Effects of DNA Methylation on Cardiovascular Disease, Target Organ Damage, and their Risk Factors in African Americans

by

Farah Ammous

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Epidemiological Science)
in the University of Michigan
2021

Doctoral Committee:

Associate Professor Jennifer A. Smith, Chair
Associate Research Scientist Lawrence F. Bielak
Professor Sharon L.R. Kardia
Assistant Research Scientist Wei Zhao
Associate Professor Xiang Zhou

Farah Ammous

fammous@umich.edu

ORCID iD: 0000-0002-8273-4539

© Farah Ammous 2021

DEDICATION

This dissertation is dedicated to my parents Rashad and Alia.

ACKNOWLEDGEMENTS

I would like to extend my deepest gratitude to the many individuals who have supported me during this dissertation process. First and foremost, I would like to thank my chair, Jennifer Smith, for her incredible mentorship and commitment to my success. From helping to me think critically about my research to providing regular feedback on my writing, Jen has been a source of endless support and guidance. She has continuously challenged me intellectually and pushed my comfort zone as a researcher. I have learnt a lot from Jen in the past four years and this dissertation would certainly not be possible without her mentorship. I am forever grateful for her support and for taking me under her wing.

I am also thankful to the other members of my dissertation committee for their support, patience, and expertise. Thanks to Wei Zhao for being so generous with her time, for helping me hone my coding skills, and providing critical feedback which improved my dissertation immensely. I thank Xiang Zhou for his valuable statistical insights and for being always available to answer my questions on statistical methods and analysis. I thank Lawrence Bielak for his valuable feedback and for challenging me to consider different perspectives. Finally, I thank Sharon Kardia for her expertise and for helping me to engage more deeply with issues related to epigenetics and its implications for health. I am also indebted to the participants of the Genetic Epidemiology Network of Arteriopathy for their participating and sharing a wealth of data for that made this dissertation possible.

I am thankful to the Department of Epidemiology for taking a chance on me and accepting me to this program. Thank you to all the faculty and staff who provided me with a supportive and rich academic environment to pursue my doctorate. I am also extremely grateful to Patricia Peyser for all her support and mentorship which opened the doors to this journey and helped me throughout. I am also grateful for the many sources of funding that supported me during my studies and allowed me to attend conference and workshops to enhance my training.

I have been very fortunate to be a part of the Genomics Research Collaborative which has been a wonderful intellectual space to connect with fellow students and staff: Dima Chaar, MJ Khou, Scott Ratliff, Lulu Shang, Miao Yu, Yi Zhe Wang, and others. I also feel lucky to have been part of the Center for Social Epidemiology and Population Health (CSEPH) and work alongside other CSEPHers. I also want to acknowledge my amazing cohort, Epidocs 2017 and the many friends I crossed paths with during this journey: Kristi Allgood, Kayla Getz, Zoey Laskaris, Julia Porth, and Erica Twardzik. Thank you for your friendship and encouragement.

Finally, I want to thank my family for their endless love and support throughout this dissertation process. I am indebted to my husband and best friend, Abbas, for his patience and encouragement. I am grateful for his support of my professional goals and for celebrating my accomplishments no matter how small. Words cannot express the gratitude and love I feel for him and our two sons, Yousef and Sami. Thank you for my sister, Hiba, for her love and friendship. And above all, I am forever grateful for my parents for all their encouragement, guidance, and endless sacrifices. I thank my Dad for always believing in me, encouraging me to work hard, and reach for the top. I thank my Mom for her endless sacrifices, inspiring me to be fearless in pursuing my dreams, and words of encouragement when I doubted myself. Thank you for everything.

PREFACE

Chapter II (Epigenome-wide Association Study Identifies DNA Methylation Sites Associated with Target Organ Damage in Older African Americans) has been published in *Epigenetics*. The list of full authors is: Farah Ammous, Wei Zhao, Scott M. Ratliff, Minjung Kho, Lulu Shang, Alana C. Jones, Ninad S. Chaudhary, Hemant K. Tiwari, Marguerite R. Irvin, Donna K Arnett, Thomas H. Mosley, Lawrence F. Bielak, Sharon L.R. Kardia, Xiang Zhou, Jennifer A. Smith.

https://doi.org/10.1080/15592294.2020.1827717

Chapter III (Epigenetic Age Acceleration in African Americans Associates with Cardiometabolic Risk Factors and Clinical Cardiovascular Disease Risk Scores) has been published in *Clinical Epigenetics*. The list of full authors is: Farah Ammous, Wei Zhao, Scott M. Ratliff, Thomas H. Mosley, Lawrence F. Bielak, Sharon L.R. Kardia, Xiang Zhou, Patricia A. Peyser, Jennifer A. Smith.

https://doi.org/10.1186/s13148-021-01035-3

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AAC	Abdominal aorta calcification
ABI	Ankle brachial index
ADM	Adrenomedullin
ASCVD	Atherosclerotic cardiovascular disease
BMI	Body mass index
CAC	Coronary artery calcification
CpG	Cytosine-phosphate-guanine
CT	Computed tomography
CVD	Cardiovascular disease
DBP	Diastolic blood pressure
DNAm	DNA methylation
EBV	Epstein-Barr virus
EEAA	Extrinsic epigenetic age acceleration
eGFR	Estimated glomerular filtration rate
ENCODE	Encyclopedia of DNA Elements
EWAS	Epigenome wide association study
FANTOM	Functional ANnoTation Of the Mammalian genome
FDR	False discovery rate
FRS	Framingham risk score
GEMMA	Genome-wide Efficient Mixed Model Association
GENOA	Genetic Epidemiology Network of Arteriopathy
GWAS	Genome wide association study
HDL-C	High density lipoprotein cholesterol
ICC	Intraclass correlation coefficient
IEAA	Intrinsic epigenetic age acceleration
LDL-C	Low density lipoprotein cholesterol
LR	Likelihood ratio
LVMI	Left ventricular mass index
MAP	Mean arterial pressure
MESA	Multi-Ethnic Study of Atherosclerosis
MR	Mendelian randomization
MRI	Magnetic resonance imaging
MRS	Methylation risk score
NRI	Net reclassification index
PAI-1	Plasminogen activator inhibitor antigen type 1

PC	Principal component
PP	Pulse pressure
ROC	Receiver operator characteristic
RWT	Relative wall thickness
SBP	Systolic blood pressure
SD	Standard deviation
SE	Standard error
SNP	Single nucleotide polymorphism
T2D	Type 2 diabetes
TC	Total cholesterol
TG	Triglycerides
TIMP-1	Tissue inhibitor metalloproteinases 1
TOD	Target organ damage
UACR	Urinary albumin-creatinine ratio
WMH	White matter hyperintensity

ABSTRACT

Cardiovascular disease (CVD) is the leading cause of mortality among US adults, and African Americans have a higher burden of CVD morbidity and mortality than any other racial group. Identification of novel CVD biomarkers is essential to better identify at-risk individuals, advance precision medicine, and inform efforts to reduce CVD burden. DNA methylation (DNAm) is an epigenetic mechanism that captures genetic influences as well as imprints of lifestyle and environmental exposures throughout the life course. DNAm patterns may help identify biological mechanisms contributing to CVD pathogenesis. This dissertation explores the effects of DNAm and DNAm-based epigenetic age acceleration (EAA) on cardiometabolic risk factors, atherosclerosis, CVD incidence, and target organ damage from hypertension (TOD) in African American participants from the Genetic Epidemiology Network of Arteriopathy (GENOA). In Aim 1, we used univariate and multivariate linear mixed models to assess the epigenome-wide association between DNAm sites (CpGs) and measures of TOD in the heart (left ventricular mass index (LVMI) and relative wall thickness (RWT)), kidneys (estimated glomerular filtration rate (eGFR) and albuminuria), and brain (white matter hyperintensity). LVMI, RWT, and albuminuria were each associated with one CpG in univariate models, and seven CpGs were associated with TOD measures in the multivariate (pleiotropy) model (false discovery rate (FDR) < 0.1). Mendelian randomization analysis provided evidence of a causal pathway between three CpGs and eGFR. In Aim 2, we assessed the associations between four measures of EAA, 10 cardiometabolic risk factors, and CVD incidence. We then evaluated

whether EAA improved the predictive accuracy of two clinically-used CVD risk scores: the Framingham risk score (FRS) and the Atherosclerotic Cardiovascular Disease risk equation (ASCVD). Increased biological aging, as assessed by EAA, was associated with worse cardiometabolic risk profile, but the associations between each of the four EAA measures differed across cardiometabolic risk factors. GrimAge acceleration (GrimAA) was associated with CVD incidence (hazard ratio per 5-year increase 1.47, 95% CI: 1.05 – 2.01, P = 0.024) after adjusting for traditional CVD risk factors. GrimAA improved model fit over clinical risk scores using likelihood ratio tests (P = 0.013 for FRS, P = 0.008 for ASCVD), did not improve C statistics (P > 0.05), and marginally improved net reclassification index (NRI) which assesses reassignment of risk categories (NRI = 0.055, 95% CI: 0.040 – 0.071 for FRS; 0.029, 95% CI: 0.006 – 0.064 for ASCVD). In Aim 3, we evaluated in prospective analyses whether EAA measures, previously identified atherosclerosis-associated CpGs, and methylation risk scores (MRSs) derived from these CpGs are associated with single- or multi-site atherosclerosis (coronary artery calcification (CAC), abdominal aorta calcification (AAC), and ankle-brachial index (ABI)). One and six CpGs were associated with AAC and multi-site atherosclerosis, respectively (FDR < 0.1). Both a 5-year increase in GrimAA as well as a one unit increase in the MRS for carotid artery plaque were associated with a 1.6-fold increase in AAC and 0.7 units increase in multi-site atherosclerosis (score range: 0-12) after adjusting for CVD risk factors (Bonferroni-adjusted P < 0.05). Together, these studies support the premise that DNAm plays an important role in CVD and TOD and is a promising biomarker that may improve risk assessment in African Americans.

Chapter 1. Introduction

1.1 Overview

Health disparities are preventable differences in the indicators of health across different population groups.[1] In the United States, African Americans bear a disproportionate burden of disease, death, and disability for a number of health outcomes, most notable of which are cardiovascular diseases (CVDs).[2] A number of underlying factors contribute to CVD disparities including biological, social, and environmental risk factors. Disparities in cardiovascular health have been persistent over the last few decades despite the overall declining trends in CVD across all groups.[3] African Americans ages 18-49 are two times as likely to die from CVD as non-Hispanic Asian or Pacific Islander individuals.[4] Similar disparities exist in the prevalence of cardiovascular risk factors and are most pronounced in hypertension, which is associated with structural and functional changes in the end organs like the kidneys, eyes, brain, peripheral arteries, and heart.[5] African Americans over 20 years of age have the highest prevalence of hypertension, estimated to be 41% higher than non-Hispanic whites.[6]

1.1.1 Clinical Cardiovascular Diseases

Collectively, CVD includes coronary heart disease (CHD), transient ischemic attack or stroke, heart failure, and peripheral arterial disease. CVD is the number one cause of death in the US and it is estimated that 9% of the US adult population has heart disease.[7] Estimates from the National Health and Nutrition Survey (NHANES, 2013-2016) show the prevalence of total CVD to be approximately 10.5% among African American adults, with CVD accounting for about

30% of the deaths.[7] African Americans have a greater burden in the onset of stroke, peripheral vascular disease, and heart failure, but not CHD.[3,7] Nevertheless, cardiovascular mortality remains higher in African Americans compared to whites.[8]

CHD clinically manifests as stable angina, unstable angina, or myocardial infarction (heart attack). CHD alone contributed to more than 45% of the total deaths attributable to CVD, equivalent to approximately 1 of every 7 deaths in the United States in 2011.[9] CHD prevalence is estimated to be 7.7% in white males, 7.2% in African American males, 6.1% in white females, and 6.5% in African American females (NHANES 2013-2016).[7] Although CHD prevalence is not significantly higher in African Americans compared to whites, African Americans have higher overall CVD mortality.[3] Analysis of data from the Atherosclerosis Risk in Communities (ARIC) study shows that the rate of CHD decline between 1987-2000 for African Americans males is only about half the rate for white males (-3.2%/year vs. -6.5%/year) and about two thirds the rate for white females (-4.0%/year in African American females vs. 5.8%/year in white females).[10] Longitudinal follow-up in the ARIC, the Cardiovascular Health Study (CHS), and the REasons for Geographic And Racial Differences in Stroke (REGARDS) study show that African American males have similar risks of fatal CHD, lower risk of nonfatal CHD, and higher CHD case-fatality compared to whites after adjustment for social determinants of health and cardiovascular risk factors.[11]

Stroke, another major cardiovascular disease, ranks 5th in the causes of death in the US.[12] It is estimated that 7.8 millions of adults in the US have had a stroke.[13] The prevalence of stroke using NHANES (2013-2016) is 3.1% in African American males, 2.4% in white males, 3.8% in African American females, and 2.5% in white females.[7] Disparities are also most notable in stroke mortality despite overall national trends of decreased mortality and incidence.[3,14-16] In

adults ages ≥ 45 years, African American males had 54% higher age-adjusted stroke death rates than white males, and African American females had a 30% higher death rate than white females.[14]

1.1.2 Subclinical Cardiovascular Disease and Atherosclerosis

Atherosclerosis, a precursor for CVD, is a chronic inflammatory condition that develops over several decades.[17-19] It is a complex process where calcification of the intimal layer of the arterial wall occurs as a result of an inflammatory response to lipid accumulation and focal plaque formation.[20] Intimal calcium deposition is often seen in the coronary arteries surrounding the heart and peripheral arteries. Atherosclerotic lesions can lead to diminished blood flow to the organs that presents as cardiac ischemia or claudication. Atherosclerotic plaque rupture leads to arterial occlusion with subsequent blood flow interruption that manifests as heart attack, stroke or limb ischemia.[17,21-23]

Atherosclerosis starts early in life with lesions starting as only fatty streaks consisting of monocyte derived macrophages and T lymphocytes.[17,24] A number of factors may induce and promote atherogenesis such as endothelial dysfunction, elevated low-density lipoproteins, and hypertension. The continued inflammation results in further activation and migration of macrophages and lymphocytes and an increase in cytokines, chemokines and other growth factors. Through these continued cycles of inflammation and migration of cells, the fatty streak becomes enlarged and changes into a fibrous plaque with a necrotic core.[17] Calcification of the plaque may ensue where macrophages in the plaque promote osteogenic differentiation and calcium deposition.[22] Continued calcification may then stabilize the plaque and provide a wall between the plaque and the blood pool, which may decrease the risk of plaque rupturing.[22]

One imaging modality for the quantification of atherosclerosis in the coronaries and aorta is noninvasive computed tomography (CT) scans. The Agatston score is used to quantify the sum of the total calcified area and maximum density calcification.[22,25] For individual lesions, the Agatston score is derived by multiplying the lesion area with a density weighting factor (DWF) in Hounsfield units (HU) units. The DWF is derived from the maximal CT attenuation within a given calcified lesion and reflects increasing categories of Hounsfield units (Hu) (DWF: 130 to 199 HU = 1; 200 to 299 HU = 2; 300 to 399 HU = 3; and ≥ 400 HU = 4).[26] Hence, the Agatston score is weighted upward for greater CAC density.[27] The scores for all lesions is then summed over all lesions to determine the total Agatston score.[26]

$$Agatston\ score_{lesion} = Area\ \times DWF$$

$$Agatstson\ score_{total} = \sum Agatston\ score_{lesion}$$

Both coronary artery calcium (CAC) and abdominal aorta calcium (AAC) have been found to be independent predictors of CVD incidence and mortality.[28-30] Calcification in the peripheral arteries can be assessed using the ankle-brachial index (ABI), which is calculated as the ratio of systolic blood pressure at the ankles divided by the brachial pressure. Previous studies have reported ABI to be a robust predictor of peripheral arterial disease and mortality.[31-33] Abnormally high ABI (>1.3) suggests the presence of calcified vessels that may not compress normally which may occur in the setting of underlying diabetes or end-stage kidney disease.[34]

Despite having a higher burden of CVD and traditional risk factors, African Americans have a lower prevalence of CAC and AAC.[35-40] In about 6,800 MESA participants, the relative risk of CAC in African Americans compared to whites was 0.78 (95% CI 0.74–0.82) after adjusting for socioeconomic and clinical risk factors.[41] African Americans, however, are more likely to have carotid atherosclerosis and a higher carotid intima media thickness (IMT). In the

same MESA sample, African Americans were found to have the highest common carotid IMT compared to other racial groups.[42] Finally, African Americans have a higher prevalence of peripheral arterial disease that was not explained by traditional and novel risk factors.[7,43,44]

1.1.3 Cardiac and Renal Target Organ Damage of the Heart, Kidneys, and Brain

Among African Americans, the prevalence of hypertension is estimated to be the highest in the world, and although awareness and treatment rates are higher than those of other racial groups, control of high blood pressures remains lower in African Americans compared to whites.[45] In this dissertation we defined hypertension as having an average systolic blood pressure ≥ 140 mmHg or diastolic blood pressure ≥ 90 mmHg, or current anti-hypertensive medication use, following previously established guidelines. More recent guidelines by the American College of Cardiology/American Heart Association (ACC/AHA) apply a lower threshold: a systolic blood pressure ≥ 130 mmHg or a diastolic blood pressure ≥ 80 mmHg.[46] Hypertension over time contributes to structural and functional changes in organ systems, referred to as target organ damage (TOD). High blood pressure causes reduced arterial plasticity and endothelial wall damage, which in turn triggers the deposition of lipids in the wall lesions, leading to atherosclerotic plaque build-up and the subsequent obstruction of these vessels.[5,47] As the vascular resistance increases, the systolic function is increased to maintain adequate organ perfusion.[5,47] Renal injury occurs when the regulatory mechanisms of the kidneys are unable to maintain flow and pressure, resulting in an increase in glomerular hypertension, permeability, and proteinuria.[5] In the heart, the remodeling of the myocardium ensues as the cardiac muscle adapts to the increased resistance in the circulation. This results in hypertrophy and increased myocardial thickness that eventually reduces relaxation of the heart muscle.[5] In the brain, small

vessel disease is associated with decreased blood flow and hypoperfusion and results in lesions seen in the white matter.[5]

TOD in the heart can be quantified using echocardiography imaging, which provides precise measurements of the dimension of the heart such as left ventricular mass and relative wall thickness. In the kidneys, TOD can be assessed using glomerular filtration rate and albuminuria. Hyperintensity of the white matter on brain MRI is often used to assess TOD in the brain. Compared to whites, epidemiological studies have shown that African Americans have greater left ventricular mass [48,49], higher incidence of chronic kidney disease, higher incidence of end stage renal disease,[50,51] and more severe white matter hyperintensity burden,[52,53] a trend that is in parallel to their higher hypertension burden.

While subtle TOD may start as asymptomatic in the early course of hypertension, more severe organ damage typically happens after long periods of uncontrolled hypertension.[54] Clinical endpoints of TOD include symptomatic chronic kidney disease, systolic and diastolic dysfunction, dementia and transient ischemic attack. Further complications can lead to end-stage renal disease, myocardial infarction, and stroke.[54] These thrombotic effects of hypertension leading to clinical events are facilitated by damage to the vessel wall, hypercoagulability, and abnormal blood flow.[54] In addition, the well documented cardio-renal interactions, where chronic dysfunction of one organ may induce dysfunction in the other, can further contribute to the progression of TOD into more pronounced clinical and cardiovascular events.[55]

1.2 Risk Factors for Cardiovascular Diseases and Target Organ Damage

Surveillance data from NHANES and the Behavioral Risk Factor Surveillance System (BRFSS) show that marked disparities exist in both measured and self-reported cardiovascular risk factors which contribute to disparities in cardiovascular health.[1] These include traditional

risk factors such as hypertension, type 2 diabetes mellitus, dyslipidemia, and obesity as well as adverse health behaviors such as smoking, poor diet quality, and physical activity.[1,3] In addition, there is emerging evidence that social and socioeconomic factors such as educational attainment, built-environment, neighborhood environment, access to care, psychosocial stress, and discrimination are important factors that contribute to the disparities in cardiovascular morbidities and mortality.[56]

The age adjusted prevalence of high blood pressure among adults ages ≥ 20 years between 2013 and 2016 was 46% for the total US population, 57.6% among African American males, 53.2% among African American females, 46.7% among white males, and 38.8% among white females.[7] The prevalence of diabetes between 2013 and 2016 was 9.8% for the total population, 14.7% and 13.4% in African American males and females, and 9.4% and 7.3% in white males and females, respectively.[7] Longitudinal evidence from the Coronary Artery Risk Development in Young Adults (CARDIA) study showed that the higher risk of diabetes onset among African Americans appears to be mediated by biological, neighborhood, psychosocial, socioeconomic, and behavioral factors.[57] Prevalence estimates of dyslipidemia (defined as having total cholesterol levels ≥ 240 mg/dl, low density lipoprotein (LDL-C) ≥ 160 mg/dl, high density lipoprotein (HDL-C) ≤ 40 mg/dl, or use of lipid-lowering medications) in African Americans are comparable to or lower compared to whites.[3,7] These prevalence estimates contradict with those based on incidence data, where African Americans have a higher rate of dyslipidemia that is more pronounced at older ages.[58] This discrepancy could be partially explained by the higher rates of CVD mortality attributed to dyslipidemia among African Americans compared to whites.[3] Obesity, defined as body mass index (BMI) ≥ 25 kg/m^2, is also higher in African Americans across the age spectrum.[3] In adults ages ≥ 20 years, the prevalence

of obesity in African American males was 37.0% compared to 35.8% in white males, and 55.3% in African American females compared to 37.8% in white females.[7] Smoking exposure, as both active and secondhand smoke, is a strong risk factor for CVD in all groups with no differential magnitude of effect or rates of smoking in African Americans compared to whites.[3,59] However, compared to whites, African Americans have lower quit rates and are more likely to be exposed to secondhand smoke.[3] Finally, adherence to dietary and activity recommendations remains difficult and below target in all racial group including African Americans.[3]

1.3 Genetics and Epigenetics in Cardiovascular Disease and Target Organ Damage

1.3.1 Genetics of Cardiovascular Disease and Target Organ Damage

In addition to the risk factors discussed above, family history and genetics may also contribute to the increased burden of CVD in African Americans. A number of independent loci have been identified for cardiovascular diseases in African Americans, including those associated with hypertension[60,61] and left ventricular mass.[62] However, most identified genetic loci explain only a small portion of the variance of complex diseases, including CVD.[63] Hence, despite the role of genetic predisposition and family history in CVD and its risk factors, social and environmental conditions are likely to be the most important determinants in CVD risk and disparities.[1,64]

1.3.2 The Emerging Role of Epigenetics in Health and Disease

Epigenetic mechanisms are modulated by lifestyle and environmental factors over the life course and can capture the influences of early and long-term exposures on health. Their effect is translated into changes in gene expression and they are recognized as a molecular bridge between the environment and disease.[65] Epigenetic markers may be important biomarkers of early disease onset and progression and can reveal important information on biological

mechanisms and pathways underlying morbidities. Combined with other genetic and phenotypic information, epigenetic markers could be used to provide precise characterization of individual risk for health conditions, including CVD.

Currently recognized epigenetic mechanisms include DNA methylation (DNAm), histone modifications, and higher order chromatin structure.[66] DNA methylation is the covalent addition of a methyl group to a cytosine of a cytosine-guanine pair, also known as a CpG site.[67] It is the most studied and best understood epigenetic mechanism and is a useful biomarker because it can be relatively stable over time yet susceptible to change in response to environmental stimuli.[66] Histone modifications are energy dependent post translational processes of nucleosomal histones. These modifications include acetylation, methylation, phosphorylation, ubiquitylation, and sumoylation and are differently associated with gene activity, gene silencing, or insulation between active and inactive gene regions.[66] Lastly, higher order chromatin structure involves nucleosomal compaction near the nuclear membrane (heterochromatin) or nucleosomal accessibility (euchromatin).[66] The higher order chromatin structure also partitions the genome into regions that are topologically associated resulting in enhancer-promoter interaction.[66,68]

1.3.3 Epigenetic clocks

Epigenetic clocks are novel measures of biological aging that are based on DNA methylation. They are added to a growing list of biological aging measures which are used to describe the biological aging process beyond chronological age.[69] Aging in essence is defined as the progressive functional decline that is the leading risk factor for major human morbidities including cardiovascular disorders and neurodegenerative diseases.[70] The concept of biological aging is based on the observation that individuals of the same chronological age do not age biologically at the same pace.[71] Hence the different trajectories of age-related decline can be

better captured by measures of biological aging than chronological age. Biological aging measures can be divided into molecular biomarkers, such as epigenetic clocks and telomere length, or phenotypic or physiological markers which include clinical measures such as blood pressure, grip strength, and lipids.[69] Understanding the aging process and pathophysiological processes underlying age-related conditions may lead to innovative prevention and therapeutic targets.

The earliest generation of epigenetic clocks, developed by Horvath[72] and Hannum[73] were trained to select CpG sites associated with chronological age using penalized regression models. The Horvath measure, based on 353 CpGs, is a multi-tissue measure derived using DNA methylation from 8000 samples from 82 publicly available datasets.[72] The datasets had representation of four racial/ethnic groups; mainly whites, Hispanics, African Americans, and to a lesser extent East Asians.[72] The Hannum measure, based on 71 CpGs, is derived using DNA methylation in whole blood from a sample of 426 whites and 230 Hispanics.[73] For both measures, the acceleration is based on residuals from a model regressing epigenetic age against chronological age. Modified versions of these measures also account for confounding by blood cell composition. [74]. Intrinsic epigenetic age acceleration (IEAA) is the residual resulting from a multivariate regression model of Horvath epigenetic age on chronological age and white blood cell counts (naive CD8+ T cells, exhausted CD8+ T cells, plasmablasts, CD4+ T cells, natural killer cells, monocytes, and granulocytes) imputed from methylation data. Extrinsic age epigenetic age acceleration (EEAA) is based on the residuals from a model of biological age calculated using the Hannum measure after incorporating an up-weighted count of three cell types known to change with age: naïve cytotoxic T cells, exhausted cytotoxic T cells, and plasmablasts.

