U^{II}, \beta_U^{II}) d\beta$. Then, it derives a corresponding quantile (q) for p using the inverse cumulative density function of the beta values from the Type I probes, i.e., $q = F^{-1}(p; \alpha_U^I, \beta_U^I)$. The same procedure is applied to the beta values from the Type II probes in the M state, but the cumulative density function must be $1 - F$ rather than F in this case.

The final step is to modify the beta values from the Type II probes in the H state using a dilation transformation. An important point here is that the beta values in the H state have to fit between the minimum of the transformed beta values from the Type II probes in the M state and the maximum of the transformed beta values from the Type II probes in the U state.

4.2 Telomere length measurement

4.2.1 Southern blot analysis of terminal restriction fragment lengths

Southern blot analysis of TRFs [198], simply known as Southern blot, was the primary method for measuring LTL in **Paper III**. This method is a multi-step procedure that encompasses a check for DNA integrity, digestion of DNA, agarose gel electrophoresis of the resulting fragments, transfer of DNA to a membrane, hybridization, and X-ray chemiluminescence for band visualization. The following bullet points will describe the details of each step.

1. DNA extraction and evaluation

DNA was extracted from biospecimens using two methods: 1) phenol-chloroform organic extraction, and 2) a commercially available kit based on salting-out. Kimura *et al.* [198] preferred the commercially available DNA kit (Gentra Puregene DNA extraction kit) to the phenol-chloroform organic extraction.[6] DNA integrity was checked by running samples side by side on an agarose gel. An intact DNA sample appears as a single compact band/crown, whereas degraded DNA exhibits a fuzzy crown that is shifted forward and is often accompanied by a long smear.

2. Digestion of genomic DNA

This step is to cleave DNA into chromosomal fragments and telomeric repeats by using different restriction endonucleases (enzymes that cut DNA at specific nucleotide sequences). Kimura *et al.* [198] primarily used two combinations of restriction enzymes: 1) HphI/MnlI and 2) HinfI/RsaI. HphI/MnlI cleaves DNA within a subtelomeric region, whereas HinfI/RsaI cuts DNA upstream of the subtelomeric region.

3. Using DNA molecular weight ladders to gauge fragment length

Molecular weight (MW) ladders are used to determine the approximate size of DNA fragments. Kimura *et al.* [198] used two commercially available MW ladders: a 1-kb ladder ranging from 0.5 to 12 kb and a collection of λ DNA fragments digested with HindIII that span 1.25-23.1 kb.

4. Agarose gel electrophoresis

Agarose gel electrophoresis, the core technique behind Southern blotting, is designed to separate DNA fragment according to their size (or length). First, the DNA fragments of an individual are pipetted into a well located at one edge of the gel. A typical gel contains 30 wells and can thus assay 30 individuals on a single electrophoretic run. Second, an electric current flowing from the wells (negative terminal) to the opposite edge (positive terminal) causes the DNA fragments to migrate through the gel at a specific rate. Electrophoresis is stopped when the band with the lowest

[6] Chloform-phenol extraction is labor-intensive, difficult to scale up, and requires a chemical fume hood because both phenol and chloroform are hazardous compounds.

MW has reached the bottom of the gel. Because they are negatively charged in different amounts due to their length, shorter DNA fragments travel further down the gel than longer ones.

5. Probe design and labeling

To distinguish between telomeric repeats and chromosomal bodies, Southern blot uses a probe consisting of three oligonucleotide repeats that are complementary to the telomeric repeats $(TTAGGG)_3$. The probe, often referred to as the telomere probe, is also labeled with digoxigenin (DIG) at the 3' end. This probe can be identified with anti-DIG-AP antibody and chemiluminescence after binding to the telomeric DNA fragments.

6. Analysis of the X-ray film

The optical density (OD) signal is extracted from the digitalized image of the X-ray film, for each vertical position (please refer to Figure 7a and Figure 8 in [198]). The mean TRF is defined as $\sum_i(OD_i)/\sum_i(OD_i/MW_i)$. This analysis is implemented in the ImageQuant software.

The next sub-section describes the other method for measuring LTL, quantitative Polymerase Chain Reaction (qPCR), that was used in **Paper III**.

4.2.2 Quantitative Polymerase Chain Reaction (qPCR)

qPCR was used to measure LTL in the Lothian Birth Cohorts of 1921 and 1936 (LBC1921 and LBC1936) that were included in **Paper III** [199]. This method quantifies relative LTL by comparing the copy number of telomere repeats with that of a single copy (or reference) gene.

LBC1921 and LBC1936 employed glyceraldehyde 3-phosphate dehydrogenase, abbreviated as GAPDH, as the reference gene.

The method implements two qPCR: one with telomere (T) primer pairs and the other with the single-copy gene (S) primer pairs. Apart from the difference in the primer pairs, the two qPCR are identical. Each qPCR obtains the C_t, i.e., the number of cycles at which the fluorescence from amplified DNA crosses a threshold. Based on the derived C_t, the method determines the relative telomere to single-copy gene (T/S) ratio.

4.2.3 Other methods for measuring TL

The quantitative fluorescent in situ hybridization (Q-FISH) measures TL by quantifying the fluorescence intensity from a hybridized peptide nucleic acid (PNA) oligonucleotide probe, i.e., $(CCCTAA)_3$ [200]. The PNA probe binds to the telomeric repeats more strongly than the complementary sequences in DNA or RNA, because the backbone of the PNA probe lacks charged phosphate groups [200]. Once hybridization is complete, an image of a metaphase chromosome spread and that of telomeric repeats are obtained and processed using an image analysis software [201]. Q-FISH has various modified versions, e.g., interphase [202], high-throughput [203], flow cytometric FISH [204], and metaphase Q-FISH [202]. Q-FISH can measure TL at both ends of each chromosome in cells (note that the substrate here is not DNA but cells) [205, 206]. However, like Southern blot, Q-FISH is labor-intensive and is unable to detect short telomeres that are below the threshold of the PNA probe. This can lead to false positives due to the binding of the PNA probe to interstitial telomeric sequences [205, 207, 208].

Single Telomere Length Analysis (STELA) is a ligation-based method involving PCR amplification and Sothern blot analysis [209]. A 'telorette[7]' is ligated to digested (by MseI) telomeric regions from a subset of chromosomes (XpYp, 2p, 11q, 12q, and 17p) [209, 210]. While TRF and qPCR are designed to measure average TL, this method can measure short telomeres on specific chromosomes. The Universal STELA is an improved version of STELA, enabling measurements on all the chromosomes. Although STELA does not require a large amount of DNA, it is labor-intensive, limited in measuring long telomeres (>8 kilobases) [210, 211], and particularly vulnerable to interstitial telomeric sequences (ITSs) [212].

Telomere Shortest Length Assay (TeSLA) is also a ligation-based method followed by PCR amplification and TRF analysis for all chromosomes, similar to the Universal STELA, but with improved specificity and sensitivity for TL measurement [212]. TeSLA uses a combination of four restriction enzymes (*BfaI/CviAII/MseI/NdeI* that minimizes subtelomeric regions) and newly-designed telomerettes, i.e., terminal adaptors (TeSLA-T 1 to 6), which ligate to the 3' C-rich strand with increased specificity. Two double-stranded adapters containing 5' AT or TA overhangs, C3 spacers, and AP primers increase the efficiency of ligation and the specificity of PCR amplification.

TeSLA is a highly attractive method because it measures the distribution of the shortest telomeres across all chromosomes [205, 212]. This is particularly useful when the focus of a study is to compare telomere attrition across groups of individuals, where it is important to determine the average length of the shortest telomeres and not just the average length of telomeres at the 96 ends of all chromosomes. The coverage of TL measurement is 1 to 18 kb. Although the method does

[7] An annealing linker comprising seven oligonucleotides complementary to TTAGGG followed by 20 bases non-complementary to the 3' G-rich overhang.

not misclassify the ITSs as telomeric repeats and does not require a large amount of DNA (less than one microgram), it is labor-intensive and costly [205]. In addition, TeSLA cannot reliably measure exceedingly long telomeres (>18 kb), such as those from inbred strains.

4.3 Statistical analyses

4.3.1 Penalized linear regression

Penalized linear regression (PLR) has been the main statistical method behind the development of epigenetic clocks. The main reason for its popularity is that PLR readily fulfills the statistical aim of epigenetic clocks, i.e., the development of a model that is most predictive of an outcome (chronological age, phenotypic age, or gestational age) based on a large number of predictors (>450,000 CpG sites). Among the many CpG sites, PLR automatically selects a subset of CpG sites and determines a coefficient for each CpG site so that their linear combination predicts the outcome of interest accurately. The end product of PLR is nothing but a linear regression equation, but its distinction lies in the automatic selection of variables and determination of coefficients.

PLR enables an automatic selection of predictors by introducing a constraint in its log-likelihood function as follows:

$$\sum_{i}^{n} l(Y_i, \boldsymbol{X_i\beta}) \ such\ that \ \left(\frac{1-\alpha}{2}\right) * \sum_{j}^{p} \beta_j^2 + \alpha * \sum_{j}^{p} |\beta_j| < \ t,$$

Where l is the negative log-likelihood function, e.g., this is $\frac{1}{2}(Y_i - \boldsymbol{X_i\beta})^2$ for the Gaussian case, $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_p)^T$ is a vector of coefficients, Y_i is an outcome of interest in the ith sample, $\boldsymbol{X_i} =$

$\left(X_{i,1}, X_{i,2}, \dots, X_{i,p}\right)^T$ is a vector of p different exposures in the ith sample, and α is a mixture parameter. When the likelihood is maximized with respect to $\boldsymbol{\beta}$ in equivalent Lagrangian form,

$$\widehat{\boldsymbol{\beta}} = arg \min_{\boldsymbol{\beta}} \sum_{i}^{n} l(Y_i, \boldsymbol{X_i\beta}) + \lambda * \left\{ \left(\frac{1-\alpha}{2}\right) * \sum_{j}^{p} \beta_j^2 + \alpha * \sum_{j}^{p} |\beta_j| \right\}$$

The α parameter was set to be 0.5 in **Paper I** and **Paper II**, and the λ parameter was determined through 10-fold cross-validation. The 10-fold cross-validation was performed in a training set, and the final model with the determined λ was validated in the test set. Accuracy and precision metrics, e.g., MAD and the correlation coefficient between the observed and the predicted outcome, were calculated in this validation process. Along with scatterplots of the observed against the predicted outcome, the metrics are essential quantities for assessing the prediction power of the model.

4.3.2 Multiple testing – the Bonferroni correction

In **Paper III**, depending on the DNAm microarray used, roughly 450,000 to 850,000 DNAm-LTL associations were tested simultaneously. This vast number of hypotheses being tested simultaneously resulted in a substantial increase in the family-wise error rate (FWER, the probability of making at least one type I error). To control FWER at the nominal significance level of α, the Bonferroni correction was applied, which corresponds to a significance level of $\frac{\alpha}{m}$, where m is the number of independent hypothesis tests. Hence, the Bonferroni-controlled FWER can be mathematically expressed as

$$FWER = \Pr\left\{\bigcup_{i=1}^{m_0}\left(p_i < \frac{\alpha}{m}\right)\right\} \leq \sum_{i=1}^{m_0}\left\{\Pr\left(p_i \leq \frac{\alpha}{m}\right)\right\} \ by\ Boole's\ equality$$

$$= m_0 * \frac{\alpha}{m} \leq \alpha,$$

where p_i is the p-value from the ith hypothesis test, α is the significance level, m_0 is the number of true null hypotheses, and m is the total number of the hypotheses. Although the Bonferroni correction suppresses FWER at level α, it also raises the type II error and reduces statistical power. FDR was not used in **Paper III**, but this approach is a widely used alternative to the Bonferroni correction for multiple testing. These methodological considerations will be elaborated further in Section 6.3.8.

5 Summary of papers

5.1 Paper I

Three blood-based epigenetic clocks for human adults were developed based on EPIC-based DNAm data from MoBa and publicly available DNAm data from the Gene Expression Omnibus (GEO) repository.

First, an Adult Blood-based EPIC clock (ABEC) was trained on MoBa samples (n=1,592, age-span: 19 to 59 years). Although ABEC showed a high overall precision, with the Pearson correlation coefficient (r) between the observed and predicted age being above 0.93 in independent test sets, it still underestimated the age of individuals older than 45 years.

To address this shortcoming, an extended ABEC (eABEC) was trained on DNAm data from MoBa and GEO (n=2,227, age-span: 18 to 88 years). eABEC largely mitigated the underestimation of the age among the individuals older than 45 years and showed a slight but noticeable improvement in age prediction (r>0.94).

Furthermore, to examine whether the additional probes on EPIC improve age prediction of an epigenetic clock, a common ABEC (cABEC) was trained on the same training set as eABEC, but this time restricting the analysis to only those CpGs that were common to 450K and EPIC. The prediction performance of cABEC was as high as that of eABEC, indicating that the additional probes on EPIC did not improve age prediction.

5.2 Paper II

Three placental epigenetic clocks were developed, leveraging previous findings that placental DNAm changes widely with fetal gestational age.

A robust placental clock (RPC) was trained on placental DNAm data (n=1,102) from GEO. RPC was highly precise (r=0.99, MAD=0.96 weeks) in estimating the gestational age of fetuses from all trimesters regardless of their pregnancy conditions, e.g., preeclampsia, gestational diabetes, and trisomy 13, 18, or 22.

Further, a control placental clock (CPC) was developed based on “control” placental samples (i.e., those without any reported pregnancy complications, n=963) to assess whether adverse pregnancy conditions influence the epigenetic gestational age estimate. CPC was also highly predictive of the gestational age of pregnancies across all trimesters (r=0.98, MAD=1.02 weeks). However, it also revealed that pregnancy conditions, e.g., gestational diabetes, intrauterine growth restriction, trisomy 13, 16, 18, and 21, were not associated with faster or slower epigenetic aging of placentas.

The second version of RPC, referred to as the refined RPC, was trained on placental samples from “uncomplicated term” pregnancies (n=733), i.e., those without any known pregnancy complications and with gestational age>36 weeks. Although the training set for the refined RPC did not include any preterm births, the refined RPC showed high prediction power (r>0.98, MAD=1.49 weeks).

5.3 Paper III

We conducted a large-scale EWAS of LTL using seven large cohorts (n=5,713) – the Framingham Heart Study (FHS), the Jackson Heart Study (JHS), the Women’s Health Initiative (WHI), the

Bogalusa Heart Study (BHS), the Lothian Birth Cohorts (LBC) of 1921 and 1936, and the Longitudinal Study of Aging Danish Twins (LSADT).

Using a meta-analysis framework, we identified 823 CpG sites significantly associated (P<1E-7) with LTL after adjustment for age, sex, ethnicity, and imputed white blood cell counts. Functional enrichment analyses revealed that these CpG sites are located near genes known to play a key role in circadian rhythm, blood coagulation, and wound healing.

6 Discussion

6.1 Summary of key findings

The aims of this thesis were to 1) develop new epigenetic biomarkers of aging and growth with high precision, and 2) to conduct an epigenome-wide association study of LTL – a cellular replicative aging biomarker – using large-scale multi-ethnic cohorts. In response to the first aim, my colleagues and I developed the ABECs in **Paper I** and the placental clocks in **Paper II**, which, respectively, showed high precision in estimating the chronological age of adults and the gestational age of fetuses. The age prediction by the ABECs and placental epigenetic clocks was more precise than previously published epigenetic clocks. The EWAS conducted under the second aim of this thesis, and reported in **Paper III**, revealed that 823 CpG sites were associated with intrinsic LTL, i.e., LTL adjusted for age, sex, ethnicity, and cell composition.

6.2 Interpretations and implications of key findings

The ABECs in Paper I revealed that the inclusion of additional 413,743 probes unique to EPIC (including 226,915 specifically targeting regulatory regions such as DNase proximal/distal and FANTOM5) did not improve the precision of the epigenetic clock. It was shown by the analysis that eABEC (including 1,791 CpGs, 1,084 of which were only present on EPIC) and cABEC (1,892 CpGs that are present on both HM450 and EPIC) performed equally well when the sample size of the training set was the same. The additional probes on EPIC provided no advantage even in a reduced training set with fewer samples (Figure 4 in **Paper I**). This result reinforced the point made by Zhang *et al.* [213] that the sample size of a training set was critical in increasing the precision of an epigenetic clock and implied that the common probes shared by 450K and EPIC were sufficient in predicting human chronological age.

Another key implication stemming from **Paper I** is that the ABECs are novel blood-based epigenetic clocks trained on “EPIC-derived” DNAm data. The ABECs are expected to facilitate other EPIC-based research exploring the link between epigenetic age and complex human traits, given that EPIC has replaced HM450. The high precision of the ABECs is also attractive to other research fields, e.g., forensic research, where accurate age prediction is of particular importance.

The placental epigenetic clocks in Paper II were novel in their use of placental tissues. The gestational age estimated by the placental epigenetic clocks possibly reflected the maturity of the placenta and the status of fetal growth. Given that preterm placental aging is a recognized determinant of adverse pregnancy outcomes [214], it is vital to accurately assess the degree of placental epigenetic maturation and understand the biological mechanisms underlying fetal growth and placental maturation. In this context, the CpG sites included in CPC and RPC (trained on non-complicated pregnancies) may provide excellent starting points to guide more in-depth biological investigations.

Paper II also confirmed that the gestational age-related change in placental DNAm across the genome was sufficiently evident to allow the development of an epigenetic clock with high precision. Although Mayne *et al.* [117] reported a suboptimal epigenetic clock, it was unknown whether the performance of a placenta-based epigenetic clock could reach that of the cord blood-based epigenetic clock, e.g., Bohlin *et al.* [16] Cord Blood clock. The results in **Paper II** showed that the gestational age prediction of RPC in **Paper II** was as precise as that of the Bohlin *et al.* [16] Cord Blood clock[8].

The EWAS findings in Paper III provided a basis for a better understanding of the genetic influence on LTL, along with the genome-wide association study (GWAS) of LTL conducted by Codd *et al.* [184]. Rakyan *et al.* [215] and Verma [216] proposed that the integration of GWAS and EWAS can potentially explain the functional pathway from genetic variation to a phenotype of interest. For example, in our case, *RTEL1* appeared in the GWAS of LTL by Codd *et al.* [184] (rs755017, P=6.71E-09) as well as in our EWAS of LTL (cg00622799, cg03339910, cg10615591, and cg17534029, P<6.43E-07, Supplementary File 1 in **Paper III**). As suggested in Rakyan *et al.* [215], one can re-analyze the association between LTL and DNAm at the four CpGs in *RTEL1*, stratified by the genetic variants in *RTEL1*, in other cohorts with genotype and methylation data.

6.3 Potential limitations

6.3.1 Elastic net regression for epigenetic clocks

Elastic net regression, a type of penalized linear regression, has been used for developing most of the previously published epigenetic clocks (Section 2.4.5), the ABECs (**Paper I**), and placental epigenetic clocks (**Paper II**). Considering the widespread use of this method, it is worthwhile to discuss its strengths and weaknesses.

[8] Knight Cord Blood clock (introduced in Section 2.4.5) was developed to estimate gestational age, but the performance of this clock was inferior to the Bohlin Cord Blood clock.

Elastic net regression enables selecting a set of CpGs that is most predictive of chronological (or gestational) age from a large number of probes (>450,000) in microarray-based DNAm data [217, 218]. In this CpG selection process, elastic net regression allows a grouping effect[9], i.e., the combined effect of highly-correlated CpGs on aging [219]. Given that the methylation levels at neighboring CpGs are highly correlated, the grouped selection ability helps to increase the prediction accuracy [219]. However, it is worth noting that the coefficient estimates, i.e., the effect sizes of CpGs on aging, are neither completely unbiased nor statistically testable due to the penalty term included in the likelihood function [220, 221].

Additionally, the linearity assumption embedded in elastic net regression enables high interpretability [219, 222]. For example, a coefficient estimate corresponding to a CpG site is the expected change in age for every unit increase in the methylation level at the CpG site when the methylation levels at the other CpGs are fixed values. However, the linearity may not be flexible enough to reflect the relationship between CpGs and age [217]. The linearity implies that the methylation patterns change at a constant rate across the lifespan, but this may not be the case in real data. For example, to address the non-linearity problem, Horvath [6] transformed chronological age as follows:

$$F(Age) = \begin{cases} \log(Age+1) - \log(adult.age+1) & if\ Age \leq adult.age \\ \dfrac{(Age - adult.age)}{adult.age} + 1 & if\ Age > adult.age, \end{cases}$$

where adult.age refers to a specific threshold for adulthood. Horvath [6] then regressed the transformed chronological age on CpGs. According to his comments, this transformation was to

[9] Unlike elastic net regression, LASSO selects one CpG for each group of neighboring CpGs.

reflect the non-linear change in methylation levels throughout childhood, adolescence, and adulthood.

Further, linear models are likely to show a reduced prediction power compared to machine learning methods with extensive flexibility [222]. The reduced prediction power, in other words, indicates increased random errors in predicted (epigenetic) age, which is often referred to as underfitting. Along with measurement errors in a training set, the random errors by underfitting directly influence EAA, i.e., the residual from a regression of chronological age on epigenetic age. The increased random errors in EAA possibly result in reduced statistical power in an association study of EAA with a phenotype of interest. In this respect, more research is needed to distinguish random errors from putative biological signals in EAA.

6.3.2 Batch effect correction in training epigenetic clocks

Correctable batch effects in **Paper I** and **Paper II** were the potential technical variation in methylation across "cohorts", whereas batch effects often indicated technical variation across chips or plates[10]. This was because most of the publicly available DNAm data from the GEO repository did not include information about the chips or plates used. A relevant methodology for batch effect correction is the ComBat method [223] from the sva R package [224].

The training sets for the ABECs and placental epigenetic clocks were not corrected for batch effects. Rather, the raw beta values after background correction and normalization were used for training the epigenetic clocks. The main reason for this was the different (gestational) age

[10] In the case of EPIC, eight samples can be placed on a chip, and 12 chips can be mounted on a plate.

distributions across cohorts (please refer to Table 1 in **Paper I** and Table 1 in **Paper II**). As pointed out in previous publications [195, 225, 226], a batch correction might introduce a bias in the DNAm data when study samples were not evenly distributed across batches. A batch correction method might have removed the methylation difference across batches regardless of whether it is due to technical aspects or a genuine aging process.

6.3.3 Selection bias in the epigenetic clocks

The subjects from MoBa, specifically STudy of Assisted Reproductive Technology (START), in **Paper I** may introduce a selection bias. This is because MoBa is a pregnancy cohort with a small number of individuals aged 45 years and above [227]. The different age distribution of mothers and fathers and the different sex ratios in the older age groups possibly led to the suboptimal age prediction by ABEC (Figure 2 in **Paper I**). This weakness was addressed by adding publicly available DNAm data on older subjects from the study by Curtis *et al.* [228], which resulted in the development of eABEC. However, a selection bias may persist because all the mothers in MoBa were pregnant at the time of enrollment into the study. Additionally, the MoBa parents appeared to have healthier lifestyles, e.g., a lower rate of diabetes, hypertension, and smoking, compared to the general Norwegian population [229].

Gruzieva *et al.* [230] reported that 196 CpG sites showed longitudinal changes in DNAm level before pregnancy and in gestational weeks 10 to 15 (FDR<0.05). Although their sample size was small (n=21 Swedish women), some of the pregnancy-associated CpGs were still significant after the Bonferroni correction. However, it is arguable whether the pregnancy-associated CpGs lower the predictive power of the ABECs in **Paper I** directly. It is thus necessary to verify whether nonpregnant women epigenetically age slower or faster than pregnant ones, i.e., testing if epigenetic drift[11] can be seen at the pregnancy-associated CpGs (see Figure 2 in Teschendorff *et al.* [231]).

Given that the ABECs were trained on the MoBa subjects with healthy lifestyles, they may overestimate the chronological age, i.e., positive epigenetic age acceleration, of individuals with chronic diseases (data not shown). There is previous evidence of patients with cardiovascular disease or cancer showing a distinct epigenetic drift compared to healthy individuals [232, 233]. The Hannum *et al.* [5] Blood-based clock, Horvath [6] Pan-tissue clock, Levine *et al.* [8] PhenoAge clock, and the Lu *et al.* [143] GrimAge clock also reported positive EAA in the diseased subjects [60, 234].

However, such epigenetic clocks resulting in positive EAA in individuals with chronic diseases can be considered as desirable if the research purpose is to develop a valid biomarker of aging. The positive EAA in diseased subjects implies that the epigenetic clock captures functional decline accurately (satisfying the second condition for being a valid biomarker of aging, Section 2.3). This tendency is more apparent when epigenetic clocks are trained solely on healthy individuals (note that the Horvath [6] Pan-tissue clock was trained exclusively on healthy tissues).

With this in mind, **Paper II** included CPC and the refined RPC in addition to RPC. RPC showed high precision regardless of pregnancy conditions, provided that the clock was trained on control (uncomplicated) pregnancies as well as complicated pregnancies[12]. However, the high precision of RPC across all pregnancies indicated that RPC might not be successful in differentiating pregnancies with and without complications. To increase the predictability of pregnancy

[11] Epigenetic drift is a phenomenon whereby two groups (usually healthy and diseased) show different rates of methylation change according to chronological age.

[12] Anencephaly, chorioamnionitis, confined placental mosaicism, diandric triploid, gestational diabetes, in-vitro fertilization, intrauterine growth restriction, preeclampsia, spinal bifida, and trisomy 16.

complications, CPC was trained on the control (uncomplicated) term and preterm pregnancies, and the refined RPC was trained on the uncomplicated term (gestational age>36 weeks) pregnancies.

Nevertheless, the association test between gestational age acceleration (by CPC) and pregnancy complications (Figure 3 in **Paper II**) was not optimal for the following reasons: 1) the small sample size for each type of pregnancy complication and 2) the absence of gestational-age-matched controls. In particular, it was challenging to find a GEO study that included a large number of complicated pregnancies and gestational age-matched controls. Some controls and preeclampsia cases were included in several GEO datasets[13] and the Robinson Lab data. We would have been able to use them in either the association test or the development of CPC (n=963), but opted only to include them in the development of CPC to avoid underfitting.

6.3.4 Trade-off between predictability of chronological age and age-related conditions

Another concern about the ABECs is that their high precision in age prediction may reduce the predictability of age-related conditions [213]. The high precision in age prediction among all individuals means that EAA would only amount to technical noise, i.e., absence of the signals for functional decline, and EAA can thus no longer be used to classify individuals with and without the age-related disease. However, this is not without exceptions, as exemplified by the Horvath *et al.* [7] Skin & Blood clock. Here, the epigenetic age estimated by this clock was highly correlated with not only chronological age (median absolute error=2.5 years, r>0.98 in subjects aged 20-80 years) but also all-cause mortality, Down syndrome, and multiple lifestyles/dietary factors.

[13] GSE100197, GSE98224, and GSE44667

It was unclear whether the ABECs in **Paper I** were predictive of age-related diseases, mortality, and morbidity. Considering that our epigenetic clocks were developed on EPIC-based DNAm data and that large cohort studies of older subjects have recently migrated from HM450 to EPIC, it may take some time before the ABECs are validated in other cohorts with older subjects.

6.3.5 Reliability of methylation measures

The reliability of methylation measures often refers to the reproducibility of the probes on the Illumina arrays [235]. A common way of assessing reproducibility is to compare the methylation values quantified twice from each DNA sample [235, 236]. Previous studies have reported high correlations (>0.9) between the two quantifications across all the probes included in the Illumina array [118, 237, 238]. In contrast, the correlation[14] at each probe varied widely (the median correlation: 0.3, 25% percentile: 0.11, and 75% percentile: 0.63) [237].

However, the probes included in epigenetic clocks appeared to be more reproducible in terms of methylation values than ostensibly unrelated probes. The median correlations at the probes for published epigenetic clocks were more than 0.49, according to the report by Bose *et al.* [237]. The probes included in the ABECs (**Paper I**) also showed decent median correlations (>0.46)[15].

The EWAS of LTL in **Paper III** might be subjected to some probes with not only low reproducibility but also low compatibility between HM450 and EPIC. This is because the meta-analyses were restricted to the loci overlapping HM450 and EPIC, and there is evidence of low

[14] Many studies used the intraclass correlation coefficient (ICC) for assessing probe reproducibility.

[15] Here, the probes unique to EPIC were not taken into account because Bose and colleagues provided ICC values for the probes on HM450. No study has yet derived ICCs for all the probes on EPIC.

correlation across the two platforms at many of the loci [235, 239]. This low reliability might be related to the low variability in the methylation measures at the loci, which is merely a reflection of random error/noise. Logue *et al.* [239] thus recommended that loci evaluated in analyses be limited to those with sufficient variability. Among the 823 LTL-associated loci in **Paper III**, there were only four loci with low variability in FHS, six in JHS, zero in WHI, six in BHS, two in LBC1921, zero in LBC1936, and three in LSADT. No CpGs with low variability across all the cohorts were detected.

6.3.6 Selection bias in the EWAS of LTL (Paper III)

Survival bias might affect the EWAS of LTL (**Paper III**). All the cohorts except BHS might be susceptible to survival bias due to the inclusion of individuals aged 70-90 years (please refer to Table 2 in **Paper III**). As shorter LTL is associated with increased risk of mortality and age-related diseases [158, 160, 161, 166], those individuals who managed to reach old age are more likely to have longer LTL than those who did not make it to old age.

6.3.7 TL measurement: Southern blot versus qPCR

The Southern blot analysis of TRF is the current gold standard for LTL measurement. Southern blot provides the distribution of absolute LTL in kilobases. It shows high reproducibility across laboratories and low inter-assay coefficient of variation (<2%) [205]. For these reasons, many cross-sectional studies, such as FHS, JHS, WHI, and BHS in **Paper III**, have used Southern blot for LTL measurement. Despite its popularity, however, Southern blot has several practical shortcomings. Besides requiring a large amount of DNA (>6 micrograms), it is also time-consuming and labor-intensive [198, 205].

The measurement of LTL by qPCR has rapidly gained popularity in recent years [199, 240] due to its low cost, low amounts of DNA required, and suitability for high-throughput processing of samples [205]. Among the seven cohorts included in **Paper III**, LBC1921 and LBC1936 used the qPCR method. Even though it is a more practical alternative than Southern blot, the qPCR-based method has several critical shortcomings. Aside from low reproducibility, the measurement metric (the T/S ratio; see Section 4.2.2) precludes a formal comparison across studies/labs [241, 242]. The low reproducibility implies a higher likelihood of measurement errors. Moreover, a difference in the relative T/S ratio between healthy controls and an at-risk group does not identify any at-risk group *per se* and is therefore not useful clinically.

The EWAS results from LBC1921 and LBC1936 in **Paper III** might be influenced by the measurement errors of the qPCR. To examine if the EWAS results from the cohorts using Southern blot versus qPCR were consistent, we narrowed down the top 823 CpG sites from the global meta-analysis and compared the Z scores from the cohort-specific meta-analysis (Supplementary Figure 6, Supplementary File 2, **Paper III**). We found consistent LTL-DNAm associations across LBC1921, LBC1936, and FHS. It may be judicious to compare the Z score from these three cohorts, as they comprised individuals of European ancestry. Although the directions of the LTL-DNAm associations, i.e., the Z scores, in LBC 1921 mismatched those in FHS at several CpGs, it is difficult to conclude that the subtle inconsistency was due to a methodological discrepancy in LTL measurement.

6.3.8 FDR: Benjamini-Hochberg procedure

The Bonferroni correction mentioned in Section 4.3.2 was designed to control for FWER, i.e., to guard against the probability of reporting at least one false discovery, at a chosen level of α. In the setting of a contingency table of true/false discovery/non-discovery (**Table 3**), FWER can be denoted as $\Pr(V \geq 1)$ [243]. A threshold induced by the Bonferroni correction guarantees that $\Pr(V \geq 1) = 1 - \Pr(V = 0) = \alpha$. Although this simple method effectively controls FWER at a specific degree of significance (α), it decreases the statistical power substantially, i.e., it lowers the probability of rejecting false null hypotheses. Further, the implication of FWER, i.e., not making any Type I errors, might be too stringent for some research questions, e.g., the overall decision of whether a new drug outperforms an existing one.

Table 3. The number of true/false and discovery/non-discovery when testing *m* hypotheses.

	Not significant	Significant	Total
True H0	U	V	m_0
False H0	T	S	m_1
Total	$m - R$	R	m

Source: the notations from Table 1 in Benjamini and Hochberg [243].

To address the shortcomings of the Bonferroni method for controlling FWER, Benjamini and Hochberg [243] suggested focusing on the expected proportion of false discoveries among the rejected (significant) hypotheses, i.e., $FDR = E(Q) = E(V/R)$ (**Table 3**). The idea behind FDR is that it is less stringent than FWER, as it takes into account the rejected hypotheses rather than all the hypotheses tested.

Benjamini and Hochberg [243] also elaborated on a procedure controlling the FDR at q^*. Given that m null hypotheses, $H_{0,1}, H_{0,2} \ldots, H_{0,m}$ are tested, the corresponding p-values can be denoted as $p_1, p_2, \ldots, p_m$. The first step is to sort the p-values such that $p_{(1)} \leq p_{(2)} \leq \cdots \leq p_{(m)}$ and denote the corresponding null hypotheses $H_{0,(1)}, H_{0,(2)}, \ldots, H_{0,(m)}$. The next step is to define $k = argmax(p_{(i)} \leq \frac{i}{m} q^*)$ and reject all $H_{(1)}, H_{(2)}, \ldots, H_{(k)}$. The detailed proof of how this procedure controls FDR at q^* has been described in great detail in Appendix A of Benjamini and Hochberg [243].

In **Paper III**, we reported 823 LTL-associated CpGs after the Bonferroni correction, not the Benjamini and Hochberg procedure. The reason behind this decision was to minimize the likelihood of false positives as much as possible, given that the two molecular measures (DNAm and LTL) were possibly exposed to random measurement errors [244]. Furthermore, the risk of false positives was higher when it was unclear to what extent false-positive findings would affect the functional enrichment analyses. However, for researchers who wish to see the extended list of LTL-associated CpGs, we included CpGs that showed moderate (P<1E-05) association with LTL in Supplementary File 1 of **Paper III**.

6.4 Recommendations for the future use of the epigenetic clocks

As mentioned in the Results section of **Paper I**, a systematic offset may occur when epigenetic clocks are applied to newly-generated DNAm data. Although the systematic offset does not affect the association between EAA and an outcome of interest [16], it can lead to an over- or

[16] The definition of EAA is the residual term from a regression of epigenetic age on chronological age. The offset is calibrated through the regression model fitting.

underestimation of chronological age. Thus, it is crucial to calibrate the systematic offset, particularly in forensic research, where accurate age prediction is of utmost importance.

To avoid the systematic offset, it is advisable to include both target individuals and reference individuals whose chronological age is known in the same batch for DNAm measurement. It is even better if the chronological age of the reference individuals has a wide range. The likelihood of over- or underestimating the epigenetic age of the target individuals can be significantly reduced by fitting a linear (or non-linear) regression of epigenetic age on chronological age.

6.5 Future research

In the field of epigenetic clocks for adults, the focus has shifted from the precise prediction of chronological age to that of age-related conditions [217]. Provided that these two components are the core conditions for a valid biomarker of aging (also mentioned in Section 2.3), an obvious challenge is to find an optimal balance between the two core conditions. Studies such as the Levine *et al.* [8] PhenoAge and Lu *et al.* [143] GrimAge have, to some extent, found a good balance (please refer to Section 2.4.5 for methodological details) and were thus classified as 'second-generation clocks.' These second-generation clocks reduce precision in age prediction but increase the predictability of mortality, healthspan, cardiovascular disease, and cancer.

However, this progress has not yet occurred in the context of epigenetic clocks for newborns, nor has the validation of these clocks been carried out widely. Gestational age acceleration has been reported to be associated with maternal risk factors in several studies [147, 148, 245, 246] but was not associated with critical postnatal conditions such as neonatal death and neurodevelopmental disorders. To pursue this validation study, the following data are required: 1) DNAm data from cord blood or placental tissue collected at birth and 2) follow-up information after birth. In this regard, MoBa or the Avon Longitudinal Study of Parents and Children (ALSPAC) would be valuable data sources.

For a better understanding of the genetic determinants of LTL, it would be judicious to integrate the existing GWAS of LTL [184] and our EWAS of LTL. More specifically, a future study may focus on *RTEL1* that is featured in both of the GWAS and EWAS of LTL. The relevant SNP (rs755017), CpGs (cg00622799, cg03339910, cg10615591, and cg17534029) and SNP-CpG interactions can be investigated in relation to LTL. Several large cohorts such as FHS, JHS, WHI, and MoBa are suitable candidates for such an in-depth investigation, given that these cohorts have already generated all the substrates for such an analysis: genotype, DNAm, and LTL data.

6.6 Conclusion

Three blood-based epigenetic clocks were developed for the precise estimation of adults' chronological age using EPIC-derived DNAm data. The clocks achieved high precision in age prediction in independent cohorts. The highly precise age prediction was not explained by the broader genomic coverage of EPIC but rather by the large training set with a wide age-span.

Three placenta-based epigenetic clocks were subsequently developed for estimating fetal gestational age using a mixture of publicly available DNAm data. These placental clocks were highly accurate estimators of GA based on placental tissue regardless of pregnancy conditions.

The EWAS of LTL identified 823 CpG sites significantly associated ($P<1E-7$) with LTL after adjustment for age, sex, ethnicity, and imputed white blood cell counts. Functional enrichment analyses revealed that these CpG sites are near genes that play a role in circadian rhythm, blood coagulation, and wound healing. This study revealed significant relationships between the two recognized hallmarks of aging: TL and DNAm.

7 Reference

1. Jaul, E. and J. Barron, *Age-Related Diseases and Clinical and Public Health Implications for the 85 Years Old and Over Population.* Front Public Health, 2017. **5**: p. 335.
2. Chang, A.Y., et al., *Measuring population ageing: an analysis of the Global Burden of Disease Study 2017.* Lancet Public Health, 2019. **4**(3): p. e159-e167.
3. Singh, A.R., *Modern Medicine: Towards Prevention, Cure, Well-being and Longevity.* Mens Sana Monogr, 2010. **8**(1): p. 17-29.
4. Baker, G.T., 3rd and R.L. Sprott, *Biomarkers of aging.* Exp Gerontol, 1988. **23**(4-5): p. 223-39.
5. Hannum, G., et al., *Genome-wide methylation profiles reveal quantitative views of human aging rates.* Mol Cell, 2013. **49**(2): p. 359-367.
6. Horvath, S., *DNA methylation age of human tissues and cell types.* Genome Biol, 2013. **14**(10): p. R115.
7. Horvath, S., et al., *Epigenetic clock for skin and blood cells applied to Hutchinson Gilford Progeria Syndrome and ex vivo studies.* Aging (Albany NY), 2018. **10**(7): p. 1758-1775.
8. Levine, M.E., et al., *An epigenetic biomarker of aging for lifespan and healthspan.* Aging (Albany NY), 2018. **10**(4): p. 573-591.
9. Armanios, M. and E.H. Blackburn, *The telomere syndromes.* Nat Rev Genet, 2012. **13**(10): p. 693-704.
10. Boonekamp, J.J., et al., *Telomere length behaves as biomarker of somatic redundancy rather than biological age.* Aging Cell, 2013. **12**(2): p. 330-2.
11. Jagger, A., et al., *Regulatory T cells and the immune aging process: a mini-review.* Gerontology, 2014. **60**(2): p. 130-7.
12. Johnson, L.C., et al., *The plasma metabolome as a predictor of biological aging in humans.* Geroscience, 2019.
13. Lehallier, B., et al., *Undulating changes in human plasma proteome profiles across the lifespan.* Nat Med, 2019. **25**(12): p. 1843-1850.
14. Odamaki, T., et al., *Age-related changes in gut microbiota composition from newborn to centenarian: a cross-sectional study.* BMC Microbiol, 2016. **16**: p. 90.
15. Raiten, D.J., R. Raghavan, and K. Kraemer, *Biomarkers in growth.* Ann Nutr Metab, 2013. **63**(4): p. 293-7.
16. Bohlin, J., et al., *Prediction of gestational age based on genome-wide differentially methylated regions.* Genome Biol, 2016. **17**(1): p. 207.
17. Knight, A.K., et al., *An epigenetic clock for gestational age at birth based on blood methylation data.* Genome Biol, 2016. **17**(1): p. 206.
18. Sidorov, I., et al., *Leukocyte telomere dynamics and human hematopoietic stem cell kinetics during somatic growth.* Exp Hematol, 2009. **37**(4): p. 514-24.
19. Quigley, R., *Developmental changes in renal function.* Curr Opin Pediatr, 2012. **24**(2): p. 184-90.
20. England, K., et al., *Age- and sex-related reference ranges of alanine aminotransferase levels in children: European paediatric HCV network.* J Pediatr Gastroenterol Nutr, 2009. **49**(1): p. 71-7.
21. Kerr, M.A., et al., *Folate, related B vitamins, and homocysteine in childhood and adolescence: potential implications for disease risk in later life.* Pediatrics, 2009. **123**(2): p. 627-35.
22. Kirkwood, T.B., *Understanding the odd science of aging.* Cell, 2005. **120**(4): p. 437-47.
23. Rodriguez Valiente, A., et al., *Extended high-frequency (9-20 kHz) audiometry reference thresholds in 645 healthy subjects.* Int J Audiol, 2014. **53**(8): p. 531-45.
24. Gan, D.C. and R.D. Sinclair, *Prevalence of male and female pattern hair loss in Maryborough.* J Investig Dermatol Symp Proc, 2005. **10**(3): p. 184-9.
25. Xue, Q.L., *The frailty syndrome: definition and natural history.* Clin Geriatr Med, 2011. **27**(1): p. 1-15.
26. Vollenhoven, B. and S. Hunt, *Ovarian ageing and the impact on female fertility.* F1000Res, 2018. **7**.
27. Wang, J.C. and M. Bennett, *Aging and atherosclerosis: mechanisms, functional consequences, and potential therapeutics for cellular senescence.* Circ Res, 2012. **111**(2): p. 245-59.
28. Yakaryilmaz, F.D. and Z.A. Ozturk, *Treatment of type 2 diabetes mellitus in the elderly.* World J Diabetes, 2017. **8**(6): p. 278-285.
29. Lionakis, N., et al., *Hypertension in the elderly.* World J Cardiol, 2012. **4**(5): p. 135-47.
30. Ganguli, M. and E. Rodriguez, *Age, Alzheimer's disease, and the big picture.* Int Psychogeriatr, 2011. **23**(10): p. 1531-4.
31. Kempster, P.A., et al., *Relationships between age and late progression of Parkinson's disease:*

a clinico-pathological study. Brain, 2010. **133**(Pt 6): p. 1755-62.

32. Benjamin, E.J., et al., *Heart Disease and Stroke Statistics-2017 Update: A Report From the American Heart Association.* Circulation, 2017. **135**(10): p. e146-e603.
33. Buist, A.S., et al., *International variation in the prevalence of COPD (the BOLD Study): a population-based prevalence study.* Lancet, 2007. **370**(9589): p. 741-50.
34. World Health Organization. *The top 10 causes of death*. 2018 [cited 2018 24 May]; Available from: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death.
35. Tanner, J.M., *Foetus into man : physical growth from conception to maturity*. 1990, Cambridge, Mass.: Harvard University Press.
36. Hepper, P.G. and B.S. Shahidullah, *Development of fetal hearing.* Arch Dis Child Fetal Neonatal Ed, 1994. **71**(2): p. F81-7.
37. Warburton, D., et al., *Lung organogenesis.* Curr Top Dev Biol, 2010. **90**: p. 73-158.
38. Papageorghiou, A.T., et al., *International standards for fetal growth based on serial ultrasound measurements: the Fetal Growth Longitudinal Study of the INTERGROWTH-21st Project.* Lancet, 2014. **384**(9946): p. 869-79.
39. Villar, J., et al., *International standards for newborn weight, length, and head circumference by gestational age and sex: the Newborn Cross-Sectional Study of the INTERGROWTH-21st Project.* Lancet, 2014. **384**(9946): p. 857-68.
40. de Onis, M., et al., *Development of a WHO growth reference for school-aged children and adolescents.* Bull World Health Organ, 2007. **85**(9): p. 660-7.
41. Natale, V. and A. Rajagopalan, *Worldwide variation in human growth and the World Health Organization growth standards: a systematic review.* BMJ Open, 2014. **4**(1): p. e003735.
42. Villar, J., et al., *Postnatal growth standards for preterm infants: the Preterm Postnatal Follow-up Study of the INTERGROWTH-21(st) Project.* Lancet Glob Health, 2015. **3**(11): p. e681-91.
43. Group, W.H.O.M.G.R.S., *WHO Motor Development Study: windows of achievement for six gross motor development milestones.* Acta Paediatr Suppl, 2006. **450**: p. 86-95.
44. Workalemahu, T., et al., *Genetic and Environmental Influences on Fetal Growth Vary during Sensitive Periods in Pregnancy.* Sci Rep, 2018. **8**(1): p. 7274.
45. Gardosi, J., *Fetal growth and ethnic variation.* Lancet Diabetes Endocrinol, 2014. **2**(10): p. 773-4.
46. Boyle, E.M., et al., *Effects of gestational age at birth on health outcomes at 3 and 5 years of age: population based cohort study.* BMJ, 2012. **344**: p. e896.
47. Been, J.V., et al., *Preterm birth and childhood wheezing disorders: a systematic review and meta-analysis.* PLoS Med, 2014. **11**(1): p. e1001596.
48. Blencowe, H., et al., *Preterm birth-associated neurodevelopmental impairment estimates at regional and global levels for 2010.* Pediatr Res, 2013. **74 Suppl 1**: p. 17-34.
49. Figlio, D.N., et al., *Long-term Cognitive and Health Outcomes of School-Aged Children Who Were Born Late-Term vs Full-Term.* JAMA Pediatr, 2016. **170**(8): p. 758-64.
50. D'Onofrio, B.M., et al., *Preterm birth and mortality and morbidity: a population-based quasi-experimental study.* JAMA Psychiatry, 2013. **70**(11): p. 1231-40.
51. Group, E., et al., *One-year survival of extremely preterm infants after active perinatal care in Sweden.* JAMA, 2009. **301**(21): p. 2225-33.
52. Derraik, J.G.B., et al., *Large-for-gestational-age phenotypes and obesity risk in adulthood: a study of 195,936 women.* Sci Rep, 2020. **10**(1): p. 2157.
53. Meas, T., et al., *Consequences of being born small for gestational age on body composition: an 8-year follow-up study.* J Clin Endocrinol Metab, 2008. **93**(10): p. 3804-9.
54. Niu, C., et al., *Lifestyle Behaviors in Elderly Cancer Survivors: A Comparison With Middle-Age Cancer Survivors.* J Oncol Pract, 2015. **11**(4): p. e450-9.
55. Rodgers, J.L., et al., *Cardiovascular Risks Associated with Gender and Aging.* J Cardiovasc Dev Dis, 2019. **6**(2): p. 19.
56. Bell, C.G., et al., *DNA methylation aging clocks: challenges and recommendations.* Genome Biol, 2019. **20**(1): p. 249.
57. Christensen, K., et al., *"Looking old for your age": genetics and mortality.* Epidemiology, 2004. **15**(2): p. 251-2.
58. Simm, A., et al., *Potential biomarkers of ageing.* Biol Chem, 2008. **389**(3): p. 257-65.
59. Johnson, T.E., *Recent results: biomarkers of aging.* Exp Gerontol, 2006. **41**(12): p. 1243-6.
60. Horvath, S. and K. Raj, *DNA methylation-based biomarkers and the epigenetic clock theory of ageing.* Nat Rev Genet, 2018. **19**(6): p. 371-384.

61. Jylhava, J., N.L. Pedersen, and S. Hagg, *Biological Age Predictors.* EBioMedicine, 2017. **21**: p. 29-36.
62. Luo, Y., X. Lu, and H. Xie, *Dynamic Alu methylation during normal development, aging, and tumorigenesis.* Biomed Res Int, 2014. **2014**: p. 784706.
63. Laurent, L., et al., *Dynamic changes in the human methylome during differentiation.* Genome Res, 2010. **20**(3): p. 320-31.
64. Lister, R., et al., *Human DNA methylomes at base resolution show widespread epigenomic differences.* Nature, 2009. **462**(7271): p. 315-22.
65. Lister, R., et al., *Global epigenomic reconfiguration during mammalian brain development.* Science, 2013. **341**(6146): p. 1237905.
66. Guo, J.U., et al., *Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain.* Nat Neurosci, 2014. **17**(2): p. 215-22.
67. Xiao, C.L., et al., *N(6)-Methyladenine DNA Modification in the Human Genome.* Mol Cell, 2018. **71**(2): p. 306-318 e7.
68. Greenberg, M.V.C. and D. Bourc'his, *The diverse roles of DNA methylation in mammalian development and disease.* Nat Rev Mol Cell Biol, 2019. **20**(10): p. 590-607.
69. Saxonov, S., P. Berg, and D.L. Brutlag, *A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters.* Proc Natl Acad Sci U S A, 2006. **103**(5): p. 1412-7.
70. Gardiner-Garden, M. and M. Frommer, *CpG islands in vertebrate genomes.* J Mol Biol, 1987. **196**(2): p. 261-82.
71. Lander, E.S., et al., *Initial sequencing and analysis of the human genome.* Nature, 2001. **409**(6822): p. 860-921.
72. Deaton, A.M. and A. Bird, *CpG islands and the regulation of transcription.* Genes Dev, 2011. **25**(10): p. 1010-22.
73. Sarda, S., et al., *Distal CpG islands can serve as alternative promoters to transcribe genes with silenced proximal promoters.* Genome Res, 2017. **27**(4): p. 553-566.
74. Jeziorska, D.M., et al., *DNA methylation of intragenic CpG islands depends on their transcriptional activity during differentiation and disease.* Proc Natl Acad Sci U S A, 2017. **114**(36): p. E7526-E7535.
75. Schubeler, D., *Function and information content of DNA methylation.* Nature, 2015. **517**(7534): p. 321-6.
76. Lyko, F., *The DNA methyltransferase family: a versatile toolkit for epigenetic regulation.* Nat Rev Genet, 2018. **19**(2): p. 81-92.
77. Okano, M., et al., *DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development.* Cell, 1999. **99**(3): p. 247-57.
78. Okano, M., S. Xie, and E. Li, *Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases.* Nat Genet, 1998. **19**(3): p. 219-20.
79. Robertson, K.D., et al., *The human DNA methyltransferases (DNMTs) 1, 3a and 3b: coordinate mRNA expression in normal tissues and overexpression in tumors.* Nucleic Acids Res, 1999. **27**(11): p. 2291-8.
80. Li, E., T.H. Bestor, and R. Jaenisch, *Targeted mutation of the DNA methyltransferase gene results in embryonic lethality.* Cell, 1992. **69**(6): p. 915-26.
81. Ben-Hattar, J. and J. Jiricny, *Methylation of single CpG dinucleotides within a promoter element of the Herpes simplex virus tk gene reduces its transcription in vivo.* Gene, 1988. **65**(2): p. 219-27.
82. Watt, F. and P.L. Molloy, *Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter.* Genes Dev, 1988. **2**(9): p. 1136-43.
83. Iguchi-Ariga, S.M. and W. Schaffner, *CpG methylation of the cAMP-responsive enhancer/promoter sequence TGACGTCA abolishes specific factor binding as well as transcriptional activation.* Genes Dev, 1989. **3**(5): p. 612-9.
84. Wagner, J.R., et al., *The relationship between DNA methylation, genetic and expression inter-individual variation in untransformed human fibroblasts.* Genome Biol, 2014. **15**(2): p. R37.
85. Kulis, M., et al., *Epigenomic analysis detects widespread gene-body DNA hypomethylation in chronic lymphocytic leukemia.* Nat Genet, 2012. **44**(11): p. 1236-42.
86. Jjingo, D., et al., *On the presence and role of human gene-body DNA methylation.* Oncotarget, 2012. **3**(4): p. 462-74.
87. Varley, K.E., et al., *Dynamic DNA methylation across diverse human cell lines and tissues.* Genome Res, 2013. **23**(3): p. 555-67.
88. Yang, X., et al., *Gene body methylation can alter gene expression and is a therapeutic target in cancer.* Cancer Cell, 2014. **26**(4): p. 577-90.
89. Zhu, H., G. Wang, and J. Qian, *Transcription factors as readers and effectors of DNA*

methylation. Nat Rev Genet, 2016. **17**(9): p. 551-65.

90. Bird, A., *DNA methylation patterns and epigenetic memory.* Genes Dev, 2002. **16**(1): p. 6-21.
91. Ferguson-Smith, A.C., *Genomic imprinting: the emergence of an epigenetic paradigm.* Nat Rev Genet, 2011. **12**(8): p. 565-75.
92. Monk, D., et al., *Genomic imprinting disorders: lessons on how genome, epigenome and environment interact.* Nat Rev Genet, 2019. **20**(4): p. 235-248.
93. Elhamamsy, A.R., *Role of DNA methylation in imprinting disorders: an updated review.* J Assist Reprod Genet, 2017. **34**(5): p. 549-562.
94. Edwards, C.A. and A.C. Ferguson-Smith, *Mechanisms regulating imprinted genes in clusters.* Curr Opin Cell Biol, 2007. **19**(3): p. 281-9.
95. Sharp, A.J., et al., *DNA methylation profiles of human active and inactive X chromosomes.* Genome Res, 2011. **21**(10): p. 1592-600.
96. Cotton, A.M., et al., *Landscape of DNA methylation on the X chromosome reflects CpG density, functional chromatin state and X-chromosome inactivation.* Hum Mol Genet, 2015. **24**(6): p. 1528-39.
97. Jin, Z. and Y. Liu, *DNA methylation in human diseases.* Genes Dis, 2018. **5**(1): p. 1-8.
98. Kulis, M. and M. Esteller, *DNA methylation and cancer.* Adv Genet, 2010. **70**: p. 27-56.
99. Paul, D.S., et al., *Increased DNA methylation variability in type 1 diabetes across three immune effector cell types.* Nat Commun, 2016. **7**: p. 13555.
100. Shao, X., et al., *Rheumatoid arthritis-relevant DNA methylation changes identified in ACPA-positive asymptomatic individuals using methylome capture sequencing.* Clin Epigenetics, 2019. **11**(1): p. 110.
101. Dayeh, T., et al., *Genome-wide DNA methylation analysis of human pancreatic islets from type 2 diabetic and non-diabetic donors identifies candidate genes that influence insulin secretion.* PLoS Genet, 2014. **10**(3): p. e1004160.
102. Willmer, T., et al., *Blood-Based DNA Methylation Biomarkers for Type 2 Diabetes: Potential for Clinical Applications.* Front Endocrinol (Lausanne), 2018. **9**: p. 744.
103. Guay, S.P., et al., *Epigenome-wide analysis in familial hypercholesterolemia identified new loci associated with high-density lipoprotein cholesterol concentration.* Epigenomics, 2012. **4**(6): p. 623-39.
104. Braun, K.V.E., et al., *Epigenome-wide association study (EWAS) on lipids: the Rotterdam Study.* Clin Epigenetics, 2017. **9**: p. 15.
105. Guay, S.P., et al., *Epigenetic and genetic variations at the TNNT1 gene locus are associated with HDL-C levels and coronary artery disease.* Epigenomics, 2016. **8**(3): p. 359-71.
106. Drinkwater, R.D., et al., *Human lymphocytes aged in vivo have reduced levels of methylation in transcriptionally active and inactive DNA.* Mutat Res, 1989. **219**(1): p. 29-37.
107. Kwabi-Addo, B., et al., *Age-related DNA methylation changes in normal human prostate tissues.* Clin Cancer Res, 2007. **13**(13): p. 3796-802.
108. Christensen, B.C., et al., *Aging and environmental exposures alter tissue-specific DNA methylation dependent upon CpG island context.* PLoS Genet, 2009. **5**(8): p. e1000602.
109. Boks, M.P., et al., *The relationship of DNA methylation with age, gender and genotype in twins and healthy controls.* PLoS One, 2009. **4**(8): p. e6767.
110. Alisch, R.S., et al., *Age-associated DNA methylation in pediatric populations.* Genome Res, 2012. **22**(4): p. 623-32.
111. Cooney, C.A., *Are somatic cells inherently deficient in methylation metabolism? A proposed mechanism for DNA methylation loss, senescence and aging.* Growth Dev Aging, 1993. **57**(4): p. 261-73.
112. Merid, S.K., et al., *Epigenome-wide meta-analysis of blood DNA methylation in newborns and children identifies numerous loci related to gestational age.* Genome Med, 2020. **12**(1): p. 25.
113. Schroeder, J.W., et al., *Neonatal DNA methylation patterns associate with gestational age.* Epigenetics, 2011. **6**(12): p. 1498-504.
114. Parets, S.E., et al., *Fetal DNA Methylation Associates with Early Spontaneous Preterm Birth and Gestational Age.* PLoS One, 2013. **8**(6): p. e67489.
115. Simpkin, A.J., et al., *Longitudinal analysis of DNA methylation associated with birth weight and gestational age.* Hum Mol Genet, 2015. **24**(13): p. 3752-63.
116. Novakovic, B., et al., *Evidence for widespread changes in promoter methylation profile in human placenta in response to increasing gestational age and environmental/stochastic factors.* BMC Genomics, 2011. **12**: p. 529.
117. Mayne, B.T., et al., *Accelerated placental aging in early onset preeclampsia pregnancies*

identified by DNA methylation. Epigenomics, 2017. **9**(3): p. 279-289.

118. Pidsley, R., et al., *Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling.* Genome Biol, 2016. **17**(1): p. 208.
119. Bibikova, M., et al., *Genome-wide DNA methylation profiling using Infinium(R) assay.* Epigenomics, 2009. **1**(1): p. 177-200.
120. Pruitt, K.D., et al., *The consensus coding sequence (CCDS) project: Identifying a common protein-coding gene set for the human and mouse genomes.* Genome Res, 2009. **19**(7): p. 1316-23.
121. Rakyan, V.K., et al., *Human aging-associated DNA hypermethylation occurs preferentially at bivalent chromatin domains.* Genome Res, 2010. **20**(4): p. 434-9.
122. Bell, C.G., et al., *Genome-wide DNA methylation analysis for diabetic nephropathy in type 1 diabetes mellitus.* BMC Med Genomics, 2010. **3**: p. 33.
123. Wolber, L.E., et al., *Epigenome-wide DNA methylation in hearing ability: new mechanisms for an old problem.* PLoS One, 2014. **9**(9): p. e105729.
124. Minning, C., et al., *Exploring breast carcinogenesis through integrative genomics and epigenomics analyses.* Int J Oncol, 2014. **45**(5): p. 1959-68.
125. Qiu, W., et al., *The impact of genetic variation and cigarette smoke on DNA methylation in current and former smokers from the COPDGene study.* Epigenetics, 2015. **10**(11): p. 1064-73.
126. Nishioka, M., et al., *Comprehensive DNA methylation analysis of peripheral blood cells derived from patients with first-episode schizophrenia.* J Hum Genet, 2013. **58**(2): p. 91-7.
127. Li, S.C., et al., *Major methylation alterations on the CpG markers of inflammatory immune associated genes after IVIG treatment in Kawasaki disease.* BMC Med Genomics, 2016. **9 Suppl 1**: p. 37.
128. Illumina, I. 2020 [cited 2020 8 Sep]; Available from: https://support.illumina.com/array/array_kits/infinium-methylationepic-beadchip-kit/downloads.html.
129. Aryee, M.J., et al., *Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays.* Bioinformatics, 2014. **30**(10): p. 1363-9.
130. Muller, F., et al., *RnBeads 2.0: comprehensive analysis of DNA methylation data.* Genome Biol, 2019. **20**(1): p. 55.
131. Bell, J.T., et al., *Epigenome-wide scans identify differentially methylated regions for age and age-related phenotypes in a healthy ageing population.* PLoS Genet, 2012. **8**(4): p. e1002629.
132. Bocklandt, S., et al., *Epigenetic predictor of age.* PLoS One, 2011. **6**(6): p. e14821.
133. Bollati, V., et al., *Decline in genomic DNA methylation through aging in a cohort of elderly subjects.* Mech Ageing Dev, 2009. **130**(4): p. 234-9.
134. Garagnani, P., et al., *Methylation of ELOVL2 gene as a new epigenetic marker of age.* Aging Cell, 2012. **11**(6): p. 1132-4.
135. Bacalini, M.G., et al., *Systemic Age-Associated DNA Hypermethylation of ELOVL2 Gene: In Vivo and In Vitro Evidences of a Cell Replication Process.* J Gerontol A Biol Sci Med Sci, 2017. **72**(8): p. 1015-1023.
136. Simpkin, A.J., et al., *Prenatal and early life influences on epigenetic age in children: a study of mother-offspring pairs from two cohort studies.* Hum Mol Genet, 2016. **25**(1): p. 191-201.
137. Simpkin, A.J., et al., *The epigenetic clock and physical development during childhood and adolescence: longitudinal analysis from a UK birth cohort.* Int J Epidemiol, 2017. **46**(2): p. 549-558.
138. Soriano-Tarraga, C., et al., *Biological age is better than chronological as predictor of 3-month outcome in ischemic stroke.* Neurology, 2017. **89**(8): p. 830-836.
139. Lind, L., et al., *Methylation-based estimated biological age and cardiovascular disease.* Eur J Clin Invest, 2018. **48**(2): p. e12872.
140. Soriano-Tarraga, C., et al., *Ischemic stroke patients are biologically older than their chronological age.* Aging (Albany NY), 2016. **8**(11): p. 2655-2666.
141. Zhang, Y., et al., *DNA methylation signatures in peripheral blood strongly predict all-cause mortality.* Nat Commun, 2017. **8**: p. 14617.
142. Fahy, G.M., et al., *Reversal of epigenetic aging and immunosenescent trends in humans.* Aging Cell, 2019. **18**(6): p. e13028.
143. Lu, A.T., et al., *DNA methylation GrimAge strongly predicts lifespan and healthspan.* Aging (Albany NY), 2019. **11**(2): p. 303-327.
144. Hillary, R.F., et al., *An epigenetic predictor of death captures multi-modal measures of brain health.* Mol Psychiatry, 2019: p. 1-11.

145. Simpkin, A.J., M. Suderman, and L.D. Howe, *Epigenetic clocks for gestational age: statistical and study design considerations.* Clin Epigenetics, 2017. **9**: p. 100.
146. Relton, C.L., et al., *Data Resource Profile: Accessible Resource for Integrated Epigenomic Studies (ARIES).* Int J Epidemiol, 2015. **44**(4): p. 1181-90.
147. Khouja, J.N., et al., *Epigenetic gestational age acceleration: a prospective cohort study investigating associations with familial, sociodemographic and birth characteristics.* Clin Epigenetics, 2018. **10**: p. 86.
148. Bright, H.D., et al., *Epigenetic gestational age and trajectories of weight and height during childhood: a prospective cohort study.* Clin Epigenetics, 2019. **11**(1): p. 194.
149. Blackburn, E.H., *Telomeres and telomerase: their mechanisms of action and the effects of altering their functions.* FEBS Lett, 2005. **579**(4): p. 859-62.
150. Chiu, C.P., et al., *Differential expression of telomerase activity in hematopoietic progenitors from adult human bone marrow.* Stem Cells, 1996. **14**(2): p. 239-48.
151. Yui, J., C.P. Chiu, and P.M. Lansdorp, *Telomerase activity in candidate stem cells from fetal liver and adult bone marrow.* Blood, 1998. **91**(9): p. 3255-62.
152. Shay, J.W. and W.E. Wright, *The reactivation of telomerase activity in cancer progression.* Trends Genet, 1996. **12**(4): p. 129-31.
153. Gomes, N.M., et al., *Comparative biology of mammalian telomeres: hypotheses on ancestral states and the roles of telomeres in longevity determination.* Aging Cell, 2011. **10**(5): p. 761-8.
154. Lee, Y., et al., *Epigenome-wide association study of leukocyte telomere length.* Aging (Albany NY), 2019. **11**(16): p. 5876-5894.
155. Okuda, K., et al., *Telomere length in the newborn.* Pediatr Res, 2002. **52**(3): p. 377-81.
156. Arai, Y., et al., *Inflammation, But Not Telomere Length, Predicts Successful Ageing at Extreme Old Age: A Longitudinal Study of Semi-supercentenarians.* EBioMedicine, 2015. **2**(10): p. 1549-58.
157. Stone, R.C., et al., *Telomere Length and the Cancer-Atherosclerosis Trade-Off.* PLoS Genet, 2016. **12**(7): p. e1006144.
158. Deelen, J., et al., *Leukocyte telomere length associates with prospective mortality independent of immune-related parameters and known genetic markers.* Int J Epidemiol, 2014. **43**(3): p. 878-86.
159. Schoenmaker, M., et al., *Evidence of genetic enrichment for exceptional survival using a family approach: the Leiden Longevity Study.* Eur J Hum Genet, 2006. **14**(1): p. 79-84.
160. Bakaysa, S.L., et al., *Telomere length predicts survival independent of genetic influences.* Aging Cell, 2007. **6**(6): p. 769-74.
161. Kimura, M., et al., *Telomere length and mortality: a study of leukocytes in elderly Danish twins.* Am J Epidemiol, 2008. **167**(7): p. 799-806.
162. Fitzpatrick, A.L., et al., *Leukocyte telomere length and cardiovascular disease in the cardiovascular health study.* Am J Epidemiol, 2007. **165**(1): p. 14-21.
163. Fried, L.P., et al., *The Cardiovascular Health Study: design and rationale.* Ann Epidemiol, 1991. **1**(3): p. 263-76.
164. Mwasongwe, S., et al., *Leukocyte telomere length and cardiovascular disease in African Americans: The Jackson Heart Study.* Atherosclerosis, 2017. **266**: p. 41-47.
165. Taylor, H.A., Jr., et al., *Toward resolution of cardiovascular health disparities in African Americans: design and methods of the Jackson Heart Study.* Ethn Dis, 2005. **15**(4 Suppl 6): p. S6-4-17.
166. Arbeev, K.G., et al., *Association of Leukocyte Telomere Length With Mortality Among Adult Participants in 3 Longitudinal Studies.* JAMA Netw Open, 2020. **3**(2): p. e200023.
167. Nan, H., et al., *Shorter telomeres associate with a reduced risk of melanoma development.* Cancer Res, 2011. **71**(21): p. 6758-63.
168. Anic, G.M., et al., *Telomere length and risk of melanoma, squamous cell carcinoma, and basal cell carcinoma.* Cancer Epidemiol, 2013. **37**(4): p. 434-9.
169. Sanchez-Espiridion, B., et al., *Telomere length in peripheral blood leukocytes and lung cancer risk: a large case-control study in Caucasians.* Cancer Res, 2014. **74**(9): p. 2476-86.
170. Seow, W.J., et al., *Telomere length in white blood cell DNA and lung cancer: a pooled analysis of three prospective cohorts.* Cancer Res, 2014. **74**(15): p. 4090-8.
171. Pellatt, A.J., et al., *Telomere length, telomere-related genes, and breast cancer risk: the breast cancer health disparities study.* Genes Chromosomes Cancer, 2013. **52**(7): p. 595-609.
172. Qu, S., et al., *Association of leukocyte telomere length with breast cancer risk: nested case-control findings from the Shanghai Women's Health Study.* Am J Epidemiol, 2013. **177**(7): p. 617-24.

173. Lynch, S.M., et al., *A prospective analysis of telomere length and pancreatic cancer in the alpha-tocopherol beta-carotene cancer (ATBC) prevention study.* Int J Cancer, 2013. **133**(11): p. 2672-80.
174. Julin, B., et al., *Circulating leukocyte telomere length and risk of overall and aggressive prostate cancer.* Br J Cancer, 2015. **112**(4): p. 769-76.
175. Telomeres Mendelian Randomization, C., et al., *Association Between Telomere Length and Risk of Cancer and Non-Neoplastic Diseases: A Mendelian Randomization Study.* JAMA Oncol, 2017. **3**(5): p. 636-651.
176. Aviv, A. and J.W. Shay, *Reflections on telomere dynamics and ageing-related diseases in humans.* Philos Trans R Soc Lond B Biol Sci, 2018. **373**(1741).
177. Opresko, P.L., et al., *Oxidative damage in telomeric DNA disrupts recognition by TRF1 and TRF2.* Nucleic Acids Res, 2005. **33**(4): p. 1230-9.
178. Lee, H.T., et al., *Molecular mechanisms by which oxidative DNA damage promotes telomerase activity.* Nucleic Acids Res, 2017. **45**(20): p. 11752-11765.
179. Singh, A., et al., *Oxidative Stress: Role and Response of Short Guanine Tracts at Genomic Locations.* Int J Mol Sci, 2019. **20**(17): p. 4258.
180. Pignatelli, P., et al., *Oxidative stress and cardiovascular disease: new insights.* Kardiol Pol, 2018. **76**(4): p. 713-722.
181. Mangino, M., et al., *Genome-wide meta-analysis points to CTC1 and ZNF676 as genes regulating telomere homeostasis in humans.* Hum Mol Genet, 2012. **21**(24): p. 5385-94.
182. Mangino, M., et al., *DCAF4, a novel gene associated with leucocyte telomere length.* J Med Genet, 2015. **52**(3): p. 157-62.
183. Levy, D., et al., *Genome-wide association identifies OBFC1 as a locus involved in human leukocyte telomere biology.* Proc Natl Acad Sci U S A, 2010. **107**(20): p. 9293-8.
184. Codd, V., et al., *Identification of seven loci affecting mean telomere length and their association with disease.* Nat Genet, 2013. **45**(4): p. 422-7.
185. Bibikova, M., et al., *High density DNA methylation array with single CpG site resolution.* Genomics, 2011. **98**(4): p. 288-95.
186. Nakabayashi, K., *Illumina HumanMethylation BeadChip for Genome-Wide DNA Methylation Profiling: Advantages and Limitations*, in *Handbook of Nutrition, Diet, and Epigenetics*. 2017, Handbook of Nutrition, Diet, Epigenetics. p. 1-15.
187. Triche, T.J., Jr., et al., *Low-level processing of Illumina Infinium DNA Methylation BeadArrays.* Nucleic Acids Res, 2013. **41**(7): p. e90.
188. Xie, Y., X. Wang, and M. Story, *Statistical methods of background correction for Illumina BeadArray data.* Bioinformatics, 2009. **25**(6): p. 751-7.
189. Bolstad, B.M., et al., *A comparison of normalization methods for high density oligonucleotide array data based on variance and bias.* Bioinformatics, 2003. **19**(2): p. 185-93.
190. Touleimat, N. and J. Tost, *Complete pipeline for Infinium((R)) Human Methylation 450K BeadChip data processing using subset quantile normalization for accurate DNA methylation estimation.* Epigenomics, 2012. **4**(3): p. 325-41.
191. Hickey, P. and K.D. Hansen. 2018 13 Jun; Available from: https://github.com/hansenlab/minfi/blob/master/R/preprocessQuantile.R.
192. Maksimovic, J., L. Gordon, and A. Oshlack, *SWAN: Subset-quantile within array normalization for illumina infinium HumanMethylation450 BeadChips.* Genome Biol, 2012. **13**(6): p. R44.
193. Fortin, J.P., et al., *Functional normalization of 450k methylation array data improves replication in large cancer studies.* Genome Biol, 2014. **15**(12): p. 503.
194. Akulenko, R., M. Merl, and V. Helms, *BEclear: Batch Effect Detection and Adjustment in DNA Methylation Data.* PLoS One, 2016. **11**(8): p. e0159921.
195. Price, E.M. and W.P. Robinson, *Adjusting for Batch Effects in DNA Methylation Microarray Data, a Lesson Learned.* Front Genet, 2018. **9**: p. 83.
196. Teschendorff, A.E., et al., *A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data.* Bioinformatics, 2013. **29**(2): p. 189-96.
197. Ji, Y., et al., *Applications of beta-mixture models in bioinformatics.* Bioinformatics, 2005. **21**(9): p. 2118-22.
198. Kimura, M., et al., *Measurement of telomere length by the Southern blot analysis of terminal restriction fragment lengths.* Nat Protoc, 2010. **5**(9): p. 1596-607.
199. Cawthon, R.M., *Telomere measurement by quantitative PCR.* Nucleic Acids Res, 2002. **30**(10): p. e47.
200. Poon, S.S.S. and P.M. Lansdorp, *Quantitative fluorescence in situ hybridization (Q-FISH).*

Curr Protoc Cell Biol, 2001. **Chapter 18**: p. 18 4 1-18 4 21.

201. Slijepcevic, P., *Telomere length measurement by Q-FISH.* Methods Cell Sci, 2001. **23**(1-3): p. 17-22.
202. Lansdorp, P.M., et al., *Heterogeneity in telomere length of human chromosomes.* Hum Mol Genet, 1996. **5**(5): p. 685-91.
203. Canela, A., et al., *High-throughput telomere length quantification by FISH and its application to human population studies.* Proc Natl Acad Sci U S A, 2007. **104**(13): p. 5300-5.
204. Baerlocher, G.M., et al., *Flow cytometry and FISH to measure the average length of telomeres (flow FISH).* Nat Protoc, 2006. **1**(5): p. 2365-76.
205. Lai, T.P., W.E. Wright, and J.W. Shay, *Comparison of telomere length measurement methods.* Philos Trans R Soc Lond B Biol Sci, 2018. **373**(1741).
206. Montpetit, A.J., et al., *Telomere length: a review of methods for measurement.* Nurs Res, 2014. **63**(4): p. 289-99.
207. Flores, I., et al., *The longest telomeres: a general signature of adult stem cell compartments.* Genes Dev, 2008. **22**(5): p. 654-67.
208. Meyne, J., et al., *Distribution of non-telomeric sites of the (TTAGGG)n telomeric sequence in vertebrate chromosomes.* Chromosoma, 1990. **99**(1): p. 3-10.
209. Baird, D.M., et al., *Extensive allelic variation and ultrashort telomeres in senescent human cells.* Nat Genet, 2003. **33**(2): p. 203-7.
210. Aubert, G., M. Hills, and P.M. Lansdorp, *Telomere length measurement-caveats and a critical assessment of the available technologies and tools.* Mutat Res, 2012. **730**(1-2): p. 59-67.
211. Bendix, L., et al., *The load of short telomeres, estimated by a new method, Universal STELA, correlates with number of senescent cells.* Aging Cell, 2010. **9**(3): p. 383-97.
212. Lai, T.P., et al., *A method for measuring the distribution of the shortest telomeres in cells and tissues.* Nat Commun, 2017. **8**(1): p. 1356.
213. Zhang, Q., et al., *Improved precision of epigenetic clock estimates across tissues and its implication for biological ageing.* Genome Med, 2019. **11**(1): p. 54.
214. Mastrolia, S.A., et al., *Placental calcifications: a clue for the identification of high-risk fetuses in the low-risk pregnant population?* J Matern Fetal Neonatal Med, 2016. **29**(6): p. 921-7.
215. Rakyan, V.K., et al., *Epigenome-wide association studies for common human diseases.* Nat Rev Genet, 2011. **12**(8): p. 529-41.
216. Verma, M., *Genome-wide association studies and epigenome-wide association studies go together in cancer control.* Future Oncol, 2016. **12**(13): p. 1645-64.
217. Field, A.E., et al., *DNA Methylation Clocks in Aging: Categories, Causes, and Consequences.* Mol Cell, 2018. **71**(6): p. 882-895.
218. Guevara, E.E. and R.R. Lawler, *Epigenetic Clocks.* Evol Anthropol, 2018. **27**(6): p. 256-260.
219. Zou, H. and T. Hastie, *Regularization and variable selection via the elastic net.* Journal of the Royal Statistical Society: Series B (Statistical Methodology), 2005. **67**(2): p. 301-320.
220. Engebretsen, S. and J. Bohlin, *Statistical predictions with glmnet.* Clin Epigenetics, 2019. **11**(1): p. 123.
221. Hastie, T., R. Tibshirani, and J.H. Friedman, *The elements of statistical learning : data mining, inference, and prediction*. 2nd ed. Springer series in statistics,. 2009, New York, NY: Springer. xxii, 745 p.
222. James, G., et al., *An introduction to statistical learning*. Vol. 112. 2013: Springer.
223. Johnson, W.E., C. Li, and A. Rabinovic, *Adjusting batch effects in microarray expression data using empirical Bayes methods.* Biostatistics, 2007. **8**(1): p. 118-27.
224. Leek, J.T., et al., *The sva package for removing batch effects and other unwanted variation in high-throughput experiments.* Bioinformatics, 2012. **28**(6): p. 882-3.
225. Nygaard, V., E.A. Rodland, and E. Hovig, *Methods that remove batch effects while retaining group differences may lead to exaggerated confidence in downstream analyses.* Biostatistics, 2016. **17**(1): p. 29-39.
226. Zindler, T., et al., *Simulating ComBat: how batch correction can lead to the systematic introduction of false positive results in DNA methylation microarray studies.* BMC Bioinformatics, 2020. **21**(1): p. 271.
227. Magnus, P., et al., *Cohort Profile Update: The Norwegian Mother and Child Cohort Study (MoBa).* Int J Epidemiol, 2016. **45**(2): p. 382-8.
228. Curtis, S.W., et al., *Exposure to polybrominated biphenyl (PBB) associates with genome-wide DNA methylation differences in peripheral blood.* Epigenetics, 2019. **14**(1): p. 52-66.
229. Nilsen, R.M., et al., *Self-selection and bias in a large prospective pregnancy cohort in Norway.* Paediatr Perinat Epidemiol, 2009. **23**(6): p. 597-608.