The second generation of the epigenetic aging measures, including PhenoAge and GrimAge, account for physiological dysfunction among individuals of the same chronological age in their selection of CpGs. PhenoAge was developed using NHANES III as the training sample where a proportional mortality hazard penalized regression model was used to narrow 42 biomarkers to 9 biomarkers and chronological age.[75] The 9 selected biomarkers were then used to derive phenotypic age. PhenoAge was then derived by regressing phenotypic age on blood DNA methylation data from the Invecchiare in Chianti, Aging in the Chianti Area (InCHIANTI) study. This produced an estimate of DNAm PhenoAge based on 513 CpGs, and PhenoAge was further validated in multiracial cohorts (Women's health initiative (WHI), Jackson Heart Study (JHS), Normative Aging Study (NAS) and Framingham Heart Study (FHS). GrimAge was derived by first defining DNA methylation-based surrogate markers of 88 plasma proteins and smoking pack-years.[76] Out of those, only 12 of the plasma proteins were correlated with the surrogate measure ($r > 0.35$). Next, time-to-death due to all-cause mortality was regressed on chronological age and the DNA methylation-based biomarkers for the 12 plasma proteins and smoking.[76] An elastic net regression selected 7 surrogate plasma protein and smoking pack-years in addition to chronological age and sex. GrimAge, based on 1030 CpGs, was then derived as the linear combination of these covariate values and transformed to be in units of years. GrimAge was trained on samples from the FHS and validated in an independent FHS cohort, WHI, JHS, and INChianti.[76] For both GrimAge and PhenoAge, an acceleration measure is derived as the residual from a model regressing the epigenetic aging measure against chronological age.

1.3.4 Epigenetics of Cardiovascular Diseases and Target Organ Damage

Epidemiological studies of DNA methylation signatures measured in proxy tissues, such as blood, typically employ an epigenome wide approach where differential DNA methylation at

individual CpGs is examined in relation to the exposure or outcome of interest. Given the immune system involvement and inflammatory responses underlying atherosclerosis and TOD, DNA methylation patterns in white blood cells are suitable proxy tissue to study the epigenetic signature of the onset and progression of these outcomes. Other methodological approaches, such as adaptation of approaches similar to polygenic risk scores to access more variance of complex trait, i.e. methylation risk scores, are still limited in use. This is mostly because DNA methylation data is more sensitive to confounding by age and tissue and it is difficult to find appropriate external weights.[77] Conversely, epigenetic age acceleration measures have been used more extensively and have been shown to be robust markers of aging and a number of aging related diseases.

To date, a limited number of studies have examined the association between genome wide DNA methylation and hypertension and TOD in the kidneys.[78,79] More studies have examined the genome-wide DNA methylation signatures of subclinical CVD[80] and clinical CVD.[81-84] However, a majority of these studies were in cohorts of European ancestry and/or were cross-sectional. Other studies have examined the association between epigenetic age acceleration and TOD, [85,86] CVD, and their risk factors[86-97], but the overall evidence remains inconclusive likely due to heterogeneity in study design, outcomes, and the epigenetic aging measures used. In two recent studies of participants of European ancestry, GrimAge acceleration was found to be associated with CVD incidence beyond traditional risk factors.[86,96] It remains unclear whether similar association exits in African Americans.

1.4 Motivation for further studies

CVD disparities in the US remain pervasive, and the causes of these disparities are complex and elusive. Epigenetic mechanisms are increasingly recognized as important genomic

regulators that may influence or be biomarkers of CVD disease risk and progression. Identifying and quantifying key epigenetic modifications may help to improve risk stratification of CVD and its risk factors. Furthermore, epigenetic studies can help identify underlying disease mechanisms and key regulatory processes and gene-environment interactions that may occur over long periods of time and contribute to CVD risk.[65] Further studies are needed to better characterize these associations, replicate previous findings and increase representation of highly burdened populations. As such, this dissertation will explore the effects of DNAm and DNAm-based epigenetic age acceleration (EAA) on cardiometabolic risk factors, atherosclerosis, CVD incidence, and target organ damage (TOD) from hypertension in African American participants from the Genetic Epidemiology Network of Arteriopathy (GENOA).

1.5 The Genetic Epidemiology Network of Arteriopathy (GENOA) study

The Genetic Epidemiology Network of Arteriopathy (GENOA) is a community-based study in Rochester, MN and Jackson, MS that aims to identify genes influencing blood pressure[98]. In the first phase of GENOA (Phase I: 1996 – 2001), sibships with at least two adults with clinically diagnosed essential hypertension before age 60 were recruited, and all siblings in the sibships were invited to participate regardless of hypertension status [20]. Exclusion criteria included secondary hypertension, alcoholism or drug abuse, pregnancy, insulin-dependent diabetes mellitus, or active malignancy.

In Phase I, a total of 1,583 non-Hispanic whites (Rochester, MN) and 1,854 African Americans (Jackson, MS) were enrolled. In the second phase (Phase II: 2001 – 2005), all participants were invited for a second examination. Eighty percent of African Americans (N = 1,482) and 75% of non-Hispanic whites (N = 1,213) from Phase 1 returned. At Phase III (2009-2011), 752 African Americans returned for a third examination. Demographic information,

medical history, clinical characteristics, lifestyle factors, and blood samples were collected in each phase.

This dissertation will focus on African American participants who had their DNA methylation measured in whole blood samples collected at Phase I (N = 1,106) and/or Phase II (N = 304). For these participants, genomic DNA was extracted from stored peripheral blood leukocytes using AutoGen FlexStar (AutoGen, Holliston, MA) and DNA methylation was measured using the Infinium MethylationEPIC BeadChip. The methods of DNA methylation processing have been previously described.[99] Following quality control, a total of 1,100 samples from Phase I and 294 from Phase II were available for analysis.

1.6 Specific Aims and Hypotheses

The specific aims and hypotheses for this dissertation are as follows:

Aim 1

To assess the longitudinal association between epigenome-wide DNA methylation and five measures of target organ damage (TOD) in GENOA African Americans.

Sub aims

1. To identify CpGs associated with five measures of TOD in the heart (left ventricular mass index and relative wall thickness), kidneys (estimated glomerular filtration rate and albuminuria), and brain (white matter hyperintensity) using univariate and multivariate regression.
2. To explore whether DNA methylation at the identified CpGs is associated with proximal gene expression.

3. To evaluate the causal relationship between the identified CpGs and TOD using Mendelian randomization.

Hypothesis 1: We hypothesize that a number of CpGs will be associated with measures of TOD, with evidence of differential proximal gene expression and causality.

Aim 2

To investigate whether four epigenetic age acceleration measures (IEAA, EEAA, PhenoAA, and GrimAA) are associated with ten cardiometabolic risk factors and CVD incidence in GENOA African Americans.

Sub aims

1. To assess the cross-sectional association between four epigenetic age acceleration measures and ten cardiometabolic markers of hypertension, insulin resistance, and dyslipidemia.
2. To investigate whether four epigenetic age acceleration measures are associated with CVD incidence and improve predictive accuracy over two clinical CVD risk scores (Framingham risk score (FRS) and the atherosclerotic cardiovascular disease (ASCVD) risk equation).

Hypothesis 2: We hypothesize that increased epigenetic age acceleration, indicative of faster biological aging, will be associated with a worse cardiometabolic risk profile and higher CVD incidence, and will improve prediction of CVD incidence over the FRS and/or the ASCVD risk scores.

Aim 3

To evaluate whether four epigenetic age acceleration measures (IEAA, EEAA, PhenoAA, and GrimAA), previously identified CpGs for atherosclerosis, and/or methylation risk scores (MRSs)

derived from these CpGs are associated with single- or multi-site atherosclerosis in GENOA African Americans.

Sub aims

1. To assess the associations between four epigenetic age acceleration measures and single- or multi-site atherosclerosis (coronary artery calcification, abdominal aorta calcification, and ankle-brachial index).
2. To evaluate the associations between previously-identified CpGs for atherosclerosis, both individually and aggregated into MRSs, and single- or multi-site atherosclerosis (coronary artery calcification, abdominal aorta calcification, and ankle-brachial index).
3. To characterize the temporal stability of the potential epigenetic biomarkers for atherosclerosis (four epigenetic age acceleration measures and MRSs) using longitudinal measures of DNA methylation.

Hypothesis 3: We hypothesize that increased epigenetic age acceleration and previously-identified CpGs for atherosclerosis, both individually and aggregated into MRSs, will be associated with single- or multi-site atherosclerosis with evidence of temporal stability.

1.7 References

1. Mensah GA, Mokdad AH, Ford ES, Greenlund KJ, Croft JB. State of disparities in cardiovascular health in the United States. *Circulation.* 2005;111(10):1233-1241.
2. Centers for Disease C, Prevention. Health disparities experienced by black or African Americans--United States. *MMWR Morb Mortal Wkly Rep.* 2005;54(1):1-3.
3. Carnethon MR, Pu J, Howard G, et al. Cardiovascular Health in African Americans: A Scientific Statement From the American Heart Association. *Circulation.* 2017;136(21):e393-e423.
4. Cunningham TJ, Croft JB, Liu Y, Lu H, Eke PI, Giles WH. Vital Signs: Racial Disparities in Age-Specific Mortality Among Blacks or African Americans - United States, 1999-2015. *MMWR Morb Mortal Wkly Rep.* 2017;66(17):444-456.
5. Cohuet G, Struijker-Boudier H. Mechanisms of target organ damage caused by hypertension: therapeutic potential. *Pharmacol Ther.* 2006;111(1):81-98.
6. Gillespie CD, Hurvitz KA, Centers for Disease C, Prevention. Prevalence of hypertension and controlled hypertension - United States, 2007-2010. *MMWR Suppl.* 2013;62(3):144-148.
7. Virani SS, Alonso A, Benjamin EJ, et al. Heart Disease and Stroke Statistics-2020 Update: A Report From the American Heart Association. *Circulation.* 2020;141(9):e139-e596.
8. Shah NS, Lloyd-Jones DM, O'Flaherty M, et al. Trends in Cardiometabolic Mortality in the United States, 1999-2017. *JAMA.* 2019;322(8):780-782.
9. Mozaffarian D, Benjamin EJ, Go AS, et al. Heart disease and stroke statistics--2015 update: a report from the American Heart Association. *Circulation.* 2015;131(4):e29-322.
10. Rosamond WD, Chambless LE, Heiss G, et al. Twenty-two-year trends in incidence of myocardial infarction, coronary heart disease mortality, and case fatality in 4 US communities, 1987-2008. *Circulation.* 2012;125(15):1848-1857.
11. Colantonio LD, Gamboa CM, Richman JS, et al. Black-White Differences in Incident Fatal, Nonfatal, and Total Coronary Heart Disease. *Circulation.* 2017;136(2):152-166.
12. Underlying Cause of Death 1999-2019 on CDC WONDER Online Database, released in 2020. Data are from the Multiple Cause of Death Files, 1999-2019, as compiled from data provided by the 57 vital statistics jurisdictions through the Vital Statistics Cooperative Program. http://wonder.cdc.gov/ucd-icd10.html. Accessed January 27, 2021
13. Villarroel MAB, D.L.; Jen, A. Tables of Summary Health Statistics for U.S. Adults: 2018 National Health Interview Survey. In: Statistics NCfH, ed2019.
14. Ingram DD, Montresor-Lopez JA. Differences in Stroke Mortality Among Adults Aged 45 and Over: United States, 2010-2013. *NCHS data brief.* 2015(207):1-8.
15. Koton S, Sang Y, Schneider ALC, Rosamond WD, Gottesman RF, Coresh J. Trends in Stroke Incidence Rates in Older US Adults: An Update From the Atherosclerosis Risk in Communities (ARIC) Cohort Study. *JAMA Neurol.* 2020;77(1):109-113.

16. Koton S, Schneider AL, Rosamond WD, et al. Stroke incidence and mortality trends in US communities, 1987 to 2011. *JAMA.* 2014;312(3):259-268.

17. Ross R. Atherosclerosis--an inflammatory disease. *N Engl J Med.* 1999;340(2):115-126.

18. Libby P, Buring JE, Badimon L, et al. Atherosclerosis. *Nat Rev Dis Primers.* 2019;5(1):56.

19. Frostegard J. Immunity, atherosclerosis and cardiovascular disease. *BMC Med.* 2013;11:117.

20. Mackey RH, Venkitachalam L, Sutton-Tyrrell K. Calcifications, arterial stiffness and atherosclerosis. *Adv Cardiol.* 2007;44:234-244.

21. Berliner JA, Navab M, Fogelman AM, et al. Atherosclerosis: basic mechanisms. Oxidation, inflammation, and genetics. *Circulation.* 1995;91(9):2488-2496.

22. Wang Y, Osborne MT, Tung B, Li M, Li Y. Imaging Cardiovascular Calcification. *J Am Heart Assoc.* 2018;7(13).

23. Vengrenyuk Y, Carlier S, Xanthos S, et al. A hypothesis for vulnerable plaque rupture due to stress-induced debonding around cellular microcalcifications in thin fibrous caps. *Proc Natl Acad Sci U S A.* 2006;103(40):14678-14683.

24. Napoli C, D'Armiento FP, Mancini FP, et al. Fatty streak formation occurs in human fetal aortas and is greatly enhanced by maternal hypercholesterolemia. Intimal accumulation of low density lipoprotein and its oxidation precede monocyte recruitment into early atherosclerotic lesions. *J Clin Invest.* 1997;100(11):2680-2690.

25. Agatston AS, Janowitz WR, Hildner FJ, Zusmer NR, Viamonte M, Jr., Detrano R. Quantification of coronary artery calcium using ultrafast computed tomography. *J Am Coll Cardiol.* 1990;15(4):827-832.

26. Blaha MJ, Mortensen MB, Kianoush S, Tota-Maharaj R, Cainzos-Achirica M. Coronary Artery Calcium Scoring: Is It Time for a Change in Methodology? *JACC Cardiovasc Imaging.* 2017;10(8):923-937.

27. Criqui MH, Denenberg JO, Ix JH, et al. Calcium density of coronary artery plaque and risk of incident cardiovascular events. *JAMA.* 2014;311(3):271-278.

28. Bastos Goncalves F, Voute MT, Hoeks SE, et al. Calcification of the abdominal aorta as an independent predictor of cardiovascular events: a meta-analysis. *Heart.* 2012;98(13):988-994.

29. Bartstra JW, Mali WP, Spiering W, de Jong PA. Abdominal aortic calcification: from ancient friend to modern foe. *Eur J Prev Cardiol.* 2020:2047487320919895.

30. Criqui MH, Denenberg JO, McClelland RL, et al. Abdominal aortic calcium, coronary artery calcium, and cardiovascular morbidity and mortality in the Multi-Ethnic Study of Atherosclerosis. *Arterioscler Thromb Vasc Biol.* 2014;34(7):1574-1579.

31. Criqui MH, Langer RD, Fronek A, et al. Mortality over a period of 10 years in patients with peripheral arterial disease. *N Engl J Med.* 1992;326(6):381-386.

32. Leng GC, Fowkes FG, Lee AJ, Dunbar J, Housley E, Ruckley CV. Use of ankle brachial pressure index to predict cardiovascular events and death: a cohort study. *BMJ.* 1996;313(7070):1440-1444.

33. Hirsch AT, Haskal ZJ, Hertzer NR, et al. ACC/AHA 2005 Practice Guidelines for the management of patients with peripheral arterial disease (lower extremity, renal, mesenteric, and abdominal aortic): a collaborative report from the American Association for Vascular Surgery/Society for Vascular Surgery, Society for Cardiovascular Angiography and Interventions, Society for Vascular Medicine and Biology, Society of Interventional Radiology, and the ACC/AHA Task Force on Practice Guidelines (Writing Committee to Develop Guidelines for the Management of Patients With Peripheral Arterial Disease): endorsed by the American Association of Cardiovascular and Pulmonary Rehabilitation; National Heart, Lung, and Blood Institute; Society for Vascular Nursing; TransAtlantic Inter-Society Consensus; and Vascular Disease Foundation. *Circulation.* 2006;113(11):e463-654.

34. Wang JC, Criqui MH, Denenberg JO, McDermott MM, Golomb BA, Fronek A. Exertional leg pain in patients with and without peripheral arterial disease. *Circulation.* 2005;112(22):3501-3508.

35. Wassel CL, Pankow JS, Peralta CA, Choudhry S, Seldin MF, Arnett DK. Genetic ancestry is associated with subclinical cardiovascular disease in African-Americans and Hispanics from the multi-ethnic study of atherosclerosis. *Circ Cardiovasc Genet.* 2009;2(6):629-636.

36. Erqou S, Kip KE, Mulukutla SR, Aiyer AN, Reis SE. Racial differences in the burden of coronary artery calcium and carotid intima media thickness between Blacks and Whites. *Neth Heart J.* 2015;23(1):44-51.

37. Doherty TM, Tang W, Detrano RC. Racial differences in the significance of coronary calcium in asymptomatic black and white subjects with coronary risk factors. *J Am Coll Cardiol.* 1999;34(3):787-794.

38. Allison MA, Budoff MJ, Nasir K, et al. Ethnic-specific risks for atherosclerotic calcification of the thoracic and abdominal aorta (from the Multi-Ethnic Study of Atherosclerosis). *Am J Cardiol.* 2009;104(6):812-817.

39. Lee TC, O'Malley PG, Feuerstein I, Taylor AJ. The prevalence and severity of coronary artery calcification on coronary artery computed tomography in black and white subjects. *J Am Coll Cardiol.* 2003;41(1):39-44.

40. Gebreab SY, Riestra P, Khan RJ, et al. Genetic ancestry is associated with measures of subclinical atherosclerosis in African Americans: the Jackson Heart Study. *Arterioscler Thromb Vasc Biol.* 2015;35(5):1271-1278.

41. Bild DE, Detrano R, Peterson D, et al. Ethnic differences in coronary calcification: the Multi-Ethnic Study of Atherosclerosis (MESA). *Circulation.* 2005;111(10):1313-1320.

42. Manolio TA, Arnold AM, Post W, et al. Ethnic differences in the relationship of carotid atherosclerosis to coronary calcification: the Multi-Ethnic Study of Atherosclerosis. *Atherosclerosis.* 2008;197(1):132-138.

43. Allison MA, Criqui MH, McClelland RL, et al. The effect of novel cardiovascular risk factors on the ethnic-specific odds for peripheral arterial disease in the Multi-Ethnic Study of Atherosclerosis (MESA). *J Am Coll Cardiol.* 2006;48(6):1190-1197.

44. Criqui MH, Vargas V, Denenberg JO, et al. Ethnicity and peripheral arterial disease: the San Diego Population Study. *Circulation.* 2005;112(17):2703-2707.

45. Benjamin EJ, Virani SS, Callaway CW, et al. Heart Disease and Stroke Statistics-2018 Update: A Report From the American Heart Association. *Circulation.* 2018;137(12):e67-e492.

46. Whelton PK, Carey RM, Aronow WS, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *J Am Coll Cardiol.* 2018;71(19):e127-e248.

47. Nadar SK, Tayebjee MH, Messerli F, Lip GY. Target organ damage in hypertension: pathophysiology and implications for drug therapy. *Curr Pharm Des.* 2006;12(13):1581-1592.

48. Fox E, Taylor H, Andrew M, et al. Body mass index and blood pressure influences on left ventricular mass and geometry in African Americans: The Atherosclerotic Risk In Communities (ARIC) Study. *Hypertension.* 2004;44(1):55-60.

49. Kizer JR, Arnett DK, Bella JN, et al. Differences in left ventricular structure between black and white hypertensive adults: the Hypertension Genetic Epidemiology Network study. *Hypertension.* 2004;43(6):1182-1188.

50. Tarver-Carr ME, Powe NR, Eberhardt MS, et al. Excess risk of chronic kidney disease among African-American versus white subjects in the United States: a population-based study of potential explanatory factors. *J Am Soc Nephrol.* 2002;13(9):2363-2370.

51. Choi AI, Rodriguez RA, Bacchetti P, Bertenthal D, Hernandez GT, O'Hare AM. White/black racial differences in risk of end-stage renal disease and death. *Am J Med.* 2009;122(7):672-678.

52. Liao D, Cooper L, Cai J, et al. The prevalence and severity of white matter lesions, their relationship with age, ethnicity, gender, and cardiovascular disease risk factors: the ARIC Study. *Neuroepidemiology.* 1997;16(3):149-162.

53. Brickman AM, Schupf N, Manly JJ, et al. Brain morphology in older African Americans, Caribbean Hispanics, and whites from northern Manhattan. *Arch Neurol.* 2008;65(8):1053-1061.

54. Messerli FH, Williams B, Ritz E. Essential hypertension. *Lancet.* 2007;370(9587):591-603.

55. Ronco C, Haapio M, House AA, Anavekar N, Bellomo R. Cardiorenal syndrome. *J Am Coll Cardiol.* 2008;52(19):1527-1539.

56. Mensah GA. Eliminating disparities in cardiovascular health: six strategic imperatives and a framework for action. *Circulation.* 2005;111(10):1332-1336.

57. Bancks MP, Kershaw K, Carson AP, Gordon-Larsen P, Schreiner PJ, Carnethon MR. Association of Modifiable Risk Factors in Young Adulthood With Racial Disparity in Incident Type 2 Diabetes During Middle Adulthood. *JAMA.* 2017;318(24):2457-2465.

58. Howard G, Safford MM, Moy CS, et al. Racial Differences in the Incidence of Cardiovascular Risk Factors in Older Black and White Adults. *J Am Geriatr Soc.* 2017;65(1):83-90.

59. Huxley RR, Yatsuya H, Lutsey PL, Woodward M, Alonso A, Folsom AR. Impact of age at smoking initiation, dosage, and time since quitting on cardiovascular disease in african americans and whites: the atherosclerosis risk in communities study. *Am J Epidemiol.* 2012;175(8):816-826.

60. Fox ER, Young JH, Li Y, et al. Association of genetic variation with systolic and diastolic blood pressure among African Americans: the Candidate Gene Association Resource study. *Hum Mol Genet.* 2011;20(11):2273-2284.

61. Sung YJ, Winkler TW, de Las Fuentes L, et al. A Large-Scale Multi-ancestry Genome-wide Study Accounting for Smoking Behavior Identifies Multiple Significant Loci for Blood Pressure. *Am J Hum Genet.* 2018;102(3):375-400.

62. Fox ER, Musani SK, Barbalic M, et al. Genome-wide association study of cardiac structure and systolic function in African Americans: the Candidate Gene Association Resource (CARe) study. *Circ Cardiovasc Genet.* 2013;6(1):37-46.

63. Visscher PM, Wray NR, Zhang Q, et al. 10 Years of GWAS Discovery: Biology, Function, and Translation. *Am J Hum Genet.* 2017;101(1):5-22.

64. Kaufman JS, Dolman L, Rushani D, Cooper RS. The contribution of genomic research to explaining racial disparities in cardiovascular disease: a systematic review. *Am J Epidemiol.* 2015;181(7):464-472.

65. Cavalli G, Heard E. Advances in epigenetics link genetics to the environment and disease. *Nature.* 2019;571(7766):489-499.

66. Feinberg AP. The Key Role of Epigenetics in Human Disease Prevention and Mitigation. *N Engl J Med.* 2018;378(14):1323-1334.

67. Rozek LS, Dolinoy DC, Sartor MA, Omenn GS. Epigenetics: relevance and implications for public health. *Annu Rev Public Health.* 2014;35:105-122.

68. Lieberman-Aiden E, van Berkum NL, Williams L, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science.* 2009;326(5950):289-293.

69. Jylhava J, Pedersen NL, Hagg S. Biological Age Predictors. *EBioMedicine.* 2017;21:29-36.

70. Lopez-Otin C, Blasco MA, Partridge L, Serrano M, Kroemer G. The hallmarks of aging. *Cell.* 2013;153(6):1194-1217.

71. Khan SS, Singer BD, Vaughan DE. Molecular and physiological manifestations and measurement of aging in humans. *Aging Cell.* 2017;16(4):624-633.

72. Horvath S. DNA methylation age of human tissues and cell types. *Genome Biol.* 2013;14(10):R115.

73. Hannum G, Guinney J, Zhao L, et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol Cell.* 2013;49(2):359-367.

74. Chen BH, Marioni RE, Colicino E, et al. DNA methylation-based measures of biological age: meta-analysis predicting time to death. *Aging (Albany NY).* 2016;8(9):1844-1865.

75. Levine ME, Lu AT, Quach A, et al. An epigenetic biomarker of aging for lifespan and healthspan. *Aging (Albany NY).* 2018;10(4):573-591.

76. Lu AT, Quach A, Wilson JG, et al. DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging (Albany NY).* 2019;11(2):303-327.

77. Huls A, Czamara D. Methodological challenges in constructing DNA methylation risk scores. *Epigenetics.* 2020;15(1-2):1-11.

78. Richard MA, Huan T, Ligthart S, et al. DNA Methylation Analysis Identifies Loci for Blood Pressure Regulation. *Am J Hum Genet.* 2017;101(6):888-902.

79. Chu AY, Tin A, Schlosser P, et al. Epigenome-wide association studies identify DNA methylation associated with kidney function. *Nat Commun.* 2017;8(1):1286.

80. Liu Y, Reynolds LM, Ding J, et al. Blood monocyte transcriptome and epigenome analyses reveal loci associated with human atherosclerosis. *Nat Commun.* 2017;8(1):393.

81. Rask-Andersen M, Martinsson D, Ahsan M, et al. Epigenome-wide association study reveals differential DNA methylation in individuals with a history of myocardial infarction. *Hum Mol Genet.* 2016;25(21):4739-4748.

82. Fernández-Sanlés A, Sayols-Baixeras S, Subirana I, et al. DNA Methylation Biomarkers Of Myocardial Infarction And Cardiovascular Disease. *bioRxiv.* 2020:707315.

83. Nakatochi M, Ichihara S, Yamamoto K, et al. Epigenome-wide association of myocardial infarction with DNA methylation sites at loci related to cardiovascular disease. *Clin Epigenetics.* 2017;9:54.

84. Agha G, Mendelson MM, Ward-Caviness CK, et al. Blood Leukocyte DNA Methylation Predicts Risk of Future Myocardial Infarction and Coronary Heart Disease. *Circulation.* 2019;140(8):645-657.

85. Smith JA, Raisky J, Ratliff SM, et al. Intrinsic and extrinsic epigenetic age acceleration are associated with hypertensive target organ damage in older African Americans. *BMC Med Genomics.* 2019;12(1):141.

86. Hillary RF, Stevenson AJ, McCartney DL, et al. Epigenetic measures of ageing predict the prevalence and incidence of leading causes of death and disease burden. *Clin Epigenetics.* 2020;12(1):115.