230. Gruzieva, O., et al., *DNA Methylation Trajectories During Pregnancy.* Epigenet Insights, 2019. **12**: p. 2516865719867090.
231. Teschendorff, A.E., J. West, and S. Beck, *Age-associated epigenetic drift: implications, and a case of epigenetic thrift?* Hum Mol Genet, 2013. **22**(R1): p. R7-R15.
232. Issa, J.P., *Aging and epigenetic drift: a vicious cycle.* J Clin Invest, 2014. **124**(1): p. 24-9.
233. Langevin, S.M., et al., *Does epigenetic drift contribute to age-related increases in breast cancer risk?* Epigenomics, 2014. **6**(4): p. 367-9.
234. Fransquet, P.D., et al., *The epigenetic clock as a predictor of disease and mortality risk: a systematic review and meta-analysis.* Clin Epigenetics, 2019. **11**(1): p. 62.
235. Sugden, K., et al., *Patterns of Reliability: Assessing the Reproducibility and Integrity of DNA Methylation Measurement.* Patterns (N Y), 2020. **1**(2).
236. Dugue, P.A., et al., *Reliability of DNA methylation measures from dried blood spots and mononuclear cells using the HumanMethylation450k BeadArray.* Sci Rep, 2016. **6**: p. 30317.
237. Bose, M., et al., *Evaluation of microarray-based DNA methylation measurement using technical replicates: the Atherosclerosis Risk In Communities (ARIC) Study.* BMC Bioinformatics, 2014. **15**: p. 312.
238. Moran, S., C. Arribas, and M. Esteller, *Validation of a DNA methylation microarray for 850,000 CpG sites of the human genome enriched in enhancer sequences.* Epigenomics, 2016. **8**(3): p. 389-99.
239. Logue, M.W., et al., *The correlation of methylation levels measured using Illumina 450K and EPIC BeadChips in blood samples.* Epigenomics, 2017. **9**(11): p. 1363-1371.
240. Cawthon, R.M., *Telomere length measurement by a novel monochrome multiplex quantitative PCR method.* Nucleic Acids Res, 2009. **37**(3): p. e21.
241. Martin-Ruiz, C.M., et al., *Reproducibility of Telomere Length Assessment--An International Collaborative Study.* Int J Epidemiol, 2015. **44**(5): p. 1749-54.
242. Verhulst, S., et al., *Commentary: The reliability of telomere length measurements.* Int J Epidemiol, 2015. **44**(5): p. 1683-6.
243. Benjamini, Y. and Y. Hochberg, *Controlling the false discovery rate: a practical and powerful approach to multiple testing.* Journal of the Royal Statistical Society: Series B (Methodological), 1995. **57**(1): p. 289-300.
244. Shear, B.R. and B.D. Zumbo, *False Positives in Multiple Regression.* Educational and Psychological Measurement, 2013. **73**(5): p. 733-756.
245. Girchenko, P., et al., *Associations between maternal risk factors of adverse pregnancy and birth outcomes and the offspring epigenetic clock of gestational age at birth.* Clin Epigenetics, 2017. **9**: p. 49.
246. Chen, L., et al., *Effects of Maternal Vitamin D3 Supplementation on Offspring Epigenetic Clock of Gestational Age at Birth: A Post-hoc Analysis of a Randomized Controlled Trial.* Epigenetics, 2020. **15**(8): p. 830-840.

Paper I

Blood-based epigenetic estimators of chronological age in human adults using DNA methylation data from the Illumina MethylationEPIC array

Lee *et al. BMC Genomics* (2020) 21:747
https://doi.org/10.1186/s12864-020-07168-8

BMC Genomics

METHODOLOGY ARTICLE Open Access

Blood-based epigenetic estimators of chronological age in human adults using DNA methylation data from the Illumina MethylationEPIC array

Yunsung Lee[1,2*], Kristine L. Haftorn[1,2,3], William R. P. Denault[1,3,4], Haakon E. Nustad[3,5], Christian M. Page[3,6], Robert Lyle[3,7,8], Sindre Lee-Ødegård[9,10], Gunn-Helen Moen[11,12,13,14], Rashmi B. Prasad[15], Leif C. Groop[15,16], Line Sletner[17,18], Christine Sommer[19], Maria C. Magnus[3,14,20], Håkon K. Gjessing[3,4], Jennifer R. Harris[1,3], Per Magnus[3], Siri E. Håberg[3], Astanand Jugessur[1,3,4†] and Jon Bohlin[3,21†]

Abstract

Background: Epigenetic clocks have been recognized for their precise prediction of chronological age, age-related diseases, and all-cause mortality. Existing epigenetic clocks are based on CpGs from the Illumina HumanMethylation450 BeadChip (450 K) which has now been replaced by the latest platform, Illumina MethylationEPIC BeadChip (EPIC). Thus, it remains unclear to what extent EPIC contributes to increased precision and accuracy in the prediction of chronological age.

Results: We developed three blood-based epigenetic clocks for human adults using EPIC-based DNA methylation (DNAm) data from the Norwegian Mother, Father and Child Cohort Study (MoBa) and the Gene Expression Omnibus (GEO) public repository: 1) an Adult Blood-based EPIC Clock (ABEC) trained on DNAm data from MoBa (*n* = 1592, age-span: 19 to 59 years), 2) an extended ABEC (eABEC) trained on DNAm data from MoBa and GEO (*n* = 2227, age-span: 18 to 88 years), and 3) a common ABEC (cABEC) trained on the same training set as eABEC but restricted to CpGs common to 450 K and EPIC. Our clocks showed high precision (Pearson correlation between chronological and epigenetic age (r) > 0.94) in independent cohorts, including GSE111165 (*n* = 15), GSE115278 (*n* = 108), GSE132203 (*n* = 795), and the Epigenetics in Pregnancy (EPIPREG) study of the STORK Groruddalen Cohort (*n* = 470). This high precision is unlikely due to the use of EPIC, but rather due to the large sample size of the training set.

Conclusions: Our ABECs predicted adults' chronological age precisely in independent cohorts. As EPIC is now the dominant platform for measuring DNAm, these clocks will be useful in further predictions of chronological age, age-related diseases, and mortality.

Keywords: DNA methylation, Epigenetic age, Chronological age, Illumina MethylationEPIC BeadChip, MoBa

* Correspondence: Yunsung.Lee@fhi.no
† Astanand Jugessur and Jon Bohlin are Joint last authors
[1] Department of Genetics and Bioinformatics, Norwegian Institute of Public Health, Oslo, Norway
[2] Institute of Health and Society, Faculty of Medicine, University of Oslo, Oslo, Norway
Full list of author information is available at the end of the article

Background

Aging is a biological phenomenon that is characterized by reduced functional capacity [1, 2]. Because chronological age is an imperfect surrogate of aging [3–6], the concept of biological aging that can capture the different rate of functional deterioration across individuals has been suggested [1]. Given the significance of biological aging, a variety of predictors of biological age have been constructed based on known hallmarks of aging [6, 7], including telomere length [8], metabolic rate [9], DNA methylation (DNAm) [10], CD4+ and CD8+ T cell ratio [11], proteomic alterations [12], and gut microbiota [13]. Among these, DNAm-based estimators of chronological age (referred to as epigenetic clocks) have garnered the most interest due to their remarkable precision in estimating chronological age, age-related diseases, and all-cause mortality [4, 14–18].

Epigenetic age is a linear combination of DNAm levels at specific CpGs, which are weighted by their respective coefficients estimated through an epigenetic clock. Most of the previously published epigenetic clocks (the Hannum Blood-based clock [19], Horvath Pan-tissue clock [20], Levine PhenoAge clock [16], and Horvath Skin & Blood clock [3]) were based on specific CpGs from the Illumina HumanMethylation450 BeadChip (450 K). This platform has recently been replaced by the Illumina MethylationEPIC BeadChip (EPIC). EPIC is a major improvement over its predecessor, 450 K (> 450,000 CpGs), in terms of the number of probes (> 850,000 CpGs) and the genomic coverage of regulatory elements [21]. To our knowledge, only one EPIC-based epigenetic clock has been published (the Alsaleh EPIC clock [22]). This clock was trained on a relatively small training set and was not sufficiently validated in independent cohorts. Thus, it remains unclear to what extent EPIC contributes to increased precision and accuracy in the prediction of chronological age.

We developed three blood-based epigenetic clocks for human adults: 1) an Adult Blood-based EPIC Clock (ABEC) trained on EPIC-derived DNAm data from adult peripheral blood in a sub-study of the Norwegian Mother, Father and Child Cohort Study (MoBa) [23] called the STudy of Assisted Reproductive Technology (MoBa-START); 2) an extended ABEC (eABEC) trained on MoBa-START and publicly available DNAm data from the Gene Expression Omnibus (GEO) with the aim of improving the performance of ABEC; and 3) a common ABEC (cABEC) trained on the same training set as eABEC but restricted to CpGs common to 450 K and EPIC. The purpose of cABEC was to determine whether the additional CpGs on EPIC improved predictions of chronological age. We validated our clocks and the other published clocks (the Hannum Blood-based clock, Horvath Pan-tissue clock, Levine PhenoAge clock, Horvath Skin & Blood clock, Alsaleh EPIC clock, and Zhang clock) in EPIC-derived DNAm data from independent cohorts, including publicly available DNAm data from GEO and the Epigenetics in Pregnancy (EPIPREG) study of the STORK Groruddalen Cohort (STORK) [24].

Results

Peripheral blood-based DNA methylation

We trained an epigenetic clock using elastic net regression on DNAm data from 1592 adults who were mothers and fathers in MoBa-START (796 women and 796 men). The chronological age of these adults ranged from 19 to 59 years (19 to 46 years for women and 19 to 59 years for men). DNAm on these individuals was measured using EPIC. For the current analyses, we focused on the 770,586 autosomal CpGs that remained after quality control (see Methods). Table 1 provides additional details regarding the MoBa-START samples.

Adult blood-based EPIC clock (ABEC)

Figure 1 summarizes our analysis flow.

We developed ABEC using a blood-based DNAm dataset consisting of adults (training set $n = 1592$, Table 1, Fig. 1). We used elastic net regression [32] to select the most predictive CpGs for chronological age. The resulting regression comprised 1695 CpGs. The predicted DNAm age was calculated using the following equation:

$$DNAm\ Age_j = \hat{\beta}_{(Intercept)} + X_{cg1,j}\hat{\beta}_{cg1} + X_{cg2,j}\hat{\beta}_{cg2} + \dots + X_{cg1695,j}\hat{\beta}_{cg1695},$$

where $DNAm\ Age_j$ is the epigenetic age of the j th individual, and $X_{cgi,\ j}$ refers to the DNAm level of the j th individual at the i th CpG site. The estimated intercept and beta coefficients are provided in Supplementary File 1.

Figure 2 shows the performance of ABEC in the training set ($n = 1592$, Fig. 2a) and the test set ($n = 424$, Fig. 2b). The prediction precision was quantified using the Pearson correlation coefficient (r) between DNAm age and chronological age. The prediction accuracy was quantified using the median absolute deviation (MAD) between DNAm age and chronological age. ABEC showed high precision and accuracy in both of the training (r = 0.999, MAD = 0.14, Fig. 2a) and test set (r = 0.95, MAD = 1.13, Fig. 2b). The red line in Fig. 2a and b represents a perfect correlation between chronological age and DNAm age, and the dotted line refers to the regression of the predicted DNAm age on chronological age.

Despite its overall high precision, ABEC slightly underestimated the age of the older individuals, particularly those above 45 years of age (Fig. 2c, d). This bias is expected given that the MoBa-START dataset is a pregnancy cohort with few individuals older than 45 years. In addition, most individuals aged 45 years or older were

Table 1 Description of the peripheral whole-blood-derived DNAm data on the EPIC platform

Cohort	Tissue type	Platform	GEO submitter	N	Normalization Method[a]	Probe exclusion Criteria[b]	Age range (years)
ABEC							
Training data							
MoBa-START	Peripheral whole blood	EPIC	–	1592	BMIQ	SC, CH, DP, SNP	19–59
Test data							
MoBa-START	Peripheral whole blood	EPIC	–	424	BMIQ	SC, CH, DP, SNP	20–58
eABEC							
Training data							
MoBa-START	Peripheral whole blood	EPIC	–	1592	BMIQ	SC, CH, DP, SNP	19–59
GSE116339	Peripheral whole blood	EPIC	Curtis et al. [25]	635	Noob	SC	23–88
Test data							
MoBa-START	Peripheral whole blood	EPIC	–	424	BMIQ	SC, CH, DP, SNP	20–58
GSE111165	Peripheral whole blood	EPIC	Shinozaki et al. [26]	15	Noob	SC	24–61
GSE115278	Peripheral whole blood	EPIC	Arpon et al. [27]	108	Noob	SC	19–66
Other test data							
EPIPREG	Peripheral whole blood	EPIC	–	470	FunNorm	SC, CH, DP, SNP	19–42
GSE132203	Peripheral whole blood	EPIC	Kilaru et al. [28]	795	Noob	SC	18–76

[a] Pre-processing method for quantifying DNAm levels in the range of 0 to 1
Noob Normal-exponential out-of-band [29]
BMIQ Beta-mixture quantile dilation [30]
FunNorm Functional normalization [31]
[b] Probe exclusion criteria
SC Sex chromosome, *CH* cross-hybridizing, *DP* detection *P*-value < 0.01 and *SNP* single-nucleotide polymorphism

males, which may introduce a sex-bias in the prediction of chronological age.

Extended adult blood EPIC clock (eABEC)

To reduce the underestimation bias and improve the precision of ABEC among older individuals in the MoBa-START dataset, we developed an extended ABEC (eABEC) by adding a publicly available DNAm dataset, GSE116339 ($n = 635$) [25], from the GEO data repository (https://www.ncbi.nlm.nih.gov/geo/) [33] to the original training set for ABEC (Fig. 3). This increased the total sample size of the new training set to 2227. Elastic net regression was used in the same manner as for ABEC above, and for this training set, the number of selected CpGs was 1791.

We validated eABEC in an extended test set consisting of the test set for ABEC and two independent cohorts (GSE111165 and GSE115278) from GEO. We selected these GEO datasets because they were EPIC-derived blood-based DNAm data with a wide age span (20 to 70 years). The inclusion of GSE116339 substantially improved the prediction in individuals aged 45 years and above (Fig. 3a, b), but there was a slight underestimation of age among individuals aged 65 years or older in both the training and test set (Fig. 3c, d).

Advantage of EPIC in developing epigenetic clocks

One major difference between our epigenetic clocks (ABEC and eABEC) and the previously published clocks was the use of EPIC for the training set. The training set of the other epigenetic clocks was mostly based on 450 K, except for the Horvath Skin & Blood clock which used both 450 K and EPIC-derived DNAm data. To assess whether EPIC-derived DNAm data yield a more accurate and precise clock, we trained a third epigenetic clock using the same training set as for eABEC but using only the 397,473 autosomal CpG sites that are in common between EPIC and 450 K. We refer to this third clock as 'common' ABEC (cABEC) hereafter. Elastic net regression selected 1892 CpG sites.

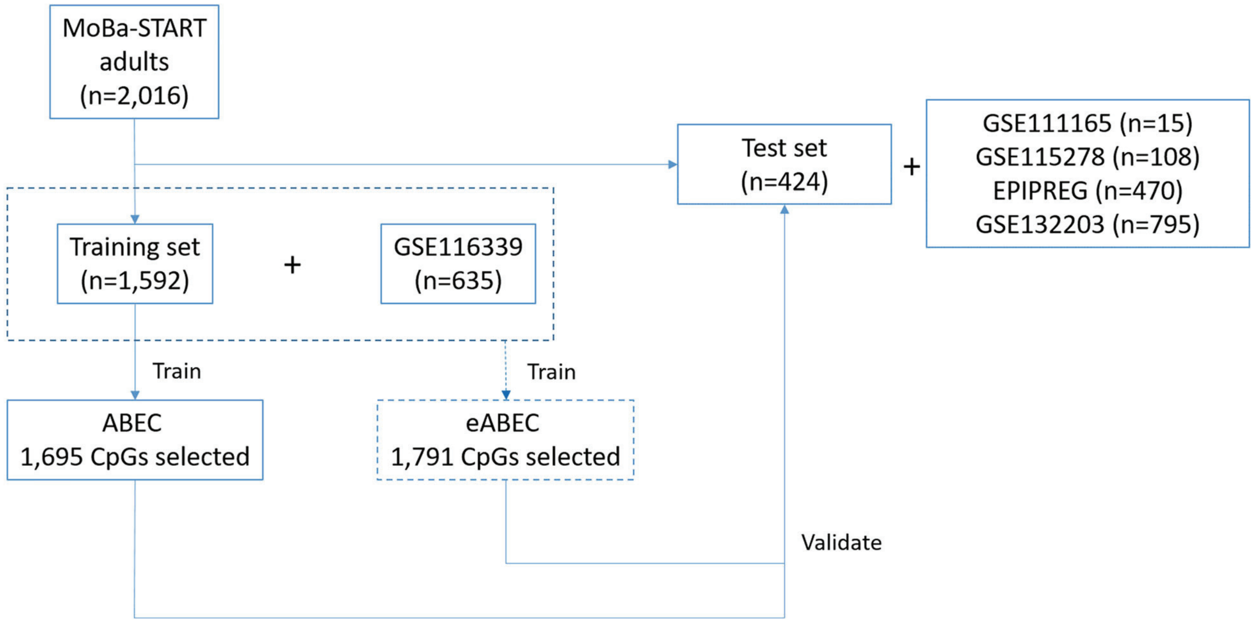


Fig. 1 Analysis flow. MoBa-START adults were randomly assigned to a training and a test set

cABEC showed a high prediction performance, similar to eABEC (Supplementary File 2, S-Figure 1). The precision metric (r) of cABEC was identical to that of eABEC. However, compared to eABEC, the accuracy of cABEC in the test set was slightly diminished (MAD = 1.25 → 1.3).

We hypothesized that the denser EPIC array might be beneficial in developing an epigenetic clock with a smaller training set. To address this point, two types of epigenetic clocks (one using all the CpGs on EPIC and the other using the CpGs common to EPIC and 450 K) were trained on random subsets of the training set of eABEC and validated in the test sample of eABEC (see Methods for further details). Both types of epigenetic clocks showed a remarkable improvement in precision and accuracy as the sample size of the training set increased (Fig. 4). However, across all the reduced training sets, the epigenetic clock based on all the CpGs on EPIC did not outperform the other clock based on the CpGs common to EPIC and 450 K (Fig. 4). This indicates that the additional CpGs on EPIC do not enhance the accuracy or precision of the epigenetic clocks when the training set is reduced.

Validation of ABECs and other epigenetic clocks

Using an independent cohort from GEO ($n = 123$), we evaluated the performance of ABEC, eABEC, and cABEC against six published epigenetic clocks: the Hannum Blood-based clock [19], Horvath Pan-tissue clock [20], Levine PhenoAge clock [16], Horvath Skin & Blood clock [3], Alsaleh EPIC clock [22], and Zhang clock [34]. The independent test set consisted of GSE111165 [26] and GSE115278 [27] from the GEO database (see Table 1 for details). None of these GEO datasets have previously been used to train any epigenetic clocks.

Figure 5 summarizes the results of epigenetic age prediction by ABEC, eABEC, cABEC, and the six published epigenetic clocks mentioned above. Our eABEC and the Zhang clock showed the highest precision (r = 0.96), followed by ABEC (r = 0.95), cABEC (r = 0.95), the Horvath Skin & Blood clock (r = 0.94), and the Hannum Blood-based epigenetic clock (r = 0.87). The 95% confidence intervals of the r values can be found in Supplementary File 2 (S-Table 1). Here, we note that only the precision metric (r) was presented in Fig. 5 because the dots in the scatter plots could deviate systematically from the 45-degree line (so-called systematic offset) but still form a very tight prediction, e.g., panel (D) in Fig. 5. In such cases where high precision and relatively low accuracy are present, the systematic offset can be calibrated using a linear transformation, or, if necessary, a non-linear transformation.

An important distinction of ABECs from the other published clocks is that they are based on an ethnically homogeneous training set (MoBa-START and GSE116339 comprised individuals of European ancestry). We validated ABEC, eABEC, cABEC, and the other published epigenetic clocks in the EPIC-derived blood-based DNAm data from EPIPREG ($n = 470$; 305 European women and 165 South Asian women, Fig. 6), a sub-study of the STORK Groruddalen Cohort [24]. ABEC, eABEC, cABEC, the Horvath Skin & Blood clock, and Zhang clock showed the highest precisions (r > 0.9). More interestingly, eABEC

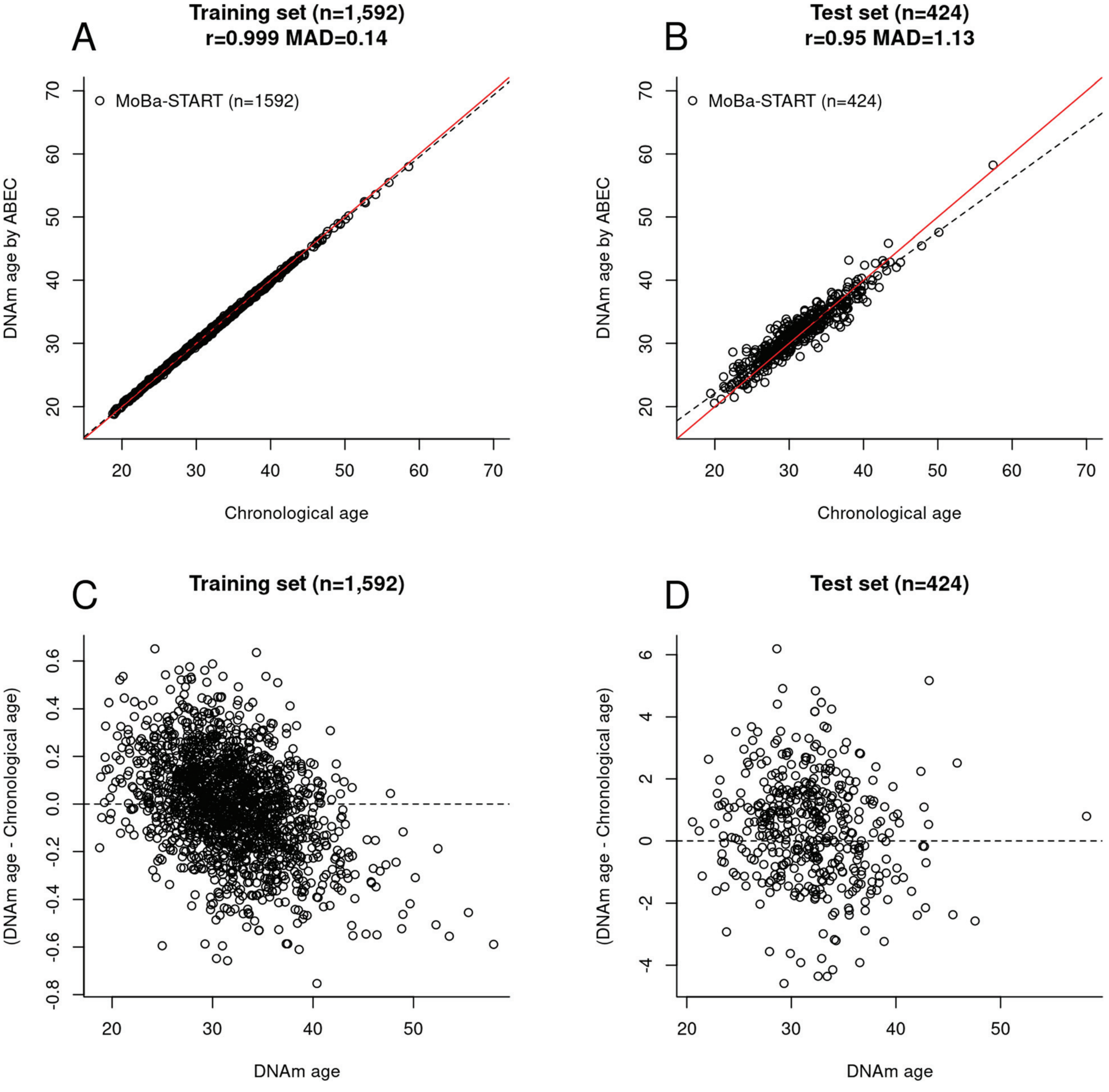


Fig. 2 Chronological age estimation by ABEC. **a** Scatter plot of chronological age against DNAm age estimated by ABEC in the training set. **b** Scatter plot of chronological age against DNAm age estimated by ABEC in the test set. **c** Residual plot in the training set. **d** Residual plot in the test set. The red line in panels (**a**) and (**b**) represents a perfect correlation between chronological age and DNAm age, and the dotted line is the regression of DNAm age on chronological age

showed that the epigenetic age acceleration (EAA; residuals from the regression of DNAm age on chronological age) was higher in South Asian women than in Norwegian women (+ 0.51 years, $P = 0.0015$, Supplementary File 2, S-Figure 2A). EAA derived by the Alsaleh EPIC clock was also elevated in South Asians compared to Norwegians (+ 0.25 years, $P = 4\text{E-}04$, Supplementary File 2, S-Figure 2B). However, EAAs derived by ABEC, cABEC, and the other published clocks did not show any difference between the two groups.

Given that ABEC, eABEC, and cABEC were trained on the ethnically homogeneous training set of Europeans, they may be sub-optimal for predicting chronological age in other ethnicities. To explore this further, we applied ABEC, eABEC, cABEC, and the other published epigenetic clocks to a GEO dataset comprising African Americans (GSE132203 [28]; $n = 795$, Supplementary File 2, S-Figure 3). All the clocks, except for the Alsaleh EPIC clock, showed high correlations between chronological age and epigenetic age (r > 0.86). The 95%

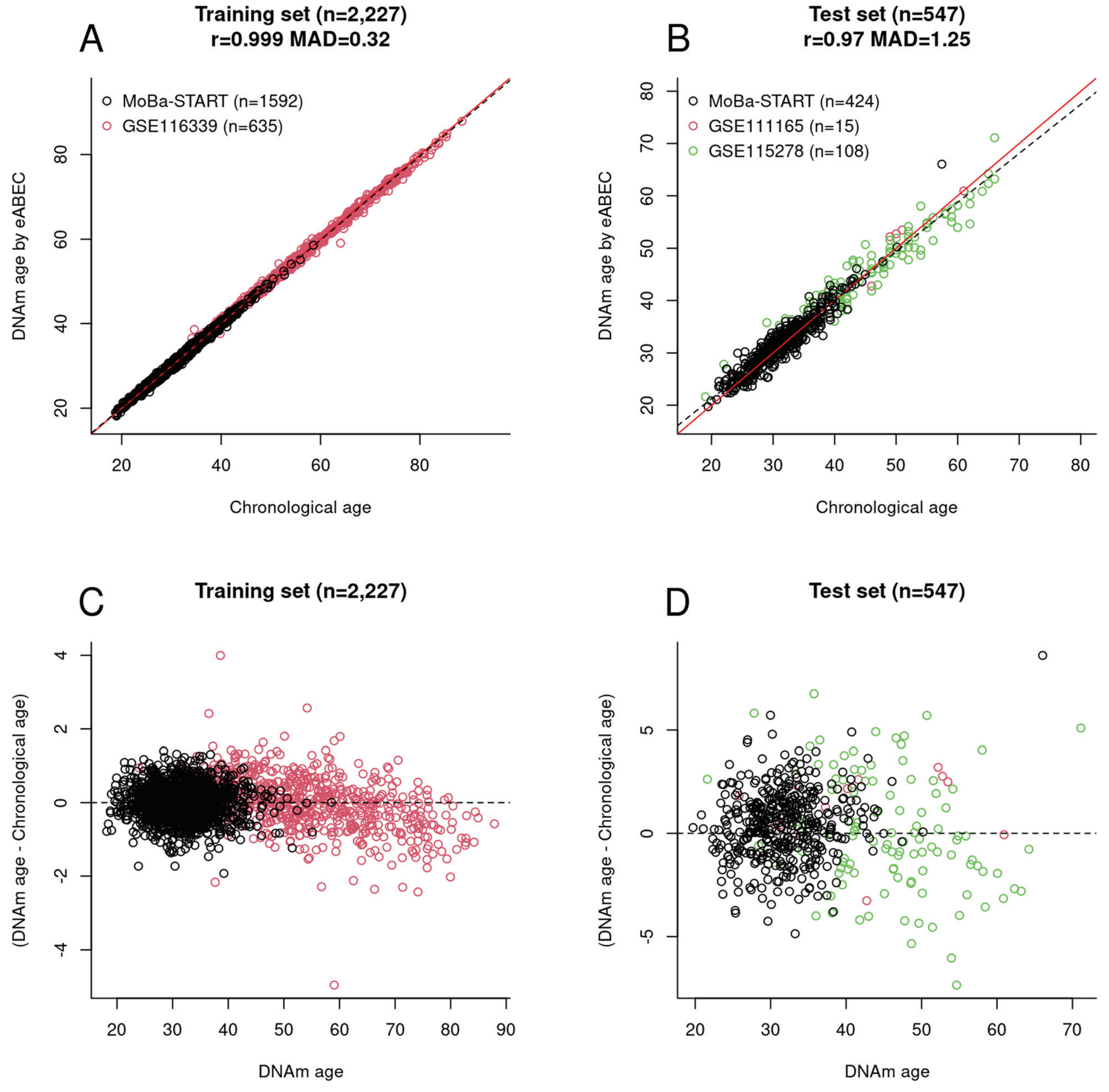


Fig. 3 Chronological age estimation by eABEC. **a** Scatter plot of chronological age against DNAm age estimated by eABEC in the extensive training set. **b** Scatter plot of chronological age against DNAm age estimated by eABEC in the test set. **c** Residual plot in the training set. **d** Residual plot in the test set. The red line in panels (**a**) and (**b**) represents a perfect correlation between chronological age and DNAm age, and the dotted line is the regression of DNAm age on chronological age

confidence intervals of the r values can be found in Supplementary File 2 (S-Table 1). eABEC, cABEC, and the Zhang clock showed the highest r of 0.96, and ABEC and the Horvath Skin & Blood clock showed the second-highest r of 0.95.

Discussion

We developed precise epigenetic clocks (ABEC and eABEC) using blood-based DNAm data from EPIC. Our epigenetic clocks showed a more precise chronological age prediction than existing blood-based epigenetic clocks (e.g., the Hannum Blood-based clock and Horvath Skin & Blood clock; Fig. 5). The reason for the higher precision is more likely due to the large training set (*n* = 2227, Table 1) and the wide age-span of the samples (19 to 88 years for the training set of eABEC, Table 1), which is consistent with the findings by Zhang and colleagues [34]. Compared to eABEC, both Hannum Blood-

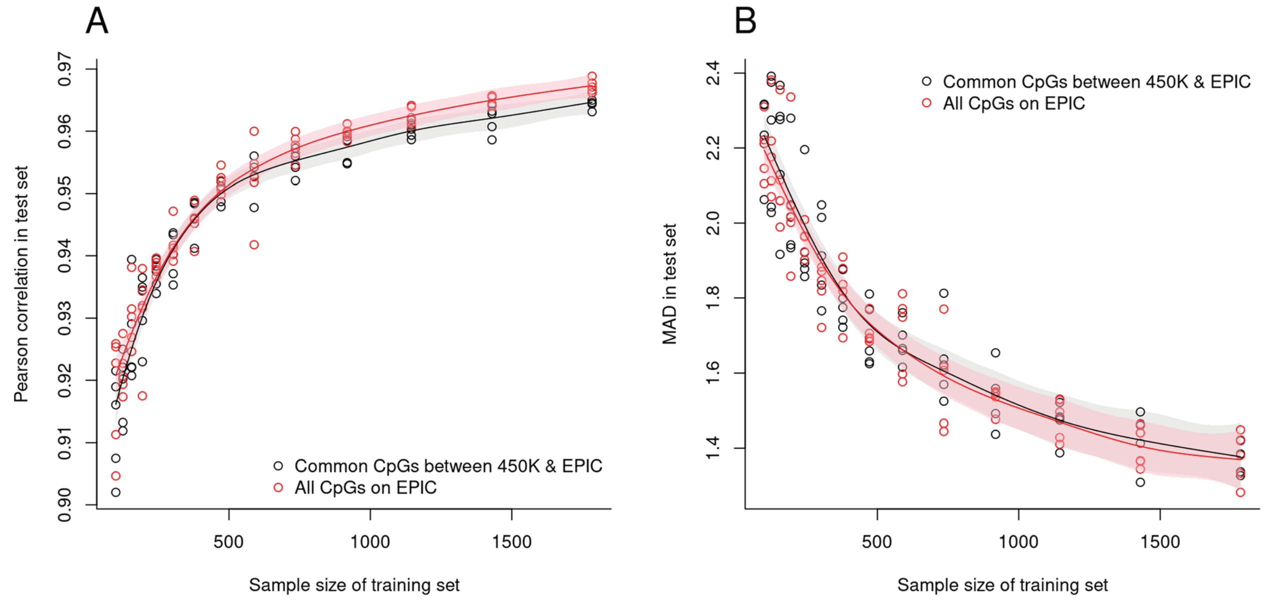


Fig. 4 Comparison of precision and accuracy between a clock based on the CpGs common to 450 K and EPIC and a clock on all the CpGs on EPIC. **a** Scatter plot of the Pearson correlation (r) in the test set against the sample size of the training set. **b** Scatter plot of MAD in the test set against the sample size of the training set. In panel (**a**), we fit the smoothing splines of the Fisher's Z-transformed r values on the sample size, derived the confidence intervals, and inverse-transformed them. In panel (**b**), we fit the smoothing splines of MAD values on the sample size without transformation. The black dots refer to the clock based on the CpGs common to 450 K and EPIC, and the red dots refer to the clock based on all the CpGs on EPIC

based clock and Horvath Skin & Blood clock were trained on fewer samples ($n = 656$ and $n = 896$, respectively) that had a wider age-span (19 to 101 years and 0 to 94 years, respectively) [3, 19]. Other clocks (the Horvath Pan-tissue clock and Levine PhenoAge clock) may not be directly comparable to eABEC for chronological age prediction. For instance, the Horvath Pan-tissue clock was designed to measure epigenetic aging not only in blood but in multiple tissues [20], and the Levine PhenoAge was designed to predict phenotypic age (estimated using 10 clinical biomarkers, e.g., albumin, creatinine, serum glucose, and seven others) based on DNAm [16].

To develop eABEC, we added GSE116339 to the training set of ABEC. GSE116339 is from a study by Curtis et al. [25] that used EPIC to measure DNAm in peripheral blood samples collected from 658 individuals of European ancestry (638 non-Hispanic and 20 Hispanic) in Michigan, USA. These individuals had been exposed to the endocrine-disrupting chemical polybrominated biphenyl when an agricultural accident introduced it into the food supply in the 1970s. We selected 635 individuals from the control group whose total PBB (PBB-153, PBB-101, PBB-77, and PBB180) exposure was lower than 5 pg/ml. The distribution of the total PBB exposure was highly right-skewed.

The high precision of eABEC cannot be attributed solely to the use of the EPIC platform as the additional 413,743 CpGs on EPIC did not improve age prediction noticeably (Fig. 4). Although the 1791 CpGs selected by eABEC included 1084 CpGs that only exist on EPIC, eABEC did not outperform cABEC that used the CpGs common to 450 K and EPIC. This indicates that 226,915 probes (out of 413,743) that are designed to cover regulatory regions (DNase proximal/distal [35] and FANTOM5 [36]) did not increase the precision of the epigenetic clocks significantly [21]. Yet, Pidsley et al. [21] reported that probes on EPIC cover 58% of FANTOM5 enhancers, 7% of distal, and 27% of proximal ENCODE regulatory regions, suggesting that the coverage of regulatory regions is still low. Thus, it is difficult to dismiss the possibility that other regulatory CpGs not currently included on EPIC might improve age prediction.

Underestimation and overestimation of epigenetic clocks should be carefully assessed using residual plots instead of scatter plots. As we regressed chronological age on DNAm levels (chronological age = DNAm levels + error), a scatter plot that displays chronological age on the x-axis and DNAm age on the y-axis may lead to the misconception that DNAm age is overestimated in the oldest age group and underestimated in the youngest age group (Supplementary File, S-Figure 4). In contrast, residual plots that display DNAm age on the x-axis and residuals (DNAm age minus chronological age) on the y-axis would enable a fair evaluation of prediction models. The strength of the current scatter plots lies in the visualization of EAA (the residuals of the regression of

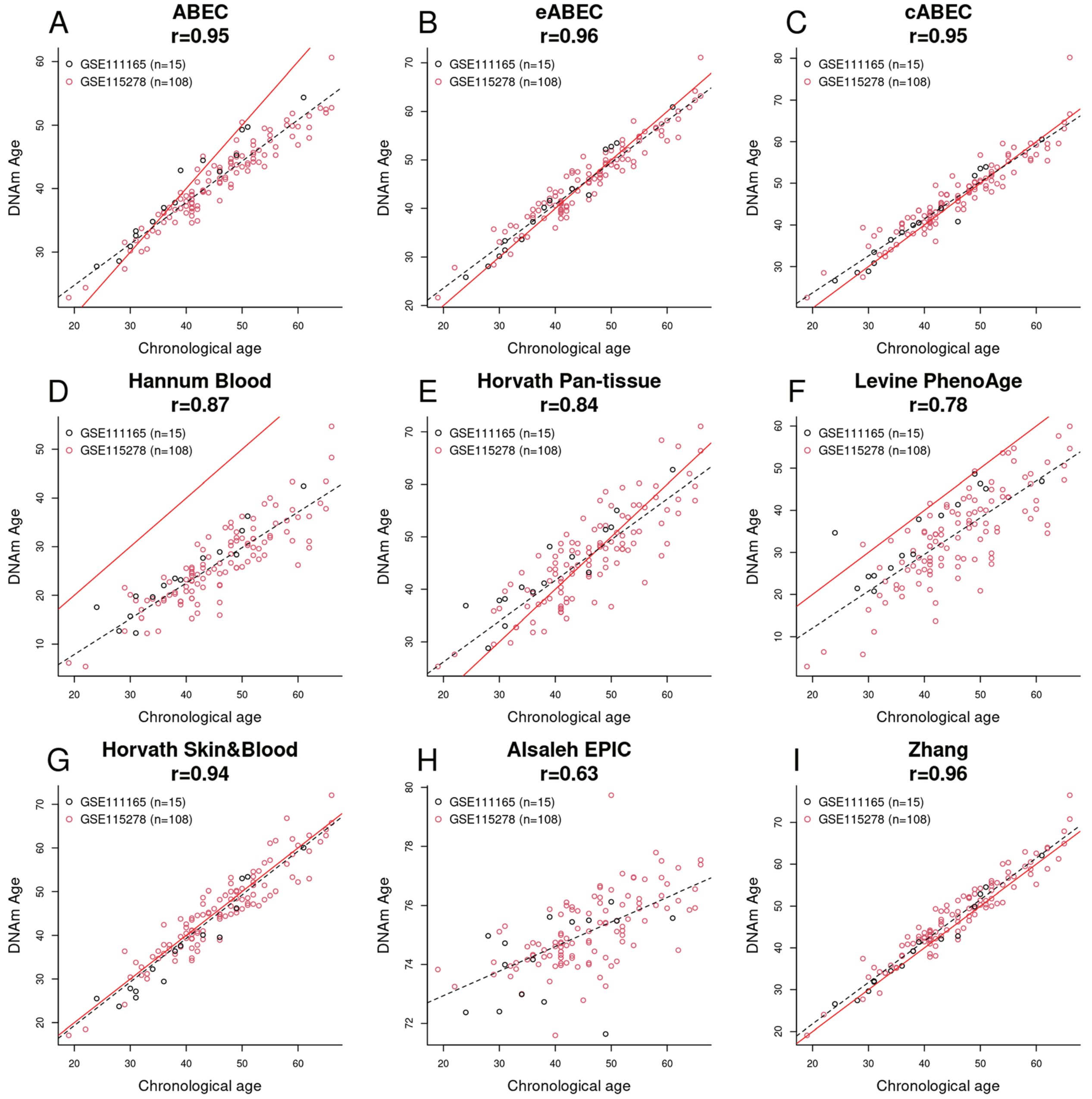


Fig. 5 Chronological age estimation by ABEC, eABEC, and the other published epigenetic age estimators. **a** ABEC, **b** eABEC, **c** Hannum Blood-based clock, **d** Horvath Pan-tissue clock, **e** Levine PhenoAge clock, **f** Horvath Skin & blood clock, **g** Alsaleh Blood-based EPIC clock (the stepwise regression), and **h** Zhang clock (elastic net regression). The red line in the panels represents a perfect correlation between chronological age and DNAm age, and the dotted line is the regression of DNAm age on chronological age

DNAm age on chronological age; i.e., the vertical distance between each dot and the dotted line in Figs. 2 and 3).