87. Ryan J, Wrigglesworth J, Loong J, Fransquet PD, Woods RL. A systematic review and meta-analysis of environmental, lifestyle and health factors associated with DNA methylation age. *J Gerontol A Biol Sci Med Sci.* 2019.

88. Irvin MR, Aslibekyan S, Do A, et al. Metabolic and inflammatory biomarkers are associated with epigenetic aging acceleration estimates in the GOLDN study. *Clin Epigenetics.* 2018;10:56.

89. Grant CD, Jafari N, Hou L, et al. A longitudinal study of DNA methylation as a potential mediator of age-related diabetes risk. *Geroscience.* 2017;39(5-6):475-489.

90. Nannini DR, Joyce BT, Zheng Y, et al. Epigenetic age acceleration and metabolic syndrome in the coronary artery risk development in young adults study. *Clin Epigenetics.* 2019;11(1):160.

91. Dugue PA, Bassett JK, Joo JE, et al. Association of DNA Methylation-Based Biological Age With Health Risk Factors and Overall and Cause-Specific Mortality. *Am J Epidemiol.* 2018;187(3):529-538.

92. Roetker NS, Pankow JS, Bressler J, Morrison AC, Boerwinkle E. Prospective Study of Epigenetic Age Acceleration and Incidence of Cardiovascular Disease Outcomes in the ARIC Study (Atherosclerosis Risk in Communities). *Circ Genom Precis Med.* 2018;11(3):e001937.

93. Fransquet PD, Wrigglesworth J, Woods RL, Ernst ME, Ryan J. The epigenetic clock as a predictor of disease and mortality risk: a systematic review and meta-analysis. *Clin Epigenetics.* 2019;11(1):62.

94. Lind L, Ingelsson E, Sundstrom J, Siegbahn A, Lampa E. Methylation-based estimated biological age and cardiovascular disease. *Eur J Clin Invest.* 2018;48(2).

95. Horvath S, Gurven M, Levine ME, et al. An epigenetic clock analysis of race/ethnicity, sex, and coronary heart disease. *Genome Biol.* 2016;17(1):171.

96. Wang C, Ni W, Yao Y, et al. DNA methylation-based biomarkers of age acceleration and all-cause death, myocardial infarction, stroke, and cancer in two cohorts: The NAS, and KORA F4. *EBioMedicine.* 2020;63:103151.

97. Perna L, Zhang Y, Mons U, Holleczek B, Saum KU, Brenner H. Epigenetic age acceleration predicts cancer, cardiovascular, and all-cause mortality in a German case cohort. *Clin Epigenetics.* 2016;8:64.

98. Daniels PR, Kardia SL, Hanis CL, et al. Familial aggregation of hypertension treatment and control in the Genetic Epidemiology Network of Arteriopathy (GENOA) study. *Am J Med.* 2004;116(10):676-681.

99. Ammous F, Zhao W, Ratliff SM, et al. Epigenome-wide association study identifies DNA methylation sites associated with target organ damage in older African Americans. *Epigenetics.* 2020:1-14.

Chapter 2. Epigenome-wide Association Study Identifies DNA Methylation Sites Associated with Target Organ Damage in Older African Americans

2.1 Abstract

Target organ damage (TOD) manifests as vascular injuries in the body organ systems associated with long standing hypertension. DNA methylation in peripheral blood leukocytes can capture inflammatory processes and gene expression changes underlying TOD. We investigated the association between epigenome-wide DNA methylation and 5 measures of TOD (estimated glomerular filtration rate (eGFR), urinary albumin-creatinine ratio (UACR), left ventricular mass index (LVMI), relative wall thickness (RWT), and white matter hyperintensity (WMH)) in 961 African Americans from hypertensive sibships. A multivariate (multi-trait) model of eGFR, UACR, LVMI, and RWT identified 7 CpGs associated with at least one of the traits (cg21134922, cg04816311 near *C7orf50*, cg09155024, cg10254690 near *OAT*, cg07660512, cg12661888 near *IFT43*, and cg02264946 near *CATSPERD*) at FDR q<0.1. Adjusting for blood pressure attenuated the association for cg04816311, and further adjustment for body mass index and type 2 diabetes attenuated the association for 3 additional CpG sites (cg21134922, cg09155024, and cg10254690). Although DNA methylation was associated with *cis*-gene expression for some CpGs, no significant evidence of mediation by gene expression was detected. Mendelian randomization analyses suggested causality between three CpGs and eGFR (cg04816311, cg10254690, and cg07660512). We also assessed whether the identified CpG sites were associated with TOD in 614 African Americans in the Hypertension Genetic Epidemiology Network (HyperGEN) study. Out of the 3 CpG sites available for replication, cg04816311 was

significantly associated with eGFR (p=0.0003), LVMI (p=0.0003), and RWT (p=0.002). This study found evidence of an association between DNA methylation and TOD in African Americans and highlights the utility of using a multivariate-based model that leverages information across related traits in epigenome-wide association studies.

2.2 Introduction

More than 40% of US adults have hypertension, with African Americans having higher prevalence than that of any other racial group.[1] According to the 2011-2014 National Health and Nutrition Examination Survey, the prevalence of hypertension was 59% in African American men and 56% in African American women, compared to a prevalence of 47% and 41% in non-Hispanic White men and women.[1] Due to long standing hypertension, target organ damage (TOD) manifests as subclinical or clinical changes to the micro and macro vascular systems of the heart, brain, eyes, and kidneys.[2] Oxidative stress, endothelial dysfunction, extracellular matrix formation, and innate and adaptive immune cells activation and invasion are some of the mechanisms recognized to play a role in TOD onset and progression.[2,3] Previous research indicates that measures of TOD are independent and strong predictors of cardiovascular morbidity and mortality and all-cause mortality.[4-7] However, there are gaps in our knowledge about the progressive effects of hypertension over the life course and which risk factors contribute to the development of subclinical changes and symptomatic diseases related to TOD. African Americans are especially susceptible to manifestations of target organ damage from hypertension leading to a higher risk of adverse cardiovascular and renal outcomes and mortality.[8,9] Previous studies have shown that African Americans have greater left ventricular mass,[10,11] higher incidence of chronic kidney disease, and end stage renal disease compared to non-Hispanic Whites.[12,13] Better understanding of the pathogenesis of TOD could elucidate and improve our prediction of related morbidities and mortality, especially in populations with the highest burden of hypertension.

A complex interplay of genetics and the environment contributes to the risk of TOD. Heritability studies point to a substantial genetic component, with heritability estimates ranging

from 0.17 to 0.76.[14-18] Familial clustering patterns in Framingham Heart Study reveal that individuals who have at least one parent with TOD had an increase in the odds of any type of TOD, even after controlling for hypertension status.[19] Epigenetic mechanisms may shed light on the ways that genetics, the environment, and traditional risk factors contribute to the progression of TOD.[20,21] To date, only a handful of well-powered epigenome-wide association studies (EWAS) have investigated the association between DNA methylation and blood pressure or related organ damage.[22-24]

In this study, we investigated the association between DNA methylation, interrogated in peripheral blood leukocytes, and 5 TOD measures in a cohort of older African Americans using longitudinally collected data. We hypothesized that biological pathways involved in the association between DNA methylation and TOD could be unique to each trait or could act in a pleiotropic manner. Hence, to gain a better understanding of these biological mechanisms, we employed both univariate (single-trait) and multivariate (multi-trait) models. Additionally, we assessed whether the effects of CpG sites on TOD were mediated by the expression of nearby genes and we used Mendelian randomization (MR) to investigate causality.

2.3 Methods

Study sample

Genetic Epidemiology Network of Arteriopathy (GENOA) is a community-based study in Rochester, MN and Jackson, MS that aims to identify genes influencing blood pressure [25]. In the first phase of GENOA (Phase I: 1996 – 2001), sibships with at least two adults with clinically diagnosed essential hypertension before age 60 were recruited, and all siblings in the sibship were invited to participate regardless of hypertension status [20]. Exclusion criteria included secondary hypertension, alcoholism or drug abuse, pregnancy, insulin-dependent diabetes

mellitus, or active malignancy. In Phase I, a total of 1,583 non-Hispanic whites (Rochester, MN) and 1,854 African Americans (Jackson, MS) were enrolled. In the second phase (Phase II: 2001 – 2005), all participants were invited for a second examination. Eighty percent of African Americans (N = 1,482) and 75% of non-Hispanic whites (N = 1,213) from Phase 1 returned. Demographic information, medical history, clinical characteristics, lifestyle factors, and blood samples were collected in each phase.

This study includes African American participants who had their DNA methylation profiles measured in whole blood samples collected at Phase I. Measures of TOD were collected at Phase II and/or in an ancillary study of brain magnetic resonance imaging (MRI) conducted shortly after Phase II. Gene expression profiles were measured from lymphoblastoid cell lines made from blood samples collected after Phase 1 (median time from Phase 1 was 5.8 years). A total of 961 participants who had DNA methylation measurements at Phase I and returned for Phase II were included in the current study. Written informed consent was obtained from all participants and approval was granted by participating institutional review boards (University of Michigan, University of Mississippi Medical Center, and Mayo Clinic).

Target organ damage measures

Five TOD measurements were selected for the current analyses. Two traits, estimated glomerular filtration rate (eGFR) and urinary-albumin-creatinine ratio (UACR), measured diminishing kidney function. Left ventricular mass index (LVMI) and relative wall thickness (RWT) captured structural remodeling of the heart. White matter hyperintensity (WMH) captured altered areas of white matter in the brain using MRI. eGFR, UACR, LVMI, and RWT were assessed at Phase II. WMH was assessed in an ancillary study shortly after Phase II.

Estimated glomerular filtration rate (eGFR) and urine albumin-to-creatinine ratio (UACR)

Blood was drawn on the morning of the study visit after an overnight fast of at least 8 hours. Serum creatinine was assessed with enzymatic assays on a Hitachi 911 Chemistry Analyzer (Roche Diagnostics, Indianapolis, IN). The CKD-EPI formula was used to calculate the estimated glomerular filtration rate (eGFR).[26] A first-morning sample urine was collected on the morning of the study visit. Urine creatinine and urine albumin were assessed with enzymatic assays on a Hitachi 911 Chemistry Analyzer (Roche Diagnostics, Indianapolis, IN).

Left ventricular mass index (LVMI) and relative wall thickness (RWT)

Left ventricular mass (LVM) was estimated as previously described.[27] Briefly, Doppler, two-dimensional (2D) and M-mode (2D-guided) echocardiograms were performed following a standardized protocol.[28] Measurements were made at the echocardiography reading center using a computerized review station equipped with a digitizing tablet and monitor overlay used for calibration and quantification (Digisonics, Inc., Houston, Texas). LVM was calculated using end-diastolic dimensions by an anatomically validated formula and LVMI was derived by indexing LVM to the height raised to the power of 2.7.[29] RWT was calculated as twice the posterior wall thickness divided by the left ventricular internal dimension.[28]

White matter hyperintensity (WMH)

Brain magnetic resonance imaging was performed using Signa 1.5 T MRI scanners (GE Medical Systems, Waukesha, WI, USA) and images were processed at Mayo Clinic.[30] Total brain and WMH volume in the corona-radiata and periventricular zone were determined from axial fluid-attenuated inversion recovery (FLAIR) images.[31] Brain scans with cortical infarctions were excluded from the analyses because of the distortion of the WMH volume estimates that would be introduced in the automated segmentation algorithm. For additional details, see Smith et al.[32] Models assessing WMH were adjusted for total intracranial volume (TIV).

Methylation measures

Genomic DNA from 1,106 African American participants from Phase I was extracted from stored peripheral blood leukocytes using AutoGen FlexStar (AutoGen, Holliston, MA). Bisulfite conversion was done using with the EZ DNA Methylation Kit (Zymo Research, Irvine, CA) and then DNA methylation was measured using the Infinium MethylationEPIC BeadChip. IDAT files were imported using the Minfi R package[33] and sex mismatches and outliers were excluded using the shinyMethyl R package.[34] Probes with detection *p*-value < 10^{-16} were considered to be successfully detected[35] and both samples and probes that failed a detection rate of at least 10% were removed. Samples with incomplete bisulfite conversion identified using the QCinfo function in the ENmix R package[36] were removed. Sample identity was checked using the 59 SNP probes included in the EPIC BeadChip and mismatched samples were removed. Afterwards, the Noob method was used for individual background and dye-bias normalization.[37] Since two types of probes were present on the EPIC BeadChip (Infinium I and Infinium II), we used the Regression on Correlated Probes (RCP) method to adjust for the probe-type bias in the data.[38] Cross-hybridizing probes and those on sex chromosomes were removed using DMRcate R package.[39]

Methylation beta values were changed to M-values using logit transformation ($log_2[Beta/1 - Beta]$).[40] Sample plate, sentrix ID and sample row were identified as batch effect variables using principle variance component analysis. White blood cell type proportions within the blood sample were estimated using Houseman's method.[41] Methylation M-values were adjusted for white blood cell type counts and batch effects using linear mixed modeling and the residuals were added to the mean. After quality control, a total of 1,100 participants remained for further analysis.

Gene expression measures

Gene expression levels in GENOA African American participants were measured in Epstein-Barr virus (EBV) transformed lymphoblastoid cell lines using the Affymetrix Human Transcriptome Array 2.0. The Affymetrix Expression Console provided by Affymetrix was used for array quality control and all array images passed visual inspection. Affymetrix CEL files were normalized using the Robust Multichip Average algorithm in the Affymetrix Power Tool software.[42] The Brainarray custom CDF version 19 was used to map the probes to genes[43] and Combat[44] was used to adjust batch effects and other technical covariates. A total of 17,616 autosomal protein coding genes were available for analysis. A total of 1,205 samples remained after quality control.

Genotyping and imputation information for MR analysis

GENOA African American samples were genotyped on the Affymetrix® Genome-Wide Human SNP Array 6.0 platform, Illumina® Human1M-Duo, or Human660W-Quad BeadChips. Participants were excluded if they had a missing SNP call rate ≥ 0.05 or were an outlier ≥ 6 standard deviations (SD) from the mean of the first 10 genome-wide principal components from genotype data. SNPs were excluded if they had unknown chromosomal location, < 95% call rate, or minor allele frequency less than 0.01. Imputation was performed separately by chip (Affymetrix or Illumina) using the 1000 Genomes Phase3 v5 reference panel and post-imputation comparison between the two groups revealed no substantial differences in genotype and allele frequencies. SNPs with imputation quality < 0.8 were removed before MR analysis.

Covariates

Height was measured by stadiometer and weight by electronic balance. Body mass index (BMI) was calculated as weight in kilograms divided by the square of height in meters. Resting

systolic blood pressure (SBP) and diastolic blood pressure (DBP) were measured by a random zero sphygmomanometer and a cuff appropriate for arm size. The second and third of three readings, after the participant sat for at least 5 minutes, were averaged for analysis.[45] Information on current anti-hypertensive medication use was collected. Smoking was categorized as current, former, or never smokers. Type 2 diabetes (T2D) was defined as fasting serum glucose concentration >126 mg/dl or self-reported physician-diagnosed diabetes and current anti-diabetes medication use (insulin or hypoglycemic agents).

Statistical analyses

Epigenome-wide association analysis of TOD

TOD traits were rank-based inverse normalized to ensure they follow a Gaussian distribution and outliers at 4 standard deviations from the mean were removed. Analyses were performed using the multivariate linear mixed model implemented in the Genome-wide Efficient Mixed Model Association (GEMMA) software.[46] Random effects in the linear mixed models were used to adjust for familial relationship by using the genetic relatedness matrix. SNP-based heritability (h^2) of the TOD traits was additionally estimated using GEMMA.[47] TOD traits were investigated individually using univariate (single-trait) linear mixed models and jointly using a multivariate (multi-trait) linear mixed model. The multivariate method tests a hypothesis of a given CpG being associated with at least one of the TOD traits. We hypothesized that given the correlation between the TOD traits and potentially similar underlying biological mechanisms, some CpG sites will have true pleiotropic effects on the different TOD traits and using this approach will provide a better model fit. Additionally, using a multivariate approach will increase our statistical power, even if the methylation sites are not associated with all of the TOD traits at the univariate level.[47] The multivariate model included eGFR, UACR, LVMI and RWT

as outcome variables. WMH was not included as the sample size is significantly smaller (N = 539). Following multivariate regression, we tested the association between the significant sites identified in the multivariate model and WMH.

Univariate and multivariate models were adjusted for age at trait measurement, sex, time between measurements, smoking status, and 4 genetic principal components (PCs) (model 1). The minimally adjusted multivariate model (model 1) was considered our main analysis model. For significant CpG sites identified in the main analysis model, we examined the scatterplots of the CpGs against TOD to identify and potentially remove any extreme outliers or leverage points. We investigated attenuation of effects for the CpG sites identified in the main analysis by adjusting for SBP, DBP and antihypertensive medication use (model 2) and T2D status and BMI (model 3). To account for multiple comparisons, the Benjamini-Hochberg procedure was applied to control the false discovery rate (FDR) at a threshold of q<10%.[48]

Illumina documentation was used to annotate significant CpG sites with functional and regulatory features that mapped their genomic location relative to CpG islands and nearby genes. Additionally, we examined any overlap with regulatory elements reported by the Encyclopedia of DNA Elements (ENCODE)[49] and the Functional ANnoTation Of the Mammalian genome (FANTOM).[50]

Mediation of CpG-TOD associations by gene expression levels

For significant CpG sites identified in the main analysis model, we selected genes within a 250 Kb range with expression levels associated with the corresponding CpG at p-value < 0.1 for formal mediation analysis. Linear mixed models were adjusted for familial relationship, age, sex, time between the measures, and 4 genetic PCs. UACR was natural log transformed as ln(UACR+1). The *mediation* R package was then used to test for mediation by proximal gene

expression levels between each CpG site and TOD association.[51,52] To account for multiple testing, an FDR $q < 0.1$ was applied across all CpG-mediator-TOD associations.

Mendelian randomization analysis of CpG-TOD associations

To assess whether the significant CpG sites identified contributed causally to TOD measures, we performed one-sample Mendelian randomization (MR) analysis. We opted to use one-sample MR given the limited availability of public methylation quantitative trait locus (mQTL) databases and TOD genome wide association studies (GWAS) in participants of African ancestry. After excluding SNPs with minor allele frequency < 0.05, instrumental variables (IV) were drawn from SNPs within ± 1 Mb of each corresponding significant CpG site. RVTEST was used to assess the CpG-SNP association and identify eligible SNPs (FDR $q < 0.1$).[53] Clumping for independence using the Phase 3, version 5 of the 1000 Genomes dataset for Africans (AFR), at r^2 of 0.1 and physical distance threshold of 100kb, was done using PLINK v1.07.[54] Causal estimates were obtained using the inverse-variance weighted (IVW) model in the *MendelianRandomization* R package[55] for each TOD trait. UACR was natural log transformed as ln(UACR+1). Models were adjusted for familial relationship, age at baseline, sex, and 4 genetic PCs. A *p*-value < 0.05 was used. Additionally, we applied the MR Egger test to examine pleiotropy because IVW MR is invalid in the presence of horizontal pleiotropic effects of IVs.[56] Finally, since both the IVW and Egger methods rely on summary data, we also calculated the causal estimates for the significant results using two-stage least squares regression.[57]

Replication analysis

For the 7 CpG sites identified in the multivariate model 1 in GENOA, we sought to test the association between each CpG site and each TOD trait univariately in 614 African Americans from the Hypertension Genetic Epidemiology Network (HyperGEN) study. Like GENOA,

families in HyperGEN were selected if sibships had ≥2 siblings who had been diagnosed with hypertension before age 60. DNA methylation was assayed from peripheral blood leukocytes (buffy coat) using the Infinium HumanMethylation450 BeadChip (450K). Further details about HyperGEN can be found in the **Supplemental Text**.

2.4 Results

Descriptive statistics

Sample characteristics are described in **Table 1**. The mean age of the participants was 57.7 years at baseline and the majority of the participants were females. The mean follow-up time between phases was approximately 5 years. Mean SBP and DBP was 137.8 mmHg and 79.3 mmHg, respectively. Of the sample, about 12% were current smokers and 60% were never smokers. TOD heritability estimates ranged between 0.214 for WMH and 0.523 for LVMI. TOD traits were significantly correlated at *p*-value < 0.05 (*r* range: -0.21 to 0.36, **Table S1**) with the exception of WMH and UACR (r=0.08).

Epigenome-wide association analysis of TOD

Results from the minimally adjusted univariate and multivariate (4-trait: eGFR, UACR, LVMI, and RWT) epigenome-wide association analyses are shown in **Table 2.** The corresponding QQ plots are shown in the **Figure S1.** RWT, LVMI, and UACR models each identified a single CpG site at FDR q < 0.1, while no site reached the statistical significance threshold for eGFR. Our initial analysis identified 8 significant CpG sites associated with TOD at FDR q < 0.1 for our main multivariate model. However, upon inspection of scatterplots, we identified one CpG site (cg07235511) with an outlier that drove the association. This site was excluded from our report of findings, leaving a total of 7 CpG sites in the minimally adjusted model (model 1).

We extended the multivariate model 1 by adjusting for SBP, DBP, antihypertensive medication use, BMI, and T2D status. **Table 3** shows the results of the adjusted models for the 7 CpG sites previously identified in model 1 (q<0.1). Adjusting for hypertension (Model 2) attenuated the association for one CpG site (cg04816311) while the remaining 6 CpG sites remained significant. Adjustment for BMI and T2D further attenuated the association for 3 more CpG sites (cg21134922, cg09155024, and cg10254690). In addition, one new CpG site was found significant at the same threshold (Model 3, cg02204965, *p*-value = 5.02×10^{-7}, FDR q = 0.098).

We next performed follow-up analysis for the 7 identified CpG sites from multivariate Model 1 to: [1] assess their association with the individual TOD traits univariately, [2] identify which combination of traits were driving the multivariate association, and [3] examine the association between the identified CpG sites and WMH. The results of the univariate models for each of these sites with each TOD trait are shown in the **Table S2**. In model 1, 4 of the 7 identified sites were significantly associated with two TOD traits using a Bonferroni adjusted cutoff of *p*-value < 0.007, while the other 3 were associated with only LVMI. Both cg04816311 and cg12661888 had a consistent effect in which increased methylation was associated with worse TOD outcomes (higher UACR, lower eGFR and/or higher LVMI). To aid in the clinical interpretation of the results, equivalent linear mixed models were run for each significant CpG site identified in the main analysis model and each of the TOD traits without normalization (**Table S3)**. Individual TOD values were calculated at the mean - 1SD and mean + 1SD M-value methylation levels for each CpG site. Adjusting for SBP, DBP, and antihypertensive medication use did not substantively change the results from Model 1, except that cg02264946 became associated with RWT in addition to LVMI. Further adjustment for BMI and T2D attenuated the

association between cg04816311 and LVMI but not UACR. None of the 7 CpG sites identified were associated with WMH at *p*-value < 0.05 **(Table S4)**.

Bioinformatic characterization of CpG sites and mediation of CpG-TOD associations by gene expression levels

We characterized significant CpG sites bioinformatically and by examining their association with proximal gene expression. Five out of the 7 sites identified in the multivariate model were located in DNase hypersensitivity sites **(Table S5)**, but no overlap was found with any FANTOM sites. Gene expression was derived from cell lines created from blood samples that were taken a minimum of 1.9 years after methylation measurement. Within the identified range of ± 250 Kb of the 7 CpG sites identified, a total of 14 genes were marginally associated with nearby CpG methylation levels at *p*-value < 0.1 **(Table 4)**. These genes were selected for the mediation analysis. Methylation at the examined CpG sites was associated with decreased expression of the *SAFB2*, *UNCX*, *MORN3*, *HPD*, and *OAT* genes as well as a number of long noncoding RNA molecules. Formal mediation analyses did not identify mediation by gene expression for any of the CpG-TOD traits associations at FDR q < 0.1; however, post hoc power calculation showed that statistical power for the mediation analysis was low (<50%).

Mendelian randomization analysis of CpG-TOD associations

Table 5 shows the results of inverse-variance weighted models to assess whether the identified CpG sites were casually associated with each of the TOD traits. For 4 of the 7 CpG sites, we identified significant independent SNPs within ±1 Mb that had a CpG-SNP FDR q < 0.1 (**Table S6**). eGFR was inversely influenced by methylation at cg10254690 near *OAT* (effect estimate = -5.75 mL/min/1.73m^2 and *p-value* = 0.027) and cg07660512 (effect estimate = -14.91 mL/min/1.73m^2 and *p-value* = 0.042). Methylation at cg04816311 near *C7orf50* positively

influenced eGFR (effect estimate = 8.84 mL/min/1.73m^2 and *p-value* = 0.014). The effect estimates were consistent in direction, but of greater magnitude, than those estimated by our baseline association analysis at the univariate level. **Figure S2 (A-C)** shows the plots of these causal estimates using the IVW method, which shows an overall consistent effect of the SNPs. The MR Egger test suggested no evidence of horizontal pleiotropy for all CpGs with the exception of the effect of cg10254690 on eGFR, where the MR Egger test showed no evidence of causality. Effect estimates and standard errors using two stage least squares regression were slightly larger than with the summary methods, but the p-values were similar **(Table S7**). The associations were only significant at a *p*-value threshold of 0.05, and none would reach statistical significance if corrected for multiple testing.

Replication analysis

Sample characteristics of the HyperGEN replication cohort are shown in **Table S8**. HyperGEN participants were younger than GENOA participants with a mean age of 48.4 (SD = 11.1) years. Mean BMI and antihypertensive medication use were similar in both cohorts. HyperGEN had a lower percentage of diabetic participants and a higher percentage of smokers. Both UACR and LVMI were higher in HyperGEN compared to GENOA. Out of 7 CpG sites significant at FDR q < 0.1 in GENOA, only 3 sites (cg04816311, cg09155024, and cg10254690) were present on the 450K chip used in HyperGEN. **Table S9** shows the univariate associations for each of the 3 CpG sites with each of the 4 TOD traits. In the minimally adjusted model (Model 1), cg04816311 was associated with eGFR, LVMI, and RWT at *p*-value < 0.05. Effect directions were consistent in that increased methylation was associated with increased TOD: higher LVMI in both cohorts, higher RWT in HyperGEN, and higher UACR in GENOA. However, it was also associated with lower eGFR in HyperGEN (an indicator of less TOD). The

second CpG, cg09155024 was not associated with any of the TOD measures in HyperGEN. The third CpG, cg10254690, which was associated with LVMI in GENOA, had the same beta coefficient direction in HyperGEN but did not reach statistical significance (*p*-value = 0.18).

2.5 Discussion

The study identified 7 CpG sites associated with 4 TOD measures in a cohort of older African Americans using a multivariate approach. The multivariate model had more power to detect significant CpG sites than the univariate models which detected a single CpG site for 3 of the TOD traits. This is in line with evidence from the literature on the power gains associated with multivariate over standard univariate analyses in genetic studies.[46,47,58-60] Testing the significant CpG sites identified in the multivariate model at the univariate level showed that not all of the sites reached statistical significance. While counterintuitive, this is expected as unassociated traits in the multivariate analysis increase power if they are correlated with the associated trait.[46,47] Our initial hypothesis was that DNA methylation would have the same direction of effect across the different TOD traits; however, our findings show that this may not be true for all CpG sites. None of the sites identified were associated with WMH, which may indicate a different underlying mechanism for that trait or may be due to the relatively low variability of WMH in this sample.

Five out of the 7 sites identified in the multivariate model were located in DNase hypersensitivity sites, which are regions of chromatin that are not highly condensed, rendering the chromatin exposed and accessible for transcription. After adjusting for hypertension-related covariates, 6 of these associations remained significant, and 3 remained significant when BMI and T2D were included. Although there was evidence for an association between the identified CpG sites and *cis*-gene expression, including both protein coding and long non-coding RNA

molecules, none were statistically significant in formal mediation analyses. MR analysis provided some evidence of causality between at least two CpG sites and eGFR.

The CpG site that was identified in GENOA and replicated in HyperGEN, cg04816311 near *C7orf50*, was previously reported to be associated with T2D in Mexican-Americans and sub-Saharan Africans.[61,62] cg04816311 was also associated with BMI in African Americans.[63] The direction of effect reported in the study by Meeks et al. showed that hypermethylation was observed at cg04816311 among T2D cases compared to the controls.[62] In our study, a change in M-value methylation (mean -1SD to mean +1SD) was associated with approximately 13 mg/g increase in UACR and 2.6 g/height$^{2.7}$ increase in LVMI in GENOA. In the model adjusting for BMI and T2D, the methylation effect was attenuated and was no longer significant. Genetic variants in this locus were also found to be associated with lipid levels,[64] blood pressure,[65,66] and longevity[67] in GWAS studies.

Methylation at the cg10254690 site, which maps near the *OAT* gene promoter region, was associated with significantly decreased expression of *OAT*. Additionally, MR analysis using IVW suggested a causal effect equivalent to approximately 6 mL/min/1.73m^2 units decrease in eGFR for each 1 unit M-value increase in DNA methylation. The *OAT* gene codes for ornithine aminotransferase, a key mitochondrial enzyme found in the liver, intestine, brain, and kidney that converts arginine and ornithine into glutamate and GABA.[68] Ornithine is involved in the urea cycle and synthesis of nitric oxide (NO). It is an important messenger molecule that regulates blood vessel dilation and has other thrombotic and inflammatory effects.[69,70] Studies suggest that *OAT* is involved in controlling the proliferation of several cell lines, including vascular smooth muscle cells, and it acts as a modulator of collagen synthesis and of extracellular matrix formation.[68,71] In addition, there is evidence that *OAT* is involved in metabolic reprogramming in

activated T-cells by providing ornithine and α-KG.[72] One previous study found a nominal association between DNA methylation at cg10254690 and cortisol stress reactivity,[73] and genetic variants in *OAT* were also associated with diastolic blood pressure.[74]

Other genes identified in the multivariate model were *IFT43* and *CATSPERD. IFT43*, near cg12661888, encodes a subunit of the intraflagellar transport complex A, a multi-protein complex involved in in cilia assembly and maintenance. This subunit is essential in regulating the Sonic Hedgehog signaling pathway, which is involved in regulating the growth, differentiation, and patterning of cells, especially during embryonic development.[75] *CATSPERD*, near cg02264946, encodes an auxiliary subunit of sperm calcium channel pore-forming proteins required for the motility of spermatozoa and male fertility.[76] cg03042953 near *SSBP3*, a single-stranded DNA-binding protein 3, was only significant for the RWT trait model. Genetic loci in the *SSBP3* locus were associated with P wave duration,[77] blood urea nitrogen,[78] and BMI.[79,80] The single trait model for UACR identified cg04816311 near *C7orf50*, which was also significant in the multivariate outcome model. cg21134922, found significant for LVMI, was not near a gene.

Changes in methylation may not affect the expression of the genes immediately proximal to the CpG site, and CpG regulatory effects could be more distal or in trans (i.e. affecting genes on different chromosomes).[81] When we evaluated whether CpG sites were associated with transcriptional changes, a number of genes within the range of 250 Kb showed decreased expression levels with increasing methylation. Previous literature identified DNA sequence variants in one of these genes, *UNCX* near cg04816311, to be associated with eGFR and other kidney-related traits in GWAS studies.[82,83] However, our study did not identify any significant mediation effects using formal mediation analysis. This could be attributed to the lack of

statistical power and/or the different cell types used for the methylation-gene expression measures (EBV-transformed lymphocytes cell line for gene expression versus peripheral blood leukocytes for DNA methylation). There is conflicting evidence on whether the methylation patterns of these cell lines are similar, whether the EBV transformation preserves the gene expression profile, and whether that translates to a similar gene expression profile in both.[84-87]

This study is among the very few studies that assessed the association between DNA methylation and TOD. To the best of our knowledge, it the first study to employ a pleiotropy-informed analysis to examine epigenome-wide DNA methylation sites associated with correlated TOD traits. However, this study has a number of limitations. First, although we employed a longitudinal design, it is difficult to rule out reverse causality, especially for TOD traits where disease onset and duration is difficult to detect. In addition, although MR analyses allowed us to characterize the direction of the association, findings did not reach statistical significance when accounting for multiple testing, which could be attributed to small sample size and low power. None of the sites identified in this study replicated epigenome-wide significant CpG sites from the limited epigenome-wide studies available for TOD traits.[22-24]

In conclusion, our study addresses an important gap in the literature on the role of DNA methylation in TOD in African Americans. CpG sites mapped to important genes that can further our understanding of biological mechanisms underlying these conditions. Future studies are needed to replicate these findings and further investigate gene expression and epigenetic profiles in more relevant tissue types such as the kidney and the heart, in addition to investigating the associations in different racial and ethnic groups. Findings from such research may direct future efforts that target early detection and intervention.

2.6 References

1. Whelton PK, Carey RM, Aronow WS, Casey DE, Jr., Collins KJ, Dennison Himmelfarb C, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. J Am Coll Cardiol. 2018;71(19):e127-e248.
2. Cohuet G, Struijker-Boudier H. Mechanisms of target organ damage caused by hypertension: therapeutic potential. Pharmacol Ther. 2006;111(1):81-98.
3. Norlander AE, Madhur MS, Harrison DG. The immunology of hypertension. J Exp Med. 2018;215(1):21-33.
4. Sehestedt T, Jeppesen J, Hansen TW, Wachtell K, Ibsen H, Torp-Pedersen C, et al. Risk prediction is improved by adding markers of subclinical organ damage to SCORE. Eur Heart J. 2010;31(7):883-91.
5. Harbaoui B, Courand PY, Defforges A, Khettab F, Milon H, Girerd N, et al. Cumulative Effects of Several Target Organ Damages in Risk Assessment in Hypertension. Am J Hypertens. 2016;29(2):234-44.
6. Ankle Brachial Index C, Fowkes FG, Murray GD, Butcher I, Heald CL, Lee RJ, et al. Ankle brachial index combined with Framingham Risk Score to predict cardiovascular events and mortality: a meta-analysis. JAMA. 2008;300(2):197-208.
7. Sundstrom J, Lind L, Arnlov J, Zethelius B, Andren B, Lithell HO. Echocardiographic and electrocardiographic diagnoses of left ventricular hypertrophy predict mortality independently of each other in a population of elderly men. Circulation. 2001;103(19):2346-51.
8. Carnethon MR, Pu J, Howard G, Albert MA, Anderson CAM, Bertoni AG, et al. Cardiovascular Health in African Americans: A Scientific Statement From the American Heart Association. Circulation. 2017;136(21):e393-e423.
9. Virani SS, Alonso A, Benjamin EJ, Bittencourt MS, Callaway CW, Carson AP, et al. Heart Disease and Stroke Statistics-2020 Update: A Report From the American Heart Association. Circulation. 2020;141(9):e139-e596.
10. Fox E, Taylor H, Andrew M, Han H, Mohamed E, Garrison R, et al. Body mass index and blood pressure influences on left ventricular mass and geometry in African Americans: The Atherosclerotic Risk In Communities (ARIC) Study. Hypertension. 2004;44(1):55-60.
11. Kizer JR, Arnett DK, Bella JN, Paranicas M, Rao DC, Province MA, et al. Differences in left ventricular structure between black and white hypertensive adults: the Hypertension Genetic Epidemiology Network study. Hypertension. 2004;43(6):1182-8.
12. Tarver-Carr ME, Powe NR, Eberhardt MS, LaVeist TA, Kington RS, Coresh J, et al. Excess risk of chronic kidney disease among African-American versus white subjects in the United States: a population-based study of potential explanatory factors. J Am Soc Nephrol. 2002;13(9):2363-70.

13. Choi AI, Rodriguez RA, Bacchetti P, Bertenthal D, Hernandez GT, O'Hare AM. White/black racial differences in risk of end-stage renal disease and death. Am J Med. 2009;122(7):672-8.

14. Freedman BI, Beck SR, Rich SS, Heiss G, Lewis CE, Turner S, et al. A genome-wide scan for urinary albumin excretion in hypertensive families. Hypertension. 2003;42(3):291-6.

15. Sachdev PS, Thalamuthu A, Mather KA, Ames D, Wright MJ, Wen W, et al. White Matter Hyperintensities Are Under Strong Genetic Influence. Stroke. 2016;47(6):1422-8.

16. Assimes TL, Narasimhan B, Seto TB, Yoon S, Curb JD, Olshen RA, et al. Heritability of left ventricular mass in Japanese families living in Hawaii: the SAPPHIRe Study. J Hypertens. 2007;25(5):985-92.

17. Peterson VR, Norton GR, Redelinghuys M, Libhaber CD, Maseko MJ, Majane OH, et al. Intrafamilial aggregation and heritability of left ventricular geometric remodeling is independent of cardiac mass in families of African ancestry. Am J Hypertens. 2015;28(5):657-63.

18. Fox CS, Yang Q, Cupples LA, Guo CY, Larson MG, Leip EP, et al. Genomewide linkage analysis to serum creatinine, GFR, and creatinine clearance in a community-based population: the Framingham Heart Study. J Am Soc Nephrol. 2004;15(9):2457-61.

19. Niiranen TJ, Lin H, Larson MG, Vasan RS. Familial clustering of hypertensive target organ damage in the community. J Hypertens. 2018;36(5):1086-93.

20. Zhang W, Song M, Qu J, Liu GH. Epigenetic Modifications in Cardiovascular Aging and Diseases. Circ Res. 2018;123(7):773-86.

21. Zhong J, Agha G, Baccarelli AA. The Role of DNA Methylation in Cardiovascular Risk and Disease: Methodological Aspects, Study Design, and Data Analysis for Epidemiological Studies. Circ Res. 2016;118(1):119-31.

22. Richard MA, Huan T, Ligthart S, Gondalia R, Jhun MA, Brody JA, et al. DNA Methylation Analysis Identifies Loci for Blood Pressure Regulation. Am J Hum Genet. 2017;101(6):888-902.

23. Chu AY, Tin A, Schlosser P, Ko YA, Qiu C, Yao C, et al. Epigenome-wide association studies identify DNA methylation associated with kidney function. Nat Commun. 2017;8(1):1286.

24. Chouliaras L, Pishva E, Haapakoski R, Zsoldos E, Mahmood A, Filippini N, et al. Peripheral DNA methylation, cognitive decline and brain aging: pilot findings from the Whitehall II imaging study. Epigenomics. 2018;10(5):585-95.

25. Daniels PR, Kardia SL, Hanis CL, Brown CA, Hutchinson R, Boerwinkle E, et al. Familial aggregation of hypertension treatment and control in the Genetic Epidemiology Network of Arteriopathy (GENOA) study. Am J Med. 2004;116(10):676-81.

26. Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF, 3rd, Feldman HI, et al. A new equation to estimate glomerular filtration rate. Ann Intern Med. 2009;150(9):604-12.

27. Arnett DK, Meyers KJ, Devereux RB, Tiwari HK, Gu CC, Vaughan LK, et al. Genetic variation in NCAM1 contributes to left ventricular wall thickness in hypertensive families. Circ Res. 2011;108(3):279-83.

28. Devereux RB, Roman MJ, Ganau A, de Simone G, Okin PM, Kligfield P. Cardiac and arterial hypertrophy and atherosclerosis in hypertension. Hypertension. 1994;23(6 Pt 1):802-9.

29. Lang RM, Badano LP, Mor-Avi V, Afilalo J, Armstrong A, Ernande L, et al. Recommendations for cardiac chamber quantification by echocardiography in adults: an update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. J Am Soc Echocardiogr. 2015;28(1):1-39 e14.

30. Jack CR, Jr., Twomey CK, Zinsmeister AR, Sharbrough FW, Petersen RC, Cascino GD. Anterior temporal lobes and hippocampal formations: normative volumetric measurements from MR images in young adults. Radiology. 1989;172(2):549-54.

31. Jack CR, Jr., O'Brien PC, Rettman DW, Shiung MM, Xu Y, Muthupillai R, et al. FLAIR histogram segmentation for measurement of leukoaraiosis volume. J Magn Reson Imaging. 2001;14(6):668-76.

32. Smith JA, Turner ST, Sun YV, Fornage M, Kelly RJ, Mosley TH, et al. Complexity in the genetic architecture of leukoaraiosis in hypertensive sibships from the GENOA Study. BMC Med Genomics. 2009;2:16.

33. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics. 2014;30(10):1363-9.

34. Fortin JP, Fertig E, Hansen K. shinyMethyl: interactive quality control of Illumina 450k DNA methylation arrays in R. F1000Res. 2014;3:175.

35. Lehne B, Drong AW, Loh M, Zhang W, Scott WR, Tan ST, et al. A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies. Genome Biol. 2015;16:37.

36. Xu Z, Niu L, Li L, Taylor JA. ENmix: a novel background correction method for Illumina HumanMethylation450 BeadChip. Nucleic Acids Res. 2016;44(3):e20.

37. Fortin JP, Triche TJ, Jr., Hansen KD. Preprocessing, normalization and integration of the Illumina HumanMethylationEPIC array with minfi. Bioinformatics. 2017;33(4):558-60.

38. Niu L, Xu Z, Taylor JA. RCP: a novel probe design bias correction method for Illumina Methylation BeadChip. Bioinformatics. 2016;32(17):2659-63.

39. Peters TJ, Buckley MJ, Statham AL, Pidsley R, Samaras K, R VL, et al. De novo identification of differentially methylated regions in the human genome. Epigenetics Chromatin. 2015;8:6.

40. Du P, Zhang X, Huang CC, Jafari N, Kibbe WA, Hou L, et al. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. BMC Bioinformatics. 2010;11:587.

41. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinformatics. 2012;13:86.

42. Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. Summaries of Affymetrix GeneChip probe level data. Nucleic Acids Res. 2003;31(4):e15.

43. Dai M, Wang P, Boyd AD, Kostov G, Athey B, Jones EG, et al. Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. Nucleic Acids Res. 2005;33(20):e175.

44. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics. 2007;8(1):118-27.

45. Turner ST, Kardia SL, Mosley TH, Rule AD, Boerwinkle E, de Andrade M. Influence of genomic loci on measures of chronic kidney disease in hypertensive sibships. J Am Soc Nephrol. 2006;17(7):2048-55.

46. Zhou X, Stephens M. Efficient multivariate linear mixed model algorithms for genome-wide association studies. Nat Methods. 2014;11(4):407-9.

47. Stephens M. A unified framework for association analysis with multiple related phenotypes. PloS one. 2013;8(7):e65245.

48. Benjamini Y, Hochberg Y. 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):289-300.

49. Consortium EP. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012;489(7414):57-74.

50. Severin J, Waterhouse AM, Kawaji H, Lassmann T, van Nimwegen E, Balwierz PJ, et al. FANTOM4 EdgeExpressDB: an integrated database of promoters, genes, microRNAs, expression dynamics and regulatory interactions. Genome Biol. 2009;10(4):R39.

51. Tingley D, Yamamoto T, Hirose K, Keele L, Imai K. mediation: R Package for Causal Mediation Analysis. 2014. 2014;59(5):38.

52. Preacher KJ, Hayes AF. SPSS and SAS procedures for estimating indirect effects in simple mediation models. Behav Res Methods Instrum Comput. 2004;36(4):717-31.

53. Zhan X, Hu Y, Li B, Abecasis GR, Liu DJ. RVTESTS: an efficient and comprehensive tool for rare variant association analysis using sequence data. Bioinformatics. 2016;32(9):1423-6.

54. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81(3):559-75.

55. Yavorska OO, Burgess S. MendelianRandomization: an R package for performing Mendelian randomization analyses using summarized data. Int J Epidemiol. 2017;46(6):1734-9.

56. Hemani G, Bowden J, Davey Smith G. Evaluating the potential role of pleiotropy in Mendelian randomization studies. Hum Mol Genet. 2018;27(R2):R195-R208.

57. Kang H, Jiang Y, Zhao Q, Small DS. ivmodel: An R Package for Inference and Sensitivity Analysis of Instrumental Variables Models with One Endogenous Variable. arXiv preprint arXiv:200208457. 2020.

58. Korte A, Vilhjalmsson BJ, Segura V, Platt A, Long Q, Nordborg M. A mixed-model approach for genome-wide association studies of correlated traits in structured populations. Nat Genet. 2012;44(9):1066-71.

59. O'Reilly PF, Hoggart CJ, Pomyen Y, Calboli FC, Elliott P, Jarvelin MR, et al. MultiPhen: joint model of multiple phenotypes can increase discovery in GWAS. PLoS One. 2012;7(5):e34861.

60. Ferreira MA, Purcell SM. A multivariate test of association. Bioinformatics. 2009;25(1):132-3.

61. Kulkarni H, Kos MZ, Neary J, Dyer TD, Kent JW, Jr., Goring HH, et al. Novel epigenetic determinants of type 2 diabetes in Mexican-American families. Hum Mol Genet. 2015;24(18):5330-44.

62. Meeks KAC, Henneman P, Venema A, Addo J, Bahendeka S, Burr T, et al. Epigenome-wide association study in whole blood on type 2 diabetes among sub-Saharan African individuals: findings from the RODAM study. Int J Epidemiol. 2019;48(1):58-70.

63. Demerath EW, Guan W, Grove ML, Aslibekyan S, Mendelson M, Zhou YH, et al. Epigenome-wide association study (EWAS) of BMI, BMI change and waist circumference in African American adults identifies multiple replicated loci. Hum Mol Genet. 2015;24(15):4464-79.

64. Ligthart S, Vaez A, Hsu YH, Inflammation Working Group of the CC, Pmi Wg XCP, LifeLines Cohort S, et al. Bivariate genome-wide association study identifies novel pleiotropic loci for lipids and inflammation. BMC Genomics. 2016;17:443.

65. Evangelou E, Warren HR, Mosen-Ansorena D, Mifsud B, Pazoki R, Gao H, et al. Genetic analysis of over 1 million people identifies 535 new loci associated with blood pressure traits. Nat Genet. 2018;50(10):1412-25.

66. Takeuchi F, Akiyama M, Matoba N, Katsuya T, Nakatochi M, Tabara Y, et al. Interethnic analyses of blood pressure loci in populations of East Asian and European descent. Nat Commun. 2018;9(1):5052.

67. Yashin AI, Wu D, Arbeev KG, Ukraintseva SV. Joint influence of small-effect genetic variants on human longevity. Aging (Albany NY). 2010;2(9):612-20.

68. Ginguay A, Cynober L, Curis E, Nicolis I. Ornithine Aminotransferase, an Important Glutamate-Metabolizing Enzyme at the Crossroads of Multiple Metabolic Pathways. Biology (Basel). 2017;6(1).

69. Mori M, Gotoh T, Nagasaki A, Takiguchi M, Sonoki T. Regulation of the urea cycle enzyme genes in nitric oxide synthesis. J Inherit Metab Dis. 1998;21 Suppl 1:59-71.

70. Huynh NN, Chin-Dusting J. Amino acids, arginase and nitric oxide in vascular health. Clin Exp Pharmacol Physiol. 2006;33(1-2):1-8.

71. Durante W, Liao L, Reyna SV, Peyton KJ, Schafer AI. Transforming growth factor-beta(1) stimulates L-arginine transport and metabolism in vascular smooth muscle cells: role in polyamine and collagen synthesis. Circulation. 2001;103(8):1121-7.

72. Wang R, Dillon CP, Shi LZ, Milasta S, Carter R, Finkelstein D, et al. The transcription factor Myc controls metabolic reprogramming upon T lymphocyte activation. Immunity. 2011;35(6):871-82.

73. Houtepen LC, Vinkers CH, Carrillo-Roa T, Hiemstra M, van Lier PA, Meeus W, et al. Genome-wide DNA methylation levels and altered cortisol stress reactivity following childhood trauma in humans. Nat Commun. 2016;7:10967.

74. Parmar PG, Taal HR, Timpson NJ, Thiering E, Lehtimaki T, Marinelli M, et al. International Genome-Wide Association Study Consortium Identifies Novel Loci Associated With Blood Pressure in Children and Adolescents. Circ Cardiovasc Genet. 2016;9(3):266-78.

75. Liem KF, Jr., Ashe A, He M, Satir P, Moran J, Beier D, et al. The IFT-A complex regulates Shh signaling through cilia structure and membrane protein trafficking. J Cell Biol. 2012;197(6):789-800.

76. Chung JJ, Navarro B, Krapivinsky G, Krapivinsky L, Clapham DE. A novel gene required for male fertility and functional CATSPER channel formation in spermatozoa. Nat Commun. 2011;2:153.

77. Christophersen IE, Magnani JW, Yin X, Barnard J, Weng LC, Arking DE, et al. Fifteen Genetic Loci Associated With the Electrocardiographic P Wave. Circ Cardiovasc Genet. 2017;10(4).

78. Wuttke M, Li Y, Li M, Sieber KB, Feitosa MF, Gorski M, et al. A catalog of genetic loci associated with kidney function from analyses of a million individuals. Nat Genet. 2019;51(6):957-72.

79. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, et al. Genetic studies of body mass index yield new insights for obesity biology. Nature. 2015;518(7538):197-206.

80. Pulit SL, Stoneman C, Morris AP, Wood AR, Glastonbury CA, Tyrrell J, et al. Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of European ancestry. Hum Mol Genet. 2019;28(1):166-74.

81. Kennedy EM, Goehring GN, Nichols MH, Robins C, Mehta D, Klengel T, et al. An integrated -omics analysis of the epigenetic landscape of gene expression in human blood cells. BMC Genomics. 2018;19(1):476.

82. Pattaro C, Teumer A, Gorski M, Chu AY, Li M, Mijatovic V, et al. Genetic associations at 53 loci highlight cell types and biological pathways relevant for kidney function. Nat Commun. 2016;7:10023.

83. Hellwege JN, Velez Edwards DR, Giri A, Qiu C, Park J, Torstenson ES, et al. Mapping eGFR loci to the renal transcriptome and phenome in the VA Million Veteran Program. Nat Commun. 2019;10(1):3842.

84. Sun YV, Turner ST, Smith JA, Hammond PI, Lazarus A, Van De Rostyne JL, et al. Comparison of the DNA methylation profiles of human peripheral blood cells and transformed B-lymphocytes. Hum Genet. 2010;127(6):651-8.

85. Aberg K, Khachane AN, Rudolf G, Nerella S, Fugman DA, Tischfield JA, et al. Methylome-wide comparison of human genomic DNA extracted from whole blood and from EBV-transformed lymphocyte cell lines. Eur J Hum Genet. 2012;20(9):953-5.

86. Hansen KD, Sabunciyan S, Langmead B, Nagy N, Curley R, Klein G, et al. Large-scale hypomethylated blocks associated with Epstein-Barr virus–induced B-cell immortalization. Genome research. 2014;24(2):177-84.

87. Çalışkan M, Cusanovich DA, Ober C, Gilad Y. The effects of EBV transformation on gene expression levels and methylation profiles. Human molecular genetics. 2011;20(8):1643-52.