Our clocks, particularly eABEC, showed a systematic underestimation in older subjects, as was the case with the Horvath Pan-tissue clock and Hannum Blood-based clock in GSE132203 [37]. The systematic underestimation may be corrected by 1) adding more DNAm data of older subjects to the training set or 2) calibrating epigenetic clocks using a non-linear transformation (e.g., piecewise cubic regression (with a knot at 70) or smoothing spline of chronological age on DNAm age). However, we could not add more EPIC-derived DNAm data from older subjects (preferably subjects of European ancestry aged 70 to 80 years) to the training set for eABEC. We note that the underestimation in older subjects can cause EAA

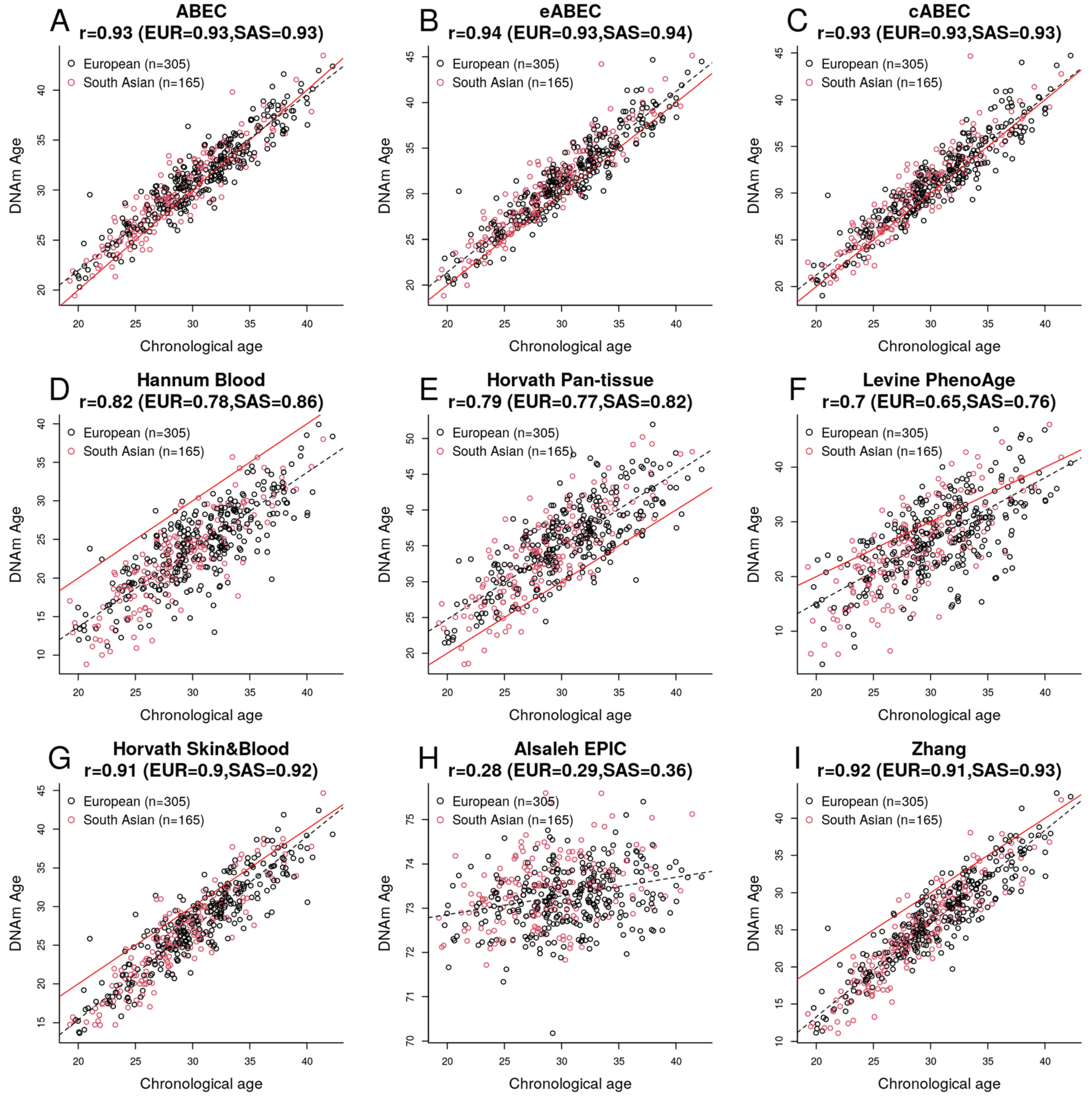


Fig. 6 Application of ABEC, eABEC, and other epigenetic clocks to DNAm data in the EPIPREG sub-study of the STORK Groruddalen cohort. The title of each panel displays the overall r as well as the ethnicity-specific r. EUR indicates the r between chronological age and DNAm age in 305 women of European ancestry, whereas SAS refers to the r between chronological age and DNAm age in 165 women of South Asian ancestry

to be dependent on chronological age. Therefore, for other researchers who are interested in the association between EAA and a given phenotype, we recommend redefining EAA (e.g., regressing DNAm age on chronological age using a piecewise cubic regression or a smoothing spline rather than an ordinary linear regression) so that EAA is independent of chronological age.

Our eABEC may result in subtle differences in EAA across different ethnic groups, e.g., Supplementary File 2, S-Figure 2A. A hypothesis explaining this bias is that the CpGs included in eABEC may be located near SNPs with a low minor allele frequency [38]. The SNPs may influence the DNAm level at the CpGs if the minor allele frequencies at the SNPs differ across ethnicities. To address this point, we added

SNP annotations generated by Zhou et al. [38] and McCartney et al. [39] to Supplementary File 1.

Conclusion

Three blood-based epigenetic clocks were developed to estimate adults' chronological age using EPIC-derived DNAm data. The precision of these clocks was high (r > 0.94) when validated in independent cohorts. The high level of precision was not explained by the broader genomic coverage of EPIC (> 850,000 CpG sites) but rather by the large training set (n = 2227) with a wide age-span (19 to 88 years).

Methods

Study population

MoBa is a nationwide pregnancy cohort study in which approximately 95,000 mothers, 75,000 fathers, and 114, 000 children were recruited from 1998 to 2008 across Norway [23]. The participants completed a series of questionnaires that are also linked to information from the Medical Birth Registry of Norway [23]. Peripheral whole-blood samples were collected from the mothers at the 17th week of gestation and at birth and from the fathers at the 17th week of gestation. Cord-blood samples were collected from newborns at birth [40, 41]. The precise chronological age in days at blood draw was calculated for the fathers and mothers. Further details on MoBa have been described in previous publications [23, 40–42]. We used data from a sub-study of MoBa (MoBa-START) with blood-based DNAm data on 2016 adults (mothers and fathers who were randomly selected among complete mother-father-newborn trios in MoBa).

GSE116339 is an epigenome-wide association study (EWAS) of polybrominated biphenyl in peripheral blood [25]. GSE111165 explored the difference in genome-wide DNAm between brain and peripheral tissues (buccal, saliva, and blood) from epilepsy patients [26]. GSE115278 is an EWAS of insulin resistance, obesity, and metabolic complications [27, 43–45]. GSE132203 examined the association between DNAm and psychiatric or stress-related symptoms [28].

EPIPREG is nested within the STORK Groruddalen Cohort study (a population-based cohort, n = 823, [24]). EPIPREG quantified DNAm in white blood cells, collected at the 28th week of gestation, from 480 women (312 of European ancestry and 168 of South Asian ancestry), using EPIC. In this study, we focused on 470 women (305 of European ancestry and 165 of South Asian ancestry) after excluding eight samples with low quality and two samples with an absolute EAA larger than 15 years. Further details of EPIPREG are described in Supplementary File 3 (S-Figure 7).

The age distributions of all the individuals included in the training and test sets can be found in Supplementary File 2 (S-Figure 5 and 6).

Pre-processing of DNA methylation

For MoBa-START, 500 nanograms of DNA stored in the MoBa Biobank (see Paltiel et al. [41] for further details of the storage of the biological samples) were shipped to LIFE & BRAIN GmbH (Bonn, Germany). The samples were bisulfite converted and processed using the EZ-96DNA methylation-Lightning™MagPrep kit (Zymo Research, Irvine, USA) according to the manufacturer's instructions. The raw iDAT files were imported and processed using the RnBeads R package [46]. 44,210 probes with cross-hybridization [39], high detection p-value (> 0.01), and 16,117 probes near single-nucleotide polymorphisms (filtering.snp = "3") were excluded. The data were run in four batches and the exclusion criteria for removing probes were applied to each batch separately. Probes that were excluded from one batch were removed from all batches. The DNAm signals at the remaining probes were control-normalized and corrected for background noise using the *wm.nasen* and *methylumi.noob* options. Additionally, among a total of 2034 non-replicated samples, we excluded 18 samples that displayed low signal intensities and deviated (outliers) from the clusters formed by principal component analysis. The two probe chemistries (Type I and Type II probes) were normalized using Beta-mixture quantile normalization (BMIQ, [30]) using the wateRmelon R package [47]. In summary, the number of remaining probes was 790,213 (770,586 from autosomes and 19, 627 from sex-chromosomes).

For the DNAm data from GEO, we downloaded the iDAT files and used normal-exponential out-of-band (Noob, [29]) normalization in the minfi R package [48]. For the DNAm data from EPIPREG, we performed functional normalization (FunNorm, [31]) using the meffil R package [49]. Further details of the DNA extraction and quality control process of EPIPREG can be found in Supplementary File 3.

Elastic net regression

Penalized regressions (glmnet R package [50]) were used to develop the three ABECs. Chronological age in days was regressed on 770,586 autosomal CpGs that remained after quality control. The mixing parameter (alpha) was set to 0.5 and the shrinkage parameter (lambda) leading to the minimum mean square error was selected after 10-fold cross-validation in the training set. Supplementary File 3 (S-Figure 8) includes cross-validation curves for lambda and alpha values. ABEC, eABEC and cABEC selected 1695 CpG sites (lambda = 0.02884886), 1791 CpG sites (lambda = 0.05281471), and 1892 CpG sites (lambda = 0.0438477), respectively. Supplementary File 1 lists these CpG sites, their corresponding coefficients for ABEC, eABEC, and cABEC, and SNP annotations generated by Zhou et al. [38] and McCartney et al. [39].

Comparison between EPIC-CpG clock and common CpG clock

The implementation resembles bootstrapping conceptually. For each of the reduced sample sizes ($n = 100$, 125, 156, 194, 243, 303, 378, 472, 589, 735, 918, 1145, 1430 and 1784; the determination of these values is detailed in Supplementary File 3), we first constructed five training sets by randomly selecting subjects from the full training set of eABEC ($n = 2227$). We made the sequence of the reduced sample sizes denser around 100 and sparser around 2227 because epigenetic clocks gradually improved their precision and accuracy when the training set was larger than 1145. On each training set, we trained two types of epigenetic clocks: one using all the CpGs on EPIC and the other using the CpGs common to EPIC and 450 K. Next, we validated these clocks in the test set of eABEC ($n = 485$) and calculated r and MAD accordingly. The mgcv R package [51] was used to fit the smoothing splines in Fig. 4. Particularly, in Fig. 4a, we fit the smoothing splines of the Fisher's Z-transformed r values ($F(r) = 0.5* \log(\frac{1+r}{1-r})$) on the sample size, derived the confidence intervals and inverse-transformed them.

Availability of epigenetic clocks

The estimated intercepts and coefficients for ABEC, eABEC, and cABEC can be found in Supplementary File 1.

The ABECs can be readily applied to any DNAm data using the following procedure: 1) generate a matrix of beta values (n individuals by p CpG sites) using a background correction method, e.g., Noob (preferably) without any batch adjustment (Supplementary File 3), 2) select the CpG sites for the ABECs (Supplementary File 1) out of the matrix of beta values, 3) calculate the linear combination of the beta values at the selected CpG sites, and 4) add the estimated intercept (Supplementary File 1) to the linear combination.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-020-07168-8.

Additional file 1. This file includes CpG sites for ABEC, eABEC, and cABEC, their corresponding coefficients, overlap with the other published clocks, genomic locations, neighboring genes, presence in the Illumina HumanMethylation450K and 27 K array, and the SNP annotations generated by Zhou et al. [38] (with the suffix of "Zhou") and McCartney et al. [39] (with the suffix of "McCartney").

Additional file 2. This file includes 1) a figure displaying the age prediction of cABEC, 2) a table containing the bootstrapped 95% confidence intervals for the r values in Figs. 4, 5 and 6) figures displaying the age prediction of the ABECs and the other published clocks in EPIPREG and GSE132203, 4) a figure illustrating the regression-to-the-mean effect and 5) histograms displaying the age distribution of individuals in each cohort.

Additional file 3. This file includes 1) further details (sample selection, DNA extraction, and quality control) of EPIPREG, 2) cross-validation curves of mean squared error over lambda and alpha values for eABEC, 3) determination of the reduced sample sizes for Fig. 4, and 4) further information regarding batch adjustment in developing the ABECs.

Abbreviations

DNAm: DNA methylation; ABEC: Adult Blood EPIC Clock; eABEC: extended Adult Blood EPIC Clock; cABEC: common Adult Blood EPIC Clock; GEO: Gene Expression Omnibus; EPIC: Illumina MethylationEPIC BeadChip; 450 K: Illumina HumanMethylation 450 K BeadChip; r: The Pearson correlation coefficient; MAD: Median absolute deviation; BMIQ: Beta-mixture quantile dilation; MoBa: Norwegian Mother, Father and Child Cohort Study; START: Study of Assisted Reproductive Technology; EPIPREG: Epigenetics in Pregnancy; EAA: Epigenetic age acceleration

Acknowledgements

MoBa is supported by the Norwegian Ministry of Health and Care Services and the Ministry of Education and Research. We are grateful to all the families in Norway who take part in this on-going cohort study. We also appreciate all the GEO submitters for making their DNAm data publically available.

Authors' contributions

YL developed the three ABECs and wrote the manuscript. HEN, CP, and RL performed quality control on the DNAm data from MoBa-START. CS initiated EPIPREG and leads the EPIPREG group. GHM performed the quality control and normalization of the EPIPREG samples. SL-Ø validated all the epigenetic clocks in the DNAm data from EPIPREG. LS participated in data acquisition in the STORK Groruddalen study. RBP and LCG performed chip analysis for EPIPREG. HKG provided statistical interpretations. SEH and PM initiated START and MoBa, respectively. All the authors contributed to data interpretation, provided overall scientific input, and revised the manuscript. The author(s) read and approved the final manuscript.

Funding

This work is supported in part by a grant from the Research Council of Norway (RCN) to AJ, JRH, PM, and HKG (grant 262043). The work of KLH, HEN, CP, RL, MCM, HKG, JRH, PM, SEH, AJ, and JB was supported in part by RCN through its Centres of Excellence funding scheme, project number 262700 (awarded to SEH and PM). GHM is supported by RCN (Postdoctoral mobility research grant 287198). The funding body played no role in the design of the study, analysis or interpretation of data, nor in writing the manuscript.

Availability of data and materials

The MoBa data can be accessed by applying directly to the Norwegian Institute of Public Health, http://www.fhi.no/en/. The EPIPREG data can be accessed by contacting Dr. Christine Sommer, Oslo University Hospital, https://www.oslodiabetes.no/christine-sommer. The publicly available DNAm data in this study (accession numbers: GSE116339, GSE111165, GSE115278, and GSE132203) are accessible on the GEO repository, https://www.ncbi.nlm.nih.gov/geo/. Public access to the GEO repository is open, and thus administrative permission to access and use the data is not needed.

Ethics approval and consent to participate

This study was approved by the Regional Committees for Medical and Health Research Ethics (REK) South-East (2017/1362) in Norway. Data collection by MoBa was carried out in accordance with the Norwegian Data Protection Agency after securing approval from REK. The participation in the STORK Groruddalen study was based on informed written consent, and the study and its sub-study, EPIPREG, were approved by the REK South-East (2015/1035).

Consent for publication

Written consents were obtained from the MoBa and STORK participants.

Competing interests

The authors declare no conflicts of interest.

Author details

[1]Department of Genetics and Bioinformatics, Norwegian Institute of Public Health, Oslo, Norway. [2]Institute of Health and Society, Faculty of Medicine,

University of Oslo, Oslo, Norway. [3]Centre for Fertility and Health, Norwegian Institute of Public Health, Oslo, Norway. [4]Department of Global Public Health and Primary Care, University of Bergen, N-5020 Bergen, Norway. [5]Deepinsight, Karl Johans gate 8, Oslo, Norway. [6]Oslo Centre for Biostatistics and Epidemiology, Section for Research Support, Oslo University Hospital, Oslo, Norway. [7]Department of Medical Genetics, Oslo University Hospital, Oslo, Norway. [8]PharmaTox Strategic Research Initiative, School of Pharmacy, Faculty of Mathematics and Natural Sciences, University of Oslo, Oslo, Norway. [9]Department of Internal Medicine, Akershus University Hospital, Kongsvinger, Norway. [10]Department of transplantation medicine, Institute of Clinical medicine, University of Oslo, Oslo, Norway. [11]Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway. [12]The University of Queensland Diamantina Institute, University of Queensland, Woolloongabba, QLD 4102, Australia. [13]K.G. Jebsen Center for Genetic Epidemiology, Department of Public Health and Nursing, Norwegian University of Science and Technology, Trondheim, Norway. [14]Population Health Science, Bristol Medical School, University of Bristol, Bristol, UK. [15]Department of Clinical Sciences, Clinical Research Centre, Lund University, Malmö, Sweden. [16]Finnish Institute of Molecular Medicine, Helsinki University, Helsinki, Finland. [17]Department of Pediatric and Adolescents Medicine, Akershus University Hospital, Lørenskog, Norway. [18]Institute of Clinical Medicine, University of Oslo, Campus AHUS, Lørenskog, Norway. [19]Department of Endocrinology, Morbid Obesity and Preventive Medicine, Oslo University Hospital, Oslo, Norway. [20]MRC Integrative Epidemiology Unit at the University of Bristol, Bristol, UK. [21]Division for Infection Control and Environmental Health, Department of Infectious Disease Epidemiology and Modelling, Norwegian Institute of Public Health, Oslo, Norway.

Received: 12 March 2020 Accepted: 20 October 2020
Published online: 27 October 2020

References

1. Baker GT 3rd, Sprott RL. Biomarkers of aging. Exp Gerontol. 1988;23(4–5): 223–39.
2. Warner HR. Current status of efforts to measure and modulate the biological rate of aging. J Gerontol A Biol Sci Med Sci. 2004;59(7):692–6.
3. Horvath S, Oshima J, Martin GM, Lu AT, Quach A, Cohen H, Felton S, Matsuyama M, Lowe D, Kabacik S, et al. Epigenetic clock for skin and blood cells applied to Hutchinson Gilford Progeria syndrome and ex vivo studies. Aging (Albany NY). 2018;10(7):1758–75.
4. Jylhava J, Pedersen NL, Hagg S. Biological age predictors. EBioMedicine. 2017;21:29–36.
5. Melzer D, Pilling LC, Ferrucci L. The genetics of human ageing. Nat Rev Genet. 2020;21(2):88–101.
6. Bell CG, Lowe R, Adams PD, Baccarelli AA, Beck S, Bell JT, Christensen BC, Gladyshev VN, Heijmans BT, Horvath S, et al. DNA methylation aging clocks: challenges and recommendations. Genome Biol. 2019;20(1):249.
7. Lopez-Otin C, Blasco MA, Partridge L, Serrano M, Kroemer G. The hallmarks of aging. Cell. 2013;153(6):1194–217.
8. Aubert G, Lansdorp PM. Telomeres and aging. Physiol Rev. 2008;88(2):557–79.
9. Johnson LC, Parker K, Aguirre BF, Nemkov TG, D'Alessandro A, Johnson SA, Seals DR, Martens CR. The plasma metabolome as a predictor of biological aging in humans. Geroscience. 2019;41(6):895–906.
10. Bocklandt S, Lin W, Sehl ME, Sanchez FJ, Sinsheimer JS, Horvath S, Vilain E. Epigenetic predictor of age. PLoS One. 2011;6(6):e14821.
11. Jagger A, Shimojima Y, Goronzy JJ, Weyand CM. Regulatory T cells and the immune aging process: a mini-review. Gerontology. 2014;60(2):130–7.
12. Lehallier B, Gate D, Schaum N, Nanasi T, Lee SE, Yousef H, Moran Losada P, Berdnik D, Keller A, Verghese J, et al. Undulating changes in human plasma proteome profiles across the lifespan. Nat Med. 2019;25(12):1843–50.
13. Odamaki T, Kato K, Sugahara H, Hashikura N, Takahashi S, Xiao JZ, Abe F, Osawa R. Age-related changes in gut microbiota composition from newborn to centenarian: a cross-sectional study. BMC Microbiol. 2016;16:90.
14. Dugue PA, Bassett JK, Joo JE, Jung CH, Ming Wong E, Moreno-Betancur M, Schmidt D, Makalic E, Li S, Severi G, et al. DNA methylation-based biological aging and cancer risk and survival: pooled analysis of seven prospective studies. Int J Cancer. 2018;142(8):1611–9.
15. 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(9):690–700.
16. Levine ME, Lu AT, Quach A, Chen BH, Assimes TL, Bandinelli S, Hou L, Baccarelli AA, Stewart JD, Li Y, et al. An epigenetic biomarker of aging for lifespan and healthspan. Aging (Albany NY). 2018;10(4):573–91.
17. Lu AT, Quach A, Wilson JG, Reiner AP, Aviv A, Raj K, Hou L, Baccarelli AA, Li Y, Stewart JD, et al. DNA methylation GrimAge strongly predicts lifespan and healthspan. Aging (Albany NY). 2019;11(2):303–27.
18. Marioni RE, Shah S, McRae AF, Chen BH, Colicino E, Harris SE, Gibson J, Henders AK, Redmond P, Cox SR, et al. DNA methylation age of blood predicts all-cause mortality in later life. Genome Biol. 2015;16:25.
19. Hannum G, Guinney J, Zhao L, Zhang L, Hughes G, Sadda S, Klotzle B, Bibikova M, Fan JB, Gao Y, et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. Mol Cell. 2013;49(2):359–67.
20. Horvath S. DNA methylation age of human tissues and cell types. Genome Biol. 2013;14(10):R115.
21. Pidsley R, Zotenko E, Peters TJ, Lawrence MG, Risbridger GP, Molloy P, Van Djik S, Muhlhausler B, Stirzaker C, Clark SJ. Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. Genome Biol. 2016;17(1):208.
22. Alsaleh H, Haddrill PR. Identifying blood-specific age-related DNA methylation markers on the Illumina MethylationEPIC(R) BeadChip. Forensic Sci Int. 2019;303:109944.
23. Magnus P, Birke C, Vejrup K, Haugan A, Alsaker E, Daltveit AK, Handal M, Haugen M, Hoiseth G, Knudsen GP, et al. Cohort profile update: the Norwegian mother and child cohort Study (MoBa). Int J Epidemiol. 2016;45(2):382–8.
24. Jenum AK, Sletner L, Voldner N, Vangen S, Morkrid K, Andersen LF, Nakstad B, Skrivarhaug T, Rognerud-Jensen OH, Roald B, et al. The STORK Groruddalen research programme: a population-based cohort study of gestational diabetes, physical activity, and obesity in pregnancy in a multiethnic population. Rationale, methods, study population, and participation rates. Scand J Public Health. 2010;38(5 Suppl):60–70.
25. Curtis SW, Cobb DO, Kilaru V, Terrell ML, Kennedy EM, Marder ME, Barr DB, Marsit CJ, Marcus M, Conneely KN, et al. Exposure to polybrominated biphenyl (PBB) associates with genome-wide DNA methylation differences in peripheral blood. Epigenetics. 2019;14(1):52–66.
26. Braun PR, Han S, Hing B, Nagahama Y, Gaul LN, Heinzman JT, Grossbach AJ, Close L, Dlouhy BJ, Howard MA 3rd, et al. Genome-wide DNA methylation comparison between live human brain and peripheral tissues within individuals. Transl Psychiatry. 2019;9(1):47.
27. Arpon A, Milagro FI, Santos JL, Garcia-Granero M, Riezu-Boj JI, Martinez JA. Interaction among sex, aging, and epigenetic processes concerning visceral fat, insulin resistance, and Dyslipidaemia. Front Endocrinol (Lausanne). 2019;10:496.
28. Kilaru V. GSE132203, DNA Methylation (EPIC) from the Grady Trauma Project; 2019.
29. 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(7):e90.
30. Teschendorff AE, Marabita F, Lechner M, Bartlett T, Tegner J, Gomez-Cabrero D, Beck S. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. Bioinformatics. 2013;29(2):189–96.
31. 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(12):503.
32. Zou H, Hastie T. Regularization and variable selection via the elastic net. J R Stat Soc B. 2005;67(2):301–20.
33. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Holko M, et al. NCBI GEO: archive for functional genomics data sets--update. Nucleic Acids Res. 2013;41(Database issue):D991–5.
34. Zhang Q, Vallerga CL, Walker RM, Lin T, Henders AK, Montgomery GW, He J, Fan D, Fowdar J, Kennedy M, et al. Improved precision of epigenetic clock estimates across tissues and its implication for biological ageing. Genome Med. 2019;11(1):54.
35. Thurman RE, Rynes E, Humbert R, Vierstra J, Maurano MT, Haugen E, Sheffield NC, Stergachis AB, Wang H, Vernot B, et al. The accessible chromatin landscape of the human genome. Nature. 2012;489(7414):75–82.
36. Andersson R, Gebhard C, Miguel-Escalada I, Hoof I, Bornholdt J, Boyd M, Chen Y, Zhao X, Schmidl C, Suzuki T, et al. An atlas of active enhancers across human cell types and tissues. Nature. 2014;507(7493):455–61.
37. El Khoury LY, Gorrie-Stone T, Smart M, Hughes A, Bao Y, Andrayas A, Burrage J, Hannon E, Kumari M, Mill J, et al. Systematic underestimation of the epigenetic clock and age acceleration in older subjects. Genome Biol. 2019;20(1):283.

38. Zhou W, Laird PW, Shen H. Comprehensive characterization, annotation and innovative use of Infinium DNA methylation BeadChip probes. Nucleic Acids Res. 2017;45(4):e22.
39. McCartney DL, Walker RM, Morris SW, McIntosh AM, Porteous DJ, Evans KL. Identification of polymorphic and off-target probe binding sites on the Illumina Infinium MethylationEPIC BeadChip. Genom Data. 2016;9:22–4.
40. Ronningen KS, Paltiel L, Meltzer HM, Nordhagen R, Lie KK, Hovengen R, Haugen M, Nystad W, Magnus P, Hoppin JA. The biobank of the Norwegian mother and child cohort Study: a resource for the next 100 years. Eur J Epidemiol. 2006;21(8):619–25.
41. Paltiel L, Anita H, Skjerden T, Harbak K, Bække S, Kristin SN, Knudsen GP, Magnus P. The biobank of the Norwegian Mother and Child Cohort Study– present status. Norsk epidemiologi. 2014;24:1–2.
42. Magnus P, Irgens LM, Haug K, Nystad W, Skjaerven R, Stoltenberg C, MoBa Study G. Cohort profile: the Norwegian mother and child cohort Study (MoBa). Int J Epidemiol. 2006;35(5):1146–50.
43. Salas-Perez F, Ramos-Lopez O, Mansego ML, Milagro FI, Santos JL, Riezu-Boj JI, Martinez JA. DNA methylation in genes of longevity-regulating pathways: association with obesity and metabolic complications. Aging (Albany NY). 2019;11(6):1874–99.
44. Arpon A, Milagro FI, Ramos-Lopez O, Mansego ML, Santos JL, Riezu-Boj JI, Martinez JA. Epigenome-wide association study in peripheral white blood cells involving insulin resistance. Sci Rep. 2019;9(1):2445.
45. Arpon A, Milagro FI, Ramos-Lopez O, Mansego ML, Riezu-Boj JI, Martinez JA, Project M. Methylome-Wide Association Study in Peripheral White Blood Cells Focusing on Central Obesity and Inflammation. Genes (Basel). 2019; 10(6):444.
46. Muller F, Scherer M, Assenov Y, Lutsik P, Walter J, Lengauer T, Bock C. RnBeads 2.0: comprehensive analysis of DNA methylation data. Genome Biol. 2019;20(1):55.
47. Pidsley R, YW CC, Volta M, Lunnon K, Mill J, Schalkwyk LC. A data-driven approach to preprocessing Illumina 450K methylation array data. BMC Genomics. 2013;14:293.
48. 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(10):1363–9.
49. Min JL, Hemani G, Davey Smith G, Relton C, Suderman M. Meffil: efficient normalization and analysis of very large DNA methylation datasets. Bioinformatics. 2018;34(23):3983–9.
50. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. J Stat Softw. 2010;33(1):1–22.
51. Wood SN. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. J R Stat Soc B. 2011;73(1):3–36.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Paper II

Placental epigenetic clocks: estimating gestational age using placental DNA methylation levels

Research Paper

Placental epigenetic clocks: estimating gestational age using placental DNA methylation levels

Yunsung Lee[1], Sanaa Choufani[2], Rosanna Weksberg[3], Samantha L. Wilson[4,5], Victor Yuan[4,5], Amber Burt[6], Carmen Marsit[6], Ake T. Lu[7], Beate Ritz[8], Jon Bohlin[9], Håkon K. Gjessing[9,10], Jennifer R. Harris[1,9], Per Magnus[9], Alexandra M. Binder[8,1,*], Wendy P. Robinson[4,5,*], Astanand Jugessur[1,9,10,*], Steve Horvath[7,11,*]

[1]Department of Genetics and Bioinformatics, Norwegian Institute of Public Health, Oslo, Norway
[2]Genetics and Genome Biology Program, Research Institute, The Hospital for Sick Children, Toronto, Ontario, Canada
[3]Genetics and Genome Biology Program, Research Institute, The Hospital for Sick Children and Institute of Medical Science, University of Toronto, Toronto, Ontario, Canada
[4]Department of Medical Genetics, University of British Columbia, Vancouver, British Columbia, Canada
[5]B.C. Children's Hospital Research Institute, Vancouver, British Columbia, Canada
[6]Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, GA 30322, USA
[7]Department of Human Genetics, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA 90095, USA
[8]Department of Epidemiology, University of California Los Angeles, Los Angeles, CA 90095, USA
[9]Centre for Fertility and Health, Norwegian Institute of Public Health, Oslo, Norway
[10]Department of Global Public Health and Primary Care, University of Bergen, Bergen, Norway
[11]Department of Biostatistics, Fielding School of Public Health, University of California Los Angeles, Los Angeles, CA 90095, USA
*Co-senior authors

Correspondence to: Steve Horvath; **email:** shorvath@mednet.ucla.edu

Keywords: DNA methylation, epigenetic clock, placenta, gestational age
Received: April 26, 2018 **Accepted:** June 17, 2019 **Published:** June 24, 2019

ABSTRACT

The human pan-tissue epigenetic clock is widely used for estimating age across the entire lifespan, but it does not lend itself well to estimating gestational age (GA) based on placental DNAm methylation (DNAm) data. We replicate previous findings demonstrating a strong correlation between GA and genome-wide DNAm changes. Using substantially more DNAm arrays (n=1,102 in the training set) than a previous study, we present three new placental epigenetic clocks: 1) a robust placental clock (RPC) which is unaffected by common pregnancy complications (e.g., gestational diabetes, preeclampsia), 2) a control placental clock (CPC) constructed using placental samples from pregnancies without known placental pathology, and 3) a refined RPC for uncomplicated term pregnancies. These placental clocks are highly accurate estimators of GA based on placental tissue; e.g., predicted GA based on RPC is highly correlated with actual GA (r>0.95 in test data, median error less than one week). We show that epigenetic clocks derived from cord blood or other tissues do not accurately estimate GA in placental samples. While fundamentally different from Horvath's pan-tissue epigenetic clock, placental clocks closely track fetal age during development and may have interesting applications.

INTRODUCTION

Gestational age (GA) of the fetus is used to forecast the date of delivery, optimize prenatal care, and monitor the growth and development of the fetus relative to other pregnancies. Short GA at delivery impacts neonatal morbidity and mortality [1-3], as well as brain development [4-6]. Thus, accurate classification of the fetus may help predict neonatal risk. In this regard, the World Health Organization defined extremely preterm (<28 weeks of gestation), very preterm (28-32 weeks of gestation) and moderate or late preterm (32-37 weeks of gestation) birth to reflect the newborn's developmental stage [7].

Traditional methods for estimating GA include early obstetric ultrasound measures or calculations based on the last menstrual period (LMP) [8]. The early ultrasound method estimates GA based on the visible fetal size (e.g., crown-rump length during the first trimester [9-11] or biparietal diameter after the second trimester [12-15]). The LMP method calculates GA based on the time elapsed since the known first day of the LMP. The early ultrasound method is widely accepted as the gold standard due to its higher accuracy [16] but is not routinely available in low and middle-income countries. More accurate classification of GA at birth may help predict neonatal risk for adverse outcomes and measure GA more accurately than through the assessment of physical and neurological features of the newborn, especially when early ultrasound measures are lacking, or the infant is growth-restricted but not preterm.

Here, we aim to develop a new molecular estimator of GA based on placental tissue samples that is more accurate than the previous clock [17]. Earlier studies have revealed profound molecular changes in placental chorionic villi, the placental structures that project into maternal decidua and are bathed in maternal blood, during gestation [18-22]. We focus on placental DNA methylation (DNAm) data, because prior work demonstrated that accurate estimators of chronological age (epigenetic clocks) can be developed based on DNAm levels from a variety of tissues [23], that one can estimate GA based on DNAm data derived from umbilical cord blood samples [24, 25], and most pertinently that one can estimate GA based on placental methylation data (Mayne et al. 2017) [17]. Our study provides more accurate placenta-based GA estimators (i.e., placental epigenetic clocks) than those developed previously, because we use a substantially larger sample for our training set (more than six times larger than that of Mayne et al. 2017). We aim to develop three different placental epigenetic clocks: 1) a "robust placental clock" (RPC) that is largely unaffected by pregnancy conditions (e.g., preeclampsia, gestational diabetes, and trisomy), 2) a "control placental clock" (CPC), tailor-made for measuring GA in normal pregnancies, and 3) a "refined RPC", trained for uncomplicated term (GA>36) pregnancies. For the RPC, we purposely included placental samples from a variety of pregnancy complications in the training data (e.g., hypertension or diabetes) as well as congenital abnormalities (e.g., trisomy 13, 18 and 21).

RESULTS

Placental DNA methylation data

We downloaded publicly available DNAm data from Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/; Table 1) that assessed DNAm levels in placental tissues. Eighteen datasets used the Illumina HumanMethylation 450K BeadChip (450K) platform and one used the more recent Illumina Methylation EPIC BeadChip (EPIC) array. Our analyses focused on the 441,870 autosomal CpG probes that are shared between the two Illumina platforms such that the resulting GA estimators (RPC and CPC) would be applicable to data from both platforms.

Robust placental clock (RPC)

An overview of our analysis is presented in Figure 1. We developed the RPC using several placental DNAm datasets (training n=1,102, Table 1, Figure 1). We regressed GA (dependent variable) on DNAm levels of CpG sites using a penalized regression model (elastic net regression [26]). The elastic net regression model automatically selected 558 CpG sites for the RPC model (Supplementary File 1). Predicted GA is a weighted average of DNAm levels at these 558 CpGs, where the weights are the regression coefficients.

Figure 2 shows the results of a comparison between the RPC and the placental clock from Mayne et al. (2017) in independent test data (test n=187, Table 1). The predictive accuracy of the placental clocks was quantified using the median absolute error (MAE, defined as the median absolute deviation between predicted GA and observed GA), and the degree of the linear association between predicted GA and observed GA was measured using the Pearson correlation coefficient (r). According to both measures, the RPC (MAE=0.96 weeks; 95% confidence interval (CI) [0.88, 1.19], r=0.99; 95% CI [0.98, 0.99]) outperformed Mayne's placental clock (MAE=2.63 weeks; 95% CI [2.17, 3.01], r=0.94; 95% CI [0.92, 0.96]). Note that Mayne's placental clock underestimated GA in two data sets: GSE73375 (green dots) and GSE75196 (blue dots), and overestimated GA in two other data sets: GSE66210 (black) and GSE70453 (red).

Table 1. Description of the publicly available placental DNAm data.