2.7 Tables

Table 2-1: Characteristics of participants in GENOA African Americans[1] (N = 961)

	Mean (SD) or N (%)	h^2 heritability (95% CI)
Female	685 (71.3%)	
Age (years)	62.7 (9.7)	
Age at phase I (years)	57.7 (10.3)	
Time between phases I and II	5.2 (1.3)	
BMI (kg/m^2)	31.8 (6.8)	
Systolic blood pressure (mmHg)	137.8 (21.1)	
Diastolic blood pressure (mmHg)	79.3 (11.1)	
Antihypertensive medication use		
No	284 (29.6%)	
Yes	662 (68.9%)	
Missing	15 (1.6%)	
Smoking status		
Current	118 (12.3%)	
Former	257 (26.7%)	
Never	571 (59.4%)	
Missing	15 (1.6%)	
Diabetes status[2]		
No	653 (68.0%)	
Yes	289 (30.1%)	
Missing	19 (2.0%)	
Estimated glomerular filtration rate (eGFR, mL/min/1.73m^2)	89.9 (21.2)	0.31 (0.113 – 0.50)
Urine albumin-to-creatinine ratio (UACR, mg/g)	48.4 (192.9)	0.49 (0.301 – 0.68)
Left ventricular mass index (LVMI, g/height$^{2.7}$)	39.1 (10.3)	0.52 (0.352 – 0.70)
Relative wall thickness (RWT)	0.32 (0.05)	0.30 (0.123 – 0.46)
White matter hyperintensity (WMH, cm^3)	9.49 (8.0)	0.21 (0 – 0.43)
Total intracranial volume (cm^3)	1373 (134.7)	

SD: Standard deviation; CI: Confidence interval; BMI: Body mass index
Sample sizes for target organ damage measures: eGFR (n=940), UACR (n=943), LVMI (n=910), RWT (n=915), WMH (n=539)
[1] Measured at phase II unless stated otherwise
[2] Defined as glucose level ≥ 126 mg/dL or taking diabetes medications

Table 2-2: Statistically significant CpG sites associated with TOD traits at FDR q < 0.1 using univariate and multivariate models in GENOA African Americans

Model 1 – univariate								
Outcome measure	**CpG site**	**Gene**	**Chr**	**Relation to CpG site**	**Relation to gene**	***p*-value**	**FDR q**	**Beta (SE)**
RWT	cg03042953	*SSBP3*	1	Island	Body	9.53×10^{-8}	0.072	1.361 (0.253)
LVMI	cg21134922		5			1.59×10^{-8}	0.013	0.663 (0.116)
UACR	cg04816311	*C7orf50*	7	North Shore	Body	5.52×10^{-8}	0.044	0.513 (0.094)

					Model 1 – multivariate [1]					
					Beta (SE) for each TOD trait					
CpG site	**Gene**	**Chr**	**Relation to CpG site**	**Relation to gene**	**eGFR**	**UACR**	**LVMI**	**RWT**	***p*-value**	**FDR**
cg21134922		5			-0.035 (0.106)	0.124 (0.118)	0.663 (0.117)	0.004 (0.113)	6.54×10^{-7}	0.044
cg04816311	*C7orf50*	7	North Shore	Body	0.178 (0.089)	0.492 (0.097)	0.390 (0.098)	0.093 (0.094)	1.06×10^{-7}	0.042
cg09155024		10			0.188 (0.057)	-0.070 (0.063)	-0.042 (0.064)	0.269 (0.060)	3.35×10^{-7}	0.044
cg10254690	*OAT*	10	Island	TSS1500 [2]	-0.154 (0.064)	-0.044 (0.071)	-0.359 (0.070)	-0.062 (0.068)	2.75×10^{-7}	0.065
cg07660512		12			-0.527 (0.162)	-0.795 (0.179)	-0.003 (0.182)	-0.242 (0.173)	2.62×10^{-7}	0.044
cg12661888	*IFT43*	14		Body	-0.306 (0.072)	0.259 (0.080)	-0.117 (0.081)	0.019 (0.077)	4.68×10^{-7}	0.053
cg02264946	*CATSPERD*	19		Body	0.236 (0.125)	0.203 (0.137)	-0.489 (0.137)	0.341 (0.132)	6.16×10^{-8}	0.042

Chr: chromosome; FDR: false discovery rate; UACR: urinary albumin to creatinine ratio; LVMI: left ventricular mass index; RWT: relative wall thickness; eGFR: estimated glomerular filtration rate

Model 1 is adjusted for age at phase II, sex, time between measurement, smoking status, and 4 genetic principal components

[1] Multivariate model includes eGFR, UACR, LVMI, and RWT as outcome measures

[2] Within 1500 kb of the gene start site (promoter region)

Table 2-3: Adjusted multivariate models for statistically significant CpG sites from Model 1 in GENOA African Americans

		Model 2 – multivariate[1]						**Model 3 – multivariate**[1]					
		Beta (SE) for each TOD trait						**Beta (SE) for each TOD trait**					
CpG site	**Gene**	**eGFR**	**UACR**	**LVMI**	**RWT**	***p*-value**	**FDR**	**eGFR**	**UACR**	**LVMI**	**RWT**	***p*-value**	**FDR**
cg21134922		-0.022 (0.011)	0.093 (0.013)	0.611 (0.012)	0.001 (0.012)	8.54×10^{-7}	**0.096**	0.004 (0.011)	0.076 (0.012)	0.421 (0.010)	-0.022 (0.012)	6.52×10^{-4}	0.9119
cg04816311	*C7orf50*	0.179 (0.008)	0.428 (0.009)	0.312 (0.009)	0.058 (0.009)	1.74×10^{-6}	0.172	0.141 (0.008)	0.243 (0.008)	0.148 (0.007)	-0.025 (0.008)	1.49×10^{-2}	0.9904
cg09155024		0.176 (0.003)	-0.096 (0.004)	-0.063 (0.004)	0.249 (0.004)	7.49×10^{-7}	**0.096**	0.164 (0.003)	-0.123 (0.003)	0.018 (0.003)	0.242 (0.004)	3.03×10^{-6}	0.2885
cg10254690	*OAT*	-0.164 (0.004)	-0.038 (0.005)	-0.340 (0.004)	-0.060 (0.004)	6.88×10^{-7}	**0.096**	-0.166 (0.004)	-0.224 (0.004)	-0.039 (0.004)	-0.008 (0.004)	3.08×10^{-4}	0.8776
cg07660512		-0.559 (0.026)	-0.800 (0.029)	0.020 (0.029)	-0.268 (0.029)	3.49×10^{-8}	**0.028**	-0.548 (0.025)	-0.703 (0.027)	0.114 (0.024)	-0.227 (0.028)	8.00×10^{-8}	**0.0363**
cg12661888	*IFT43*	-0.312 (0.005)	0.236 (0.006)	-0.137 (0.006)	-0.002 (0.006)	4.96×10^{-7}	**0.096**	-0.319 (0.005)	0.225 (0.005)	-0.122 (0.005)	-0.006 (0.006)	6.21×10^{-7}	**0.0981**
cg02264946	*CATSPERD*	0.219 (0.015)	0.261 (0.017)	-0.416 (0.017)	0.380 (0.017)	8.48×10^{-8}	**0.034**	0.224 (0.015)	0.280 (0.015)	-0.338 (0.014)	0.396 (0.017)	9.19×10^{-7}	**0.0363**

FDR: false discovery rate; SE: standard error; eGFR: estimated glomerular filtration rate; UACR: urinary albumin to creatinine ratio; LVMI: left ventricular mass index; RWT: relative wall thickness

Model 2 is adjusted for age at phase II, sex, time between measurement, smoking status, 4 genetic principal components, systolic blood pressure, diastolic blood pressure and antihypertensive medication use

Model 3 is adjusted for Model 2 variables plus body mass index and type 2 diabetes status

Results with a significant FDR $q < 0.1$ are in bold font

[1] Multivariate model includes eGFR, UACR, LVMI, and RWT as outcome measure

Table 2-4: Associations between statistically significant CpG sites from multivariate Model 1 and nearby gene expression (± 250 Kb, p-value < 0.1) in GENOA African Americans

CpG site	Gene	Protein Type	Beta	SE	*p*-value
cg04816311	*AC073957.1*	LncRNA	-0.156	0.050	0.002
	AC073094.1	LncRNA	-0.183	0.060	0.003
	AC091729.2	LncRNA	-0.127	0.057	0.026
	UNCX	Protein-coding	-0.108	0.058	0.063
cg10254690	*OAT*	Protein-coding	-0.163	0.057	0.005
cg07660512	*MORN3*	Protein-coding	-0.096	0.044	0.031
	AC079360.1	LncRNA	-0.060	0.028	0.035
	HPD	Protein-coding	-0.085	0.043	0.050
cg02264946	*AC011444.1*	LncRNA	-0.095	0.044	0.030
	SAFB2	Protein-coding	-0.147	0.073	0.045

SE: standard error

Table 2-5: Mendelian randomization results showing the inverse-variance weighted effects of multiple SNPs used as instrumental variables in the association of CpG sites and TOD traits in GENOA African Americans

TOD trait	Chr	CpG site	Independent SNPs with CpG-SNP FDR q < 0.1	Causal effect estimate [1]	Causal effect estimate SE	Lower bound of causal estimate	Upper bound of causal estimate	*p*-value
Estimated glomerular filtration rate (eGFR), mL/min/1.73m^2	5	cg21134922	6	8.718	6.2	-3.432	20.869	0.16
	7	cg04816311	15	8.844	3.584	1.819	15.868	0.014
	10	cg10254690	10	-5.752	2.594	-10.836	-0.669	0.027
	12	cg07660512	13	-14.912	7.321	-29.262	-0.563	0.042
Urinary albumin-to-creatinine ratio (UACR), mg/g	5	cg21134922	6	0.075	0.485	-0.876	1.026	0.877
	7	cg04816311	15	0.285	0.254	-0.213	0.783	0.261
	10	cg10254690	10	0.359	0.222	-0.076	0.793	0.106
	12	cg07660512	13	-0.57	0.571	-1.69	0.55	0.319
Left ventricular mass index (LVMI), g/height$^{2.7}$	5	cg21134922	6	6.292	3.299	-0.173	12.757	0.056
	7	cg04816311	15	1.816	1.919	-1.946	5.578	0.344
	10	cg10254690	10	-2.485	1.378	-5.186	0.216	0.071
	12	cg07660512	13	-5.694	3.689	-12.925	1.537	0.123
Relative wall thickness (RWT)	5	cg21134922	6	0.004	0.016	-0.027	0.036	0.780
	7	cg04816311	15	0.015	0.009	-0.002	0.031	0.084
	10	cg10254690	10	-0.002	0.007	-0.015	0.011	0.798
	12	cg07660512	13	0	0.02	-0.04	0.039	0.994

Chr: chromosome; SE: standard error

[1]The causal effect estimate is the pooled estimate calculated using inverse-variance weighted models from the Mendelian randomization analyses and is interpreted as the effect per 1 unit change in DNA methylation (M-value) using genetic variants on target organ damage trait

2.8 Supplementary Material

Table SM 2-1: Pearson correlation coefficients of TOD traits in GENOA African Americans

	Estimated glomerular filtration rate (eGFR)	Urinary albumin-creatinine ratio (UACR)	Left ventricular mass index (LVMI)	Relative wall thickness (RWT)
Estimated glomerular filtration rate (eGFR)				
Urinary albumin-creatinine ratio (UACR)	-0.20****			
Left ventricular mass index (LVMI)	-0.15****	0.17****		
Relative wall thickness (RWT)	-0.18****	0.09**	0.36****	
White matter hyperintensity (WMH)	-0.21****	0.08	0.14**	0.11*

p-value < 0.0001 '****', *p-value* < 0.001 '***', *p-value* < 0.01 '**', *p-value* < 0.05 '*'

Table SM 2-2: Univariate associations between CpG sites1 and each TOD trait in GENOA African Americans[1]

	CpG site	eGFR			UACR			LVMI			RWT		
		Beta	**SE**	***p*-value**	**Beta**	**SE**	***p*-value**	**Beta**	**SE**	***p*-value**	**Beta**	**SE**	***p*-value**
Model 1	cg21134922	-0.008	0.104	9.42E-01	0.142	0.115	2.20E-01	0.663	0.116	**1.59E-08**	0.009	0.112	9.36E-01
	cg04816311	0.179	0.086	3.66E-02	0.513	0.094	**5.52E-08**	0.389	0.098	**8.01E-05**	0.100	0.094	2.86E-01
	cg09155024	0.181	0.056	**1.19E-03**	-0.077	0.062	2.14E-01	-0.043	0.063	5.01E-01	0.264	0.060	**1.29E-05**
	cg10254690	-0.149	0.062	1.74E-02	-0.030	0.069	6.61E-01	-0.362	0.070	**2.64E-07**	-0.059	0.068	3.85E-01
	cg07660512	-0.443	0.159	**5.43E-03**	-0.754	0.175	**1.79E-05**	-0.009	0.181	9.60E-01	-0.248	0.172	1.50E-01
	cg12661888	-0.317	0.070	**6.10E-06**	0.246	0.077	**1.54E-03**	-0.117	0.081	1.49E-01	0.016	0.077	8.38E-01
	cg02264946	0.236	0.123	5.56E-02	0.173	0.135	2.02E-01	-0.478	0.136	**4.64E-04**	0.340	0.131	9.59E-03
Model 2	cg21134922	0.007	0.103	9.45E-01	0.104	0.110	3.47E-01	0.608	0.110	**3.84E-08**	0.000	0.111	9.97E-01
	cg04816311	0.184	0.085	3.07E-02	0.452	0.089	**5.38E-07**	0.310	0.093	**8.36E-04**	0.069	0.092	4.53E-01
	cg09155024	0.171	0.055	**2.11E-03**	-0.103	0.059	8.08E-02	-0.065	0.060	2.75E-01	0.242	0.059	**4.81E-05**
	cg10254690	-0.156	0.062	1.24E-02	-0.025	0.066	7.01E-01	-0.346	0.066	**1.88E-07**	-0.059	0.066	3.73E-01
	cg07660512	-0.473	0.158	**2.81E-03**	-0.779	0.166	**3.27E-06**	0.004	0.170	9.80E-01	-0.281	0.169	9.58E-02
	cg12661888	-0.327	0.069	**2.62E-06**	0.223	0.074	**2.57E-03**	-0.136	0.076	7.41E-02	-0.006	0.076	9.34E-01
	cg02264946	0.218	0.122	7.53E-02	0.225	0.129	8.09E-02	-0.400	0.129	**1.96E-03**	0.371	0.129	**4.06E-03**
Model 3	cg21134922	0.035	0.103	7.36E-01	0.080	0.106	4.53E-01	0.411	0.101	**4.98E-05**	-0.025	0.110	8.23E-01
	cg04816311	0.157	0.085	6.46E-02	0.271	0.087	**1.96E-03**	0.142	0.085	9.48E-02	-0.019	0.092	8.38E-01
	cg09155024	0.166	0.056	**2.89E-03**	-0.120	0.057	3.59E-02	-0.019	0.055	7.26E-01	0.240	0.059	**5.22E-05**
	cg10254690	-0.159	0.062	1.03E-02	-0.007	0.064	9.06E-01	-0.225	0.060	**2.10E-04**	-0.038	0.066	5.63E-01
	cg07660512	-0.473	0.157	**2.66E-03**	-0.684	0.161	**2.50E-05**	0.097	0.156	5.32E-01	-0.233	0.168	1.66E-01
	cg12661888	-0.335	0.069	**1.31E-06**	0.211	0.071	**3.11E-03**	-0.122	0.069	7.86E-02	-0.004	0.075	9.61E-01
	cg02264946	0.226	0.122	6.34E-02	0.267	0.124	3.19E-02	-0.325	0.118	**5.94E-03**	0.392	0.128	**2.32E-03**

SE: standard error; eGFR: estimated glomerular filtration rate; UACR: urinary albumin to creatinine ratio; LVMI: left ventricular mass index; RWT: relative wall thickness
Model 1 is adjusted for age at phase II, sex, time between measurements, smoking status, and 4 genetic principal components
Model 2 is adjusted for Model 1 variables plus systolic blood pressure, diastolic blood pressure and antihypertensive medication use
Model 3 is adjusted for Model 2 variables plus body mass index and type 2 diabetes status
Results significant at Bonferroni adjusted *p*-value are in bold font
[1] Only CpG sites that were statistically significant in multivariate Model 1 were included in this analysis

Table SM 2-3: Predicted change in TOD associated with CpG sites1 using linear mixed models and TOD modeled without normalization in GENOA African Americans[1]

CpG site	CpG mean (SD) in M-values	eGFR (mL/min/1.73m^2)	UACR[2] (mg/g)	LVMI (g/height$^{2.7}$)	RWT
cg21134922	1.61 (0.27)	-0.215	2.698	**3.328**	0.001
cg04816311	1.10 (0.33)	2.160	**13.111**	**2.559**	0.004
cg09155024	4.41 (0.50)	**3.453**	-2.438	0.024	**0.012**
cg10254690	-4.01 (0.45)	-2.778	-0.427	**-2.871**	-0.002
cg07660512	-0.51 (0.18)	**-3.209**	**-12.176**	-0.049	-0.004
cg12661888	4.11 (0.40)	**-5.214**	**6.786**	-0.967	0.002
cg02264946	5.01 (0.23)	2.110	3.246	**-2.096**	0.008

SD: standard deviation; eGFR: estimated glomerular filtration rate; UACR: urinary albumin to creatinine ratio; LVMI: left ventricular mass index; RWT: relative wall thickness

Estimates are calculated from linear mixed regression models of DNA methylation (M-values) on each of the TOD traits after adjusting for age at phase II, sex, time between measurements, smoking status, 4 genetic principal components, and pedigree information as a random effect. Shown values of TOD were calculated based on the difference in predicted TOD between those with less (-1SD) and those with more (+1SD) from mean DNA methylation M-values.

Results significant at Bonferroni adjusted *p*-value in the univariate Model 1 are in bold font

[1] Only CpG sites that were statistically significant in multivariate Model 1 were included in this analysis

[2] UACR was log-transformed, results shown were calculated using exponentiated coefficients

Table SM 2-4: Association between CpG sites1 and white matter hyperintensity in GENOA African Americans[1]

CpG site	Beta	SE	*p*-value
cg21134922	-0.203	0.130	0.118
cg04816311	0.195	0.104	0.063
cg09155024	-0.039	0.067	0.554
cg10254690	-0.116	0.077	0.131
cg07660512	-0.028	0.192	0.884
cg12661888	-0.026	0.085	0.757
cg02264946	0.131	0.150	0.383

SE: standard error

Model is adjusted for age at white matter hyperintensity measurement, sex, time between measurements, smoking status, 4 genetic principal components, and total intracranial volume

[1] Only CpG sites that were statistically significant in multivariate Model 1 were included in this analysis

Table SM 2-5: DNase hypersensitivity annotation for statistically significant CpG sites in GENOA African Americans[1]

CpG site	DNase Hypersensitivity
cg21134922	chr5:60586580-60587255
cg04816311	chr7:1066385-1067595
cg09155024	
cg10254690	chr10:126107765-126108135
cg07660512	chr12:122087245-122087650
cg12661888	
cg02264946	chr19:5775685-5776015

[1] Only CpG sites that were statistically significant in multivariate Model 1 were included in this analysis

Table SM 2-6: Independent SNPs identified as instrumental variables at FDR q < 0.1 for Mendelian randomization analyses in GENOA African Americans

Chromosome	**CpG site**	**Number of SNPs**	**rs ID[1]**	
5	cg21134922	6	rs6864367 rs372083232 rs111793128	rs11382326 rs7712470 rs6449494
7	cg04816311	15	rs2362529 rs74652290 rs79808627 rs12702456 rs4723350 rs2030958 rs183727057 rs112493548	rs111875889 rs12701824 rs11764167 rs35174522 rs4723974 rs4916968 rs4916971
10	cg10254690	10	rs9422806 rs11815286 rs138831926 rs17151852 rs6597838	rs11819582 rs2674335 rs11812766 rs201985927 rs115375556
12	cg07660512	13	rs11043181 rs11833075 rs115316599 rs151282886 rs28613497 rs144083908 rs145352774	rs140202929 rs370447554 rs138657431 rs10773287 rs143897726 rs56785986

[1] Independent based on clumping (r^2=0.1 and physical distance threshold=100kb) using the Phase 3, version 5 of the 1000 Genomes dataset for Africans (AFR) and CpG-SNP FDR q < 0.1

Table SM 2-7: Significant Mendelian randomization results with causal estimates of CpG sites and TOD traits calculated using the two-stage least squares regression method in GENOA African Americans

TOD trait	Chr	CpG site	Independent SNPs with CpG-SNP FDR q < 0.1	Causal effect estimate	Causal effect estimate SE	*p*-value
Estimated glomerular filtration rate (eGFR), mL/min/1.73m^2	7	cg04816311	15	11.836	4.591	0.010
	10	cg10254690	10	-6.764	3.011	0.025
	12	cg07660512	13	-15.879	8.251	0.055

Chr: chromosome; SE: standard error

Table SM 2-8: Characteristics of participants in HyperGEN (N = 614)

	Mean (SD) or N (%)
Female	409 (66.6%)
Age (years)	48.4 (11.1)
BMI (kg/m^2)	32.4 (8.2)
Systolic blood pressure (mmHg)	131.6 (23.6)
Diastolic blood pressure (mmHg)	75.5 (12.6)
Antihypertensive medication use	
No	175 (28.6%)
Yes	438 (71.5%)
Missing	1 (0.0%)
Smoking status	
Current	193 (31.4%)
Former	149 (24.3%)
Never	262 (42.7%)
Missing	10 (1.6%)
Diabetes status[1]	
No	492 (80.1%)
Yes	122 (19.9%)
Missing	0 (0.0%)
Estimated glomerular filtration rate (eGFR, mL/min/1.73m^2)	92.1 (21.9)
Urine albumin-to-creatinine ratio (UACR, mg/g)	81.6 (403.3)
Left ventricular mass index (LVMI, g/height$^{2.7}$)	45.0 (16.1)
Relative wall thickness (RWT)	0.3 (0.1)

SD: Standard deviation; CI: Confidence interval; BMI: Body mass index
Sample sizes for target organ damage measures: eGFR (N=613), UACR (N=606), LVMI (N=611), RWT (N=613)
[1] Defined as glucose level ≥ 126 mg/dL or taking diabetes medication

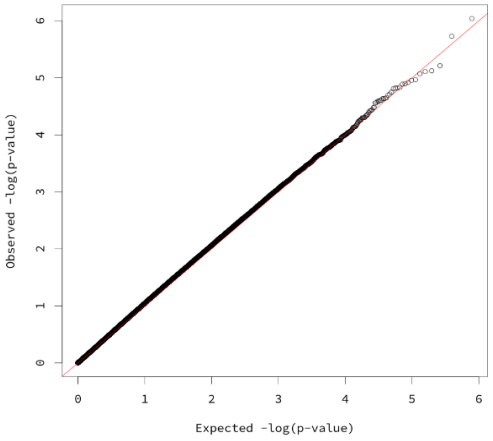
Table SM 2-9: Univariate associations between CpG sites[1] and each TOD trait in HyperGEN

	CpG site	eGFR			UACR			LVMI			RWT		
		Beta	**SE**	***p*-value**	**Beta**	**SE**	***p*-value**	**Beta**	**SE**	***p*-value**	**Beta**	**SE**	***p*-value**
Model 1	cg04816311	0.473	0.131	**3.07E-04**	0.141	0.139	0.32	0.402	0.004	**3.33E-04**	0.422	0.137	**2.13E-03**
	cg09155024	0.047	0.126	0.71	-0.162	0.134	0.23	-0.168	0.133	0.21	-0.041	0.133	0.76
	cg10254690	0.181	0.110	9.79E-02	-0.147	0.115	0.20	-0.153	0.114	0.18	-0.187	0.115	0.10
Model 2	cg04816311	0.496	0.132	**1.78E-04**	0.051	0.131	0.69	0.244	0.123	**4.63E-02**	0.349	0.133	**8.62E-03**
	cg09155024	0.041	0.127	0.75	-0.146	0.125	0.24	-0.153	0.116	0.19	-0.029	0.127	0.82
	cg10254690	0.172	0.111	0.12	-0.144	0.108	0.18	-0.086	0.101	0.39	-0.170	0.111	0.12
Model 3	cg04816311	0.502	0.135	**2.08E-04**	-0.008	0.130	0.54	0.070	0.113	0.54	0.337	0.135	**1.27E-02**
	cg09155024	0.057	0.127	0.65	-0.085	0.122	0.48	-0.113	0.105	0.28	-0.015	0.127	0.90
	cg10254690	0.170	0.111	0.13	-0.120	0.105	0.26	-0.051	0.091	0.58	-0.175	0.110	0.11

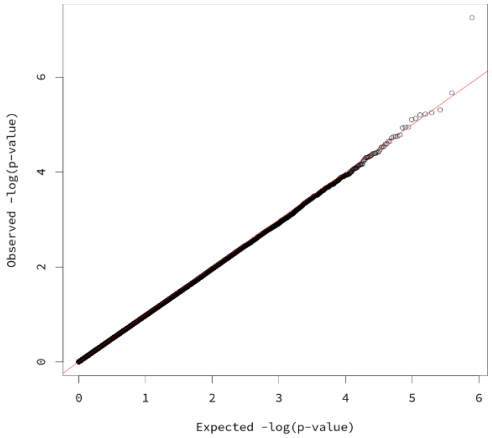
Model 1 is adjusted for age at phase II, sex, smoking status, 4 genetic principal components, and family (random)
Model 2 is adjusted for Model 1 variables plus systolic blood pressure, diastolic blood pressure and antihypertensive medication use
Model 3 is adjusted for Model 2 variables plus body mass index and type 2 diabetes status
Results significant at *p*-value <0.05 are in bold font
[1] Only CpG sites that were statistically significant in multivariate Model 1 in GENOA were included in this replication analysis

Figure SM 2-1: QQ plots of the association between DNA methylation and TOD using univariate and multivariate model 1 in GENOA African Americans

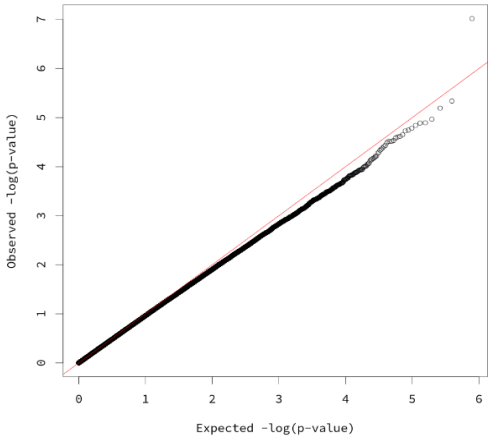
A: Estimated glomerular filtration rate (eGFR)



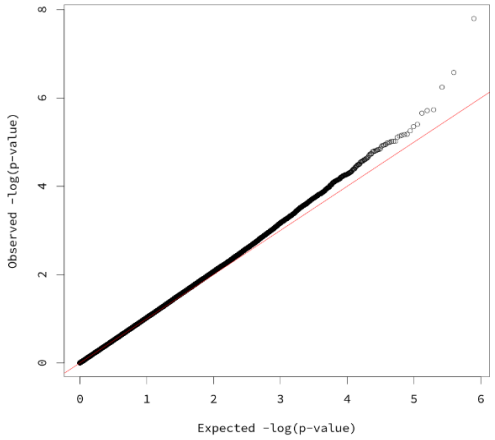
B: Urinary albumin-creatinine ratio (UACR)



C: Relative wall thickness (RWT)



D: Left ventricular mass index (LVMI)



E: Multivariate model (eGFR, UACR, RWT, LVMI)