GEO Number	Placental tissue type	GEO submitter	N	Platform	Normalization method	Probe exclusion criteria[11]	GA range (weeks)
Training data							
GSE71678	Fetal side, near the cord insertion	Marsit et al.	343	450K[2]	funNorm[4]	SC, CH, SNP, DP	30-42
GSE75248	Fetal side	Marsit et al.	334	450K[2]	funNorm[4]	SC, CH, SNP, DP	37-42
GSE71719	Fetal side, near the cord insertion	Marsit et al.	44	450K[2]	noob[5]	SC, CH, SNP, DP	37-41
RL[1]	Fetal side, chorionic villi	-	121	450K[2]	funNorm[4]	SC	14-42
GSE100197	Fetal side, chorionic villi	Robinson et al.	16	450K[2]	SWAN[6]	SC, SNP, DP, MB	26-39
GSE108567	Fetal side, chorionic villi	Robinson et al.	7	450K[2]	SWAN[6]	SC, CH, SNP, DP, BR	29-38
GSE69502	Fetal side, chorionic villi	Robinson et al.	7	450K[2]	SWAN[6]	SC, CH, SNP, DP, BR	16-24
GSE74738	Fetal side, chorionic villi	Robinson et al.	8	450K[2]	SWAN[6]	SC, CH, SNP, DP, BR	6-13
GSE115508	Fetal side, chorionic villi	Robinson et al.	44	EPIC[3]	funNorm[4]	SC, CH, SNP, DP, BR	28-37
GSE44667	Fetal side, chorionic villi	Robinson et al.	27	450K[2]	SWAN[6]	SC, SNP, DP, MB	25-37
GSE49343	Fetal side, chorionic villi	Robinson et al.	13	450K[2]	SWAN[6]	SC, SNP, DP	5-39
GSE42409	Fetal side, chorionic villi	Robinson et al.	4	450K[2]	SWAN[6]	SC, SNP, DP	26-33
GSE120250	Fetal side, near the cord insertion	Weksberg et al.	86	450K[2]	GenomeStudioNorm[7]	SC, SNP, DP	35-41
GSE98224	Fetal side	Cox et al.	48	450K[2]	SWAN[6]	SC, SNP, DP, MB	27-41
Test data							
GSE70453	Maternal side, decidua near the cord	Binder et al.	82	450K[2]	BMIQ[8]	SC, CR, SNP	35-42
GSE73375	Fetal side	Fry et al.	36	450K[2]	quanNorm[9]	DP	22-41
GSE75196	Fetal side	Chiu et al.	24	450K[2]	dasen[10]	SC, SNP, DP, BR	32-40
GSE76641	Fetal side, chorionic villi	Slieker et al.	4	450K[2]	funNorm[4]	SC, SNP, DP, BR	9-22
GSE66210	Fetal side, chorionic villi	Bojesen et al.	41	450K[2]	GenomeStudioNorm[7]	-	11-15

[1] Placental DNAm data generated from the Robinson laboratory at the University of British Columbia (Vancouver, BC, Canada); The data for which is publicly available as part of the GEO data sets listed below.

[2] 450K: Illumina Infinium HumanMethylation450 BeadChip

[3] EPIC: Infinium MethylationEPIC BeadChip

[4] funNorm: Functional normalization [27]

[5] noob: Normal-exponential out-of-band [29]

[6] SWAN: Subset-quantile within array normalization [28]

[7] GenomeStudioNorm: Genome Studio normalization (details available in the GenomeStudio Methylation Module v1.8 User Guide, https://www.illumina.com/content/dam/illumina-support/documents/documentation/software_documentation/genomestudio/genomestudio-2011-1/genomestudio-methylation-v1-8-user-guide-11319130-b.pdf)

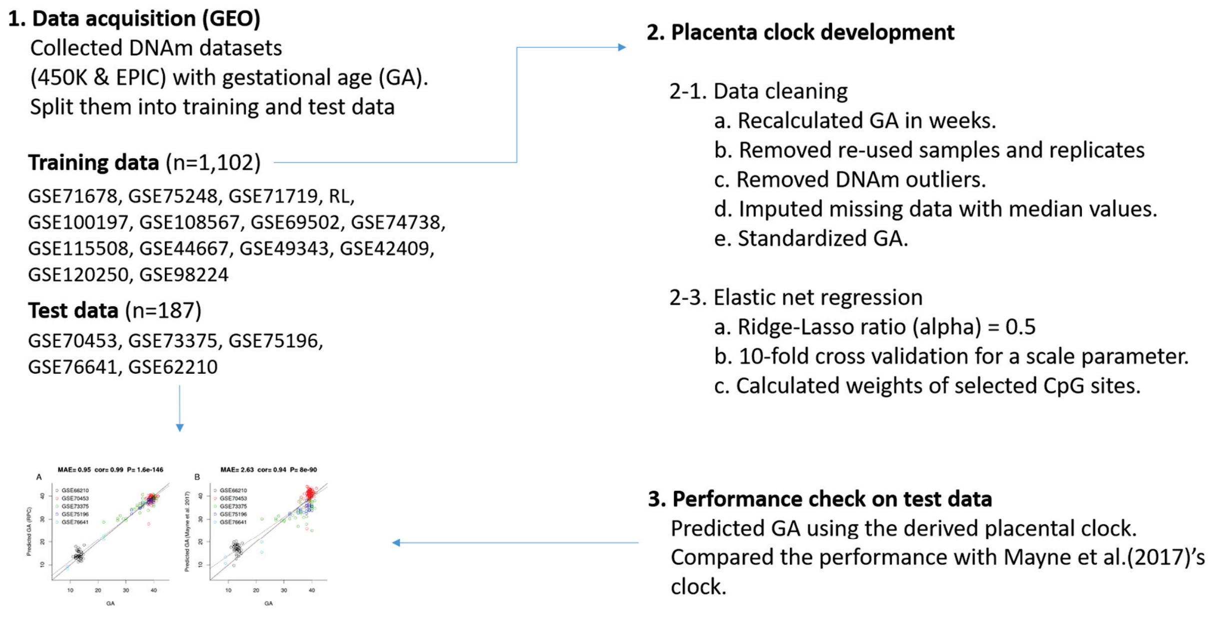
[8] BMIQ: Beta-mixture quantile dilation [30]

[9] quanNorm: Quantile normalization [31, 32]

[10] dasen: Data-driven separate normalization [33]

[11] Probe exclusion criteria

SC: Sex chromosome, CH: Cross-hybridizing, SNP: Single nucleotide polymorphism, DP: Detection P-value < 0.01, MB: Missing beta > 5%, and BR: Bead replicates < 3.



1. Data acquisition (GEO)
Collected DNAm datasets
(450K & EPIC) with gestational age (GA).
Split them into training and test data

Training data (n=1,102)
GSE71678, GSE75248, GSE71719, RL,
GSE100197, GSE108567, GSE69502, GSE74738,
GSE115508, GSE44667, GSE49343, GSE42409,
GSE120250, GSE98224

Test data (n=187)
GSE70453, GSE73375, GSE75196,
GSE76641, GSE62210

2. Placenta clock development

2-1. Data cleaning
a. Recalculated GA in weeks.
b. Removed re-used samples and replicates
c. Removed DNAm outliers.
d. Imputed missing data with median values.
e. Standardized GA.

2-3. Elastic net regression
a. Ridge-Lasso ratio (alpha) = 0.5
b. 10-fold cross validation for a scale parameter.
c. Calculated weights of selected CpG sites.

3. Performance check on test data
Predicted GA using the derived placental clock.
Compared the performance with Mayne et al.(2017)'s clock.

A
MAE= 0.95 cor= 0.99 P= 1.6e-146
GSE66210
GSE70453
GSE73375
GSE75196
GSE76641
Predicted GA (RPC)
GA
10
20
30
40
B
MAE= 2.63 cor= 0.94 P= 8e-90
Predicted GA (Mayne et al. 2017)

Figure 1. Flow chart of the RPC development.

The advantage of the RPC is particularly pronounced in later gestation, e.g., when restricting the analysis to placental samples with GA > 25 weeks, the RPC (MAE=0.89 [0.73, 1.02], r=0.82 [0.76, 0.87]) greatly outperforms Mayne's clock (MAE=2.25 [1.9, 2.63], r=0.61 [0.05,0.71], Figure 2C and 2D).

As expected by its construction, the RPC predicted GA accurately even in placental samples with adverse pregnancy conditions such as preeclampsia, gestational diabetes, and trisomy 13, 18 or 21 (Supplementary Figure S1). However, Mayne's placental clock underestimated GA in placental samples from pre-eclampsia cases and overestimated GA in cases of gestational diabetes and trisomy (Supplementary Figure S1). In case of trisomy, the RPC (MAE=2.26 [1.63, 2.88], r=0.12 [-0.25, 0.46]) was more accurate than Mayne's clock (MAE=3.99 [3.35, 5.4], r=0.02 [-0.34, 0.39]) but still showed a slight overestimation. The RPC's GA estimate was not associated with fetal sex (Supplementary Figure S2). We could not evaluate the effect of ethnicity because our test data did not include ethnic information except for GSE73375 (n=36, Supplementary Figure S3).

The training data used in the construction of the RPC employed several different normalization methods: functional normalization (funNorm, [27]), subset-quantiles within arrays (SWAN, [28]) and the normal-exponential out-of-band (noob, [29]) approach. This lack of uniformity in normalization methods in the training data has a statistical advantage: it makes it more likely that the RPC will be robust with respect to different normalization methods. In support of this, we found that the RPC validated in test data that were normalized using various methods: beta-mixture quantile dilation (BMIQ, [30]), quantile normalization (quanNorm, [31, 32]), data-driven separate normalization (dasen, [33]) as detailed in Table 1.

Control placental clock (CPC)

We trained the CPC on placental samples (training n=963, Table 1) that had been designated as "control" samples. Hence, placental samples with higher GA were probably from relatively normal pregnancies. However, placental samples with lower GA might contain samples that would be considered abnormal (i.e., premature rupture of membranes, spontaneous premature labor) but minimal placental pathology relative to pre-eclampsia cases. The analysis flow was identical as for the RPC, except for the composition of the training and test sets (Supplementary Figure S4). The elastic net regression model used for the CPC automatically selected 546 CpG sites (Supplementary File 1).

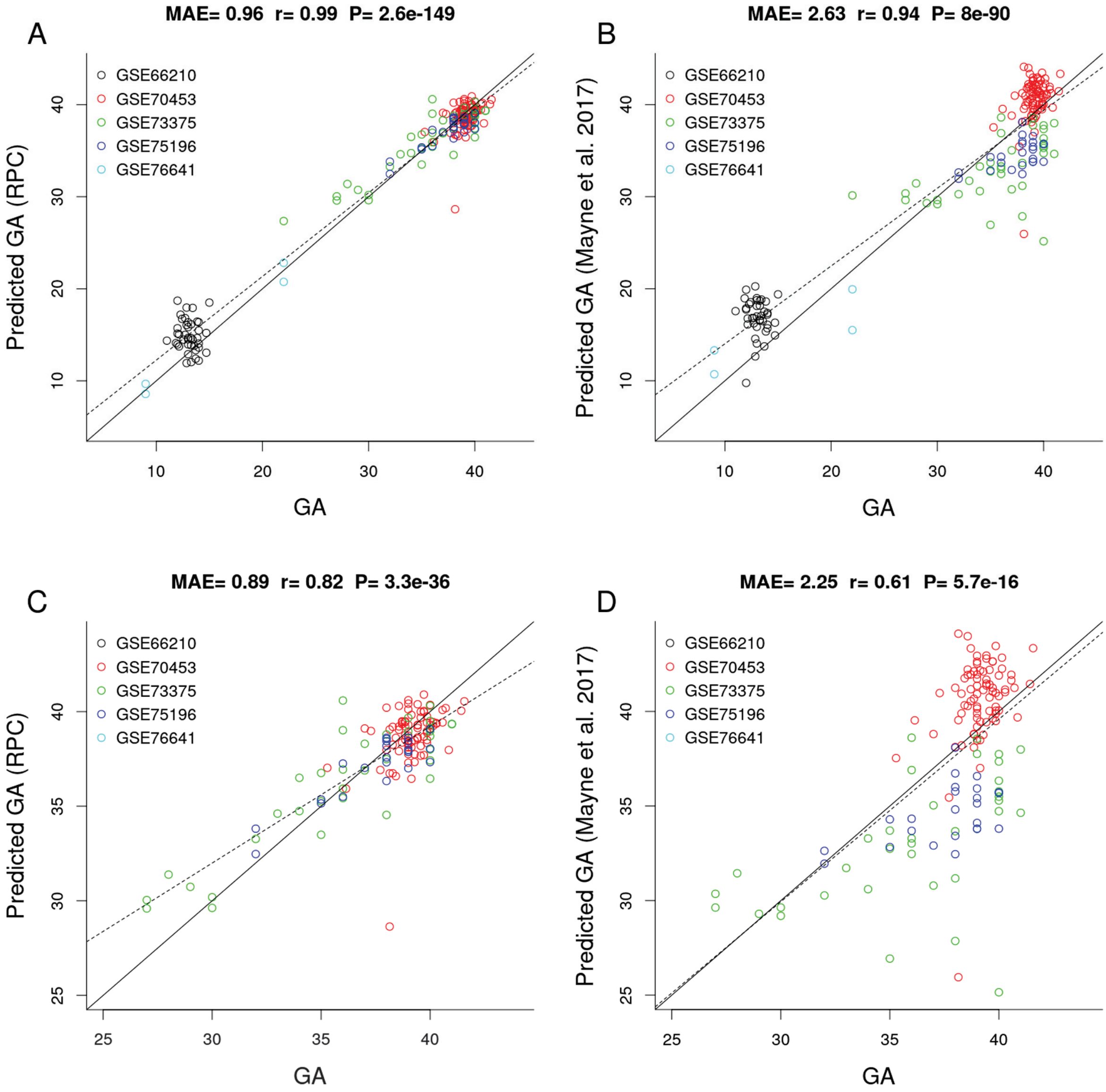


Figure 2. Gestational age estimation of the RPC and Mayne et al. (2017)'s placental clock. (**A**) Scatter plot between observed GA and DNAm-predicted GA (RPC) across all trimesters. (**B**) Scatter plot between observed GA and DNAm-predicted GA (Mayne et al. 2017) across all trimesters. (**C**) Zoom-in on panel **A** restricting GA > 25 weeks. (**D**) Zoom-in on panel **B** restricting GA > 25 weeks.

To assess whether adverse pregnancy conditions influence the epigenetic GA estimate, we applied the CPC to placental samples associated with chromosomal abnormalities (confined placental mosaicism, diandric triploidy, trisomy 13, 16, 18 and 21), neural tube defects (anencephaly and spinal bifida), intrauterine growth restriction, maternal complications (gestational diabetes and preeclampsia), and chorioamnionitis (test, n=326). Interestingly, the CPC accurately predicted the GA of fetuses with the above-mentioned conditions (MAE=1.02, r=0.98, Figure 3A) even though the CPC was constructed using unaffected control samples only.

To test whether pregnancy conditions are associated with faster/slower epigenetic aging, we used epigenetic measures of GA acceleration that were formally defined as raw residuals resulting from regressing the DNAm GA estimate on observed GA. By definition, this residual-based measure of GA acceleration is not correlated with true GA (r=0). GA acceleration did not significantly deviate from zero for any pregnancy conditions mentioned above (Figure 3B), but we acknowledge the small sample sizes for diandric triploidy (n=3) and trisomy 16 (n=3). When restricting the analysis to placental samples from the first trimester (weeks 1 to 12), we found the CPC's GA estimates to be slightly inaccurate, which was due to the small training set (only n=7 fetuses with GA < 12 weeks).

Refined robust placental clock for uncomplicated term pregnancies

For researchers who are particularly interested in uncomplicated term pregnancies, we also developed a second version of the RPC using placental samples from "uncomplicated term" pregnancies (defined as GA > 36 weeks) without any known pregnancy condition.

Toward this end, we selected "uncomplicated" term placental samples (n=733) from the training set used for the original RPC. Further, we restricted the penalized regression model analysis to the 558 CpGs that make up the original RPC. The penalized regression model automatically selected 395 CpG sites out of the 558 sites (Supplementary File 1). We find that the "refined" RPC for uncomplicated term pregnancies leads to highly accurate GA estimates (MAE=1.49, r=0.98, Figure 4A) in the RPC's test set (n=187).

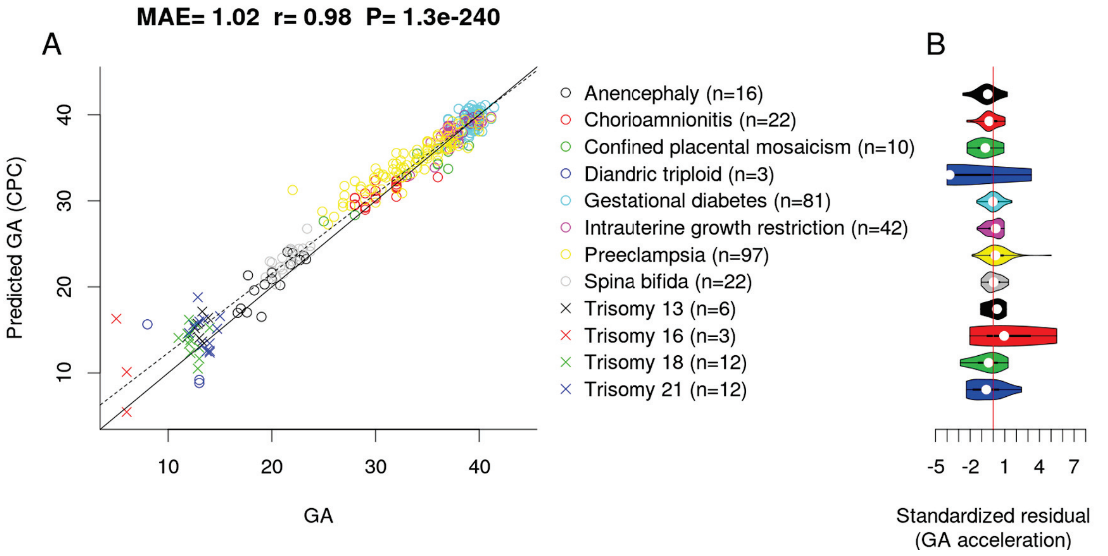


Figure 3. Effect of pregnancy condition on the GA estimate by CPC. (**A**) Scatter plot between GA and DNAm-predicted GA (CPC) across all trimesters. (**B**) Violin plot of GA acceleration (standardized residual) for each pregnancy condition.

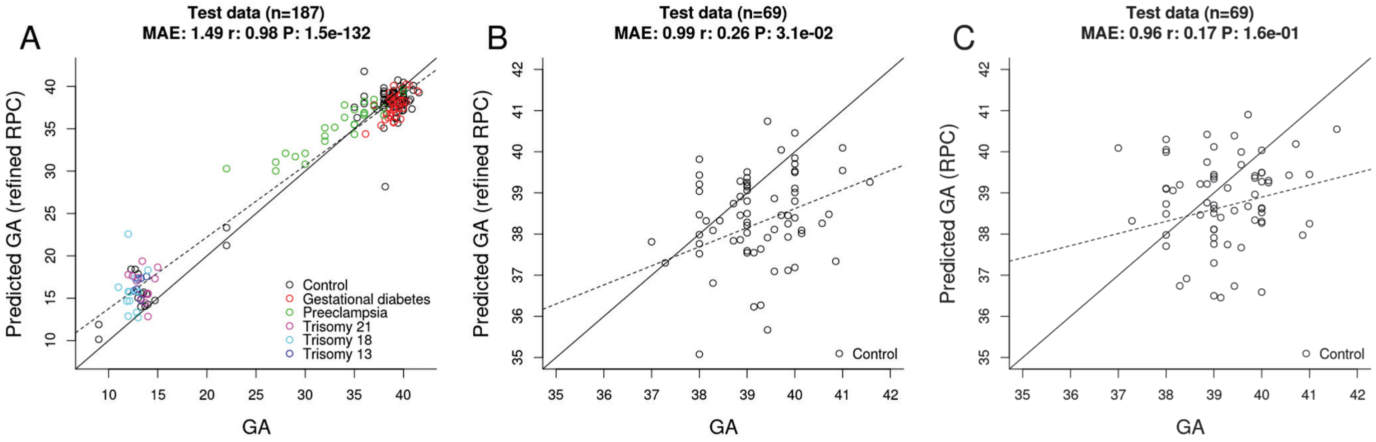


Figure 4. Gestational age estimation by the refined RPC and the RPC. (**A**) Scatter plot between observed GA and DNAm-predicted GA (by the refined RPC) – all samples from the RPC's test data (n=187). (**B**) Scatter plot between observed GA and DNAm-predicted GA (by the refined RPC) - uncomplicated term samples from the RPC's test data (n=69). (**C**) Scatter plot between observed GA and DNAm-predicted GA (by the RPC) - uncomplicated term samples from the RPC's test data (n=69).

Evaluating other epigenetic clocks

Using RPC’s test data (n=187), we found that previously published epigenetic clocks derived from cord blood samples or other tissues do not apply to the estimation of GA based on placental samples.

No significant correlation between GA and predicted DNAm age could be observed for clocks by Hannum (2013) [34], Horvath (2013) [23], Levine (2018) [35], and Horvath (2018) [36] (Supplementary Figure S5). However, the DNAm age estimate is close to zero for Horvath’s pan-tissue clock and the more recently developed Skin & Blood clock. Similarly, GA estimators for cord blood (Bohlin's cord blood clock [24], Knight’s cord blood clock [25]) failed to accurately predict GA in placental samples (Supplementary Figure S6). Overall, these studies demonstrate that the placenta is quite distinct from other tissues regarding the development and application of DNAm based age estimators.

Fetal sex-classifier based on DNAm

Several GEO datasets did not report fetal sex (e.g., GSE70453, GSE73375 and GSE76641) and CpGs present on sex chromosomes. Therefore, we developed a fetal sex-classifier based RPC’s training data (n=1,102) using CpGs that are present on autosomes. Toward this end, we regressed fetal sex (binary outcome) on 441,870 autosomal CpG sites using an elastic net implemented in the glmnet R package [37]. The elastic net automatically selected 220 autosomal CpG sites. The classification accuracy was 100% for the placental test data from GSE75196 (n=24). Interestingly, the placental sex classifier turns out to be highly accurate, when applied to blood-based DNAm data from adults (e.g., an accuracy of 96% in the data

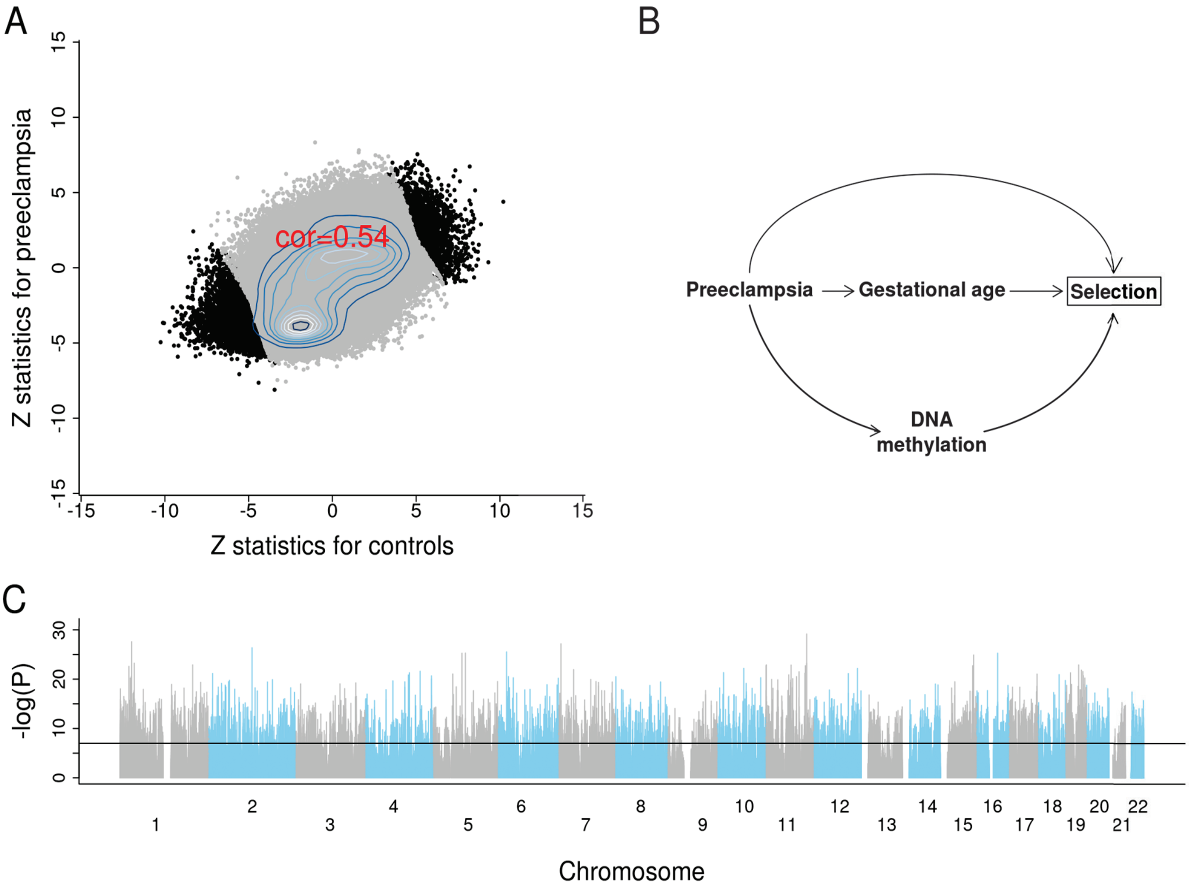


Figure 5. Results of EWAS and potential confounding between DNA methylation and gestational age due to selection bias. (**A**) Scatter plots between Z scores from controls and Z scores from preeclampsia. (**B**) The depicted minimal causal diagram under the null hypothesis of no effect of GA on DNAm. Here, the pregnancy condition (preeclampsia) would induce a spurious association between DNAm and GA, because preeclampsia could prompt earlier delivery (shorter GA) and influence DNAm. Note that the association between GA and DNAm is not due to a direct causal relationship between DNAm and GA. Rather, the association is confounded by preeclampsia. If the selection criteria differ substantially across studies, the placental clock models may not perform well. (**C**) EWAS Manhattan plot of GA.

from the Framingham Heart Study, n=2,356). As the sex of the fetus is typically identical to the sex of its placenta (except for rare cases of chimerism or sex-chromosome mosaicism), the sex-classifier was used to impute fetal sex in GSE66210, GSE70453, GSE73375 and GSE76641.

Epigenome-wide association studies of gestational age

We briefly report the results from an epigenome-wide association study (EWAS) of GA to demonstrate the profound effect of GA on placental DNAm levels. To protect against confounding by preeclampsia, we conducted EWAS in two separate strata: first, for placental samples from control pregnancies (n=831); second, for placental samples from pregnancies with preeclampsia (n=70). We combined the summary statistics from the two EWAS using Stouffer's method for meta-analysis [38]. The two EWAS summary statistics presented consistent DNAm-GA correlations across 441,870 autosomal CpGs (Figure 5A).

Strikingly, 10,827 CpG sites exhibit a genome-wide significant correlation with GA (P<1E-07; Figure 5C, Supplementary File 2). Among these sites, 5,940 were in CpGs islands, 262 were in the north shelf, 1,165 in the north shore, 241 in the south shelf, and 902 in the south shore. The top four genes with the largest number of significant CpG sites were *MAD1L1* (17 CpGs), *BRD2* (13 CpGs), *INPP5A* (12 CpGs) and *RPTOR* (9 CpGs). The top 25 CpG sites and their nearest gene(s) are reported in Table 2.

The RPC had 36 epigenome-wide significant (P<1E-07) CpG sites, the CPC had 39, and the refined RPC had 32.

DISCUSSION

Using the largest placental training set to date (n=1,102), we developed highly robust molecular estimators of GA. The robust placental epigenetic clock (RPC) is expected to perform well, even when applied to cases with adverse fetal outcomes or pregnancy complications. We developed this clock using a placenta-based training set that included several adverse conditions, including chromosomal abnormalities (trisomy and triploidy), neural tube defects (anencephaly and spinal bifida), intrauterine growth restriction, maternal complications (gestational diabetes and preeclampsia), and chorioamnionitis.

Table 2. The top 25 CpG sites associated with GA.

CpG	Gene	Chr	Relation to UCSC CpG Island	UCSC RefGene Group	Meta Z (P) (n=901)	Z (P) of Control (n=831)	Z (P) of Preeclampsia (n=70)
cg23034799	*CADM1*	11	Island	TSS200	-11.4 (7E-30)	-10.3 (6E-23)	-4.9 (2E-06)
cg03418552	*CADM1*	11	Island	TSS200	-10.1 (6E-24)	-9.4 (8E-20)	-3.9 (2E-04)
cg21155609	*FAM167B*	1	N_Shore	1stExon	11. (3E-28)	10.2 (1E-22)	4.4 (2E-05)
cg27339550	*ZNF853*	7	Island	TSS1500	-10.9 (7E-28)	-9.2 (4E-19)	-5.9 (1E-08)
cg20025003	*TFCP2L1*	2	Island	TSS200	-10.8 (4E-27)	-9.5 (6E-20)	-5.2 (6E-07)
cg02215898		6	Island		-10.6 (3E-26)	-10.1 (2E-22)	-3.7 (3E-04)
cg11544721	*CETN3*	5	Island	Body	-10.5 (5E-26)	-10. (4E-22)	-3.7 (3E-04)
cg01152986	*SETD6;SETD6*	16	Island	TSS200	-10.5 (5E-26)	-9.7 (7E-21)	-4.2 (4E-05)
cg08757742	*RASGRF2*	5	Island	TSS200	-10.5 (6E-26)	-9.2 (6E-19)	-5.2 (6E-07)
cg26662656		15	N_Shelf		10.5 (1E-25)	8.2 (1E-15)	6.7 (2E-10)
cg13458335	*BMP8B*	1	Island	TSS1500	-10.1 (6E-24)	-9. (2E-18)	-4.6 (8E-06)
cg20630277	*MRPL23*	11	Island	Body	-10. (1E-23)	-9. (2E-18)	-4.4 (2E-05)
cg21908248	*PPP1R15B*	1	Island	1stExon	-10. (1E-23)	-9. (4E-18)	-4.5 (1E-05)
cg26940573	*ZNF566*	19	Island	1stExon;5'UTR;TSS200	-10. (1E-23)	-8.8 (9E-18)	-4.7 (5E-06)
cg13242525	*FAM86C*	11	Island	TSS1500	-10. (1E-23)	-8.2 (2E-15)	-5.9 (2E-08)
cg13512138	*CHID1*	11	Island	5'UTR	-10. (2E-23)	-8.8 (1E-17)	-4.7 (4E-06)
cg05569874	*SEMA4B*	15	Island	5'UTR;1stExon	-10. (2E-23)	-9.3 (2E-19)	-3.8 (2E-04)
cg21060796	*LAYN*	11	Island	Body	-10. (2E-23)	-8.5 (1E-16)	-5.2 (6E-07)
cg01103597	*RUNX3*	1		Body	9.9 (3E-23)	8.5 (1E-16)	5.1 (7E-07)
cg12799981	*ASCC1;C10orf104*	10	N_Shore	1stExon;5'UTR;TSS1500	-9.9 (7E-23)	-9.4 (1E-19)	-3.4 (7E-04)
cg12888127	*KNTC1;RSRC2*	12	Island	TSS1500;TSS200	-9.9 (7E-23)	-9.2 (4E-19)	-3.7 (3E-04)
cg03366925	*GLI3*	7	Island	TSS1500	-9.8 (1E-22)	-8.5 (1E-16)	-4.9 (2E-06)
cg19599862	*ZNF226*	19		1stExon;5'UTR	-9.8 (1E-22)	-8.2 (1E-15)	-5.4 (2E-07)
cg16449659	*TIGD4;ARFIP1*	4	S_Shore	TSS1500;5'UTR	-9.7 (2E-22)	-9.1 (1E-18)	-3.7 (3E-04)
cg27006129	*ZNF114*	19	N_Shore	TSS1500	-9.7 (3E-22)	-7.9 (1E-14)	-5.7 (3E-08)

In contrast, the only other published placental clock by Mayne and colleagues was trained on a small training set (n=170). In our independent test set (n=187), Mayne's clock under/overestimated GA according to pregnancy conditions. (Supplementary Figure S1). These systematic deviations from Mayne's clock might reflect interesting biological effects or technical artifacts (batch effects, normalization methods). Another potential limitation of Mayne's clock is that the authors limited the eligible CpG sites to the approxi-mately 18,437 autosomal sites on the 27K and 450K bead chips. This might explain why the Mayne's clock uses only 62 CpG sites, whereas our RPC uses 558.

To infer biological processes under the 558 and 546 CpG sites, we conducted functional gene enrichment analyses using the Genomic Regions Enrichment of Annotation Tool (GREAT, v.3.0, [39]). However, we did not find any significant biological annotations associated with fetal aging. Elastic net regressions automatically select predictive CpG sites of gestational age (GA), but these CpG sites are not always biologically meaningful.

Our study had several limitations. First, the "observed" GA used for building these epigenetic clocks were estimated either by early pregnancy ultrasound or the LMP method. Although early pregnancy ultrasound based on fetal growth is the gold standard in a clinical setting, it is susceptible to variations in fetal size and leads to a systematic underestimation of GA in smaller fetuses [40-42].

There is also a concern that some of the training sets might be subject to systematic confounding due to adverse pregnancy conditions, as is the case for preeclampsia (Figure 5B). GA tends to be overestimated for placentas linked to preeclampsia, which is consistent with the associated pathology of advanced villous maturation, as well as previous reports of molecular signs of advanced aging [17, 43]. In this hypothetical example, preeclampsia confounds the association between placental DNAm and GA (Figure 5B, [44-46]). However, this type of confounding probably does not affect our placental clocks for the following reasons. First, the CPC for control samples and the refined RPC for uncomplicated term samples also accurately predicted GA even in pregnancies with known complications. Second, our EWAS of GA reveals profound associations between GA and DNA methylation levels even after stratifying the analysis by preeclampsia.

Moreover, it is possible that the RPC and the CPC might not perform well in case of non-live births, because the proportion of non-live births was extremely small amongst the third trimester samples in the training datasets, while unavoidably all first and second trimester samples are non-live births. In addition, it has been suggested that gravidity or parity may change placental physiology (e.g., higher placental weight associated with higher parity [47]) and therefore might modify the relationship between the placental epigenome and GA.

The clinical application of the RPC might be limited, because obtaining placental samples during pregnancy is highly invasive (e.g., chorionic villus sampling [48, 49]). However, the existence of a predictive placental clock – the RPC – opens the possibility to develop another epigenetic clock based on cell-free fetal DNA (cffDNA). cffDNA is fragmented from placenta trophoblasts [50, 51], and circulates in maternal blood during pregnancy [52]. If the development of a cffDNA clock is successful, clinicians readily estimate GA simply by collecting and analyzing maternal blood anytime during pregnancy.

METHODS

Study population

We collected publicly available data from Gene Expression Omnibus (GEO) using the GEOparse Python package (Python 3.6.5: Anaconda, Inc.). Table 1 details each dataset. GSE71678 examined the correlation between placental DNAm and arsenic exposures in the New Hampshire Birth Cohort Study [53]. GSE75248 examined placental DNAm in relation to newborns' neurobehavioral outcomes [54]. GSE71719 studied the association between DNA hydroxymethylation and gene expression using placental samples [55]. The Robinson laboratory (RL) at the University of British Columbia (Vancouver, BC, Canada) transferred placental DNAm data that are publicly available in the GEO database. GSE100197 and GSE98224 were studies that aimed to find placental DNAm profiles for preeclampsia and intrauterine growth restriction in women recruited at the University of British Columbia Women's and Children's Hospital (Vancouver, Canada) and at Mount Sinai Hospital (Toronto, Canada), respectively [56]. GSE108567 investigated batch effects in DNAm micro array data [57]. GSE69502 explored DNAm patterns in multi-tissue samples (placental chorionic villi, kidney, spinal cord, brain, and muscle) from fetuses that were aborted due to neural tube defects [58]. GSE74738 aimed to identify differentially-methylated imprinted regions using a genome-wide approach [59]. GSE115508 compared DNAm patterns in cases of placental inflammation (acute chorioamnionitis) with those in unaffected controls [60]. GSE44667 studied the association between placental DNAm in gene enhancer regions and early-onset preeclampsia [61]. GSE49343 investigated placental DNAm with trisomy and preeclampsia [62]. GSE42409 enhanced probe annotation of Illumina HumanMethylation 450K BeadChip to facilitate biologically meaningful data interpretation [63]. GSE120250 examined the impact of assisted reproductive technology on the placental DNA methylome [64]. GSE70453 conducted epigenome-wide and transcriptome-wide analyses of gestational diabetes [65]. GSE73375 examined DNAm in the preeclamptic placenta in relation to the transforming growth factor beta pathway [66]. GSE75196 studied different DNAm patterns in patients with preeclampsia and unaffected controls [67]. GSE76641 studied the transcriptional and DNAm trajectory of 21 organs during fetal development [68].

Measurement of DNA methylation

Either the Illumina Infinium HumanMethylation450 BeadChip or the Infinium MethylationEPIC BeadChip was used to measure DNAm level at each CpG site. The DNAm level (β-value) was the ratio of two fluorescence signals (methylated and unmethylated). The minfi R package [31] was used to preprocess all the DNAm datasets except for GS2E115508 and GSE120250 (preprocessed by Illumina's proprietary software, Genome Studio). The preprocessing methods and probe exclusion criteria differed across studies. For example, Marsit and colleagues, the largest GEO submitter, used the funNorm, whereas Robinson and colleagues mostly used the funNorm or the SWAN (Table 1). Other GEO submitters used the BMIQ, funNorm, quanNorm, dasen, or noob. Most GEO submitters excluded probes on sex-chromosomes, near single nucleotide polymorphisms, with cross-hybridization or with a detection p-value > 0.01.

Pre-processing of DNA methylation data

We ensured that all samples were included only one time in our training data. Some GEO datasets re-used the same samples or included technical replicates. For example, 154 samples were re-used in GSE100197, GSE108567, GSE69502, GSE74738, GSE44667, GSE49343 and RL data; and 15 technical replicates were found in GSE100197 and RL data. The sample size (N) in Table 1 refers to the counts after excluding the re-used samples and replicates.

We detected and removed outliers using the following steps: 1) we defined a gold standard DNAm profile as the inter-sample median value. For each CpG, we computed the median beta value across all placental samples. 2) The gold standard was correlated with each placental sample to calculate the Pearson correlation coefficient. 3) Placental samples were excluded if their correlation with the gold standard profile was lower than 0.9. Overall, only four putative outliers were removed from the analysis.

Missing DNAm levels were imputed with the gold standard DNAm levels. Thus, if the beta value of a CpG was missing, the missing value was imputed with the interpersonal median value across all samples. These imputations were only implemented in the training data.