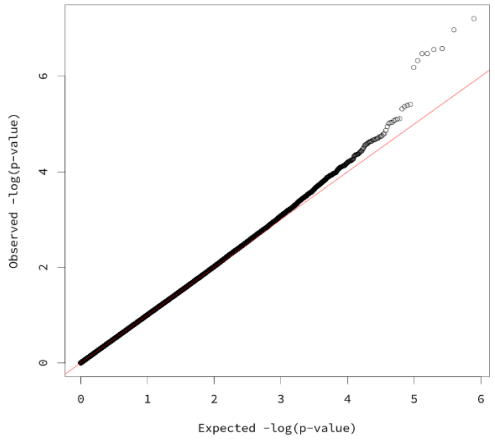
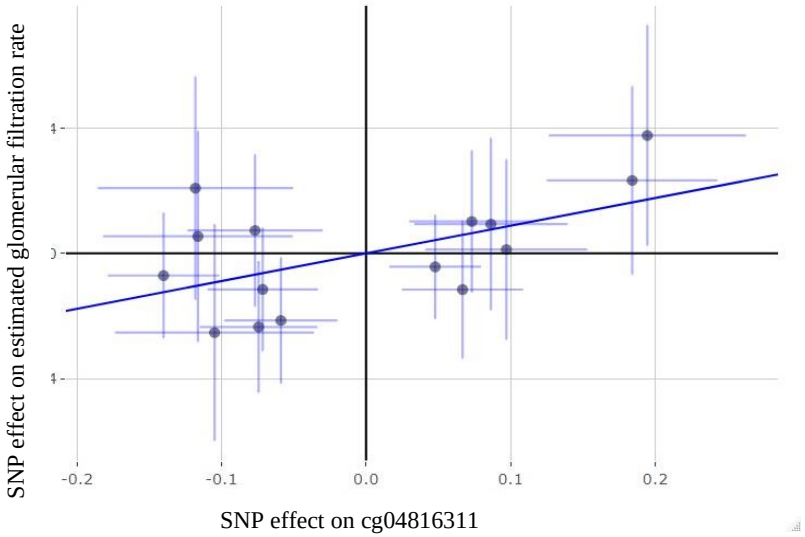
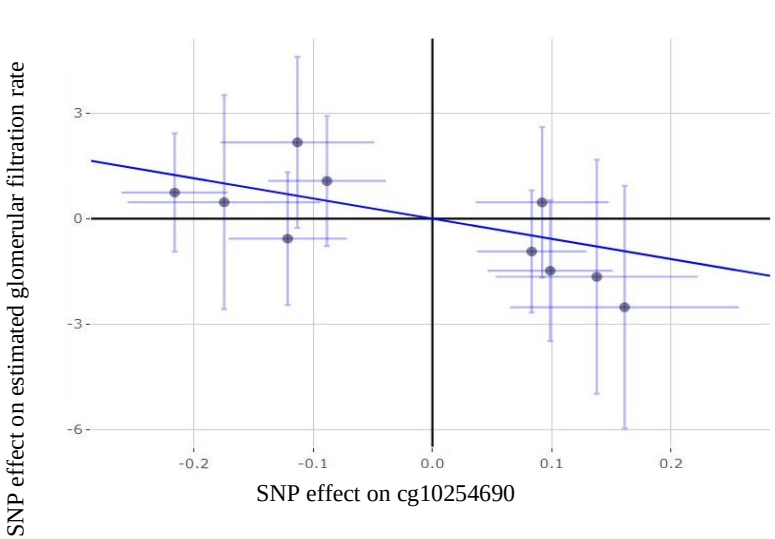


Figure SM 2-2: Mendelian randomization (MR) scatter plots of estimated glomerular filtration rate (A-C). Causal estimates are based on inverse-variance weighted (IVW) estimates using cis-SNPs within ± 1Mb of the corresponding CpG site

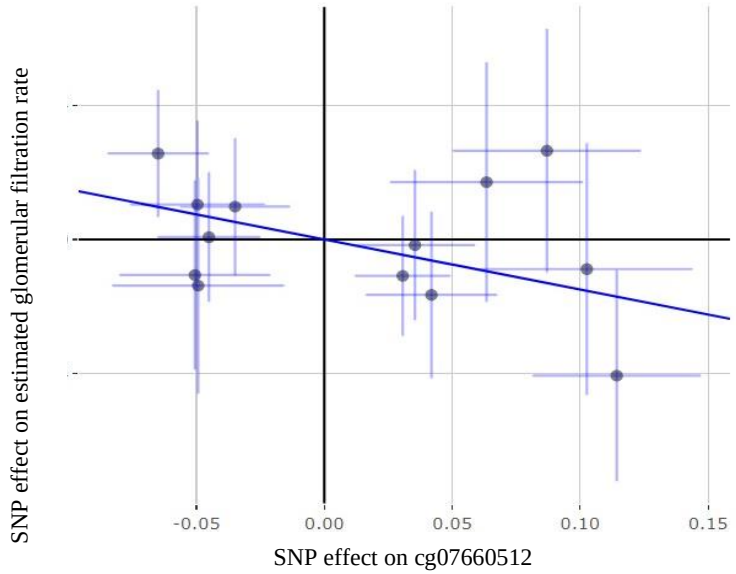
A



B



C



Chapter 3. Epigenetic Age Acceleration in African Americans Associates with Cardiometabolic Risk Factors and Clinical Cardiovascular Disease Risk Scores

3.1 Abstract

Cardiovascular disease (CVD) is the leading cause of mortality among US adults. African Americans have higher burden of CVD morbidity and mortality compared to any other racial group. Identifying biomarkers for clinical risk prediction of CVD offers an opportunity for precision prevention and earlier intervention. Using linear mixed models, we investigated the cross-sectional association between four measures of epigenetic age acceleration (intrinsic (IEAA), extrinsic (EEAA), PhenoAge (PhenoAA), and GrimAge (GrimAA)) and ten cardiometabolic markers of hypertension, insulin resistance, and dyslipidemia in 1,100 primarily hypertensive African Americans from sibships in the Genetic Epidemiology Network of Arteriopathy (GENOA). We then assessed the association between epigenetic age acceleration and time to self-reported incident CVD using frailty hazard models and investigated CVD risk prediction improvement compared to models with clinical risk scores (Framingham risk score (FRS) and the atherosclerotic cardiovascular disease (ASCVD) risk equation). After adjusting for sex and chronological age, increased epigenetic age acceleration was associated with higher systolic blood pressure (IEAA), higher pulse pressure (EEAA and GrimAA), higher fasting glucose (PhenoAA and GrimAA), higher fasting insulin (EEAA), lower low density cholesterol (GrimAA), and higher triglycerides (GrimAA). A five year increase in GrimAA was associated with CVD incidence with a hazard ratio of 1.54 (95% CI 1.22–2.01) and remained significant after adjusting for CVD risk factors. The addition of GrimAA to risk score models improved

model fit using likelihood ratio tests (P = 0.013 for FRS and P = 0.008 for ASCVD), but did not improve C statistics (P > 0.05). Net reclassification index (NRI) showed small but significant improvement in reassignment of risk categories with the addition of GrimAA to FRS (NRI: 0.055, 95% CI 0.040–0.071) and the ASCVD equation (NRI: 0.029, 95% CI 0.006–0.064). Epigenetic age acceleration measures are associated with traditional CVD risk factors in an African American cohort with a high prevalence of hypertension. GrimAA was associated with CVD incidence, and slightly improved prediction of CVD events over clinical risk scores.

3.2 Introduction

Cardiovascular disease (CVD) is the leading cause of mortality among US adults.[1] African Americans have the highest CVD morbidity and mortality burden, a trend which has been consistent over the last few decades.[2] Underlying this higher CVD prevalence is a greater burden of a number of risk factors, including hypertension, type 2 diabetes, and obesity.[3-5] Yet a focus on established risk factors and their management has failed to fully reduce the excess CVD burden among African Americans. Identification of novel biomarkers that go beyond traditional ones may help better identify at-risk individuals, advance precision medicine, and inform efforts to reduce CVD burden.

Epigenetic aging, based on DNA methylation (DNAm) at CpG dinucleotides, is a novel measure of biological aging that offers the opportunity to identify molecular markers of disease risk. The first generation of epigenetic aging measures, the HorvathAge [6] and HannumAge [7] epigenetic clocks, were trained on chronological age and are estimated based on 363 and 71 CpG sites selected using elastic net regression modeling, respectively. HorvathAge was trained using multi-tissue samples from children and adults, while HannumAge was trained using a single tissue (whole blood) from adults. Modified versions of these two measures were later derived to account for confounding by blood cell composition: intrinsic epigenetic age acceleration (IEAA) based on HorvathAge explicitly adjusts for blood cell composition, and extrinsic epigenetic age acceleration (EEAA) based on the HannumAge is a composite measure that includes a weighted average of cell counts known to vary strongly with age.[8] PhenoAge, a more recent measure based on whole blood from adults, was estimated using 513 CpG sites and was trained on a composite clinical measure of phenotypic age that is based on chronological age and nine biomarkers including albumin, creatinine, serum glucose, white blood cell counts.[9] The

biomarkers were selected for their association with the hazard of mortality using a Cox penalized regression model. GrimAge is another recent measure constructed based on the linear combination of 1030 CpG sites that represent DNAm-based surrogate measures for a number of plasma proteins and smoking pack-years.[10] Like PhenoAge, it is based on whole blood from adults. In addition to chronological age, both PhenoAge and GrimAge account for physiological dysfunction among individuals of the same chronological age in their selection of CpGs. For each of these measures, epigenetic age *acceleration* is defined as the discrepancy between epigenetic age and chronological age. These four epigenetic age acceleration measures are hypothesized to be capturing different aspects of aging and are based mostly on unique CpG sites.[11]

A growing body of literature has examined the association between epigenetic age acceleration and CVD and its risk factors, such as blood pressure and lipids, but the overall evidence remains inconclusive likely due to heterogeneity in study design, the specific outcomes examined, and the epigenetic aging measures used.[12-21] PhenoAge and GrimAge are more recently developed measures, and so validation of their associations and comparisons to the first-generation measures are in early stages. Two recent studies in participants of European ancestry show that GrimAA outperforms other measures in its association with CVD incidence after adjusting for CVD risk factors,[21,22] and additional studies report similar findings with all-cause mortality.[21-24] Yet it is unclear whether epigenetic age acceleration measures could be used to improve CVD prediction in a clinical setting.

In this study, we investigated the relationship between four epigenetic age acceleration measures and ten cardiometabolic markers of hypertension, insulin resistance, and dyslipidemia in 1,100 primarily hypertensive African Americans in the Genetic Epidemiology Network of

Arteriopathy (GENOA) study. We additionally assessed the association between four epigenetic age acceleration measures and incident CVD. Finally, we examined whether epigenetic age acceleration measures can improve the predictive accuracy of two clinically-used CVD risk scores: the Framingham risk score (FRS)[25] and the more recently developed atherosclerotic cardiovascular disease (ASCVD) risk equation.[26]

3.3 Methods

Study sample

Genetic Epidemiology Network of Arteriopathy (GENOA) is a community-based study in Rochester, MN and Jackson, MS that was established to identify genes influencing blood pressure and development of target organ disease.[44] In the first phase of GENOA (Phase I: 1996 – 2001), sibships with at least two adults with clinically diagnosed essential hypertension before age 60 were recruited, and all siblings in the sibship were invited to participate regardless of hypertension status. Exclusion criteria included secondary hypertension, alcoholism or drug abuse, pregnancy, insulin-dependent diabetes mellitus, or active malignancy. In Phase I (i.e. baseline), a total of 1,583 non-Hispanic whites (Rochester, MN) and 1,854 African Americans (Jackson, MS) were enrolled. In the second phase (Phase II: 2001 – 2005), all participants were invited for a second examination. Eighty percent of African Americans (N = 1,482) and 75% of non-Hispanic whites (N = 1,213) from Phase 1 returned. At Phase III (2009-2011), 752 African Americans returned for a third examination. This study includes African American participants who had their DNA methylation profiles measured in whole blood samples collected at Phase I. Demographic information, medical history, clinical characteristics, lifestyle factors, and blood samples were collected in each phase. Written informed consent was obtained from all

participants and approval was granted by participating institutional review boards (University of Michigan, University of Mississippi Medical Center, and Mayo Clinic).

DNA methylation and epigenetic age acceleration measures

The methods of DNA methylation processing have been previously described.[45] Briefly, genomic DNA from 1,106 African American participants from Phase I and 304 from Phase II was extracted from stored peripheral blood leukocytes using AutoGen FlexStar (AutoGen, Holliston, MA). Sex mismatches and outliers were excluded using the shinyMethyl R package,[46] probes with detection *P*-value < 10^{-16} were considered to be successfully detected[47] and both samples and probes that failed a detection rate of at least 10% were removed. The Noob method was used for individual background and dye-bias normalization[48] and the Regression on Correlated Probes method was used to adjust for the probe-type bias in the data.[49] White blood cell type proportions within the blood sample were estimated using Houseman's method.[50]

After quality control, a total of 1,100 samples from Phase I and 294 from Phase II were available for assessment of epigenetic age acceleration; however, only Phase I measures were included in this study. Methylation beta values were uploaded to the online Horvath epigenetic age calculator to calculate DNAm Age.[51] Four measures of epigenetic age (HannumAge, HorvathAge, PhenoAge and GrimAge) were estimated for the current analysis. IEAA based on the Horvath measure, are the regression residuals after adjusting for chronological age and blood cell count.[6-8] EEAA was calculated using the Hannum epigenetic age after incorporating weighted averages of three white blood cell types (naïve cytotoxic T cells, exhausted cytotoxic T cells, and plasmablasts).[7,8] PhenoAge and GrimAge are considered to be extrinsic measures of aging because they capture both cell intrinsic methylation changes as well as extracellular changes in blood cell composition.[9,10,31] We also estimated 7 DNAm based surrogate plasma

proteins (adrenomedullin (ADM), beta-2-microglobulin, cystatin C, GDF-15, leptin, plasminogen activator inhibitor antigen type 1 (PAI-1), tissue inhibitor metalloproteinases 1 (TIMP-1)), and smoking pack-years that comprise GrimAge in order to identify individual components that may drive associations or that are more predictive than the overall measure itself.[10]

Cardiometabolic risk factors

Resting systolic blood pressure (SBP) and diastolic blood pressure (DBP) were measured by a random zero sphygmomanometer and a cuff appropriate for arm size. The second and third of three readings, after the participant sat for at least 5 minutes, were averaged for analysis.[52] Mean arterial pressure (MAP) was calculated as the weighted average of SBP and DBP (1/3*SBP + 2/3*DBP) and pulse pressure (PP) was calculated as the difference between SBP and DBP (SBP – DBP). Information on current anti-hypertensive medication use and lipid-lowering statin medication use were collected. Hypertension was defined as SBP ≥ 140 mmHg, DBP ≥ 90 mmHg, or antihypertensive medication use. Smoking was categorized as current, former, or never. Blood glucose and insulin levels were measured for participants fasting for at least 10 hours. Serum total cholesterol, HDL-C, and triglycerides (TGs) were measured by standard enzymatic methods on a Hitcahi 911 Chemistry Analyzer (Roche Diagnostics, Indianapolis, IN). LDL-C was calculated using the Friedewald formula [LDL in mg/dl = TC – HDL-C – (TGs/5)][53] and individuals with triglycerides levels ≥ 400 mg/dl were excluded from LDL-C association analysis. Type 2 diabetes was defined as fasting serum glucose concentration > 126 mg/dl or self-reported physician-diagnosed diabetes and current medication use (insulin or hypoglycemic agents). Educational attainment was based on self-reported years of education. Alcohol consumption was calculated as the number of drinks per week based on aggregated

measurements of a variety of alcoholic drinks. Height was measured by stadiometer and weight by electronic balance and body mass index (BMI) was calculated as weight in kilograms divided by the square of height in meters.

CVD events and risk scores

Framingham risk score (FRS), predicting the 10-year risk of a CVD (defined as coronary death, myocardial infarction, coronary insufficiency, angina, ischemic stroke, hemorrhagic stroke, transient ischemic attack, peripheral artery disease, and heart failure), was estimated using age, sex, total cholesterol, HDL-C, anti-hypertensive medication use, SBP, smoking status, and type 2 diabetes status after limiting the sample to individuals aged between 30 – 74 years (N= 945, events = 69, person-years = 7874.9).[25] While FRS was developed in participants of European ancestry, ASCVD was developed using a pooled community-based population cohort with a higher proportion of African Americans and has been validated for prediction of clinical events in more race/ethnically-diverse cohorts. [26] It is based on the same covariates as FRS and predicts the 10-year risk of developing a first ASCVD event, defined as nonfatal myocardial infarction or coronary heart disease death or fatal or nonfatal stroke. Using sex- and race-specific parameters, we estimated the ASCVD risk equation, after limiting the sample those between the ages of 20 – 79 years (N = 988, events = 71, person-years = 8115.5). Risk scores were modeled as continuous variables and as categorical predictors where they were used to group participants into low risk (10-year risk $\leq$ 7.5%) or high risk (> 7.5%) groups.[26]

Information about CVD, as reported by participants, was collected at baseline and at each subsequent follow-up phase. An event was defined as myocardial infarction, coronary revascularization (stenting, balloon angioplasty, or coronary artery bypass grafting), stroke (ischemic or hemorrhagic events), or surgical carotid artery revascularization. Participants only

reported the year of CVD events. Time to CVD was modeled by setting the CVD event time at the mid-point of the year in which participants reported an event. For censored participants, follow-up time was set at the time point they were last interviewed.

Statistical analysis

Outliers at more than 5 standard deviations from the mean were removed for the cardiometabolic outcomes and the epigenetic age acceleration measures. Glucose, insulin, HDL-C and triglycerides were natural log-transformed as ln(measure +1). Linear mixed models that account for familial relatedness were used to assess the cross-sectional univariate association between each epigenetic age acceleration measure (predictor) and each cardiometabolic risk factors (outcome) at Phase I. Base models were adjusted for age and sex (Model 1). In subsequent models, we additionally adjusted for education, smoking status, body mass index, and alcohol consumption (Model 2) and white blood cell counts for PhenoAA and GrimAA to assess confounding by changes in cell composition (Model 3). For LDL-C and triglycerides, we performed sensitivity analyses excluding participants who were not fasting for at least 10 hours before the blood draw.

After excluding participants with baseline CVD events, associations with time to first CVD event (incident CVD) were assessed using Cox proportional hazards models, and hazard ratios (HR) and 95% confidence intervals were estimated. A simple random effects (frailty) term in the Cox model was included to take into account family structure.[54] We next adjusted for traditional CVD risk factors (age, sex, education, body mass index, alcohol consumption, total cholesterol, HDL-C, anti-hypertensive medication use, SBP, smoking status, and type 2 diabetes status). Finally, we adjusted for white blood cell counts. The proportional hazard assumption was evaluated using Schoenfeld residuals, and all models satisfied the assumption. As a sensitivity

analysis, we additionally modeled time to CVD using interval censoring using the iceReg package.[55]

Likelihood ratio (LR) testing of nested models (addition of epigenetic age acceleration to a base model with either FRS or ASCVD) was used to assess improvement in model fit. For measures with $P < 0.05$, we assessed improvement in risk prediction of incident CVD by adding the epigenetic age acceleration measure to the base model with the clinical risk scores. We assessed the improvement in CVD risk prediction using C-statistics computed from Cox proportional hazards models of time to CVD events and risk scores as continuous predictors.[56] We additionally used the net reclassification index (NRI) to assess net improvement in reassignment of the risk categories.[28] Categorized CVD risk scores were used in the base model and improvement in risk reassignment was then assessed after the addition of epigenetic acceleration measures. For this analysis, we also examined the associations excluding individuals taking lipid-lowering statin medications (N = 40).

Statistical tests were two-sided and a *P* value of < 0.05 was considered nominally significant. We also applied a Bonferroni threshold for statistical significance (0.05/10 = adjusted $P < 0.005$) to account for multiple testing in assessing the association between epigenetic acceleration measures and the 10 cardiometabolic traits. For NRI, bootstrapping (10,000 iterations) was used to compute 95% confidence intervals, and an empirical $P < 0.05$ was considered significant. Analyses were performed using R (Version 4.0.2)[57] and the following packages: lme4,[58] survival,[59,60] nricens, and DescTools.

3.4 Results

Descriptive statistics

Baseline characteristics of the participants are shown in **Table 1**. The 1,100 participants from 530 sibships had a mean age of 57.1 years, and 71% were women. About 60% were never smokers and mean alcohol consumption was 0.66 drinks per week. About 70% of the participants had hypertension and 20% had type 2 diabetes at baseline. At baseline, 91 participants had prevalent CVD and another 72 developed incident CVD over 8,161 person-years of follow-up. The mean Framingham risk score (FRS) was 14.4% and the mean of the atherosclerotic cardiovascular disease (ASCVD) risk equation was 11.6%. FRS and ASCVD were positively and significantly correlated (r = 0.94, $P = 2.2 \times 10^{-16}$). **Supplemental Figure 1** shows the scatterplots for each of the DNAm age measures with chronological age. As previously reported, all of the DNAm age measures were strongly and significantly correlated with chronological age (all r > 0.8, **Supplemental Table 1)**.[27] The means of the age acceleration measures ranged between 0.11 years for GrimAA and 0.38 years for PhenoAA. The acceleration measures were not strongly correlated with each other (r range: 0.19 – 0.50), nor where they correlated with chronological age (**Supplemental Table 1, Supplemental Figure 2).**

Association between epigenetic age acceleration and cardiometabolic risk factors

Table 2 shows the regression results from linear mixed models for the univariate associations between the epigenetic age acceleration measures and cardiometabolic risk factors with beta coefficients for 1-year increase in epigenetic age acceleration after adjusting for age, sex, and familial relatedness. Effect sizes are also reported below per 5-year increase, which is equivalent to approximately one standard deviation of the epigenetic acceleration measures. At $P < 0.05$, IEAA, EEAA, and PhenoAA were each associated with four cardiometabolic risk factors, while GrimAA was associated with five.

IEAA was associated with higher systolic blood pressure (SBP), and both EEAA and GrimAA were associated with higher pulse pressure after accounting for multiple testing. A 5-year increase in IEAA was associated with an approximately 1.85 mmHg increase in SBP (95% CI 0.55–3.14). For EEAA and GrimAA, a 5-year increase was associated with a 1.20 mmHg (95% CI 0.41–2.0) and a 1.75 mmHg (95% CI 0.73–2.72) increase in pulse pressure, respectively.

GrimAA was associated with higher fasting glucose levels and EEAA was associated with higher fasting insulin levels after accounting for multiple testing. A 5-year increase in GrimAA was associated with a 4.08% increase (95% CI 1.51%–6.18%) in glucose levels. EEAA was the only measure associated with insulin, where a 5-year increase was associated with a 5.13% increase (95% CI 2.53%–10.5%).

Only GrimAA was associated with any of the lipid traits examined after accounting for multiple testing. A 5-year increase in GrimAA was associated with a 3.85 mg/dl (95% CI -6.50 to -1.20) decrease in low density lipoprotein (LDL-C) and a 5.13% (95% CI 3.05%–8.87%) increase in triglyceride levels. The associations between the lipid measures and GrimAA remained significant after excluding participants who were not fasting for at least 10 hours (β = -1.01, P = 0.001, N = 863 for LDL-C and β = 0.011, P = 0.001, N = 881 for triglycerides). **Supplemental Table 2** shows the adjusted linear mixed model regression results for associations significant at $P < 0.05$ from **Table 2**. Although some of the nominally significant associations fully attenuated after adjusting for education, smoking status, body mass index (BMI), and alcohol consumption (Model 2), all of the associations that were significant after multiple testing in the base model (Bonferroni-corrected $P < 0.05$) remained significant at $P < 0.05$. When we further adjusted PhenoAA and GrimAA associations for white blood cell counts (Model 3), all of

the associations became less significant, and the associations between PhenoAA and glucose and GrimAA and pulse pressure were fully attenuated ($P = 0.100$ and $P = 0.054$, respectively). GrimAA, however, remained significantly associated with glucose, LDL-C, and triglycerides ($P < 0.05$).

Epigenetic age acceleration associations with clinical cardiovascular risk scores and incident CVD

All epigenetic age acceleration measures were significantly associated with the FRS and the ASCVD risk equation, except for IEAA with FRS. The effect estimates from the linear mixed models were in the expected direction with increased biological aging associated with an increase in the predicted 10-year risk of CVD (**Table 3)**. The largest effect estimate was observed for GrimAA, where a 5-year increase in epigenetic age acceleration was associated with a 2.9% (95% CI 2.2%–3.6%) and a 2.2% (95% CI 1.7%–2.8%) increase in the 10-year CVD risk using FRS and ASCVD equations, respectively.

When we examined whether epigenetic acceleration measures were associated with time to first CVD event, a similar trend emerged with GrimAA showing the only significant association with incidence of CVD events. **Table 4** shows the hazard ratios (HR) and 95% confidence intervals estimated from Cox proportional hazards models with a frailty term for the associations of the four epigenetic age acceleration measures with incident CVD. A 5-year increase in GrimAA was associated with a HR of 1.54 (95% CI 1.22–2.01) in the base model (adjusted for age, sex, and family structure). Further adjusting for traditional CVD risk factors (education, alcohol consumption, body mass index, total cholesterol, HDL-C, anti-hypertensive medication use, SBP, smoking status, and type 2 diabetes status) only slightly attenuated the association (HR per 5-year increase in GrimAA: 1.47, 95% CI 1.05–2.01, $P = 0.024$). Additionally adjusting for white blood cell counts did not attenuate the association (HR per 5-

year increase in GrimAA: 1.54, 95% CI 1.10–2.19, *P* = 0.01). Findings were similar when time to CVD was modeled using interval censoring. Last, we examined the association between the individual components comprising GrimAge and incident CVD to identify components that may be driving the association between GrimAA and CVD or outperform the overall GrimAA measure itself **(Supplemental Table 3).** Adrenomedullin (ADM), smoking pack-years, and plasminogen activator inhibitor antigen type 1 (PAI-1) were associated with incident CVD (*P* < 0.05) in the base model after further adjustment for white blood cell types, with HRs only slightly lower than that of GrimAA. **Figure 1** shows the box plots of the standardized DNAm surrogate measures of the 7 plasma protein and smoking pack-years in GrimAge by incident CVD status. The means of ADM, smoking pack-years, and PAI-1 were higher among those with incident CVD.

Evaluating the performance of epigenetic age acceleration measures in CVD prediction

Likelihood ratio (LR) tests of nested models showed that GrimAA improved model fit when added to a model with age, sex, and FRS (HR per 1-year increase in GrimAA: 1.07, 95% CI 1.02–1.13, *P* for LR test of model fit = 0.013) or the ASCVD equation (HR per 1-year increase in GrimAA: 1.08, 95% CI 1.02–1.13, *P* for LR test = 0.008) **(Table 5)**. None of the other age acceleration measures improved model fit.

Since GrimAA improved model fit, we next evaluated whether it could improve CVD risk prediction compared with the FRS and ASCVD risk equations using the C-statistic and the net reclassification index (NRI). The C-statistic is the probability that a randomly selected participant who experienced the CVD event will have a higher predicted probability of having the event compared to a randomly selected participant who did not experience the event. **Table 6** shows the C-statistics for the performance of GrimAA in predicting incident CVD. The addition

of GrimAA to a model with each risk score increased the C-statistic to 0.698 for FRS and to 0.685 for the ASCVD risk equation (all $P > 0.05$). **Supplementary Figure 3** shows the receiver operator characteristic (ROC) curves for the risk scores before and after adding GrimAA to the model.

Next, we compared the classification of CVD events with and without GrimAA using the net reclassification index (NRI). The NRI is an index of the net improvement in reassignment of the risk categories.[28] The FRS categorized 36.5% of the GENOA cohort as low risk (≤ 7.5%) while the ASCVD equation categorized 47.2% of the cohort as low risk. Net reclassification for CVD was small but significant with the addition of GrimAA to a model of age, sex, and FRS (NRI: 0.055, 95% CI 0.040–0.071, $P < 0.0001$), and for a similar model of age, sex, and the ASCVD equation (NRI: 0.029, 95% CI 0.006–0.064, $P = 0.0011$). **Supplemental Figure 4** shows the reclassification tables of predicted CVD based on the NRI for models with FRS or ASCVD and GrimAA. The improvement in risk prediction was driven by the classification of CVD nonevents as low risk.

When we excluded participants taking lipid-lowering statin medications, improvement in risk prediction in models with GrimAA was almost identical to that of the full sample. GrimAA remained associated with incident CVD after adding it to a base model with FRS and ASCVD (HR = 1.08 in both models, $P = 0.004$ and $P = 0.003$, respectively). As in the full sample, addition of GrimAA to a model with FRS or the ASCVD equation increased the C-statistics, but the increases were not significant at $P < 0.05$. The NRIs with the addition of GrimAA to the risk scores were also similar (NRI = 0.052 for FRS, NRI = 0.030 for the ASCVD equation).

3.5 Discussion

In this study of primarily hypertensive African American participants from GENOA, we showed that increased biological aging is associated with a worse cardiometabolic risk profile, although the associations with specific cardiometabolic risk factors varied across the age acceleration measures. All of the epigenetic acceleration measures were correlated with risk of CVD onset as modeled by clinical CVD risk scores (FRS and ASCVD equation). GrimAA outperformed IEAA, EEAA, and PhenoAA in predicting CVD incidence, and the association remained significant after adjusting for traditional CVD risk factors. The addition of GrimAA to FRS or ASCVD did not improve the C-statistics of CVD risk prediction; however, the NRIs showed small but significant improvement in the reassignment of risk categories.

Differences in the cardiometabolic risk factor and CVD incidence associations among the various epigenetic clocks may be attributed to a number of factors. IEAA, EEAA, and PhenoAA share only between 5 and 36 CpG sites.[9] Information on the CpGs included in GrimAge are not publicly available, so we cannot assess how many CpG sites this measure shares with the other three. In addition to differences in training algorithms (chronological age for IEAA and EEAA vs. aging correlates and outcomes for PhenoAA and GrimAA), the second generation of epigenetic measures (PhenoAA and GrimAA) were trained using longitudinal data.[29] This is particularly relevant for studies assessing their prediction of aging related outcomes. The use of cross-sectional training data may have biased the algorithm as individuals with accelerated aging rates will have a higher morality burden and may have been selected out from the training samples.[29,30] Nevertheless, an analysis of the transcriptional profiles of IEAA, EEAA, and PhenoAge shows that they have relatively similar transcriptional signatures.[11]

Our study found cross-sectional associations between epigenetic age acceleration and a number of cardiometabolic risk factors. Out of the 10 cardiometabolic risk factors examined in the base model, GrimAA was associated with 4 measures, EEAA with 2 measures, and IEAA and PhenoAA with 1 measure after accounting for multiple testing. Some of the associations were unique to one specific cardiometabolic feature such as the association between GrimAA and lipid traits. For the significant associations between the acceleration measures and cardiometabolic risk factors, the effect directions were as expected with the exception of the association between GrimAA and LDL-C. Higher epigenetic acceleration, indicative of tissue aging faster than expected by chronological age, was associated with worsening outcomes as measured by cardiometabolic risk factors. Increased tissue aging in blood is accompanied by changes in cell-type composition.[31] However, the associations between GrimAA and cardiometabolic risk factors were not attenuated after adjusting for blood cell composition, with the exception of the association between GrimAA and pulse pressure. This suggests that the associations observed are not due to age-related changes in blood cell composition and that GrimAA is capturing cell-intrinsic properties or innate changes related to aging rather than changes in immune cell composition.

In our study, IEAA was associated with SBP, and both EEAA and GrimAA were associated with pulse pressure. Previous studies on the association between IEAA and EEAA and cardiometabolic phenotypes show inconsistent findings.[13,20,32,33] A previous study of IEAA and EEAA in the Women's Health initiative (WHI) found no associations with systolic or diastolic blood pressure after adjusting for diet and metabolic syndrome symptoms.[32] However, in a smaller sample of African Americans from the Bogalusa Heart Study (N=288), both IEAA and EEAA were associated with hypertension.[20] Another study of approximately 5,000 individuals

from the Generation Scotland: Scottish Family Health Study found evidence of an association between EEAA and high blood pressure, but not IEAA.[34] A previous analysis in GENOA found no association between blood pressure measured at Phase II and IEAA or EEAA,[35] although significant associations were detected in this study using concurrently measured blood pressure (Phase I). PhenoAA and GrimAA were more recently developed, so fewer studies have assessed their associations with cardiometabolic risk factors. However, in WHI, both PhenoAA and GrimAA were significantly correlated with SBP but not DBP.[9,10]

In this study, we also found evidence of associations between PhenoAA and GrimAA and glucose, and EEAA and insulin. GrimAA was the only measure associated with any of the lipid traits. In WHI, no associations between measures of insulin resistance and dyslipidemia (HDL-C and triglycerides) were detected with IEAA or EEAA, except for an association between EEAA and triglycerides ($\beta = 0.004$, P value = 0.04).[20] However, other studies have reported an inverse association between fasting HDL-C levels and EEAA[13] and IEAA.[34] Cross-sectional examination of WHI revealed correlations between both PhenoAA and GrimAA and insulin, glucose, triglycerides, and HDL-C.[9] PhenoAA, but not GrimAA, was also correlated with LDL-C.[10] As in our study, GrimAA was associated with lower total cholesterol and LDL-C in a cross-sectional analyses of 709 individuals from the Lothian Birth Cohort.[36] Another study in the Methyl Epigenome Network Association and a Spanish cohort found significant correlations between GrimAA and glucose levels, HDL-C, and triglycerides.[37]

In our analyses, higher GrimAA was the only measure associated with CVD incidence in GENOA African Americans independent of CVD risk factors. Adjustment for white blood cell counts did not attenuate the association. Neither EEAA nor IEAA were associated with incident coronary heart disease in WHI.[20] However, among Black participants from the Atherosclerosis

Risk in Communities (ARIC) study, epigenetic age acceleration based on the Horvath and Hannum measures were both associated with increased hazard of fatal coronary heart disease (HR: 1.17, 95% CI 1.02–1.33 and HR: 1.22, 95% CI 1.04–1.44, respectively).[17] A German case-cohort study reported an increase in the hazard of cardiovascular mortality associated with Horvath age acceleration,[38] while a study in the Melbourne Collaborative Cohort found no association with the Horvath or Hannum measures.[16] Increased PhenoAA, but not HorvathAA, was also associated with increased risk of cardiovascular mortality in 500 males from the US Normative Aging Study.[39]

Our findings are in line with the literature in cohorts of European ancestry showing that GrimAA outperforms other measures in its association with CVD incidence.[21,22] The effect size of GrimAA on CVD incidence appears to be remarkably similar across studies in European ancestry, and similar to our estimate in African Americans. Comparing the same four measures of epigenetic acceleration that we investigated, Hillary et al. found that over thirteen years of follow-up, GrimAA outperforms the other measures in terms of its association with incidence of heart disease (HR: 1.41, 95% CI 1.18–1.68, per 1 SD).[21] Wang et al. found that a 1 SD increase in GrimAA was associated with elevated risk of myocardial infarction (HR: 1.44, 95% CI 1.16–1.79) and stroke (HR: 1.42, 95% CI 1.06–1.91) in a study of elderly participants from the Normative Ageing Study and the Cooperative Health Research in the Region Augsburg (KORA) study.[22]

To our knowledge, no previous study has assessed the performance of the epigenetic age acceleration measures in improving the predictive accuracy of clinical risk scores of CVD. GrimAA appears to marginally improve prediction of CVD events beyond traditional risk factors when assessed using NRI but not using changes in the area of the ROC curves. The net gains in

risk prediction were mostly due to the down-classification of non-cases as low risk, with fewer up-classification of non-cases as high risk. More studies are needed to validate and replicate these findings and assess their utility in clinical settings. In addition, each of our clinical score measures predict a different set of outcomes. FRS is a predictor of total CVD, including coronary death, peripheral artery disease and heart failure. The ASCVD, on the other hand, predicts outcomes related to coronary heart disease and stroke. Further studies are needed to fully investigate the predictive performance of GrimAA in relation to different cardiovascular outcomes. However, GrimAA may be a promising biomarker since it is a composite measure of multiple plasma proteins, some of which have been shown to be independent biomarkers that can improve CVD prediction.[40-43] Additionally, for some of the components of GrimAge (PAI-1, TIMP-1, and cystatin C), DNAm-based surrogates were found to outperform the observed biomarkers.[10] Lu et al. found that DNAm smoking pack-years was a more significant predictor of lifespan than self-reported smoking and that it predicted mortality even among non-smokers. This may be related to errors in self-reporting or because DNAm pack-years may capture intrinsic variation across individuals with lasting biological damage related to smoking.[10]

Our study has a number of limitations. Our findings were based on self-reported events, with only the year of the event reported, which could be subject to recall bias. We also note that there was loss to follow-up between baseline and Phases II and III. Those lost to follow-up between baseline and Phase III were 1.63 years older and had 3.5 mmHg higher systolic blood pressure on average ($P = 0.014$). Additionally, individuals lost to follow up had higher GrimAA, FRS, and ASCVD risk scores (all $P < 0.05$). This indicates that participants at greater risk of CVD events were more likely to be lost to follow-up. Another limitation is that although we adjusted for a number of important confounders, we lacked information on dietary data in

GENOA African Americans. Finally, our sample is predominantly hypertensive and has an overrepresentation of women, so our findings may not be representative of other cohorts. A strength is that our study provides insights on the association between four different epigenetic aging measures and cardiometabolic risk factors and CVD in a relatively large cohort of older African Americans. In addition, we also explored improvement of CVD risk prediction by incorporating epigenetic aging measures in clinical risk equations and investigated potential molecular drivers of the observed associations.

Epigenetic information is an important molecular readout of lifetime exposures. We have shown that epigenetic aging measures are associated with some cardiometabolic risk factors in this relatively large cohort of African Americans. GrimAge acceleration was the only measure associated with CVD incidence after adjusting for CVD risk factors. Further studies are needed to replicate and further investigate potential improvement of clinical risk prediction using GrimAge acceleration.

3.6 References

1. Shah NS, Lloyd-Jones DM, O'Flaherty M, Capewell S, Kershaw KN, Carnethon M, et al. Trends in Cardiometabolic Mortality in the United States, 1999-2017. JAMA. 2019;322(8):780-2.
2. Carnethon MR, Pu J, Howard G, Albert MA, Anderson CAM, Bertoni AG, et al. Cardiovascular Health in African Americans: A Scientific Statement From the American Heart Association. Circulation. 2017;136(21):e393-e423.
3. Whelton PK, Carey RM, Aronow WS, Casey DE, Jr., Collins KJ, Dennison Himmelfarb C, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. J Am Coll Cardiol. 2018;71(19):e127-e248.
4. Kung HC, Xu J. Hypertension-related Mortality in the United States, 2000-2013. NCHS Data Brief. 2015(193):1-8.
5. Zhang Q, Wang Y, Huang ES. Changes in racial/ethnic disparities in the prevalence of Type 2 diabetes by obesity level among US adults. Ethn Health. 2009;14(5):439-57.
6. Horvath S. DNA methylation age of human tissues and cell types. Genome Biol. 2013;14(10):R115.
7. Hannum G, Guinney J, Zhao L, Zhang L, Hughes G, Sadda S, et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. Mol Cell. 2013;49(2):359-67.
8. Chen BH, Marioni RE, Colicino E, Peters MJ, Ward-Caviness CK, Tsai PC, et al. DNA methylation-based measures of biological age: meta-analysis predicting time to death. Aging (Albany NY). 2016;8(9):1844-65.
9. Levine ME, Lu AT, Quach A, Chen BH, Assimes TL, Bandinelli S, et al. An epigenetic biomarker of aging for lifespan and healthspan. Aging (Albany NY). 2018;10(4):573-91.
10. Lu AT, Quach A, Wilson JG, Reiner AP, Aviv A, Raj K, et al. DNA methylation GrimAge strongly predicts lifespan and healthspan. Aging (Albany NY). 2019;11(2):303-27.
11. Liu Z, Leung D, Thrush K, Zhao W, Ratliff S, Tanaka T, et al. Underlying features of epigenetic aging clocks in vivo and in vitro. Aging Cell. 2020;19(10):e13229.
12. Ryan J, Wrigglesworth J, Loong J, Fransquet PD, Woods RL. A systematic review and meta-analysis of environmental, lifestyle and health factors associated with DNA methylation age. J Gerontol A Biol Sci Med Sci. 2019.
13. Irvin MR, Aslibekyan S, Do A, Zhi D, Hidalgo B, Claas SA, et al. Metabolic and inflammatory biomarkers are associated with epigenetic aging acceleration estimates in the GOLDN study. Clin Epigenetics. 2018;10:56.

14. Grant CD, Jafari N, Hou L, Li Y, Stewart JD, Zhang G, et al. A longitudinal study of DNA methylation as a potential mediator of age-related diabetes risk. Geroscience. 2017;39(5-6):475-89.

15. Nannini DR, Joyce BT, Zheng Y, Gao T, Liu L, Yoon G, et al. Epigenetic age acceleration and metabolic syndrome in the coronary artery risk development in young adults study. Clin Epigenetics. 2019;11(1):160.

16. Dugue PA, Bassett JK, Joo JE, Baglietto L, Jung CH, Wong EM, et al. Association of DNA Methylation-Based Biological Age With Health Risk Factors and Overall and Cause-Specific Mortality. Am J Epidemiol. 2018;187(3):529-38.

17. Roetker NS, Pankow JS, Bressler J, Morrison AC, Boerwinkle E. Prospective Study of Epigenetic Age Acceleration and Incidence of Cardiovascular Disease Outcomes in the ARIC Study (Atherosclerosis Risk in Communities). Circ Genom Precis Med. 2018;11(3):e001937.

18. Fransquet PD, Wrigglesworth J, Woods RL, Ernst ME, Ryan J. The epigenetic clock as a predictor of disease and mortality risk: a systematic review and meta-analysis. Clin Epigenetics. 2019;11(1):62.

19. Lind L, Ingelsson E, Sundstrom J, Siegbahn A, Lampa E. Methylation-based estimated biological age and cardiovascular disease. Eur J Clin Invest. 2018;48(2).

20. Horvath S, Gurven M, Levine ME, Trumble BC, Kaplan H, Allayee H, et al. An epigenetic clock analysis of race/ethnicity, sex, and coronary heart disease. Genome Biol. 2016;17(1):171.

21. Hillary RF, Stevenson AJ, McCartney DL, Campbell A, Walker RM, Howard DM, et al. Epigenetic measures of ageing predict the prevalence and incidence of leading causes of death and disease burden. Clin Epigenetics. 2020;12(1):115.

22. Wang C, Ni W, Yao Y, Just A, Heiss J, Wei Y, et al. DNA methylation-based biomarkers of age acceleration and all-cause death, myocardial infarction, stroke, and cancer in two cohorts: The NAS, and KORA F4. EBioMedicine. 2020;63:103151.

23. Li X, Ploner A, Wang Y, Magnusson PK, Reynolds C, Finkel D, et al. Longitudinal trajectories, correlations and mortality associations of nine biological ages across 20-years follow-up. Elife. 2020;9.

24. McCrory C, Fiorito G, Hernandez B, Polidoro S, O'Halloran AM, Hever A, et al. GrimAge outperforms other epigenetic clocks in the prediction of age-related clinical phenotypes and all-cause mortality. J Gerontol A Biol Sci Med Sci. 2020.

25. D'Agostino RB, Sr., Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, et al. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. Circulation. 2008;117(6):743-53.

26. Goff DC, Jr., Lloyd-Jones DM, Bennett G, Coady S, D'Agostino RB, Gibbons R, et al. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. Circulation. 2014;129(25 Suppl 2):S49-73.

27. Zhao W, Ammous F, Ratliff S, Liu J, Yu M, Mosley TH, et al. Education and Lifestyle Factors Are Associated with DNA Methylation Clocks in Older African Americans. Int J Environ Res Public Health. 2019;16(17).

28. Pencina MJ, D'Agostino RB, Sr., D'Agostino RB, Jr., Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med. 2008;27(2):157-72; discussion 207-12.

29. Levine ME. Assessment of Epigenetic Clocks as Biomarkers of Aging in Basic and Population Research. J Gerontol A Biol Sci Med Sci. 2020;75(3):463-5.

30. Nelson PG, Promislow DEL, Masel J. Biomarkers for Aging Identified in Cross-sectional Studies Tend to Be Non-causative. J Gerontol A Biol Sci Med Sci. 2020;75(3):466-72.

31. Horvath S, Raj K. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. Nat Rev Genet. 2018;19(6):371-84.

32. Quach A, Levine ME, Tanaka T, Lu AT, Chen BH, Ferrucci L, et al. Epigenetic clock analysis of diet, exercise, education, and lifestyle factors. Aging (Albany NY). 2017;9(2):419-46.

33. Huang RC, Lillycrop KA, Beilin LJ, Godfrey KM, Anderson D, Mori TA, et al. Epigenetic Age Acceleration in Adolescence Associates With BMI, Inflammation, and Risk Score for Middle Age Cardiovascular Disease. J Clin Endocrinol Metab. 2019;104(7):3012-24.

34. McCartney DL, Stevenson AJ, Walker RM, Gibson J, Morris SW, Campbell A, et al. Investigating the relationship between DNA methylation age acceleration and risk factors for Alzheimer's disease. Alzheimers Dement (Amst). 2018;10:429-37.

35. Smith JA, Raisky J, Ratliff SM, Liu J, Kardia SLR, Turner ST, et al. Intrinsic and extrinsic epigenetic age acceleration are associated with hypertensive target organ damage in older African Americans. BMC Med Genomics. 2019;12(1):141.

36. Hillary RF, Stevenson AJ, Cox SR, McCartney DL, Harris SE, Seeboth A, et al. An epigenetic predictor of death captures multi-modal measures of brain health. Mol Psychiatry. 2019.

37. 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.

38. Perna L, Zhang Y, Mons U, Holleczek B, Saum KU, Brenner H. Epigenetic age acceleration predicts cancer, cardiovascular, and all-cause mortality in a German case cohort. Clin Epigenetics. 2016;8:64.

39. Gao X, Colicino E, Shen J, Just AC, Nwanaji-Enwerem JC, Wang C, et al. Comparative validation of an epigenetic mortality risk score with three aging biomarkers for predicting mortality risks among older adult males. Int J Epidemiol. 2019;48(6):1958-71.

40. Jung RG, Motazedian P, Ramirez FD, Simard T, Di Santo P, Visintini S, et al. Association between plasminogen activator inhibitor-1 and cardiovascular events: a systematic review and meta-analysis. Thromb J. 2018;16:12.

41. Tofler GH, Massaro J, O'Donnell CJ, Wilson PWF, Vasan RS, Sutherland PA, et al. Plasminogen activator inhibitor and the risk of cardiovascular disease: The Framingham Heart Study. Thromb Res. 2016;140:30-5.

42. Nishida H, Horio T, Suzuki Y, Iwashima Y, Kamide K, Kangawa K, et al. Plasma adrenomedullin as an independent predictor of future cardiovascular events in high-risk patients: comparison with C-reactive protein and adiponectin. Peptides. 2008;29(4):599-605.

43. Khan SQ, O'Brien RJ, Struck J, Quinn P, Morgenthaler N, Squire I, et al. Prognostic value of midregional pro-adrenomedullin in patients with acute myocardial infarction: the LAMP (Leicester Acute Myocardial Infarction Peptide) study. J Am Coll Cardiol. 2007;49(14):1525-32.

44. Daniels PR, Kardia SL, Hanis CL, Brown CA, Hutchinson R, Boerwinkle E, et al. Familial aggregation of hypertension treatment and control in the Genetic Epidemiology Network of Arteriopathy (GENOA) study. Am J Med. 2004;116(10):676-81.

45. Ammous F, Zhao W, Ratliff SM, Kho M, Shang L, Jones AC, et al. Epigenome-wide association study identifies DNA methylation sites associated with target organ damage in older African Americans. Epigenetics. 2020:1-14.

46. Fortin JP, Fertig E, Hansen K. shinyMethyl: interactive quality control of Illumina 450k DNA methylation arrays in R. F1000Res. 2014;3:175.

47. Lehne B, Drong AW, Loh M, Zhang W, Scott WR, Tan ST, et al. A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies. Genome Biol. 2015;16:37.

48. Fortin JP, Triche TJ, Jr., Hansen KD. Preprocessing, normalization and integration of the Illumina HumanMethylationEPIC array with minfi. Bioinformatics. 2017;33(4):558-60.

49. Niu L, Xu Z, Taylor JA. RCP: a novel probe design bias correction method for Illumina Methylation BeadChip. Bioinformatics. 2016;32(17):2659-63.

50. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinformatics. 2012;13:86.

51. DNA Methylation Age Calculator. https://dnamage.genetics.ucla.edu/ [updated 06 Nov 2020.

52. Turner ST, Kardia SL, Mosley TH, Rule AD, Boerwinkle E, de Andrade M. Influence of genomic loci on measures of chronic kidney disease in hypertensive sibships. J Am Soc Nephrol. 2006;17(7):2048-55.

53. Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. Clin Chem. 1972;18(6):499-502.

54. Therneau TM, Grambsch PM, Pankratz VS. Penalized survival models and frailty. Journal of computational and graphical statistics. 2003;12(1):156-75.

55. Anderson-Bergman C. icenReg: Regression Models for Interval Censored Data in R. 2017. 2017;81(12):23.

56. Harrell FE, Jr., Califf RM, Pryor DB, Lee KL, Rosati RA. Evaluating the yield of medical tests. JAMA. 1982;247(18):2543-6.

57. Team RC. R: A language and environment for statistical computing. Vienna, Austria; 2019.

58. Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using lme4. 2015. 2015;67(1):48.

59. Højsgaard S, Halekoh U, Yan J. The R Package geepack for Generalized Estimating Equations. 2005. 2005;15(2):11.

60. Yan J, Fine J. Estimating equations for association structures. Stat Med. 2004;23(6):859-74; discussion 75-7,79-80.

3.7 Tables

Table 3-1: Descriptive characteristics of GENOA African Americans (N = 1,100) †

	Mean (SD) or N (%)
Female	781 (71.0%)
Age (years)	57.1 (10.6)
Education (years)	12.3 (3.5)
Smoking status	
Never	666 (60.5%)
Former	255 (23.2%)
Current	179 (16.3%)
Alcohol consumption (drinks/week)	0.66 (2.6)
BMI (kg/m^2)	31.20 (6.6)
Type 2 diabetes	216 (19.6%)
Anti-hypertensive medication use	649 (59.0%)
Hypertension	771 (70.1%)
Epigenetic age acceleration	
IEAA (years) (N = 1099)	0.15 (4.8)
EEAA (years)	0.27 (5.9)
PhenoAA (years) (N = 1099)	0.38 (7.2)
GrimAA (years) (N = 1099)	0.11 (5.0)
Cardiometabolic parameters	
Systolic blood pressure (mmHg)	133.8 (21.6)
Diastolic blood pressure (mmHg)	77.7 (11.9)
Mean arterial pulse pressure (mmHg)	96.4 (13.5)
Pulse pressure (mmHg)	56.2 (17.7)
Glucose (mg/dl) (N = 883)	109.2 (42.1)
Insulin (mIU/l) (N = 882)	11.5 (13.4)
Total cholesterol (mg/dl) (N = 1098)	204.3 (45.2)
HDL-C (mg/dl)	55.2 (17.9)
LDL-C (mg/dl) (N = 1076)	120.8 (41.5)
Triglycerides (mg/dl) (N = 1099)	146.0 (82.1)
CVD 10-year risk scores	
Framingham risk score (%) (N = 945)	14.4 (12.7)
ASCVD risk equation (%) (N = 988)	11.6 (11.1)
Prevalent CVD at baseline	91 (8.3%)
Incident CVD at follow-up	72 (8.8 per 1000 person-years)

Abbreviations: SD, standard deviation; IEAA, intrinsic epigenetic age acceleration; EEAA, extrinsic epigenetic age acceleration; BMI, body mass index; HDL-C, high density lipoprotein; LDL-C, low density lipoprotein; CVD, cardiovascular disease; ASCVD, Atherosclerotic cardiovascular disease. Cardiovascular disease (CVD) was defined as self-reported myocardial infarction, coronary artery revascularization, cerebrovascular events, or surgical carotid artery revascularization.
†Total N = 991 at Phase II and N = 496 at Phase III with DNAm measures.