Elastic net regression of gestational age

We fit a penalized regression model using the "glmnet" R package [37]. GA was regressed on 441,870 CpG sites that are shared between the 450K and the EPIC array. The glmnet mixing parameter alpha was set to 0.5 (specifying elastic net regression), and the shrinkage parameter, lambda resulting in the minimum mean square error, was chosen using 10-fold cross-validation in the training data. The RPC automatically selected 558 CpG sites (lambda=0.0936), the CPC did 546 CpG sites (lambda=0.0892), the refined RPC did 395 CpG sites (lambda=0.0116), and the fetal sex-classifier did 220 CpGs (lambda=0.0073). The number of overlapping CpGs between the RPC and CPC was 199. Supplementary File 1 includes CpG sites and their corresponding coefficients for the RPC, CPC, refined RPC and fetal sex-classifier.

Epigenome-wide association study of gestational age

We used the R function "standardScreening-NumericTrait" from the weighted gene co-expression network analysis R package (WGCNA; [69]) to carry out a robust correlation test (based on the biweight midcorrelation) between each CpG and GA. We conducted two separate EWAS of GA: one in control placental samples (n=831) and the other in placental samples from preeclampsia cases (n=70). We computed biweight midcorrelations between DNAm levels and GA, and the corresponding Z statistics and p-values in each stratum. The Z statistics of the two sets of EWAS were combined using the weighted Stouffer's method [38] as: $\sum Z_i w_i / \sqrt{\sum w_i^2}$, where w_i is the square root of the sample size in the ith stratum. The corresponding p-values were computed as $2\left(1-\Phi(|Z_{\text{meta}}|)\right)$. The EWAS was limited on the 411,870 autosomal probes available on both the 450K and the EPIC array platform.

Software availability

The coefficient values of the placental clocks and the fetal sex classifier can be found in Supplementary File 1.

Abbreviations

GA: gestational age; DNAm: DNA methylation; LMP: last menstrual period; RPC: robust placental clock; CPC: control placental clock; GEO: Gene Expression Omnibus; 450K: Illumina HumanMethylation 450K BeadChip; EPIC: Illumina MethylationEPIC BeadChip; MAE: median absolute error; funNorm: functional normalization; SWAN: subset-quantiles within arrays; noob: normal-exponential out-of-band; BMIQ: beta-mixture quantile dilation; quanNorm: quantile normalization; dasen: data-driven separate normalization; WGCNA: weighted gene co-expression network analysis; EWAS: epigenome-wide association study; cffDNA: cell-free fetal DNA.

AUTHOR CONTRIBUTIONS

YL and SH developed the placental clocks and wrote the manuscript. The remaining authors contributed data, edited the manuscript, and interpreted the results.

ACKNOWLEDGEMENTS

We appreciate all the placenta donors and GEO submitters for making their placental DNAm data publicly available.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

FUNDING

This work is partly supported by a grant from the Norwegian Research Council (NRC) to AJ (project number 262043) and additional funding to YL from the NRC through a Personal Overseas Research Grant (project number 262043/F20).

SH acknowledges support from the National Institutes of Health (NIH) (1U01AG060908 – 01).

Research reported in this publication was supported by the National Cancer Institute of the National Institutes of Health under Award Number K07CA225856 (to AMB). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

The work of JB, HKG, JRH, PM, AJ was supported by the Research Council of Norway Center of Excellence funding scheme (Grant 262700). The funding body played no role in the design of the study, analysis or interpretation of data, nor in writing the manuscript.

The work of VY and WPR was supported by the National Institutes of Health (to WPR; RFN 5R01HD089713-04).

REFERENCES

1. Engle WA. Morbidity and mortality in late preterm and early term newborns: a continuum. Clin Perinatol. 2011; 38:493–516. https://doi.org/10.1016/j.clp.2011.06.009 PMID:21890021

2. Hansen AK, Wisborg K, Uldbjerg N, Henriksen TB. Risk of respiratory morbidity in term infants delivered by elective caesarean section: cohort study. BMJ. 2008; 336:85–87. https://doi.org/10.1136/bmj.39405.539282.BE PMID:18077440

3. Young PC, Glasgow TS, Li X, Guest-Warnick G, Stoddard G. Mortality of late-preterm (near-term) newborns in Utah. Pediatrics. 2007; 119:e659–65. https://doi.org/10.1542/peds.2006-2486 PMID:17332185

4. Davis EP, Buss C, Muftuler LT, Head K, Hasso A, Wing DA, Hobel C, Sandman CA. Children's Brain Development Benefits from Longer Gestation. Front Psychol. 2011; 2:1. https://doi.org/10.3389/fpsyg.2011.00001 PMID:21713130

5. Parikh LI, Reddy UM, Männistö T, Mendola P, Sjaarda L, Hinkle S, Chen Z, Lu Z, Laughon SK. Neonatal outcomes in early term birth. Am J Obstet Gynecol. 2014; 211:265.e1–11. https://doi.org/10.1016/j.ajog.2014.03.021 PMID:24631438

6. Yang S, Platt RW, Kramer MS. Variation in child cognitive ability by week of gestation among healthy term births. Am J Epidemiol. 2010; 171:399–406. https://doi.org/10.1093/aje/kwp413 PMID:20080810

7. Organization WH. Born too soon: the global action report on preterm birth. 2012.

8. Lynch CD, Zhang J. The research implications of the selection of a gestational age estimation method. Paediatr Perinat Epidemiol. 2007 (Suppl 2); 21:86–96. https://doi.org/10.1111/j.1365-3016.2007.00865.x PMID:17803622

9. Robinson HP, Fleming JE. A critical evaluation of sonar "crown-rump length" measurements. Br J Obstet Gynaecol. 1975; 82:702–10. https://doi.org/10.1111/j.1471-0528.1975.tb00710.x PMID:1182090

10. Papageorghiou AT, Kemp B, Stones W, Ohuma EO, Kennedy SH, Purwar M, Salomon LJ, Altman DG,

Noble JA, Bertino E, Gravett MG, Pang R, Cheikh Ismail L, et al, and International Fetal and Newborn Growth Consortium for the 21st Century (INTERGROWTH-21st). Ultrasound-based gestational-age estimation in late pregnancy. Ultrasound Obstet Gynecol. 2016; 48:719–26. https://doi.org/10.1002/uog.15894 PMID:26924421

11. Papageorghiou AT, Kennedy SH, Salomon LJ, Ohuma EO, Cheikh Ismail L, Barros FC, Lambert A, Carvalho M, Jaffer YA, Bertino E, Gravett MG, Altman DG, Purwar M, et al, and International Fetal and Newborn Growth Consortium for the 21st Century (INTERGROWTH-21st). International standards for early fetal size and pregnancy dating based on ultrasound measurement of crown-rump length in the first trimester of pregnancy. Ultrasound Obstet Gynecol. 2014; 44:641–48. https://doi.org/10.1002/uog.13448 PMID:25044000

12. Campbell S, Newman GB. Growth of the fetal biparietal diameter during normal pregnancy. J Obstet Gynaecol Br Commonw. 1971; 78:513–19. https://doi.org/10.1111/j.1471-0528.1971.tb00309.x PMID:5559266

13. Hadlock FP, Deter RL, Harrist RB, Park SK. Fetal biparietal diameter: a critical re-evaluation of the relation to menstrual age by means of real-time ultrasound. J Ultrasound Med. 1982; 1:97–104. https://doi.org/10.7863/jum.1982.1.3.97 PMID:6152941

14. Kurtz AB, Wapner RJ, Kurtz RJ, Dershaw DD, Rubin CS, Cole-Beuglet C, Goldberg BB. Analysis of biparietal diameter as an accurate indicator of gestational age. J Clin Ultrasound. 1980; 8:319–26. https://doi.org/10.1002/jcu.1870080406 PMID:6772680

15. Campbell S, Thoms A. Ultrasound measurement of the fetal head to abdomen circumference ratio in the assessment of growth retardation. Br J Obstet Gynaecol. 1977; 84:165–74. https://doi.org/10.1111/j.1471-0528.1977.tb12550.x PMID:843490

16. Mongelli M, Wilcox M, Gardosi J. Estimating the date of confinement: ultrasonographic biometry versus certain menstrual dates. Am J Obstet Gynecol. 1996; 174:278–81. https://doi.org/10.1016/S0002-9378(96)70408-8 PMID:8572021

17. Mayne BT, Leemaqz SY, Smith AK, Breen J, Roberts CT, Bianco-Miotto T. Accelerated placental aging in early onset preeclampsia pregnancies identified by DNA methylation. Epigenomics. 2017; 9:279–89. https://doi.org/10.2217/epi-2016-0103 PMID:27894195

18. Mikheev AM, Nabekura T, Kaddoumi A, Bammler TK, Govindarajan R, Hebert MF, Unadkat JD. Profiling gene expression in human placentae of different gestational ages: an OPRU Network and UW SCOR Study. Reprod Sci. 2008; 15:866–77. https://doi.org/10.1177/1933719108322425 PMID:19050320

19. Sitras V, Fenton C, Paulssen R, Vårtun Å, Acharya G. Differences in gene expression between first and third trimester human placenta: a microarray study. PLoS One. 2012; 7:e33294. https://doi.org/10.1371/journal.pone.0033294 PMID:22442682

20. Uusküla L, Männik J, Rull K, Minajeva A, Kõks S, Vaas P, Teesalu P, Reimand J, Laan M. Mid-gestational gene expression profile in placenta and link to pregnancy complications. PLoS One. 2012; 7:e49248. https://doi.org/10.1371/journal.pone.0049248 PMID:23145134

21. Winn VD, Haimov-Kochman R, Paquet AC, Yang YJ, Madhusudhan MS, Gormley M, Feng KT, Bernlohr DA, McDonagh S, Pereira L, Sali A, Fisher SJ. Gene expression profiling of the human maternal-fetal interface reveals dramatic changes between midgestation and term. Endocrinology. 2007; 148:1059–79. https://doi.org/10.1210/en.2006-0683 PMID:17170095

22. Novakovic B, Yuen RK, Gordon L, Penaherrera MS, Sharkey A, Moffett A, Craig JM, Robinson WP, Saffery R. Evidence for widespread changes in promoter methylation profile in human placenta in response to increasing gestational age and environmental/stochastic factors. BMC Genomics. 2011; 12:529. https://doi.org/10.1186/1471-2164-12-529 PMID:22032438

23. 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

24. Bohlin J, Håberg SE, Magnus P, Reese SE, Gjessing HK, Magnus MC, Parr CL, Page CM, London SJ, Nystad W. Prediction of gestational age based on genome-wide differentially methylated regions. Genome Biol. 2016; 17:207. https://doi.org/10.1186/s13059-016-1063-4 PMID:27717397

25. Knight AK, Craig JM, Theda C, Bækvad-Hansen M, Bybjerg-Grauholm J, Hansen CS, Hollegaard MV, Hougaard DM, Mortensen PB, Weinsheimer SM, Werge TM, Brennan PA, Cubells JF, et al. An epigenetic clock for gestational age at birth based on blood methylation data. Genome Biol. 2016; 17:206. https://doi.org/10.1186/s13059-016-1068-z PMID:27717399

26. Zou H, Hastie T. Regularization and variable selection via the elastic net. J R Stat Soc Series B Stat Methodol. 2005; 67:301–20. https://doi.org/10.1111/j.1467-9868.2005.00503.x

27. 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

28. Maksimovic J, Gordon L, Oshlack A. SWAN: subset-quantile within array normalization for illumina infinium HumanMethylation450 BeadChips. Genome Biol. 2012; 13:R44. https://doi.org/10.1186/gb-2012-13-6-r44 PMID:22703947

29. 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

30. Teschendorff AE, Marabita F, Lechner M, Bartlett T, Tegner J, Gomez-Cabrero D, Beck S. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. Bioinformatics. 2013; 29:189–96. https://doi.org/10.1093/bioinformatics/bts680 PMID:23175756

31. 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

32. Touleimat N, Tost J. Complete pipeline for Infinium(®) Human Methylation 450K BeadChip data processing using subset quantile normalization for accurate DNA methylation estimation. Epigenomics. 2012; 4:325–41. https://doi.org/10.2217/epi.12.21 PMID:22690668

33. Pidsley R, Y Wong CC, Volta M, Lunnon K, Mill J, Schalkwyk LC. A data-driven approach to pre-processing Illumina 450K methylation array data. BMC Genomics. 2013; 14:293. https://doi.org/10.1186/1471-2164-14-293 PMID:23631413

34. 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

35. 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

36. 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

37. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. J Stat Softw. 2010; 33:1–22. https://doi.org/10.18637/jss.v033.i01 PMID:20808728

38. Stouffer SA, Suchman EA, DeVinney LC, Star SA, Williams RM Jr. (1949). The American soldier: Adjustment during army life. (Studies in social psychology in World War II), Vol. 1.

39. McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, Wenger AM, Bejerano G. GREAT improves functional interpretation of cis-regulatory regions. Nat Biotechnol. 2010; 28:495–501. https://doi.org/10.1038/nbt.1630 PMID:20436461

40. Dietz PM, England LJ, Callaghan WM, Pearl M, Wier ML, Kharrazi M. A comparison of LMP-based and ultrasound-based estimates of gestational age using linked California livebirth and prenatal screening records. Paediatr Perinat Epidemiol. 2007 (Suppl 2); 21:62–71. https://doi.org/10.1111/j.1365-3016.2007.00862.x PMID:17803619

41. Henriksen TB, Wilcox AJ, Hedegaard M, Secher NJ. Bias in studies of preterm and postterm delivery due to ultrasound assessment of gestational age. Epidemiology. 1995; 6:533–37. https://doi.org/10.1097/00001648-199509000-00012 PMID:8562631

42. Morin I, Morin L, Zhang X, Platt RW, Blondel B, Bréart G, Usher R, Kramer MS. Determinants and consequences of discrepancies in menstrual and ultrasonographic gestational age estimates. BJOG. 2005; 112:145–52. https://doi.org/10.1111/j.1471-0528.2004.00311.x PMID:15663577

43. Leavey K, Benton SJ, Grynspan D, Bainbridge SA, Morgen EK, Cox BJ. Gene markers of normal villous maturation and their expression in placentas with maturational pathology. Placenta. 2017; 58:52–59.

https://doi.org/10.1016/j.placenta.2017.08.005 PMID:28962696

44. Jia RZ, Zhang X, Hu P, Liu XM, Hua XD, Wang X, Ding HJ. Screening for differential methylation status in human placenta in preeclampsia using a CpG island plus promoter microarray. Int J Mol Med. 2012; 30:133–41. https://doi.org/10.3892/ijmm.2012.983 PMID:22552323

45. Kulkarni A, Chavan-Gautam P, Mehendale S, Yadav H, Joshi S. Global DNA methylation patterns in placenta and its association with maternal hypertension in pre-eclampsia. DNA Cell Biol. 2011; 30:79–84. https://doi.org/10.1089/dna.2010.1084 PMID:21043832

46. Yuen RK, Peñaherrera MS, von Dadelszen P, McFadden DE, Robinson WP. DNA methylation profiling of human placentas reveals promoter hypomethylation of multiple genes in early-onset preeclampsia. Eur J Hum Genet. 2010; 18:1006–12. https://doi.org/10.1038/ejhg.2010.63 PMID:20442742

47. Wallace JM, Bhattacharya S, Horgan GW. Gestational age, gender and parity specific centile charts for placental weight for singleton deliveries in Aberdeen, UK. Placenta. 2013; 34:269–74. https://doi.org/10.1016/j.placenta.2012.12.007 PMID:23332414

48. Kazy Z, Rozovsky IS, Bakharev VA. Chorion biopsy in early pregnancy: A method of early prenatal diagnosis for inherited disorders. Prenat Diagn. 1982; 2:39–45. https://doi.org/10.1002/pd.1970020107

49. Ward RH, Modell B, Petrou M, Karagözlu F, Douratsos E. Method of sampling chorionic villi in first trimester of pregnancy under guidance of real time ultrasound. Br Med J (Clin Res Ed). 1983; 286:1542–44. https://doi.org/10.1136/bmj.286.6377.1542 PMID:6405878

50. Alberry M, Maddocks D, Jones M, Abdel Hadi M, Abdel-Fattah S, Avent N, Soothill PW. Free fetal DNA in maternal plasma in anembryonic pregnancies: confirmation that the origin is the trophoblast. Prenat Diagn. 2007; 27:415–18. https://doi.org/10.1002/pd.1700 PMID:17286310

51. Gupta AK, Holzgreve W, Huppertz B, Malek A, Schneider H, Hahn S. Detection of fetal DNA and RNA in placenta-derived syncytiotrophoblast microparticles generated in vitro. Clin Chem. 2004; 50:2187–90. https://doi.org/10.1373/clinchem.2004.040196 PMID:15502097

52. Lo YM, Tein MS, Lau TK, Haines CJ, Leung TN, Poon PM, Wainscoat JS, Johnson PJ, Chang AM, Hjelm NM. Quantitative analysis of fetal DNA in maternal plasma and serum: implications for noninvasive prenatal diagnosis. Am J Hum Genet. 1998; 62:768–75. https://doi.org/10.1086/301800 PMID:9529358

53. Green BB, Karagas MR, Punshon T, Jackson BP, Robbins DJ, Houseman EA, Marsit CJ. Epigenome-Wide Assessment of DNA Methylation in the Placenta and Arsenic Exposure in the New Hampshire Birth Cohort Study (USA). Environ Health Perspect. 2016; 124:1253–60. https://doi.org/10.1289/ehp.1510437 PMID:26771251

54. Paquette AG, Houseman EA, Green BB, Lesseur C, Armstrong DA, Lester B, Marsit CJ. Regions of variable DNA methylation in human placenta associated with newborn neurobehavior. Epigenetics. 2016; 11:603–13. https://doi.org/10.1080/15592294.2016.1195534 PMID:27366929

55. Green BB, Houseman EA, Johnson KC, Guerin DJ, Armstrong DA, Christensen BC, Marsit CJ. Hydroxymethylation is uniquely distributed within term placenta, and is associated with gene expression. FASEB J. 2016; 30:2874–84. https://doi.org/10.1096/fj.201600310R PMID:27118675

56. Wilson SL, Leavey K, Cox BJ, Robinson WP. Mining DNA methylation alterations towards a classification of placental pathologies. Hum Mol Genet. 2018; 27:135–46. https://doi.org/10.1093/hmg/ddx391 PMID:29092053

57. Price EM, Robinson WP. Adjusting for Batch Effects in DNA Methylation Microarray Data, a Lesson Learned. Front Genet. 2018; 9:83. https://doi.org/10.3389/fgene.2018.00083 PMID:29616078

58. Price EM, Peñaherrera MS, Portales-Casamar E, Pavlidis P, Van Allen MI, McFadden DE, Robinson WP. Profiling placental and fetal DNA methylation in human neural tube defects. Epigenetics Chromatin. 2016; 9:6. https://doi.org/10.1186/s13072-016-0054-8 PMID:26889207

59. Hanna CW, Peñaherrera MS, Saadeh H, Andrews S, McFadden DE, Kelsey G, Robinson WP. Pervasive polymorphic imprinted methylation in the human placenta. Genome Res. 2016; 26:756–67. https://doi.org/10.1101/gr.196139.115 PMID:26769960

60. Konwar C, Price EM, Wang LQ, Wilson SL, Terry J, Robinson WP. DNA methylation profiling of acute chorioamnionitis-associated placentas and fetal membranes: insights into epigenetic variation in spontaneous preterm births. Epigenetics Chromatin.

2018; 11:63. https://doi.org/10.1186/s13072-018-0234-9 PMID:30373633

61. Blair JD, Yuen RK, Lim BK, McFadden DE, von Dadelszen P, Robinson WP. Widespread DNA hypomethylation at gene enhancer regions in placentas associated with early-onset pre-eclampsia. Mol Hum Reprod. 2013; 19:697–708. https://doi.org/10.1093/molehr/gat044 PMID:23770704

62. Blair JD, Langlois S, McFadden DE, Robinson WP. Overlapping DNA methylation profile between placentas with trisomy 16 and early-onset preeclampsia. Placenta. 2014; 35:216–22. https://doi.org/10.1016/j.placenta.2014.01.001 PMID:24462402

63. Price ME, Cotton AM, Lam LL, Farré P, Emberly E, Brown CJ, Robinson WP, Kobor MS. Additional annotation enhances potential for biologically-relevant analysis of the Illumina Infinium HumanMethylation450 BeadChip array. Epigenetics Chromatin. 2013; 6:4. https://doi.org/10.1186/1756-8935-6-4 PMID:23452981

64. Choufani S, Turinsky AL, Melamed N, Greenblatt E, Brudno M, Bérard A, Fraser WD, Weksberg R, Trasler J, Monnier P, Fraser WD, Audibert F, Dubois L, et al, and 3D cohort study group. Impact of assisted reproduction, infertility, sex and paternal factors on the placental DNA methylome. Hum Mol Genet. 2019; 28:372–85. https://doi.org/10.1093/hmg/ddy321 PMID:30239726

65. Binder AM, LaRocca J, Lesseur C, Marsit CJ, Michels KB. Epigenome-wide and transcriptome-wide analyses reveal gestational diabetes is associated with alterations in the human leukocyte antigen complex. Clin Epigenetics. 2015; 7:79. https://doi.org/10.1186/s13148-015-0116-y PMID:26244062

66. Martin E, Ray PD, Smeester L, Grace MR, Boggess K, Fry RC. Epigenetics and Preeclampsia: Defining Functional Epimutations in the Preeclamptic Placenta Related to the TGF-β Pathway. PLoS One. 2015; 10:e0141294. https://doi.org/10.1371/journal.pone.0141294 PMID:26510177

67. Yeung KR, Chiu CL, Pidsley R, Makris A, Hennessy A, Lind JM. DNA methylation profiles in preeclampsia and healthy control placentas. Am J Physiol Heart Circ Physiol. 2016; 310:H1295–303. https://doi.org/10.1152/ajpheart.00958.2015 PMID:26968548

68. Roost MS, Slieker RC, Bialecka M, van Iperen L, Gomes Fernandes MM, He N, Suchiman HE, Szuhai K, Carlotti F, de Koning EJ, Mummery CL, Heijmans BT, Chuva de Sousa Lopes SM. DNA methylation and transcriptional trajectories during human development and reprogramming of isogenic pluripotent stem cells. Nat Commun. 2017; 8:908. https://doi.org/10.1038/s41467-017-01077-3 PMID:29030611

69. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008; 9:559. https://doi.org/10.1186/1471-2105-9-559 PMID:19114008

SUPPLEMENTARY MATERIAL

Supplementary File 1

This file includes CpG sites and their corresponding coefficients used for the RPC, CPC, refined RPC and fetal sex-classifier.

Supplementary File 2

This file includes part of summary statistics of EWAS of GA (the 10,827 CpG sites with meta P<1E-07).

Paper III

Epigenome-wide association study of leukocyte telomere length

Research Paper

Epigenome-wide association study of leukocyte telomere length

Yunsung Lee[1,#], Dianjianyi Sun[2,3,#], Anil P.S. Ori[4], Ake T. Lu[5], Anne Seeboth[6], Sarah E. Harris[7,8], Ian J. Deary[7,8], Riccardo E. Marioni[6,7], Mette Soerensen[9,10,11], Jonas Mengel-From[9,10], Jacob Hjelmborg[9], Kaare Christensen[9,10], James G. Wilson[12,13], Daniel Levy[14,15], Alex P. Reiner[16], Wei Chen[3], Shengxu Li[17], Jennifer R. Harris[1,18], Per Magnus[18], Abraham Aviv[19,*], Astanand Jugessur[1,18,20,*], Steve Horvath[5,21,*]

[1]Department of Genetics and Bioinformatics, Norwegian Institute of Public Health, Oslo, Norway
[2]Department of Epidemiology and Biostatistics, School of Public Health, Peking University Health Science Center, Beijing, China
[3]Department of Epidemiology, Tulane University, New Orleans, LA 70118, USA
[4]Center for Neurobehavioral Genetics, Semel Institute for Neuroscience and Human Behavior, University of California Los Angeles, Los Angeles, CA 90095, USA
[5]Department of Human Genetics, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA 90095, USA
[6]Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK
[7]Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, UK
[8]Department of Psychology, University of Edinburgh, Edinburgh, UK
[9]Epidemiology, Biostatistics and Biodemography, Department of Public Health, University of Southern Denmark, Odense C, Denmark
[10]Department of Clinical Genetics, Odense University Hospital, Odense C, Denmark
[11]Center for Individualized Medicine in Arterial Diseases, Department of Clinical Biochemistry and Pharmacology, Odense University Hospital, Odense C, Denmark
[12]Department of Physiology and Biophysics, University of Mississippi Medical Center, Jackson, MS 39216, USA
[13]Department of Cardiology, Beth Israel Deaconess Medical Center, Boston, MA 02215, USA
[14]The Framingham Heart Study, Framingham, MA 01702, USA
[15]Population Sciences Branch, Division of Intramural Research, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD 20892, USA
[16]Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA
[17]Children's Minnesota Research Institute, Children's Hospitals and Clinics of Minnesota, Minneapolis, MN 55404, USA
[18]Centre for Fertility and Health, Norwegian Institute of Public Health, Oslo, Norway
[19]Center of Development and Aging, New Jersey Medical School, Rutgers State University of New Jersey, Newark, NJ 07103, USA
[20]Department of Global Public Health and Primary Care, University of Bergen, Bergen, Norway
[21]Department of Biostatistics, Fielding School of Public Health, University of California Los Angeles, Los Angeles, CA 90095, USA
[#]Co-first authors
[*]Co-last authors

Correspondence to: Steve Horvath; **email:** shorvath@mednet.ucla.edu

Keywords: DNA methylation, leukocyte telomere length, multi-ancestry
Received: June 3, 2019 **Accepted:** August 18, 2019 **Published:** August 26, 2019

ABSTRACT

Telomere length is associated with age-related diseases and is highly heritable. It is unclear, however, to what extent epigenetic modifications are associated with leukocyte telomere length (LTL). In this study, we conducted a large-scale epigenome-wide association study (EWAS) of LTL using seven large cohorts (n=5,713) – the Framingham Heart Study, the Jackson Heart Study, the Women's Health Initiative, the Bogalusa Heart Study, the Lothian Birth Cohorts of 1921 and 1936, and the Longitudinal Study of Aging Danish Twins. Our stratified analysis suggests that EWAS findings for women of African ancestry may be distinct from those of three other groups: males of African ancestry, and males and females of European ancestry. Using a meta-analysis framework, we identified DNA methylation (DNAm) levels at 823 CpG sites to be significantly associated (P<1E-7) with LTL after adjusting for age, sex, ethnicity, and imputed white blood cell counts. Functional enrichment analyses revealed that these CpG sites are near genes that play a role in circadian rhythm, blood coagulation, and wound healing. Weighted correlation network analysis identified four co-methylation modules associated with LTL, age, and blood cell counts. Overall, this study reveals highly significant relationships between two hallmarks of aging: telomere biology and epigenetic changes.

INTRODUCTION

Telomeres are the $(TTAGGG)_n$ repeats located at the ends of each chromosome. Their broad function is to prevent genomic instability [1]. Telomeres in adult germ cells [2], bone marrow [3, 4] and embryonic stem cells [5] are largely maintained by telomerase. After birth, however, telomeres in somatic cells gradually shorten because of the repressed activities of telomerase [3–6]. In cultured cells, when telomeres become critically short, the cell reaches replicative senescence [1, 7]. Telomere length (TL) is reported to be shorter in leukocytes of men than women, but this sex difference may depend on the measurement method [8]. In their meta-analysis of data from 36 cohorts with a total of 36,230 participants, Gardner and colleagues found longer telomeres in women only for the terminal restriction fragments (TRF) Southern blot method [8]. By contrast, no sex effect was detected for the other TL measurement methods including the widely used quantitative real-time polymerase chain reaction (qPCR) protocol originally described by Cawthon [9]. TL is also shorter in leukocytes of individuals of European ancestry than individuals of African ancestry [10, 11]. Further, leukocyte telomere length (LTL) is associated with the two disease categories that largely define longevity in contemporary humans—cancer and cardiovascular disease [12–14].

High heritability estimates for LTL have been reported irrespective of the methods used for measuring LTL; reported heritability estimates are between 36% and 82% based on Southern blot [15–18], and between 51% and 76% based on qPCR [19, 20]. Genome-wide association studies (GWAS) conducted in large observational cohorts have identified 11 loci associated with LTL [21–24]. A subset of these loci harbor telomere maintenance genes. These loci, however, explain only a small proportion of the genetic variance in LTL. Similarly, relatively little is known about epigenetic changes and LTL. Here, we focus on the relationship between LTL and DNA methylation levels in leukocytes. Epigenome-wide association studies (EWAS) have emerged as a powerful tool for evaluating genome-wide changes in DNAm for a given phenotype of interest [25]. Previous studies have explored the association between DNAm and LTL [26–28], but these studies were somewhat limited due to moderate sample sizes or the focus on specific regions in the genome. Here, we conduct the largest EWAS of LTL to date in different groups defined by sex and ethnicity.

RESULTS

Epigenome-wide association study of leukocyte telomere length

We considered two sets of adjustments for LTL confounders: 1) partially adjusted LTL for age, sex, and ethnicity and 2) fully adjusted LTL for age, sex, ethnicity, and imputed white blood cell counts (CD4+ naïve, CD8+ naïve and exhausted cytotoxic T cell). We conducted a large-scale multi-ancestry EWAS of the partially and fully adjusted LTL using seven cohorts – the Framingham Heart Study (FHS, n=874), the Jackson Heart Study (JHS, n=1,637), the Women's Health Initiative (WHI, n=818), the Bogalusa Heart Study (BHS, n=831), the Lothian Birth Cohorts (LBC1921 and LBC1936, n=403 and n=906, respectively), and the Longitudinal Study of Aging Danish Twins (LSADT, n=244). The analysis flow is depicted in Figure 1. We note that adjustment in this script indicates a mixture of data stratification and regression adjustment.

Overall, 8,716 CpG sites were significantly (P<1E-07) associated with the partially adjusted LTL in the global meta-analysis. The top four genes with the largest number of significant CpGs were *VARS* (16 CpGs), *PRDM16* (15 CpGs), *MAGI2* (14 CpGs) and *MSI2* (13 CpGs). In the group-specific meta-analyses, we found 87 significant CpGs in men of European ancestry, 14 significant CpGs in men of African ancestry, 298 significant CpGs in women of European ancestry, and 20 significant CpGs in women of African ancestry (Supplementary File 1).

We identified 823 significant (P<1E-07) CpG sites associated with the fully adjusted LTL through the global meta-analysis. Our statistical significance threshold (1E-07) corresponds to a 5% family-wise error for 450K array studies [29]. Table 1 presents the top 30 CpGs among the 823 significant CpGs and groups them by

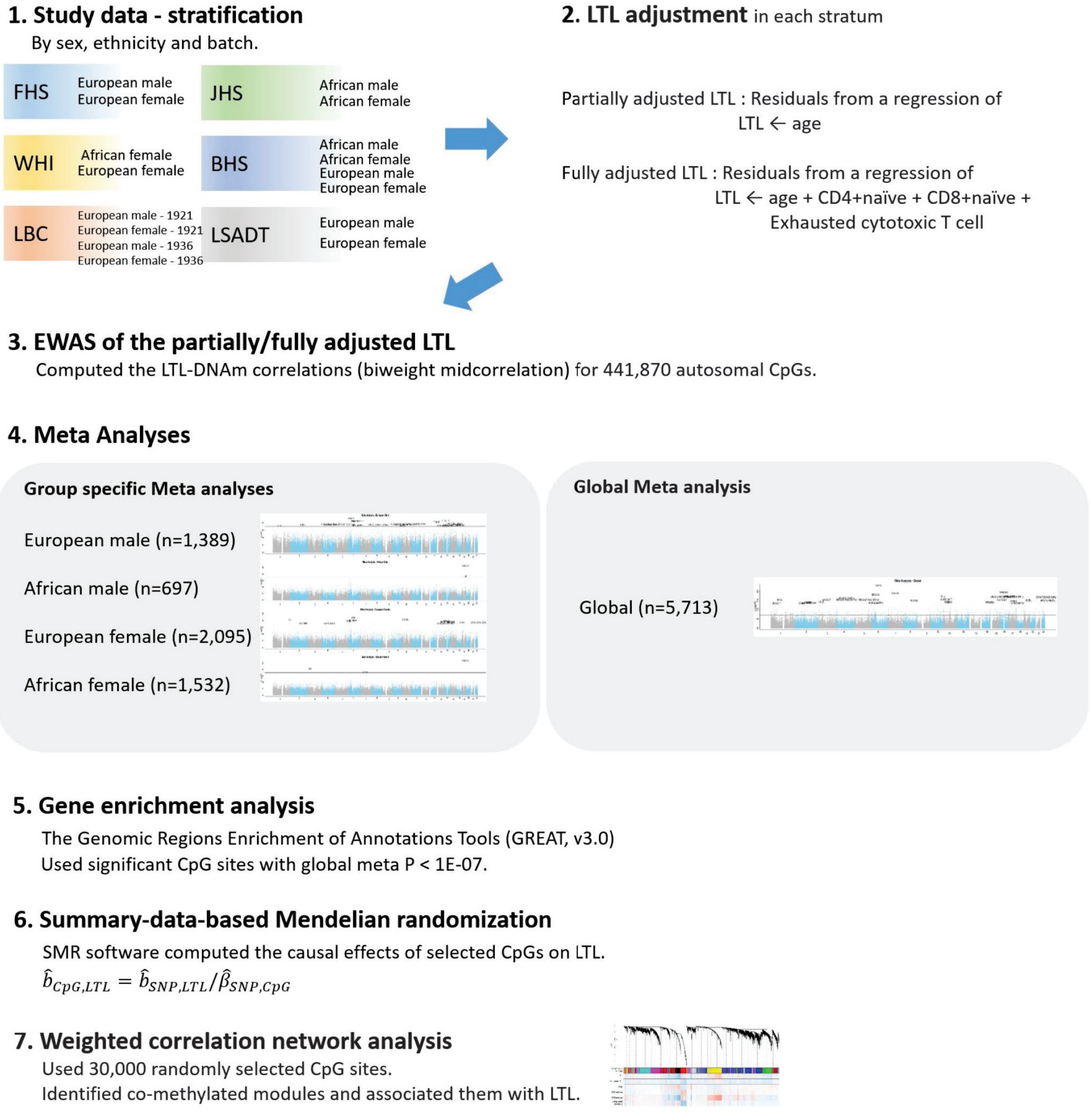

1. Study data - stratification
By sex, ethnicity and batch.
FHS
European male
European female
JHS
African male
African female
WHI
African female
European female
BHS
African male
African female
European male
European female
LBC
European male - 1921
European female - 1921
European male - 1936
European female - 1936
LSADT
European male
European female
2. LTL adjustment in each stratum
Partially adjusted LTL : Residuals from a regression of
LTL ← age
Fully adjusted LTL : Residuals from a regression of
LTL ← age + CD4+naïve + CD8+naïve +
Exhausted cytotoxic T cell
3. EWAS of the partially/fully adjusted LTL
Computed the LTL-DNAm correlations (biweight midcorrelation) for 441,870 autosomal CpGs.
4. Meta Analyses
Group specific Meta analyses
European male (n=1,389)
African male (n=697)
European female (n=2,095)
African female (n=1,532)
Global Meta analysis
Global (n=5,713)
5. Gene enrichment analysis
The Genomic Regions Enrichment of Annotations Tools (GREAT, v3.0)
Used significant CpG sites with global meta P < 1E-07.
6. Summary-data-based Mendelian randomization
SMR software computed the causal effects of selected CpGs on LTL.
$\hat{b}_{CpG,LTL} = \hat{b}_{SNP,LTL} / \hat{\beta}_{SNP,CpG}$
7. Weighted correlation network analysis
Used 30,000 randomly selected CpG sites.
Identified co-methylated modules and associated them with LTL.

Figure 1. Analysis flow chart.

Table 1. The top 30 most significant CpG sites associated with the fully adjusted LTL.

CpG	Gene	Chr	Relation to UCSC CpG island	UCSC RefGene group	Meta-Analysis				
					Global meta Z (P) n=5,713	European male Z (P) n=1,389	African male Z (P) n=697	European female Z (P) n=2,095	African female Z (P) n=1,532
cg08899667	*VARS*	6	N_Shelf	Body	-10.1 (4E-24)	-5.2 (3E-07)	-6.0 (2E-09)	-5.1 (4E-07)	-4.2 (3E-05)
cg02980249	*VARS*	6	N_Shelf	Body	-8.7 (2E-18)	-5.8 (5E-09)	-4.0 (6E-05)	-4.8 (2E-06)	-3.4 (7E-04)
cg02597894	*VARS*	6	N_Shelf	Body	-8.1 (4E-16)	-4.8 (2E-06)	-4.2 (3E-05)	-5.2 (2E-07)	-2.7 (6E-03)
cg04368724	*VARS*	6	N_Shelf	Body	-8.0 (9E-16)	-3.0 (2E-03)	-5.0 (5E-07)	-4.2 (3E-05)	-4.0 (8E-05)
cg04018738	*VARS*	6	N_Shelf	Body	-8.0 (2E-15)	-3.6 (3E-04)	-4.6 (4E-06)	-4.4 (1E-05)	-3.5 (4E-04)
cg24771152	*VARS*	6	N_Shelf	Body	-7.8 (6E-15)	-3.8 (2E-04)	-4.3 (2E-05)	-4.0 (6E-05)	-3.7 (2E-04)
cg20507228	*MAN2A2*	15	-	Body	-9.2 (5E-20)	-5.4 (8E-08)	-5.7 (2E-08)	-3.6 (3E-04)	-3.5 (4E-04)
cg08972170	*C7orf41*	7	-	Body	-9.0 (2E-19)	-3.7 (2E-04)	-4.9 (8E-07)	-4.1 (5E-05)	-5.4 (7E-08)
cg27343900*	*ERGIC1*	5	-	Body	-8.8 (1E-18)	-6.1 (8E-10)*	-5.1 (3E-07)	-4.2 (2E-05)	-2.4 (2E-02)
cg10549018	*TLL2*	10	-	Body	-8.6 (1E-17)	-5.3 (1E-07)	-3.9 (1E-04)	-4.5 (8E-06)	-4.0 (7E-05)
cg26709300*	*YPEL3*	16	N_Shore	1stExon;Body	-8.6 (1E-17)	-3.9 (8E-05)	-5.4 (6E-08)*	-2.4 (2E-02)	-4.8 (1E-06)
cg27106909*	*YPEL3*	16	N_Shore	1stExon;5′UTR;5′UTR	-8.5 (2E-17)	-5.6 (2E-08)*	-5.1 (3E-07)	-2.5 (1E-02)	-3.4 (6E-04)
cg12798040*	*XRCC3*	14	-	Body	-8.5 (2E-17)	-5.4 (8E-08)*	-5.4 (8E-08)*	-4.1 (4E-05)	-2.2 (2E-02)
cg02194129	*XRCC3*	14	-	Body	-8.3 (1E-16)	-4.9 (1E-06)	-5.0 (5E-07)	-4.3 (2E-05)	-2.6 (9E-03)
cg19841423*	*ZGPAT;LIME1*	20	S_Shore	Body;TSS1500	-8.4 (3E-17)	-5.0 (6E-07)	-5.5 (5E-08)*	-3.7 (2E-04)	-2.7 (8E-03)
cg02810967	*NCAPG;DCAF16*	4	S_Shore	Body;TSS1500	8.3 (9E-17)	4.4 (1E-05)	5.4 (9E-08)	4.1 (4E-05)	2.8 (5E-03)
cg19935065	*DNTT*	10	-	TSS1500	-8.1 (4E-16)	-3.5 (4E-04)	-4.9 (1E-06)	-5.0 (5E-07)	-3.2 (1E-03)
cg11093760	*CILP*	15	-	5′UTR;1stExon	-8.1 (5E-16)	-5.9 (4E-09)	-4.1 (5E-05)	-3.3 (1E-03)	-3.1 (2E-03)
cg19097500	*NFIA*	1	N_Shore	TSS1500	-8.1 (6E-16)	-5.4 (7E-08)	-3.7 (2E-04)	-3.7 (2E-04)	-3.6 (3E-04)
cg09626867	*EXOSC7*	3	-	Body	-8.1 (7E-16)	-5.2 (2E-07)	-4.1 (3E-05)	-4.5 (6E-06)	-2.8 (5E-03)
cg04509882	*EIF4G1*	3	-	Body;1stExon;5′UTR	-8.1 (8E-16)	-5.5 (4E-08)	-4.3 (2E-05)	-3.3 (1E-03)	-3.1 (2E-03)
cg23661483	*ILVBL*	19	S_Shelf	Body	-8.0 (9E-16)	-3.7 (2E-04)	-4.3 (2E-05)	-5.4 (7E-08)	-3.3 (1E-03)
cg01012082	*NCOA2*	8	-	3′UTR	-8.0 (1E-15)	-4.7 (3E-06)	-4.0 (7E-05)	-4.4 (1E-05)	-3.4 (8E-04)
cg21461082	*PRMT2*	21	Island	Body	8.0 (2E-15)	2.9 (4E-03)	4.4 (9E-06)	4.5 (6E-06)	4.4 (1E-05)
cg25921609	*MYH10*	17	N_Shore	Body	-7.9 (3E-15)	-5.2 (3E-07)	-3.6 (3E-04)	-4.5 (6E-06)	-3.1 (2E-03)
cg24420089*	*PTDSS2*	11	N_Shore	Body	-7.8 (8E-15)	-3.4 (7E-04)	-5.8 (7E-09)*	-2.3 (2E-02)	-3.5 (5E-04)
cg07414525	*CHL1*	3	-	Body	-7.8 (9E-15)	-3.5 (4E-04)	-3.0 (3E-03)	-3.5 (5E-04)	-5.8 (6E-09)
cg14817906	*CNNM4*	2	-	Body	-7.7 (1E-14)	-4.4 (1E-05)	-4.1 (4E-05)	-3.9 (8E-05)	-3.2 (1E-03)
cg04860432*	*PTGER2*	14	S_Shore	Body	-7.7 (2E-14)	-5.8 (7E-09)*	-4.3 (1E-05)	-2.3 (2E-02)	-2.7 (7E-03)
cg23570810	*IFITM1*	11	N_Shore	Body	7.7 (2E-14)	4.2 (3E-05)	4.2 (2E-05)	4.2 (2E-05)	3.0 (2E-03)

* The CpGs were more strongly associated with LTL in one or two sex and ethnicity specific groups than in the rest of the groups.

their annotated gene names. Among the top 30 CpGs, six were in *VARS*, two were in *YPEL2* and two were in *XRCC3*. The CpGs highlighted by an asterisk in Table 1 were more strongly associated with LTL in one or two sex and ethnicity-specific groups than in the rest of the groups. Specifically, the LTL-DNAm correlations at cg27343900 (in *ERGIC1*) and cg12798040 (in *XRCC3*) were stronger in men of European ancestry than in women of African ancestry. The LTL-DNAm correlation at cg27106909 near *YPEL3* was stronger in men of European ancestry than in women of European ancestry.

Figure 2 displays regional test statistics of LTL-associated CpGs on top of the local DNAm correlation structure for the top four genes listed in Table 1. *VARS* showed a cluster of CpGs above and right below the threshold of significance, while *MAN2A2*, *C7orf41* (current name, *MTURN*) and *ERGIC1* had one or two significant CpGs.

The clusters detected in *VARS* might be because of the high probe density on the array and the strong inter-CpG correlations.

The group-specific meta-analyses also detected several significant (P<1E-07) CpGs associated with the fully adjusted LTL. Figure 3 shows that 25 CpGs were significant in men of European ancestry, three CpGs in men of African ancestry, 19 CpGs in women of European ancestry, and four CpGs in women of African ancestry. Figure 4 displays scatter plots across the four group-specific meta-analyses. The correlation coefficient of each scatter plot was lowest between African American females and European males (r=-0.02) and highest between European females and European males (r=0.40). Population and sample size differences between strata may influence the correlations. The black dots in the panels refer to the top 30 CpG sites detected through the global meta-analysis. Across the 30 CpGs, we did observe high correlations (r≈0.92).

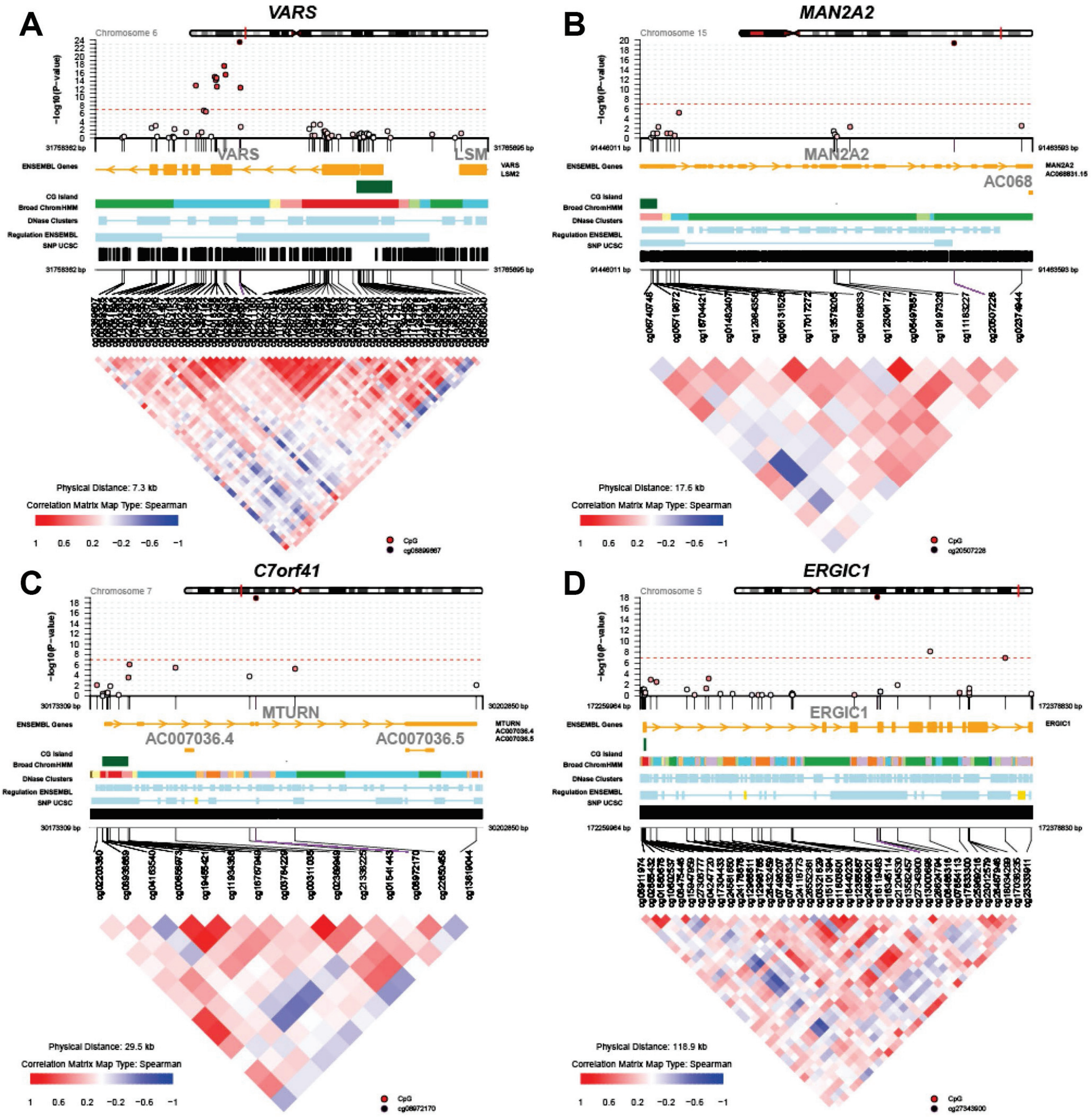


Figure 2. Regional Manhattan plots and inter-CpG correlations for the top four genes identified in the global meta-analysis. (**A**) *VARS*; (**B**) *MAN2A2*; (**C**) *C7orf41 (MTURN)*; (**D**) *ERGIC1*.

Functional enrichment analysis of LTL-associated CpG sites

To infer the biological meaning underlying LTL-associated CpG sites, the Genomic Regions Enrichment of Annotations Tool (GREAT) was used to associate differentially methylated probes (DMPs) with nearby genes of known pathway annotations. We performed both a gene-based and a region-based enrichment analysis for (1) all DMPs (n=850), (2) hypermethylated probes (n=95), and (3) hypomethylated probes (n=755).

Analyzing all DMPs, we found 11 biological annotations to be significantly enriched with both the gene-based as well as the region-based test (Supplementary File 2, Figure S1, Table S1). Of these, five annotations showed a region-fold enrichment > 1.5; the circadian clock (3.9x), blood coagulation (1.9x), hemostasis (1.9x), wound healing (1.8x), and response to wounding (1.7x). Other annotations also related to circadian rhythm, blood coagulation and wound healing, further strengthening the main observations (Supplementary File 2, Tables S1, S2).

Next, analyzing hypomethylated probes only, we found that CpGs negatively correlated with LTL mainly explain the above-mentioned functional enrichment. In contrast, hypermethylated probes led to less significant enrichment p values, a finding likely due to the lower number of CpGs (Supplementary File 3). We observed an enrichment of genes involved in mitogen-activated protein kinase phosphatase activity and immune regulation (Supplementary File 2, Figure S1). As part of a robustness/sensitivity analysis, we repeated the enrichment study after excluding CpGs with single-nucleotide polymorphisms (SNPs) in the extension base (global minor allele frequency > 1%) or probes prone to mapping to multiple regions in the genome. Across overlapping annotations (n=1,590), we found high concordance with our initial findings (r=0.97, P<2.2E-16), indicating that our results are highly robust against potentially faulty probes. Details can be found in Supplementary File 3.

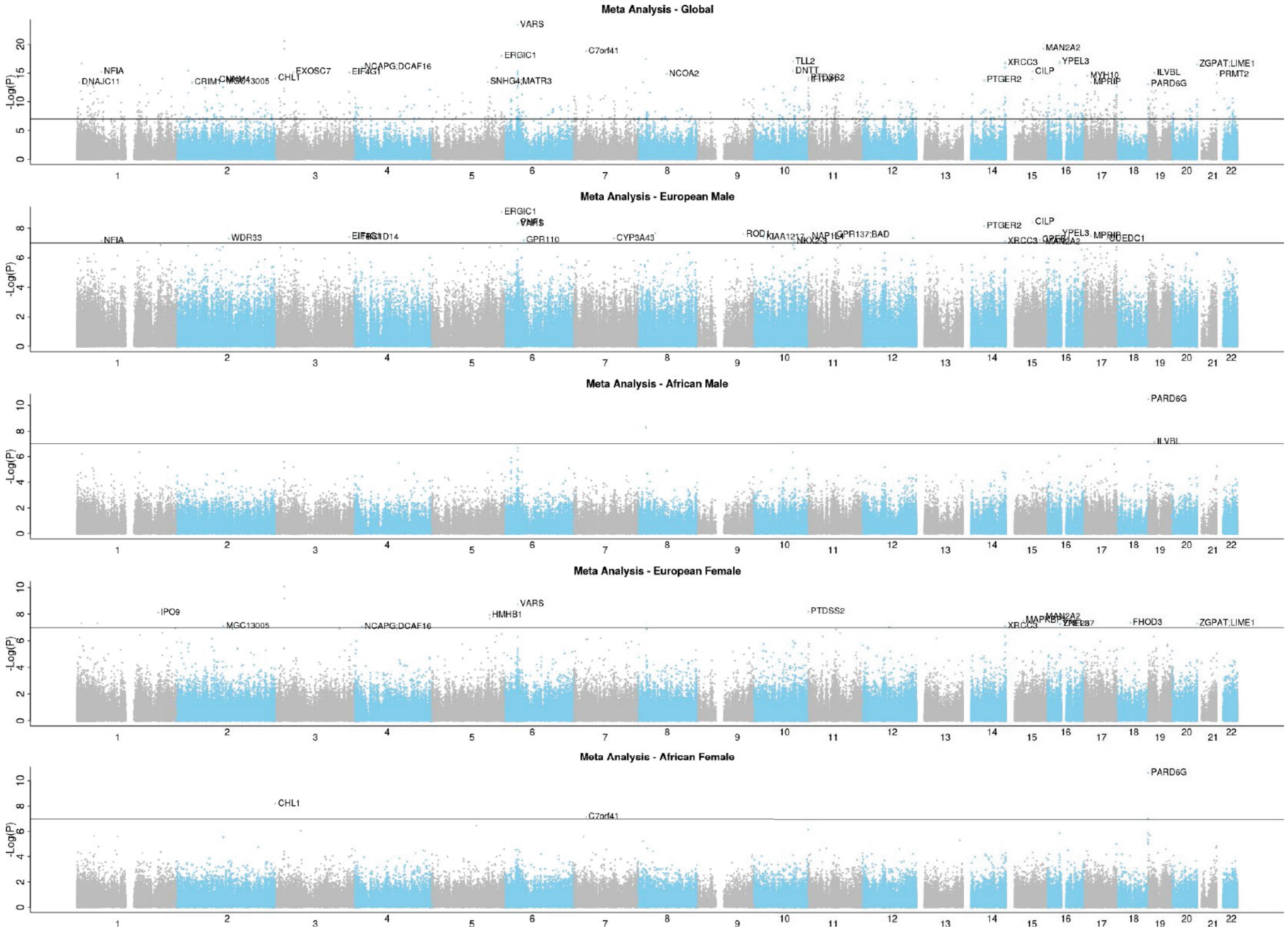


Figure 3. EWAS Manhattan plots of the fully adjusted LTL.

DNA methylation in subtelomeric regions

We observed a higher proportion of the positive LTL-DNAm correlations in subtelomeric regions than in non-subtelomeric regions when we focused on the 823 significant CpGs that were associated with the fully adjusted LTL. The proportion of the positive LTL-DNAm correlations was 17.1% in the subtelomeric regions and 9.9% in the non-subtelomeric bodies (Chi-squared test, P=0.01; Supplementary File 2, Table S3). The subtelomeric regions were defined as each chromosome's head and tail, each of which was 5% of each chromosome's length. However, this approach may not be optimal for the following reasons: 1) the inter-CpG correlations may differ between the non-subtelomeric and subtelomeric regions; 2) one cannot clearly dichotomize genomic loci into non-subtelomeric and subtelomeric regions; and 3) the LTL measurements were not chromosome-specific but averaged across all chromosomes.

Summary-data-based Mendelian randomization

We calculated the causal effects of the 823 CpGs (significantly associated with the fully adjusted LTL) on LTL using summary-data-based Mendelian randomization (SMR) [30] and found that 16 CpGs had a significant (P<0.05) causal effect on LTL (Supplementary File 2, Table S5). The causal effect of cg00622799 near *RTEL1* led to the lowest p-value (P= 6E-4) among the 823 CpGs when SNP rs909334 was used as an instrumental variable. A non-significant p-value (P=0.21) for the test for heterogeneity in independence instruments (HEIDI) is desirable because it indicates that rs909334 (instrumental variable) is the only SNP that influences LTL through the DNAm level at cg00622799. A GWAS of LTL [21] and cis methylation quantitative trait locus (cis-mQTL, a reduced GWAS of DNAm) [31] were used to obtain the SMR causal effects (betas), p-values and HEIDI p-values. The SMR p-value identifies possible methylation sites via which genetic

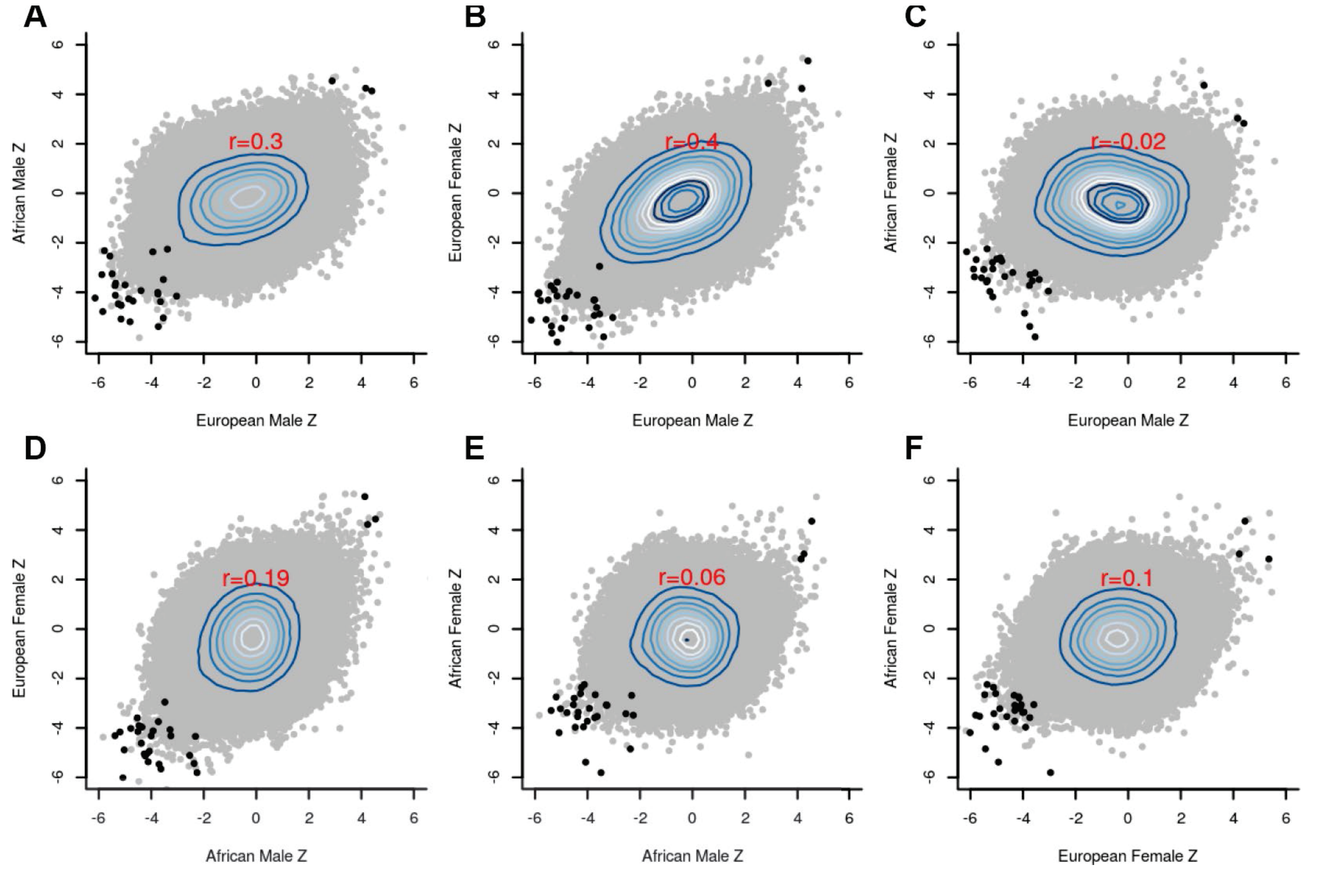


Figure 4. Scatter plots between the group-specific meta-Z scores. (**A**) European male *vs* African male; (**B**) European male *vs* European female; (**C**) European male *vs* African female; (**D**) African male *vs* European female; (**E**) African male *vs* African female; (**F**) African female *vs* European female; The black dots in the panels refer to the top 30 CpG sites detected by the global meta-analysis, whereas the grey dots indicate the remaining CpG sites. Pearson correlation coefficients (red font) reveal strong agreement (r=0.4) between males and females of European ancestry.

variants (SNPs) might be influencing LTL. The HEIDI p-value then indicates the evidence that there is (1) a single causal SNP whose effect on LTL is mediated through the methylation CpG site (HEIDI P>0.05) or (2) different SNPs linked to the methylation level and LTL (HEIDI P<0.05).

Additionally, we examined whether the 823 CpGs overlapped significantly with 54,942 known cis-methylation QTLs. Strikingly, a highly significant number of CpGs (188 CpGs out of 823 CpGs) were known cis-mQTLs (hypergeometric test P= 1.02E-16). To carry out this overlap analysis, we retrieved 188 SNPs each of which corresponded to the 188 CpGs from the cis-mQTL summary statistics. Next, we looked up each of the 188 SNPs in the most recent GWAS catalogue database (v1.02, https://www.ebi.ac.uk/gwas/docs/file-downloads). 22 SNPs were associated with complex traits (Supplementary File 2, Table S6). Among these 22 SNPs, rs2540949 in *CEP68* was associated with atrial fibrillation, and rs17708984 in *TPM4* (GWAS P=6E-16) was associated with platelet count (Supplementary File 2, Table S6). Platelet count is related to blood coagulation and wound healing, which were identified through the functional gene enrichment analysis of the LTL-associated CpGs described above.

Weighted correlation network analysis (WGCNA)

Weighted correlation network analysis (WGCNA) identified four important co-methylated modules (labeled black, red, ivory and yellow in Figure 5) using FHS, JHS and WHI (n=3,329). Hypermethylation in the black module was associated with increased age, shortened LTL, decreased CD8+ naïve T cell counts, and increased exhausted cytotoxic T cell counts, whereas hypermethylation in the red module showed opposite correlations. Elevated methylation levels in the yellow module were correlated with longer LTL and higher CD8+ naïve T cell counts. The ivory module had a pattern similar to the one in the black module. None of the modules revealed any strong correlation with the fully adjusted LTL, which is not surprising as this measure of LTL is adjusted for age and white blood cell type composition. The relationships between co-methylated module representatives and traits of interest (LTL, the partially adjusted LTL, fully adjusted LTL, age, and white blood cell counts) are displayed in Figure 6.

DISCUSSION

This multi-ethnic EWAS of LTL is the largest to date and revealed strong associations between LTL and DNAm levels in all groups defined by sex and ancestry. Our stratified analysis showed that the EWAS findings for women of African ancestry are distinct from those of three other groups: males of African ancestry, males and females of European ancestry. A detailed analysis reveals that this difference does not reflect differences in sample size, age distribution, or LTL. We analyzed 1,532 blood samples from women of African ancestry, 697 from men of African ancestry, 1,389 from men of European ancestry, and 2,095 from women of European ancestry. Although men of African ancestry had the smallest sample size, their EWAS results were consistent with those from the two European groups.

Our unadjusted meta-analysis across the groups revealed profound relationships between TL and global DNA methylation levels, which largely reflect confounding by blood cell composition. However, one can observe genome-wide significant relationships between methylation levels and LTL even after adjusting for differences in blood cell composition. In particular, we report 823 CpGs (close to or within 557 genes) that are significantly correlated with the fully adjusted LTL. More than 88 percent (730 CpGs) of these 823 significant CpG sites exhibit a negative correlation with LTL, meaning that higher methylation levels are associated with shorter LTL at these CpG sites.

Among the 823 CpGs, the top 10 CpGs were linked to seven genes/loci (*VARS, MAN2A2, C7orf41, ERGIC1, TLL2, YPEL3 and XRCC3*). *VARS* encodes the enzyme Valyl-tRNA synthetase that is critical in eukaryotic translation [32]. Mutations in *VARS* cause neurodevelopmental disorders, such as microcephaly, cortical dysgenesis, seizures, and progressive cerebral atrophy [32, 33]. *MAN2A2* encodes alpha-mannosidase 2x that is active in N-glycan biosynthesis [34]. *MAN2A2* null males were largely infertile in mouse studies [35]. *C7orf41* (current official name, *MTURN*), encodes Maturin, a protein that controls neurogenesis in the early nervous systems [36]. *ERGIC1* encodes a cycling membrane protein that contributes to membrane trafficking and selective cargo transport between intermediate compartments [37, 38]. *TLL2* encodes Tolloid-like protein 2 [39] and is associated with attention-deficit/hyperactivity disorder [40]. *YPEL3* codes for Yippee-like 3, a protein that suppresses tumor growth, proliferation and metastasis in several types of cancer [41, 42]. *XRCC3* encodes a RecA/Rad51-related protein that maintains chromosome stability and repairs DNA damage [43, 44].

Functional enrichment studies demonstrate that the significant CpG sites were located near genes that play a role in circadian clock, blood coagulation, and wound healing, respectively. A rich literature links TL to circadian rhythm. For example, cellular senescence impairs circadian rhythmicity both in vitro and in vivo [45]. Sleep disorders and shorter sleep duration are

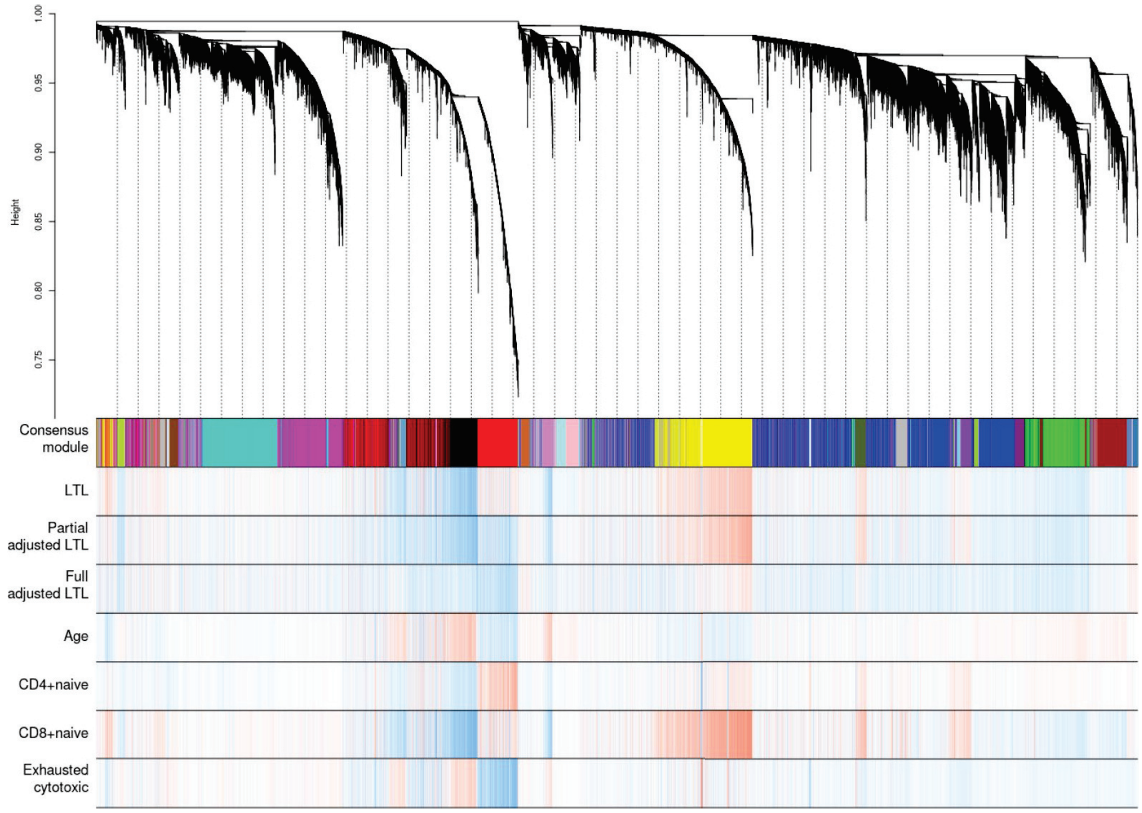
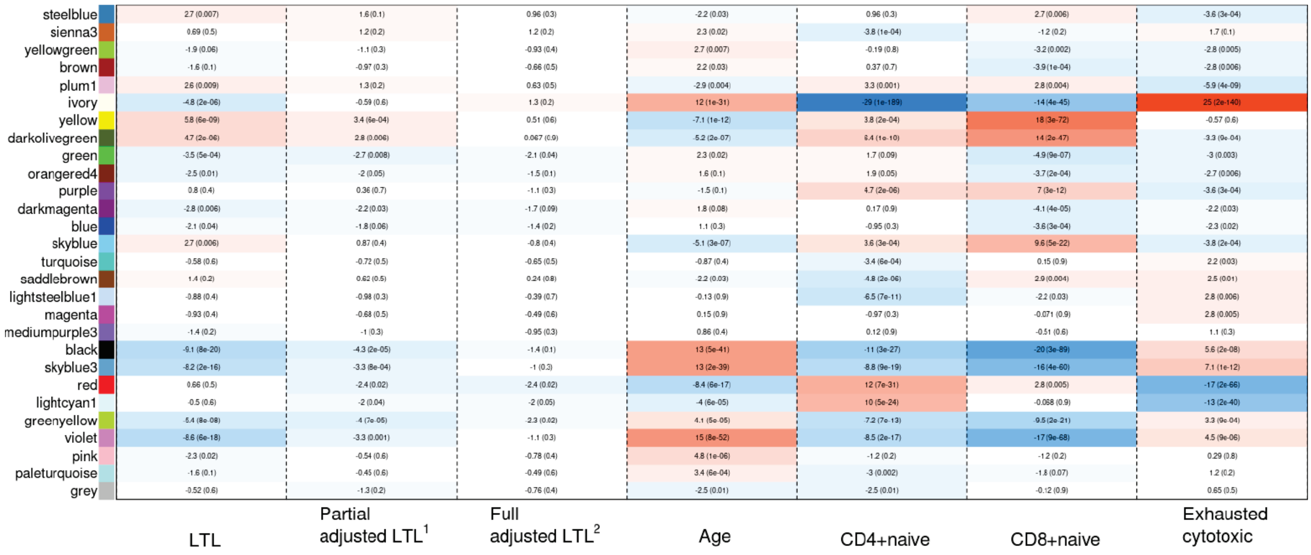


Figure 5. Hierarchical clustering of CpG sites by weighted gene co-expression network analysis (WGCNA). Each data point on the x-axis of the dendrogram refers to an individual CpG site. The color band 'Consensus module' displays co-methylated modules (clusters) in different colors. The other color bands highlight the degree of correlations between DNA methylation of CpG sites and traits of interest. Red represents a positive correlation, whereas blue represents a negative correlation.



	LTL	Partial adjusted LTL[1]	Full adjusted LTL[2]	Age	CD4+naive	CD8+naive	Exhausted cytotoxic
steelblue	2.7 (0.007)	1.6 (0.1)	0.96 (0.3)	-2.2 (0.03)	0.96 (0.3)	2.7 (0.006)	-3.6 (3e-04)
sienna3	0.69 (0.5)	1.2 (0.2)	1.2 (0.2)	2.3 (0.02)	-3.8 (1e-04)	-1.2 (0.2)	1.7 (0.1)
yellowgreen	-1.9 (0.06)	-1.1 (0.3)	-0.93 (0.4)	2.7 (0.007)	-0.19 (0.8)	-3.2 (0.002)	-2.8 (0.005)
brown	-1.6 (0.1)	-0.97 (0.3)	-0.66 (0.5)	2.2 (0.03)	0.37 (0.7)	-3.9 (1e-04)	-2.8 (0.006)
plum1	2.6 (0.009)	1.3 (0.2)	0.63 (0.5)	-2.9 (0.004)	3.3 (0.001)	2.8 (0.004)	-5.9 (4e-09)
ivory	-4.8 (2e-06)	-0.59 (0.6)	1.3 (0.2)	12 (1e-31)	-29 (1e-189)	-14 (4e-45)	25 (2e-140)
yellow	5.8 (6e-09)	3.4 (6e-04)	0.51 (0.6)	-7.1 (1e-12)	3.8 (2e-04)	18 (3e-72)	-0.57 (0.6)
darkolivegreen	4.7 (2e-06)	2.8 (0.006)	0.067 (0.9)	-5.2 (2e-07)	6.4 (1e-10)	14 (2e-47)	-3.3 (9e-04)
green	-3.5 (5e-04)	-2.7 (0.008)	-2.1 (0.04)	2.3 (0.02)	1.7 (0.09)	-4.9 (9e-07)	-3 (0.003)
orangered4	-2.5 (0.01)	-2 (0.05)	-1.5 (0.1)	1.6 (0.1)	1.9 (0.05)	-3.7 (2e-04)	-2.7 (0.006)
purple	0.8 (0.4)	0.36 (0.7)	-1.1 (0.3)	-1.5 (0.1)	4.7 (2e-06)	7 (3e-12)	-3.6 (3e-04)
darkmagenta	-2.8 (0.006)	-2.2 (0.03)	-1.7 (0.09)	1.8 (0.08)	0.17 (0.9)	-4.1 (4e-05)	-2.2 (0.03)
blue	-2.1 (0.04)	-1.8 (0.06)	-1.4 (0.2)	1.1 (0.3)	-0.95 (0.3)	-3.6 (3e-04)	-2.3 (0.02)
skyblue	2.7 (0.006)	0.87 (0.4)	-0.8 (0.4)	-5.1 (3e-07)	3.6 (3e-04)	9.6 (5e-22)	-3.8 (2e-04)
turquoise	-0.58 (0.6)	-0.72 (0.5)	-0.65 (0.5)	-0.87 (0.4)	-3.4 (6e-04)	0.15 (0.9)	2.2 (0.03)
saddlebrown	1.4 (0.2)	0.62 (0.5)	0.24 (0.8)	-2.2 (0.03)	-4.8 (2e-06)	2.9 (0.004)	2.5 (0.01)
lightsteelblue1	-0.88 (0.4)	-0.98 (0.3)	-0.39 (0.7)	-0.13 (0.9)	-6.5 (7e-11)	-2.2 (0.03)	2.8 (0.006)
magenta	-0.93 (0.4)	-0.68 (0.5)	-0.49 (0.6)	0.15 (0.9)	-0.97 (0.3)	-0.071 (0.9)	2.8 (0.005)
mediumpurple3	-1.4 (0.2)	-1 (0.3)	-0.95 (0.3)	0.86 (0.4)	0.12 (0.9)	-0.51 (0.6)	1.1 (0.3)
black	-9.1 (8e-20)	-4.3 (2e-05)	-1.4 (0.1)	13 (5e-41)	-11 (3e-27)	-20 (3e-89)	5.6 (2e-08)
skyblue3	-8.2 (2e-16)	-3.3 (8e-04)	-1 (0.3)	13 (2e-39)	-8.8 (9e-19)	-16 (4e-60)	7.1 (1e-12)
red	0.66 (0.5)	-2.4 (0.02)	-2.4 (0.02)	-8.4 (6e-17)	12 (7e-31)	2.8 (0.005)	-17 (2e-66)
lightcyan1	-0.5 (0.6)	-2 (0.04)	-2 (0.05)	-4 (6e-05)	10 (5e-24)	-0.068 (0.9)	-13 (2e-40)
greenyellow	-5.4 (8e-08)	-4 (7e-05)	-2.3 (0.02)	4.1 (5e-05)	-7.2 (7e-13)	-9.5 (2e-21)	3.3 (9e-04)
violet	-8.6 (6e-18)	-3.3 (0.001)	-1.1 (0.3)	15 (8e-52)	-8.5 (2e-17)	-17 (9e-68)	4.5 (9e-06)
pink	-2.3 (0.02)	-0.54 (0.6)	-0.78 (0.4)	4.8 (1e-06)	-1.2 (0.2)	-1.2 (0.2)	0.29 (0.8)
paleturquoise	-1.6 (0.1)	-0.45 (0.6)	-0.49 (0.6)	3.4 (6e-04)	-3 (0.002)	-1.8 (0.07)	1.2 (0.2)
grey	-0.52 (0.6)	-1.3 (0.2)	-0.76 (0.4)	-2.5 (0.01)	-2.5 (0.01)	-0.12 (0.9)	0.65 (0.5)

Figure 6. Heat map of correlations between the co-methylated module representatives and LTL, the partially adjusted LTL, the fully adjusted LTL, age, and blood cell counts. The numbers in the cells refer to meta-Z scores and their corresponding p-values. Meta-Z scores were calculated based on biweight midcorrelations between DNAm and a trait of interest in the six strata. [1]Partially adjusted LTL for age, sex and ethnicity. [2]Fully adjusted LTL for age, sex, ethnicity, CD4+ naïve, CD8+ naïve and exhausted cytotoxic T cell.

associated with shorter telomeres [46, 47]. Telomerase and TERT mRNA expression are furthermore under the control of CLOCK-BMAL1 regulation (a core component of the circadian clock) and exhibit endogenous circadian rhythms [48]. CLOCK-deficient mice display shortened TL and abnormal oscillations of telomerase activity [48]. Our results are in line with these findings and support a relationship between LTL and circadian rhythm.

TL has also been associated with wound healing and blood coagulation. For example, mice with longer telomeres show higher wound healing rates of the skin [49]. Furthermore, exogenous delivery of the human TERT gene significantly improved wound healing in an aged rabbit model [50]. In humans, poor wound healing has been reported in individuals with dyskeratosis congenita, a rare congenital disorder caused by a defect in telomere maintenance [51]. While assigning causality remains a challenge, our findings do provide evidence that telomere functioning is associated with the circadian clock, wound healing and blood coagulation through the DNA methylome in a population-based sample. Future work is needed to further understand the mechanisms by which this is regulated and how it impacts human health and diseases.

Our findings were based on a considerably larger sample size (n=5,713) than previous studies. Buxton et al. (2014) used 24 blood and 36 Epstein-Barr virus cell-line samples of 44 to 45 years old males and identified 65 and 36 TL-associated gene promoters, respectively [27]. Gadalla et al. (2012) was based on a sample of 40 cases with dyskeratosis congenita and 51 controls [28], and the authors reported a positive correlation between LTL and methylation at LINE-1 and subtelomeric sites only among the cases. Bell and colleagues performed an EWAS of age, TL and other age-related phenotypes using 172 samples of female twins [26]. Due to the small sample size, the authors could not find genome-wide significant associations between DNAm levels and TL.

We adjusted LTL for imputed blood cell composition in addition to age, sex, and ethnicity, because blood cell composition confounds the relationship between DNAm [52, 53] and LTL [54]. Consistent with previous findings, our WGCNA analyses in Figure 5 also showed that the black, red, and yellow modules were highly related to both blood cell counts and LTL. One concern was that blood cell counts might be causally influenced by DNAm and LTL (i.e., blood cell counts might be a collider between DNAm and LTL), which may introduce bias in LTL-DNAm correlations. Thus, we ran another EWAS without considering blood cell counts and compared LTL-DNAm correlations before and after adjustment for blood cell counts (Supplementary File 1). The correlations listed in Table 1 became slightly weaker after adjustment for blood cell counts but remained significant nonetheless. However, the number of associated CpG sites was greatly reduced after adjustment for blood cell counts. Cell type heterogeneity is thus an important variable to consider in studies of telomere length. Future work should be extended to cell type-specific analysis as well as to tissues beyond whole blood.

We did not adjust LTL for cigarette smoking in our main analyses because smoking had a non-significant effect on LTL (FHS: P=0.83 for never *vs* former smoker and P=0.76 for never *vs* current smoker; WHI: P=0.20 for never *vs* former smoker and P=0.24 for never *vs* current smoker), though suggestive associations could be found in JHS (P=0.08 for never *vs* former smoker and P=0.02 for never *vs* current smoker). These results pointing to a very weak effect of smoking are consistent with those from Astuti and colleagues [55] who reported that 50 of 84 studies found no association between smoking and TL, although their meta-analysis concluded that smokers may have shorter TL. Our sensitivity analyses also revealed that all the 823 CpGs remained significant regardless of smoking variables. Our EWAS summary statistics includes this sensitivity analysis with additional adjustment for smoking (see the names of columns starting with "aaa" in Supplementary File 1).

One limitation of our study is that it does not elucidate the biological pathways or mechanisms linking DNAm and LTL. In other words, our findings do not explain whether DNAm shortens or lengthens LTL, or whether LTL regulates DNAm. Second, we did not include genotypic information in our analyses. Other studies have suggested that genomic variants might regulate DNAm [31] and LTL [21–24, 56]. Third, LTL measurement is sensitive to the methods used for DNA extraction and LTL estimation [57]. Fourth, we only used EWAS and WGCNA to analyze the data. A supervised machine-learning approach for predicting TL based on DNAm levels will be described in a separate article [58].

This study represents the largest EWAS analysis of DNA methylation and LTL to date. We identified over 800 genome-wide significant CpG sites that are located in or near genes with links to circadian rhythm, blood coagulation and wound healing. These findings link two hallmarks of aging: epigenetic changes and telomere biology.

MATERIALS AND METHODS

Study population

The FHS Offspring Cohort started in 1971 to inaugurate epidemiological studies of young adults in Framingham, Massachusetts, USA. The FHS recruited 5,124 individuals and invited them to examinations at the FHS facilities [59]. The JHS recruited 5,306 African Americans from 2000 to 2004 in the Jackson metropolitan area, Mississippi, USA, to investigate risk factors for cardiovascular disease [60]. Participants provided medical history, social records and whole-blood samples. The WHI started in 1992 and enrolled 64,500 postmenopausal women aged between 50 and 79 years into either clinical trials or observational studies [61]. Among many sub-studies, WHI "Broad Agency Award 23" has provided both blood-based LTL and DNAm array data. The BHS started in 1972 and has recruited multiple waves of participants from childhood, adolescence and adulthood in Louisiana, USA [62]. The LBC1921 and LBC1936 are longitudinal studies of 550 individuals born in Scotland in 1921 and of 1091 individuals born in Scotland in 1936. The studies were set up in 1999 and 2004, respectively, with the aim of studying cognitive aging [63, 64]. The LSADT was initiated in 1995 and is a cohort-sequential study of Danish twins aged 70 years or more [65, 66]. Surviving twins were surveyed every second year until 2005. In 1997, whole blood samples were collected from 689 same-sex twins and the present study included all twin pairs who participated in the 1997 wave and for whom LTL measurements were available.

The sample size of each cohort used in this study as follows: FHS (n=874), JHS (n=1,637), WHI (n=818), BHS (n=831), LBC1921 (n=403), LBC1936 (n=906), and LSADT (n=244).

Measurement of LTL

LTL was measured by either of two methods: Southern blot [67] or qPCR [9]. All cohorts used Southern blot, except for LBC1921 and LBC1936 that used qPCR. LTL measurement by Southern blot provides the mean of TRFs, whereas qPCR provides the ratio of telomeric template to glyceraldehyde 3-phosphate dehydrogenase. The average inter-assay coefficients of variation were 2.4% in FHS, 2.0% in JHS, 2.0% in WHI, 1.4% in BHS, 5.1% in LBC (LBC1921 and LBC1936 combined), and 2.5% in LSADT. Further details on the measurement of LTL in each cohort are provided in Supplementary File 2.

Measurement of DNA methylation

DNAm data were generated on either of two different Illumina array platforms: the Illumina Infinium HumanMehtylation450 Bead-Chip (Illumina, San Diego, CA, USA) or the Illumina Infinium MethylationEPIC Bead-Chip (Illumina, San Diego, CA, USA). Beta values were computed, which quantify methylation levels between 0 and 1, with 0 being unmethylated and 1 being fully methylated. Further details on normalization and quality control of the data can be found in Supplementary File 2.

Statistical analysis

We stratified the seven cohorts (FHS, JHS, WHI, BHS, LBC1921, LBC1936 and LSADT) by sex, ethnicity and batch, which resulted in 16 strata (Table 2).

In each of the 16 strata, we applied two sets of adjustments on LTL using a regression: 1) partially adjusted for age alone, and 2) fully adjusted for age and DNAm-based estimated cell type proportions (CD4+ naïve, CD8+ naïve T cell and exhausted cytotoxic T cell). In FHS and LSADT, we used a linear mixed model to regress LTL on the adjusting variable(s) (fixed effect) and family structure (random effect). In JHS, WHI, BHS, LBC1921 and LBC1936, an ordinary linear regression was used. The blood cell type proportions were estimated using Horvath's DNAm age calculator (https://dnamage.genetics.ucla.edu/home), with the exception of LSADT where the blood cell counts were estimated using Houseman et al. (2012)'s method [68].

The ***R*** package for weighted gene co-expression network analysis (WGCNA; [69]) was used to compute epigenome-wide biweight midcorrelations between DNAm levels and adjusted LTL in each of the 16 strata. The biweight midcorrelation is an attractive method for computing correlation coefficients because 1) it is more robust than Pearson correlation and 2) unlike the Spearman correlation, it preserves the biological signal as shown in large empirical studies [70]. We focused on 441,870 autosomal probes that were shared between the 450K and the EPIC array. We combined the 16 sets of EWAS summary statistics into four group-specific or one global meta summary statistics as described in Figure 1. Meta Z values and the corresponding p-values were computed as $\sum Z_i w_i / \sqrt{\sum w_i^2}$ and $2\left(1-\Phi(|Z_{\text{meta}}|)\right)$, where w_i is the square root of the sample size in the *i*th stratum, respectively.

Genomic Regions Enrichment of Annotations Tools (GREAT, v3.0) was used to predict the biological function of DMPs by associating both proximal and distal genomic CpG sites with their putative target genes [71]. GREAT implements both a gene-based test and a region-based test using the hypergeometric and binomial test, respectively, to assess enrichment of genomic regions in biological annotations. DMPs were uploaded to the GREAT web portal (http://great.stanford.edu/public/html/) and analyses were run using the hg19 reference annotation and the whole genome as background. Genomic regions were assigned to genes if they are between 5 Kb upstream and 1 Kb downstream of the TSS, plus up to 1 Mb distal.

Table 2. Sample size of the 16 strata used in the meta-analyses.

Cohort	Stratum	Sample size	Mean age (range)	Mean LTL[2] (range)	Age-LTL correlation[3]
FHS	European female	442	57 (33-81)	7.07 (5.51-8.7)	-0.29
	European male	432	58 (36-82)	6.92 (5.59-8.52)	-0.34
JHS	African female	1034	56 (23-92)	7.22 (4.93-10.03)	-0.39
	African male	603	55 (22-93)	7.06 (5.12-9.24)	-0.45
WHI	African female	342	63 (50-80)	7.12 (5.57-9.06)	-0.24
	European female	476	68 (51-80)	6.77 (5.24-8.49)	-0.27
BHS	African female	156	44 (30-54)	7.34 (5.35-9.22)	-0.08
	African male	94	44 (33-49)	7.21 (5.60-9.47)	-0.17
	European female	315	43 (29-55)	6.82 (5.02-9.17)	-0.08
	European male	266	43 (28-52)	6.75 (5.27-8.54)	-0.18
LBC1921[1]	European female	242	79 (78-80)	3.99 (3.00-4.72)	-0.29
	European male	161	79 (78-81)	4.26 (3.46-5.31)	-0.29
LBC1936[1]	European female	448	70 (68-71)	4.05 (2.69-6.00)	0.01
	European male	458	70 (68-71)	4.33 (2.99-7.12)	0.17
LSADT	European female	172	79 (73-90)	5.79 (3.94-7.38)	-0.25
	European male	72	79 (74-87)	5.60 (4.53-6.78)	-0.17

[1] LBC recruited adults living in and around Edinburgh and who were born in 1921 and 1936.

[2] In kilobases; LTL measurement in TRF (Southern blot): FHS, JHS, WHI, BHS and LSADT; LTL measurement in T/S (qPCR): LBC1921 and LBC1936.

[3] Pearson correlation coefficients.

Pathway annotations from GO Biological Processes, GO Cellular Component, GO Molecular Function, MSigDB, and PANTHER were used to infer the biological meanings behind the DMPs that were associated with LTL. GREAT outputs statistics of the gene-based and region-based tests, which were subsequently adjusted for multiple testing using the Bonferroni correction.

The SMR executable software (https://cnsgenomics.com/software/smr/#Download) was used to calculate the causal effects of the selected CpGs on LTL [30]. The SMR obtains a causal effect estimate ($\hat{b}_{CpG,LTL}$ = $\hat{b}_{SNP,LTL}/\hat{\beta}_{SNP,CpG}$) by dividing the effect of a SNP on LTL ($\hat{b}_{SNP,LTL}$) by the effect of a SNP on CpG ($\hat{\beta}_{SNP,CpG}$). GWAS of LTL summary data by Codd and colleagues [21] was downloaded from the European Network for Genetic and Genomic Epidemiology consortium (https://downloads.lcbru.le.ac.uk/engage). The mQTL data by McRae and colleagues [31] were downloaded from the SMR website (http://cnsgenomics.com/data/SMR/LBC_BSGS_meta.tar.gz).

WGCNA performed a consensus network analysis using FHS, JHS and WHI. 30,000 randomly selected CpG sites were used to improve readability (resulting in a single cluster tree) and offset computational limitations. WGCNA hierarchically clustered the 30,000 CpGs based on their similarities. The merging threshold of clusters (modules) was 0.15. All the statistical analyses were performed using ***R*** version 3.5.1.

Abbreviations

LTL: leukocyte telomere length; TL: telomere length, DNAm: DNA methylation; TRF: terminal restriction fragment; qPCR: quantitative real-time polymerase chain reaction; GWAS: Genome-wide association study; EWAS: epigenome-wide association study; FHS: the Framingham Heart Study; JHS: the Jackson Heart Study; WHI: the Women's Health Initiative; BHS: the Bogalusa Heart Study; LBC: the Lothian Birth Cohorts; LSADT: the Longitudinal Study of Aging Danish Twins; DMP: differentially methylated probe; SNP: single-nucleotide polymorphism; SMR: summary-data-based Mendelian randomization; HEIDI: heterogeneity in independence instruments; mQTL: methylation quantitative trait locus; WGCNA: Weighted correlation network analysis; DC: dyskeratosis congenital; qPCR: quantitative real-time polymerase chain reaction; GREAT: Genomic Regions Enrichment of Annotations Tools.

AUTHOR CONTRIBUTIONS

YL and SH conducted EWAS of LTL using FHS, JHS and WHI and wrote the manuscript. DS, AS and MS conducted EWAS of LTL using BHS, LBC and LSADT, respectively. AO conducted gene enrichment analyses and wrote the biological interpretations. The remaining authors contributed data, helped with manuscript preparation, and interpreted the results.

CONFLICTS OF INTEREST

The authors have no conflicts of interest.

FUNDING

This work was partly supported by a grant from the Norwegian Research Council (NRC) to AJ (project number 262043), additional funding from NRC through a Personal Overseas Research Grant to YL (project number 262043/F20) and a grant from NIH to AA (R01HL134840-01).

SH and ATL acknowledge support from 1U01AG060908-01.

FHS is funded by the National Institute of Health (NIH) contract N01-HC-25195 and HHSN268201500001I. The laboratory work for this investigation was funded by the Division of Intramural Research, National Heart, Lung, and Blood Institute, NIH. The analytical component of this project was funded by the Division of Intramural Research, National Heart, Lung, and Blood Institute, and the Center for Information Technology, NIH, Bethesda, MD. JMM and KLL were supported by R01AG029451.

JHS is supported and conducted in collaboration with Jackson State University (HHSN268201800013I), Tougaloo College (HHSN268201800014I), the Mississippi State Department of Health (HHSN268201800015I/HHSN26800001) and the University of Mississippi Medical Center (HHSN268201800010I, HHSN268201800011I and HHSN268201800012I) contracts from the National Heart, Lung, and Blood Institute and the National Institute for Minority Health and Health Disparities. The authors also wish to thank the staff and participants of the JHS. JGW is supported by U54GM115428 from the National Institute of General Medical Sciences.

WHI program is funded by the National Heart, Lung, and Blood Institute, NIH, and the U.S. Department of Health and Human Services through contracts HHSN268201600018C, HHSN268201600001C, HHSN268201600002C, HHSN268201600003C, and HHSN268201600004C. The authors thank the WHI investigators and staff for their dedication, and the study participants for making the program possible. A full listing of WHI investigators can be found at: http://www.whi.org/researchers/Documents%20%20Write%20a%20Paper/WHI%20Investigator%20Long%20List.pdf

BHS was supported by grants R01AG016592 and R03AG060619 from National Institute of Aging. We thank the participants and investigators and staff members of the BHS for their outstanding commitment and cooperation.

LBC1921 was supported by the UK's Biotechnology and Biological Sciences Research Council (BBSRC), a Royal Society–Wolfson Research Merit Award to IJD, and the Chief Scientist Office (CSO) of the Scottish Government's Health Directorates. LBC1936 is supported by Age UK (Disconnected Mind program) and the Medical Research Council (MR/M01311/1). Methylation typing was supported by Centre for Cognitive Ageing and Cognitive Epidemiology (Pilot Fund award), Age UK, The Wellcome Trust Institutional Strategic Support Fund, The University of Edinburgh, and The University of Queensland. This work was in part conducted in the Centre for Cognitive Ageing and Cognitive Epidemiology, which is supported by the Medical Research Council and Biotechnology and Biological Sciences Research Council (MR/K026992/1). AS is supported by a Medical Research Council PhD Studentship in Precision Medicine with funding by the Medical Research Council Doctoral Training Program and the University of Edinburgh College of Medicine and Veterinary Medicine. REM is supported by the Alzheimer's Research UK major project grant ARUK-PG2017B-10. The authors thank LBC study participants and research team members who have contributed, and continue to contribute, to ongoing LBC studies.

LSADT was funded by The Danish Council for Independent Research – Medical Sciences (DFF-6110-00016), NIH (NIH-NIA P01 AG08761) and the European Union's Seventh Framework Program (FP7/2007-2011) under grant agreement number 259679.

REFERENCES

1. Blackburn EH. Telomeres and telomerase: their mechanisms of action and the effects of altering their functions. FEBS Lett. 2005; 579:859–62. https://doi.org/10.1016/j.febslet.2004.11.036 PMID:15680963

2. Wright WE, Piatyszek MA, Rainey WE, Byrd W, Shay JW. Telomerase activity in human germline and embryonic tissues and cells. Dev Genet. 1996; 18:173–79.

https://doi.org/10.1002/(SICI)1520-6408(1996)18:2<173::AID-DVG10>3.0.CO;2-3 PMID:8934879

3. Chiu CP, Dragowska W, Kim NW, Vaziri H, Yui J, Thomas TE, Harley CB, Lansdorp PM. Differential expression of telomerase activity in hematopoietic progenitors from adult human bone marrow. Stem Cells. 1996; 14:239–48. https://doi.org/10.1002/stem.140239 PMID:8991544

4. Yui J, Chiu CP, Lansdorp PM. Telomerase activity in candidate stem cells from fetal liver and adult bone marrow. Blood. 1998; 91:3255–62. PMID:9558381

5. Shay JW, Wright WE. The reactivation of telomerase activity in cancer progression. Trends Genet. 1996; 12:129–31. https://doi.org/10.1016/0168-9525(96)30018-8 PMID:8901415

6. Gomes NM, Ryder OA, Houck ML, Charter SJ, Walker W, Forsyth NR, Austad SN, Venditti C, Pagel M, Shay JW, Wright WE. Comparative biology of mammalian telomeres: hypotheses on ancestral states and the roles of telomeres in longevity determination. Aging Cell. 2011; 10:761–68. https://doi.org/10.1111/j.1474-9726.2011.00718.x PMID:21518243

7. Allsopp RC, Vaziri H, Patterson C, Goldstein S, Younglai EV, Futcher AB, Greider CW, Harley CB. Telomere length predicts replicative capacity of human fibroblasts. Proc Natl Acad Sci USA. 1992; 89:10114–18. https://doi.org/10.1073/pnas.89.21.10114 PMID:1438199

8. Gardner M, Bann D, Wiley L, Cooper R, Hardy R, Nitsch D, Martin-Ruiz C, Shiels P, Sayer AA, Barbieri M, Bekaert S, Bischoff C, Brooks-Wilson A, et al, and Halcyon study team. Gender and telomere length: systematic review and meta-analysis. Exp Gerontol. 2014; 51:15–27. https://doi.org/10.1016/j.exger.2013.12.004 PMID:24365661

9. Cawthon RM. Telomere measurement by quantitative PCR. Nucleic Acids Res. 2002; 30:e47. https://doi.org/10.1093/nar/30.10.e47 PMID:12000852

10. Hunt SC, Chen W, Gardner JP, Kimura M, Srinivasan SR, Eckfeldt JH, Berenson GS, Aviv A. Leukocyte telomeres are longer in African Americans than in whites: the National Heart, Lung, and Blood Institute Family Heart Study and the Bogalusa Heart Study. Aging Cell. 2008; 7:451–58. https://doi.org/10.1111/j.1474-9726.2008.00397.x PMID:18462274

11. Elbers CC, Garcia ME, Kimura M, Cummings SR, Nalls MA, Newman AB, Park V, Sanders JL, Tranah GJ, Tishkoff SA, Harris TB, Aviv A. Comparison between southern blots and qPCR analysis of leukocyte telomere length in the health ABC study. J Gerontol A Biol Sci Med Sci. 2014; 69:527–31. https://doi.org/10.1093/gerona/glt121 PMID:23946336

12. Aviv A, Shay JW. Reflections on telomere dynamics and ageing-related diseases in humans. Philos Trans R Soc Lond B Biol Sci. 2018; 373:373. https://doi.org/10.1098/rstb.2016.0436 PMID:29335375

13. Stone RC, Horvath K, Kark JD, Susser E, Tishkoff SA, Aviv A. Telomere Length and the Cancer-Atherosclerosis Trade-Off. PLoS Genet. 2016; 12:e1006144. https://doi.org/10.1371/journal.pgen.1006144 PMID:27386863

14. Samavat H, Xun X, Jin A, Wang R, Koh WP, Yuan JM. Association between prediagnostic leukocyte telomere length and breast cancer risk: the Singapore Chinese Health Study. Breast Cancer Res. 2019; 21:50. https://doi.org/10.1186/s13058-019-1133-0 PMID:30995937

15. Andrew T, Aviv A, Falchi M, Surdulescu GL, Gardner JP, Lu X, Kimura M, Kato BS, Valdes AM, Spector TD. Mapping genetic loci that determine leukocyte telomere length in a large sample of unselected female sibling pairs. Am J Hum Genet. 2006; 78:480–86. https://doi.org/10.1086/500052 PMID:16400618

16. Hjelmborg JB, Dalgård C, Möller S, Steenstrup T, Kimura M, Christensen K, Kyvik KO, Aviv A. The heritability of leucocyte telomere length dynamics. J Med Genet. 2015; 52:297–302. https://doi.org/10.1136/jmedgenet-2014-102736 PMID:25770094

17. Slagboom PE, Droog S, Boomsma DI. Genetic determination of telomere size in humans: a twin study of three age groups. Am J Hum Genet. 1994; 55:876–82. PMID:7977349

18. Vasa-Nicotera M, Brouilette S, Mangino M, Thompson JR, Braund P, Clemitson JR, Mason A, Bodycote CL, Raleigh SM, Louis E, Samani NJ. Mapping of a major locus that determines telomere length in humans. Am J Hum Genet. 2005; 76:147–51. https://doi.org/10.1086/426734 PMID:15520935

19. Broer L, Codd V, Nyholt DR, Deelen J, Mangino M, Willemsen G, Albrecht E, Amin N, Beekman M, de Geus EJ, Henders A, Nelson CP, Steves CJ, et al. Meta-analysis of telomere length in 19,713 subjects reveals high

heritability, stronger maternal inheritance and a paternal age effect. Eur J Hum Genet. 2013; 21:1163–68. https://doi.org/10.1038/ejhg.2012.303 PMID:23321625

20. Zhu Y, Voruganti VS, Lin J, Matsuguchi T, Blackburn E, Best LG, Lee ET, MacCluer JW, Cole SA, Zhao J. QTL mapping of leukocyte telomere length in American Indians: the Strong Heart Family Study. Aging (Albany NY). 2013; 5:704–16. https://doi.org/10.18632/aging.100600 PMID:24036517

21. Codd V, Nelson CP, Albrecht E, Mangino M, Deelen J, Buxton JL, Hottenga JJ, Fischer K, Esko T, Surakka I, Broer L, Nyholt DR, Mateo Leach I, et al, and CARDIoGRAM consortium. Identification of seven loci affecting mean telomere length and their association with disease. Nat Genet. 2013; 45:422–7, 427e1–2. https://doi.org/10.1038/ng.2528 PMID:23535734

22. Levy D, Neuhausen SL, Hunt SC, Kimura M, Hwang SJ, Chen W, Bis JC, Fitzpatrick AL, Smith E, Johnson AD, Gardner JP, Srinivasan SR, Schork N, et al. Genome-wide association identifies OBFC1 as a locus involved in human leukocyte telomere biology. Proc Natl Acad Sci USA. 2010; 107:9293–98. https://doi.org/10.1073/pnas.0911494107 PMID:20421499

23. Mangino M, Christiansen L, Stone R, Hunt SC, Horvath K, Eisenberg DT, Kimura M, Petersen I, Kark JD, Herbig U, Reiner AP, Benetos A, Codd V, et al. DCAF4, a novel gene associated with leucocyte telomere length. J Med Genet. 2015; 52:157–62. https://doi.org/10.1136/jmedgenet-2014-102681 PMID:25624462

24. Mangino M, Hwang SJ, Spector TD, Hunt SC, Kimura M, Fitzpatrick AL, Christiansen L, Petersen I, Elbers CC, Harris T, Chen W, Srinivasan SR, Kark JD, et al. Genome-wide meta-analysis points to CTC1 and ZNF676 as genes regulating telomere homeostasis in humans. Hum Mol Genet. 2012; 21:5385–94. https://doi.org/10.1093/hmg/dds382 PMID:23001564

25. Rakyan VK, Down TA, Balding DJ, Beck S. Epigenome-wide association studies for common human diseases. Nat Rev Genet. 2011; 12:529–41. https://doi.org/10.1038/nrg3000 PMID:21747404

26. Bell JT, Tsai PC, Yang TP, Pidsley R, Nisbet J, Glass D, Mangino M, Zhai G, Zhang F, Valdes A, Shin SY, Dempster EL, Murray RM, et al, and MuTHER Consortium. Epigenome-wide scans identify differentially methylated regions for age and age-related phenotypes in a healthy ageing population. PLoS Genet. 2012; 8:e1002629. https://doi.org/10.1371/journal.pgen.1002629 PMID:22532803

27. Buxton JL, Suderman M, Pappas JJ, Borghol N, McArdle W, Blakemore AI, Hertzman C, Power C, Szyf M, Pembrey M. Human leukocyte telomere length is associated with DNA methylation levels in multiple subtelomeric and imprinted loci. Sci Rep. 2014; 4:4954. https://doi.org/10.1038/srep04954 PMID:24828261

28. Gadalla SM, Katki HA, Shebl FM, Giri N, Alter BP, Savage SA. The relationship between DNA methylation and telomere length in dyskeratosis congenita. Aging Cell. 2012; 11:24–28. https://doi.org/10.1111/j.1474-9726.2011.00755.x PMID:21981348

29. Mansell G, Gorrie-Stone TJ, Bao Y, Kumari M, Schalkwyk LS, Mill J, Hannon E. Guidance for DNA methylation studies: statistical insights from the Illumina EPIC array. BMC Genomics. 2019; 20:366. https://doi.org/10.1186/s12864-019-5761-7 PMID:31088362

30. Zhu Z, Zhang F, Hu H, Bakshi A, Robinson MR, Powell JE, Montgomery GW, Goddard ME, Wray NR, Visscher PM, Yang J. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nat Genet. 2016; 48:481–87. https://doi.org/10.1038/ng.3538 PMID:27019110

31. McRae AF, Marioni RE, Shah S, Yang J, Powell JE, Harris SE, Gibson J, Henders AK, Bowdler L, Painter JN, Murphy L, Martin NG, Starr JM, et al. Identification of 55,000 Replicated DNA Methylation QTL. Sci Rep. 2018; 8:17605. https://doi.org/10.1038/s41598-018-35871-w PMID:30514905

32. Stephen J, Nampoothiri S, Banerjee A, Tolman NJ, Penninger JM, Elling U, Agu CA, Burke JD, Devadathan K, Kannan R, Huang Y, Steinbach PJ, Martinis SA, et al. Loss of function mutations in VARS encoding cytoplasmic valyl-tRNA synthetase cause microcephaly, seizures, and progressive cerebral atrophy. Hum Genet. 2018; 137:293–303. https://doi.org/10.1007/s00439-018-1882-3 PMID:29691655

33. Karaca E, Harel T, Pehlivan D, Jhangiani SN, Gambin T, Coban Akdemir Z, Gonzaga-Jauregui C, Erdin S, Bayram Y, Campbell IM, Hunter JV, Atik MM, Van Esch H, et al. Genes that Affect Brain Structure and Function Identified by Rare Variant Analyses of Mendelian Neurologic Disease. Neuron. 2015; 88:499–513. https://doi.org/10.1016/j.neuron.2015.09.048 PMID:26539891

34. Misago M, Liao YF, Kudo S, Eto S, Mattei MG, Moremen KW, Fukuda MN. Molecular cloning and expression of cDNAs encoding human alpha-mannosidase II and a previously unrecognized alpha-mannosidase IIx isozyme. Proc Natl Acad Sci USA. 1995; 92:11766–70.
https://doi.org/10.1073/pnas.92.25.11766
PMID:8524845

35. Akama TO, Nakagawa H, Sugihara K, Narisawa S, Ohyama C, Nishimura S, O'Brien DA, Moremen KW, Millan JL, Fukuda MN. Germ cell survival through carbohydrate-mediated interaction with Sertoli cells. Science. 2002; 295:124–27.
https://doi.org/10.1126/science.1065570
PMID:11778047

36. Martinez-De Luna RI, Ku RY, Lyou Y, Zuber ME. Maturin is a novel protein required for differentiation during primary neurogenesis. Dev Biol. 2013; 384:26–40.
https://doi.org/10.1016/j.ydbio.2013.09.028
PMID:24095902

37. Breuza L, Halbeisen R, Jenö P, Otte S, Barlowe C, Hong W, Hauri HP. Proteomics of endoplasmic reticulum-Golgi intermediate compartment (ERGIC) membranes from brefeldin A-treated HepG2 cells identifies ERGIC-32, a new cycling protein that interacts with human Erv46. J Biol Chem. 2004; 279:47242–53.
https://doi.org/10.1074/jbc.M406644200
PMID:15308636

38. Vainio P, Mpindi JP, Kohonen P, Fey V, Mirtti T, Alanen KA, Perälä M, Kallioniemi O, Iljin K. High-throughput transcriptomic and RNAi analysis identifies AIM1, ERGIC1, TMED3 and TPX2 as potential drug targets in prostate cancer. PLoS One. 2012; 7:e39801.
https://doi.org/10.1371/journal.pone.0039801
PMID:22761906

39. Scott IC, Clark TG, Takahara K, Hoffman GG, Eddy RL, Haley LL, Shows TB, Greenspan DS. Assignment of TLL1 and TLL2, which encode human BMP-1/Tolloid-related metalloproteases, to chromosomes 4q32—>q33 and 10q23—>q24 and assignment of murine Tll2 to chromosome 19. Cytogenet Cell Genet. 1999; 86:64–65.
https://doi.org/10.1159/000015412
PMID:10516436

40. Lesch KP, Timmesfeld N, Renner TJ, Halperin R, Röser C, Nguyen TT, Craig DW, Romanos J, Heine M, Meyer J, Freitag C, Warnke A, Romanos M, et al. Molecular genetics of adult ADHD: converging evidence from genome-wide association and extended pedigree linkage studies. J Neural Transm (Vienna). 2008; 115:1573–85.
https://doi.org/10.1007/s00702-008-0119-3
PMID:18839057

41. Tuttle R, Miller KR, Maiorano JN, Termuhlen PM, Gao Y, Berberich SJ. Novel senescence associated gene, YPEL3, is repressed by estrogen in ER+ mammary tumor cells and required for tamoxifen-induced cellular senescence. Int J Cancer. 2012; 130:2291–99.
https://doi.org/10.1002/ijc.26239
PMID:21671470

42. Tuttle R, Simon M, Hitch DC, Maiorano JN, Hellan M, Ouellette J, Termuhlen P, Berberich SJ. Senescence-associated gene YPEL3 is downregulated in human colon tumors. Ann Surg Oncol. 2011; 18:1791–96.
https://doi.org/10.1245/s10434-011-1558-x
PMID:21267786

43. Pierce AJ, Johnson RD, Thompson LH, Jasin M. XRCC3 promotes homology-directed repair of DNA damage in mammalian cells. Genes Dev. 1999; 13:2633–38.
https://doi.org/10.1101/gad.13.20.2633
PMID:10541549

44. Tebbs RS, Zhao Y, Tucker JD, Scheerer JB, Siciliano MJ, Hwang M, Liu N, Legerski RJ, Thompson LH. Correction of chromosomal instability and sensitivity to diverse mutagens by a cloned cDNA of the XRCC3 DNA repair gene. Proc Natl Acad Sci USA. 1995; 92:6354–58.
https://doi.org/10.1073/pnas.92.14.6354
PMID:7603995

45. Kunieda T, Minamino T, Katsuno T, Tateno K, Nishi J, Miyauchi H, Orimo M, Okada S, Komuro I. Cellular senescence impairs circadian expression of clock genes in vitro and in vivo. Circ Res. 2006; 98:532–39.
https://doi.org/10.1161/01.RES.0000204504.25798.a8
PMID:16424366

46. Gaspar LS, Álvaro AR, Moita J, Cavadas C. Obstructive Sleep Apnea and Hallmarks of Aging. Trends Mol Med. 2017; 23:675–92.
https://doi.org/10.1016/j.molmed.2017.06.006
PMID:28739207

47. James S, McLanahan S, Brooks-Gunn J, Mitchell C, Schneper L, Wagner B, Notterman DA. Sleep Duration and Telomere Length in Children. J Pediatr. 2017; 187:247–52.e1.
https://doi.org/10.1016/j.jpeds.2017.05.014
PMID:28602380

48. Chen WD, Wen MS, Shie SS, Lo YL, Wo HT, Wang CC, Hsieh IC, Lee TH, Wang CY. The circadian rhythm controls telomeres and telomerase activity. Biochem Biophys Res Commun. 2014; 451:408–14.
https://doi.org/10.1016/j.bbrc.2014.07.138
PMID:25109806

49. Varela E, Muñoz-Lorente MA, Tejera AM, Ortega S, Blasco MA. Generation of mice with longer and better preserved telomeres in the absence of genetic manipulations. Nat Commun. 2016; 7:11739.

https://doi.org/10.1038/ncomms11739 PMID:27252083

50. Mogford JE, Liu WR, Reid R, Chiu CP, Said H, Chen SJ, Harley CB, Mustoe TA. Adenoviral human telomerase reverse transcriptase dramatically improves ischemic wound healing without detrimental immune response in an aged rabbit model. Hum Gene Ther. 2006; 17:651–60. https://doi.org/10.1089/hum.2006.17.651 PMID:16776573

51. Marciniak RA, Johnson FB, Guarente L. Dyskeratosis congenita, telomeres and human ageing. Trends Genet. 2000; 16:193–95. https://doi.org/10.1016/S0168-9525(00)01984-3 PMID:10782108

52. 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

53. Horvath S. Erratum to: DNA methylation age of human tissues and cell types. Genome Biol. 2015; 16:96. https://doi.org/10.1186/s13059-015-0649-6 PMID:25968125

54. Weng NP, Levine BL, June CH, Hodes RJ. Human naive and memory T lymphocytes differ in telomeric length and replicative potential. Proc Natl Acad Sci USA. 1995; 92:11091–94. https://doi.org/10.1073/pnas.92.24.11091 PMID:7479943

55. Astuti Y, Wardhana A, Watkins J, Wulaningsih W, Network PR, and PILAR Research Network. Cigarette smoking and telomere length: A systematic review of 84 studies and meta-analysis. Environ Res. 2017; 158:480–89. https://doi.org/10.1016/j.envres.2017.06.038 PMID:28704792

56. Lu AT, Xue L, Salfati EL, Chen BH, Ferrucci L, Levy D, Joehanes R, Murabito JM, Kiel DP, Tsai PC, Yet I, Bell JT, Mangino M, et al. GWAS of epigenetic aging rates in blood reveals a critical role for TERT. Nat Commun. 2018; 9:387. https://doi.org/10.1038/s41467-017-02697-5 PMID:29374233

57. Barrett JH, Iles MM, Dunning AM, Pooley KA. Telomere length and common disease: study design and analytical challenges. Hum Genet. 2015; 134:679–89. https://doi.org/10.1007/s00439-015-1563-4 PMID:25986438

58. Lu AT, Seeboth A, Tsai PC, Sun D, Quach A, Reiner AP, Kooperberg C, Ferrucci L, Hou L, Baccarelli A, Li Y, Harris SE, Corley J, et al. DNA methylation-based estimator of telomere length. Aging (Albany NY). 2019. [Epub ahead of print]. https://doi.org/10.18632/aging.102173 PMID:31422385

59. Kannel WB, Feinleib M, McNamara PM, Garrison RJ, Castelli WP. An investigation of coronary heart disease in families. The Framingham offspring study. Am J Epidemiol. 1979; 110:281–90. https://doi.org/10.1093/oxfordjournals.aje.a112813 PMID:474565

60. 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:S6-4-17. PMID:16320381

61. Anderson G, Cummings S, Freedman LS, Furberg C, Henderson M, Johnson SR, Kuller L, Manson J, Oberman A, Prentice RL, Rossouw JE. 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

62. Berenson GS. Bogalusa Heart Study: a long-term community study of a rural biracial (black/white) population. Am J Med Sci. 2001; 322:267–74. https://doi.org/10.1097/00000441-200111000-00007 PMID:11721800

63. Deary IJ, Gow AJ, Pattie A, Starr JM. Cohort profile: the Lothian Birth Cohorts of 1921 and 1936. Int J Epidemiol. 2012; 41:1576–84. https://doi.org/10.1093/ije/dyr197 PMID:22253310

64. Taylor AM, Pattie A, Deary IJ. Cohort Profile Update: The Lothian Birth Cohorts of 1921 and 1936. Int J Epidemiol. 2018; 47:1042–1042-r. https://doi.org/10.1093/ije/dyy022 PMID:29546429

65. Christensen K, Holm NV, McGue M, Corder L, Vaupel JW. A Danish population-based twin study on general health in the elderly. J Aging Health. 1999; 11:49–64. https://doi.org/10.1177/089826439901100103 PMID:10848141

66. McGue M, Christensen K. Social activity and healthy aging: a study of aging Danish twins. Twin Res Hum Genet. 2007; 10:255–65. https://doi.org/10.1375/twin.10.2.255 PMID:17564515

67. Kimura M, Stone RC, Hunt SC, Skurnick J, Lu X, Cao X, Harley CB, Aviv A. Measurement of telomere length by the Southern blot analysis of terminal restriction fragment lengths. Nat Protoc. 2010; 5:1596–607. https://doi.org/10.1038/nprot.2010.124

PMID:21085125

68. 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

69. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008; 9:559. https://doi.org/10.1186/1471-2105-9-559 PMID:19114008

70. Song L, Langfelder P, Horvath S. Comparison of co-expression measures: mutual information, correlation, and model based indices. BMC Bioinformatics. 2012; 13:328. https://doi.org/10.1186/1471-2105-13-328 PMID:23217028

71. McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, Wenger AM, Bejerano G. GREAT improves functional interpretation of cis-regulatory regions. Nat Biotechnol. 2010; 28:495–501. https://doi.org/10.1038/nbt.1630 PMID:20436461

SUPPLEMENTARY MATERIALS

Please browse Full Text version to see the data of Supplementary Files 1, 2, 3.

Supplementary File 1. Part of summary statistics of EWAS of adjusted LTL (global meta P<1E-05 with full adjustment). Each row corresponds to a single CpG site. The annotations are based on the Human genome 19 (NCBI 37). The remaining columns indicate the biweight midcorrelations and their corresponding Z-scores, p-values and sample size. The suffix "a_" means that LTL was adjusted for age, sex and ethnicity. The suffix "aa_" means that LTL was adjusted for age, sex, ethnicity and blood cell counts. The suffix "aaa_" means that LTL was adjusted for age, sex, ethnicity, blood cell counts and smoking.

Supplementary File 2.Additional analyses for 1) functional enrichment analysis, 2) the LTL-DNAm correlation in subtelomeric regions, 3) summary-data-based Mendelian randomization, 4) sensitivity analyses, and 5) detailed descriptions of each study cohort.

Supplementary File 3. GREAT gene enrichment analyses.