Table 3-2: Association between epigenetic age acceleration and cardiometabolic risk factors in GENOA African Americans

Cardiometabolic Risk Factor (Outcome)	IEAA			EEAA			PhenoAA			GrimAA		
	β	**95% CI**	***P* value**	**β**	**95% CI**	***P* value**	**β**	**95% CI**	***P* value**	**β**	**95% CI**	***P* value**
SBP	**0.37**	**0.109, 0.627**	**0.005**†	**0.27**	**0.055, 0.485**	**0.014**	**0.22**	**0.050, 0.395**	**0.012**	**0.36**	**0.093, 0.627**	**0.008**
DBP	**0.16**	**0.009, 0.303**	**0.037**	0.04	-0.083, 0.160	0.534	0.07	-0.033, 0.163	0.195	0.02	-0.133, 0.170	0.809
MAP	**0.23**	**0.056, 0.393**	**0.009**	0.11	-0.026, 0.253	0.112	**0.12**	**0.004 0.229**	**0.042**	0.13	-0.042, 0.305	0.138
PP	**0.22**	**0.026, 0.414**	**0.027**	**0.24**	**0.082, 0.403**	**0.003**†	**0.16**	**0.033, 0.293**	**0.014**	**0.35**	**0.145, 0.544**	**0.001**†
Log glucose	2×10^{-3}	-0.002, 0.006	0.280	**3×10^{-3}**	**0, 0.007**	**0.049**	**4×10^{-3}**	**0.001, 0.007**	**0.004**†	**8×10^{-3}**	**0.003, 0.012**	**4.0×10^{-4}**†
Log insulin	2×10^{-3}	-0.008, 0.011	0.724	**0.01**	**0.005, 0.020**	**0.002**†	0.01	0, 0.012	0.056	2×10^{-3}	-0.008, 0.011	0.722
Total cholesterol	-0.22	-0.778, 0.332	0.430	-0.17	-0.627, 0.291	0.473	-0.06	-0.431, 0.309	0.748	-0.49	-1.057, 0.080	0.092
Log HDL-C	-3×10^{-4}	-0.004, 0.003	0.855	-1×10^{-3}	-0.004, 0.002	0.475	1×10^{-3}	-0.001, 0.003	0.444	-2×10^{-3}	-0.005, 0.002	0.328
LDL-C	-0.38	-0.894, 0.144	0.157	-0.15	-0.579, 0.281	0.498	-0.23	-0.576, 0.120	0.199	**-0.77**	**-1.303, -0.236**	**0.005**†
Log Triglycerides	3×10^{-3}	-0.002, 0.009	0.214	1×10^{-3}	-0.004, 0.005	0.758	3×10^{-3}	-0.001, 0.007	0.105	**0.01**	**0.006, 0.017**	**1.05×10^{-4}**†

Abbreviations: IEAA, intrinsic epigenetic age acceleration; EEAA, extrinsic epigenetic age acceleration; SBP, systolic blood pressure; DBP, diastolic blood pressure; MAP, mean arterial pressure; PP, pulse pressure; HDL-C, high density lipoprotein; LDL-C, low density lipoprotein.
Models are adjusted for age, sex and familial relatedness.
Effect sizes (β) correspond to the change in the cardiometabolic risk factor associated with a 1-year increase in the epigenetic age acceleration measure.
Associations significant at $P < 0.05$ are shown in bold; †Associations significant at Bonferroni adjusted $P < 0.005$.

Table 3-3: Association between epigenetic age acceleration and clinical CVD risk scores in GENOA African Americans

	Framingham risk score (FRS) (N = 945)		**Atherosclerotic cardiovascular disease equation (ASCVD)** (N = 988)	
Epigenetic Age Acceleration (Predictor)	**β (95% CI)**	***P* value**	**β (95% CI)**	***P* value**
IEAA	0.12 (-0.02 – 0.26)	0.088	**0.17 (0.06 – 0.29)**	**0.004**
EEAA	**0.21 (0.10 – 0.32)**	**3.77×10^{-4}**	**0.23 (0.13 – 0.33)**	**3.85×10^{-6}**
PhenoAA	**0.15 (0.06 – 0.24)**	**0.001**	**0.18 (0.10 – 0.26)**	**4.78×10^{-6}**
GrimAA	**0.58 (0.44 – 0.71)**	**4.53×10^{-16}**	**0.44 (0.33 – 0.56)**	**2.38×10^{-13}**

Abbreviations: IEAA, intrinsic epigenetic age acceleration; EEAA, extrinsic epigenetic age acceleration.
Models are adjusted for age, sex and familial relatedness.
FRS and ASCVD are modeled as continuous predictors.
Effect sizes (β) correspond to the change in predicted 10-year risk of CVD using the FRS or ASCVD risk equation associated with 1-year increase in the epigenetic age acceleration measure.
Cardiovascular disease (CVD) was defined as self-reported myocardial infarction, coronary artery revascularization, cerebrovascular events, or surgical carotid artery revascularization.
Associations significant at $P < 0.05$ are shown in bold.

Table 3-4: Incident CVD hazard ratios for epigenetic age acceleration in GENOA African Americans

Epigenetic Age Acceleration (Predictor)	**HR (95% CI)**	***P* value**
IEAA	0.97 (0.93 – 1.02)	0.300
EEAA	1.04 (1.00 – 1.08)	0.057
PhenoAA	1.00 (0.97 – 1.04)	0.810
GrimAA	**1.09 (1.04 – 1.15)**	**4.20×10^{-4}**

Abbreviations: CVD, cardiovascular disease; HR, hazard ratio; IEAA, intrinsic epigenetic age acceleration; EEAA, extrinsic epigenetic age acceleration.
Models are adjusted for age, sex and familial relatedness.
Hazard ratios correspond to the risk of a CVD event associated with a 1-year increase in the epigenetic age acceleration measure.
Cardiovascular disease (CVD) was defined as self-reported myocardial infarction, coronary artery revascularization, cerebrovascular events, or surgical carotid artery revascularization.
Associations significant at $P < 0.05$ are shown in bold.

Table 3-5: Incident CVD hazard ratios for GrimAA and clinical CVD risk scores in GENOA African Americans

Predictor	Adjusted HR (95% CI)			
	FRS only (N = 945)	**ASCVD only** (N = 988)	**FRS + GrimAA** (N = 945)	**ASCVD + GrimAA** (N = 988)
FRS	1.03 (1.02 - 1.05) $P = 9.5 \times 10^{-6}$	--	1.03 (1.01 – 1.05) $P = 4.7 \times 10^{-4}$	--
ASCVD	--	1.04 (1.02 – 1.06) $P = 2.7 \times 10^{-5}$	--	1.03 (1.01 – 1.05) $P = 9.8 \times 10^{-4}$
GrimAA	--	--	1.07 (1.02 – 1.13) $P = 0.011$	1.08 (1.02 – 1.13) $P = 0.007$

Abbreviations: CVD, cardiovascular disease; HR, hazard ratio; FRS, Framingham risk score; ASCVD, atherosclerotic cardiovascular disease.
Models consisted of clinical risk scores with and without GrimAA. All models were adjusted for age, sex and familial relatedness.
Adjusted hazard ratios correspond to the risk of a CVD event associated with a 1-unit increase in the clinical risk score or the epigenetic age acceleration measure.
Cardiovascular disease (CVD) was defined as self-reported myocardial infarction, coronary artery revascularization, cerebrovascular events, or surgical carotid artery revascularization.

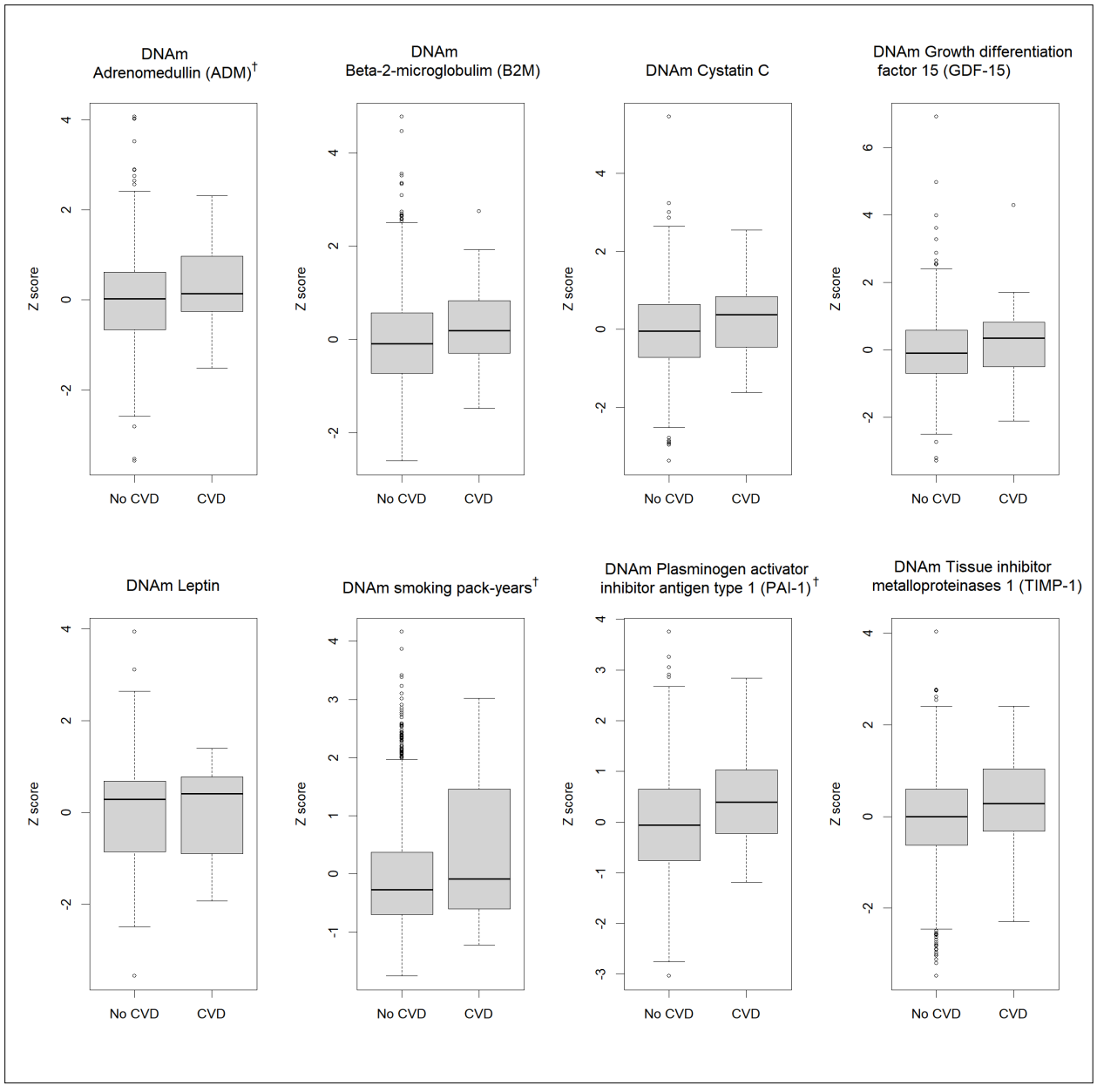
Table 3-6: C-statistics evaluating the predictive performance of GrimAA on incident CVD in GENOA African Americans

Model	**C-statistic**	**95% CI**
N = 945		
Base model (age + sex)	0.595	0.525 – 0.664
Age + sex + GrimAA	0.643	0.576 – 0.709
Age + sex + FRS	0.687	0.624 – 0.749
Age + sex + FRS + GrimAA	0.698	0.637 – 0.759
N = 988		
Base model (age + sex)	0.588	0.521 – 0.656
Age + sex + GrimAA	0.636	0.571 – 0.701
Age + sex + ASCVD	0.670	0.606 – 0.728
Age + sex + ASCVD + GrimAA	0.685	0.625 – 0.746

Abbreviations: CVD, cardiovascular disease; FRS, Framingham risk score; ASCVD, atherosclerotic cardiovascular disease.
The C statistics associated with a set of nested models for time to CVD events are shown. All models are adjusted for familial relatedness.
Cardiovascular disease (CVD) was defined as self-reported myocardial infarction, coronary artery revascularization, cerebrovascular events, or surgical carotid artery revascularization

3.8 Figures

Figure 3-1: Boxplots of standardized GrimAge components by incident CVD status



CVD, cardiovascular disease.
Cardiovascular disease (CVD) was defined as self-reported myocardial infarction, coronary artery revascularization, cerebrovascular events, or surgical carotid artery revascularization.
† GrimAge components significantly associated with incident CVD in models adjusted for age, sex, and familial relatedness ($P < 0.05$).

3.9 Supplementary Material

Table SM 3-1: Pearson correlations between age, DNA methylation age, and epigenetic age acceleration in GENOA African Americans

	Age	HorvathAge	HannumAge	PhenoAge	GrimAge	IEAA	EEAA	PhenoAA
HorvathAge	0.86***							
HannumAge	0.90***	0.90***						
PhenoAge	0.82***	0.84***	0.85***					
GrimAge	0.85***	0.78***	0.81***	0.80***				
IEAA	0.02	0.51***	0.24***	0.28***	0.12***			
EEAA	0.00	0.28***	0.43***	0.30***	0.15***	0.43***		
PhenoAA	-0.01	0.23***	0.20***	0.56***	0.16***	0.44***	0.50***	
GrimAA	-0.02	0.07*	0.07*	0.17***	0.50***	0.19***	0.27***	0.32***

*** p < .0001; ** p < .01; * p < .05
IEAA, intrinsic epigenetic age acceleration; EEAA, extrinsic epigenetic age acceleration.

Table SM 3-2: Adjusted associations between epigenetic age acceleration and cardiometabolic risk factors for associations with P < 0.05 in the base model (Model 1) in GENOA African Americans

Cardiometabolic Risk Factor (Outcome)	Epigenetic Age Acceleration (Predictor)	Model 2			Model 3		
		β	**95% CI**	***P* value**	**β**	**95% CI**	***P* value**
SBP	**IEAA**	**0.33**	**0.075, 0.594**	**0.012**			
	EEAA	**0.23**	**0.012, 0.443**	**0.039**			
	PhenoAA	0.17	-0.002, 0.347	0.053	0.14	-0.047, 0.329	0.141
	GrimAA	0.31	-0.024, 0.638	0.069	0.29	-0.070, 0.639	0.115
DBP	**IEAA**	**0.16**	**0.010, 0.305**	**0.037**			
Mean arterial pressure	**IEAA**	**0.22**	**0.0459, 0.385**	**0.013**			
	PhenoAA	0.10	-0.018, 0.210	0.100	0.10	-0.019, 0.226	0.099
Pulse pressure	**IEAA**	0.18	-0.010, 0.378	0.063			
	EEAA	**0.20**	**0.043, 0.365**	**0.013**			
	PhenoAA	0.12	-0.011, 0.25	0.073	0.06	-0.080, 0.200	0.403
	GrimAA	**0.34**	**0.095, 0.588**	**0.007**	0.26	-0.004, 0.523	0.054
Log glucose	**EEAA**	2×10^{-3}	-0.001, 0.006	0.137			
	PhenoAA	$\mathbf{3 \times 10^{-3}}$	**0, 0.005**	**0.047**	2×10^{-3}	-0.001, 0.005	0.100
	GrimAA	$\mathbf{8 \times 10^{-3}}$	**0.003, 0.013**	**0.001**	$\mathbf{8 \times 10^{-3}}$	**0.003, 0.013**	**0.004**
Log insulin	**EEAA**	**0.010**	**0.003, 0.017**	**0.004**			
LDL-C	**GrimAA**	**-1.05**	**-1.709, -0.383**	**0.002**	**-0.74**	**-1.444, -0.025**	**0.043**
Log Triglycerides	**GrimAA**	**0.01**	**0.001, 0.015**	**0.026**	**0.01**	**0.0002, 0.015**	**0.045**

IEAA, intrinsic epigenetic age acceleration; EEAA, extrinsic epigenetic age acceleration; SBP, systolic blood pressure; DBP, diastolic blood pressure; HDL-C, high density lipoprotein; LDL-C, low density lipoprotein.
Model 1 (base model) is adjusted for age, sex, and familial relatedness. Results for Model 1 are shown in Table 2. Only associations that were significant in Model 1 (P<0.05) were further evaluated and reported in this table.
Model 2 is adjusted for Model 1 and years of education, smoking status, body mass index, and alcohol consumption.
Model 3 is adjusted for Model 2 and white blood cell counts.
Effect sizes (β) correspond to the change in the cardiometabolic risk factor associated with a 1-year increase in the epigenetic age acceleration measure.
Associations significant at $P < 0.05$ are shown in bold.

Table SM 3-3: Incident CVD hazard ratios for GrimAge components in GENOA African Americans

GrimAge component (Predictor)	**HR (95% CI)**	***P* value**
Adrenomedullin (ADM)	**1.45 (1.07 – 1.98)**	**0.017**
Beta-2-microglobulim (B2M)	1.04 (0.77 – 1.41)	0.780
Cystatin C	1.38 (0.95 – 1.99)	0.089
Growth differentiation factor 15 (GDF-15)	1.21 (0.92 – 1.59)	0.170
Leptin	1.38 (0.91 – 2.10)	0.130
Smoking pack-years	**1.36 (1.09 – 1.72)**	**0.0076**
Plasminogen activator inhibitor antigen type 1 (PAI-1)	**1.48 (1.17 – 1.87)**	**0.0012**
Tissue inhibitor metalloproteinases 1 (TIMP-1)	1.59 (0.97 – 2.60)	0.068

Abbreviations: CVD, cardiovascular disease; HR, hazard ratio.
Models adjusted for age, sex, white blood cell counts, and familial relatedness.
Hazard ratios significant at $P < 0.05$ are shown in bold font.
Hazard ratios correspond to the risk of a CVD event associated with a one-standard deviation increase in the GrimAge component.
Cardiovascular disease (CVD) was defined as self-reported myocardial infarction, coronary artery revascularization, cerebrovascular events, or surgical carotid artery revascularization.
Associations significant at $P < 0.05$ are shown in bold.

Figure SM 3-1: Scatterplots of DNA methylation age measures against chronological age in GENOA African Americans

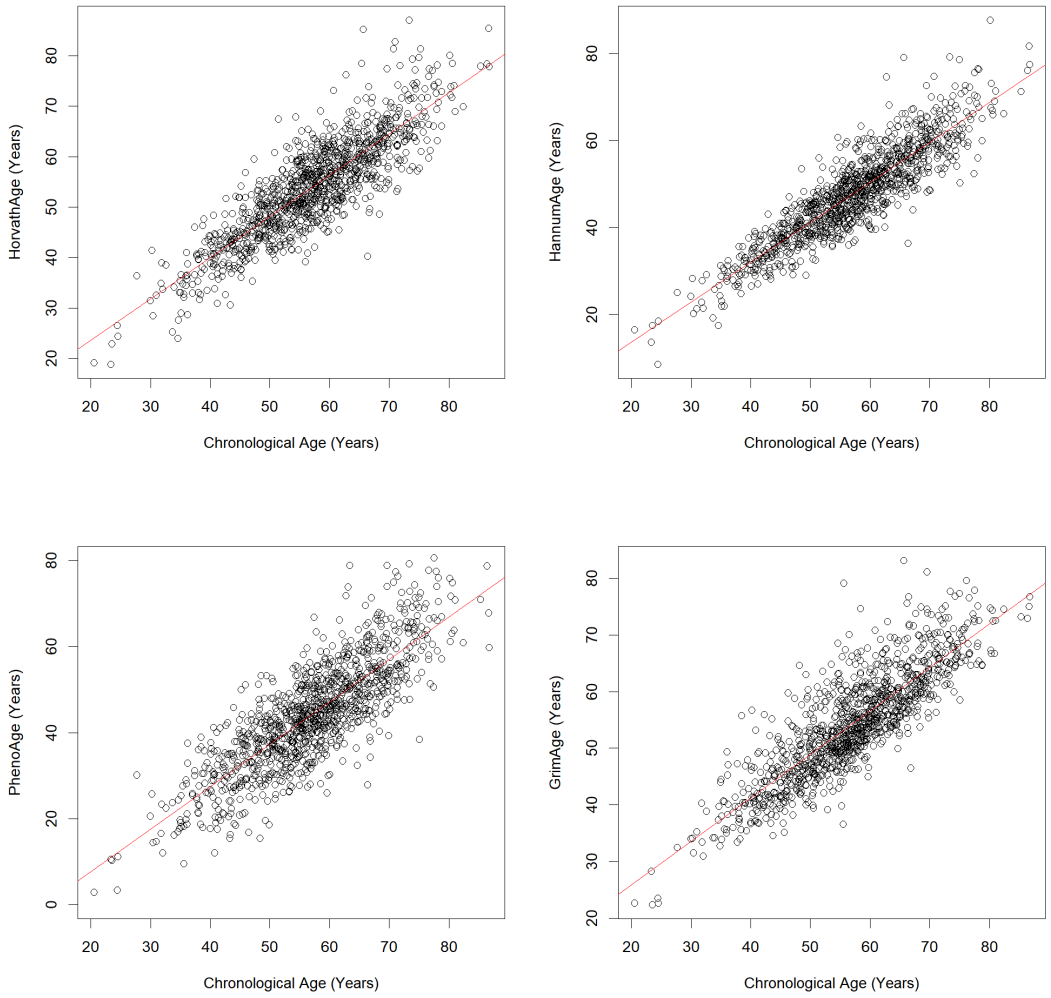


Figure SM 3-2: Scatterplots of epigenetic age acceleration measures against chronological age in GENOA African Americans

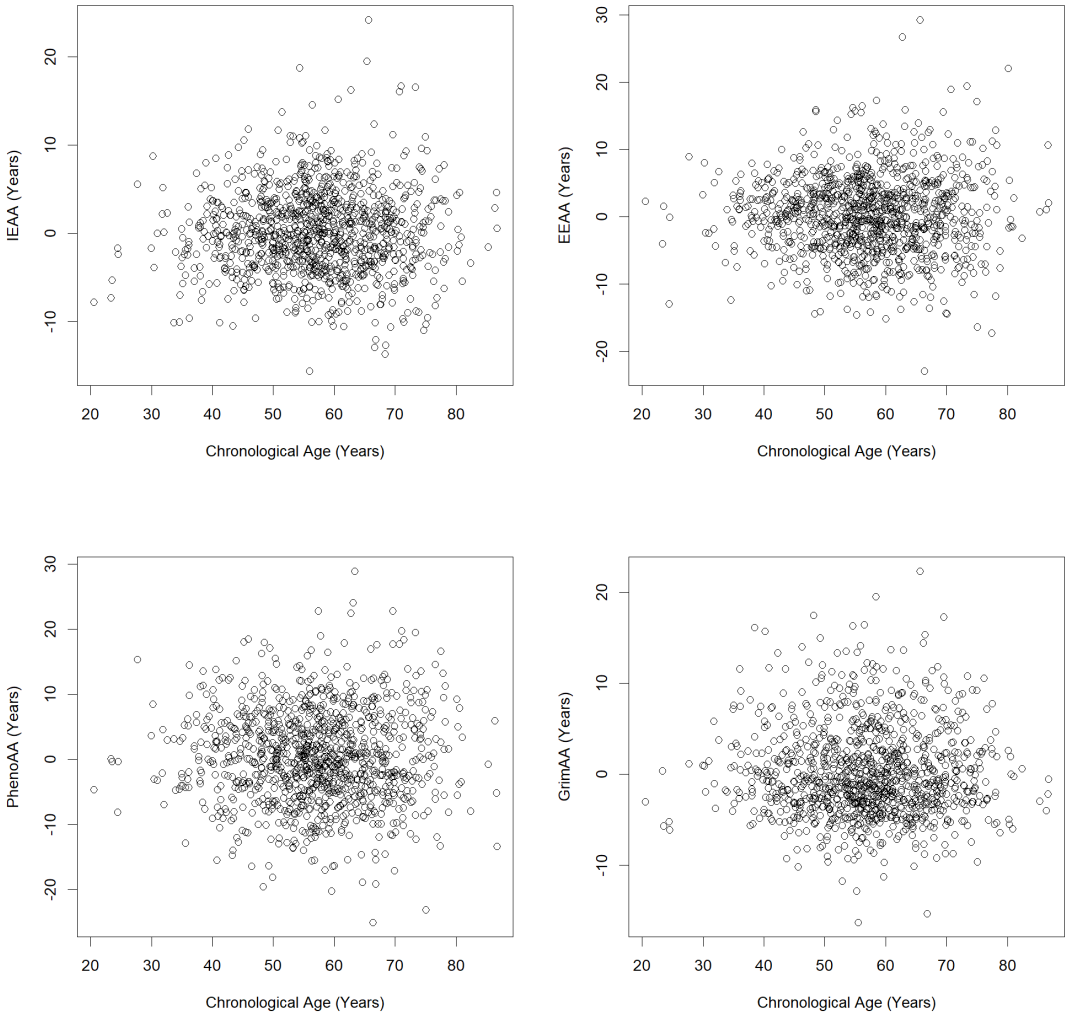
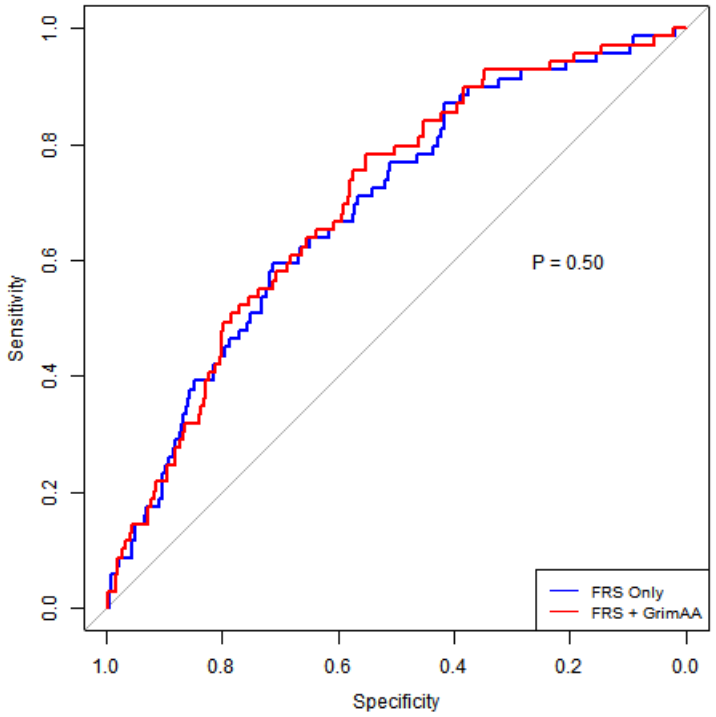
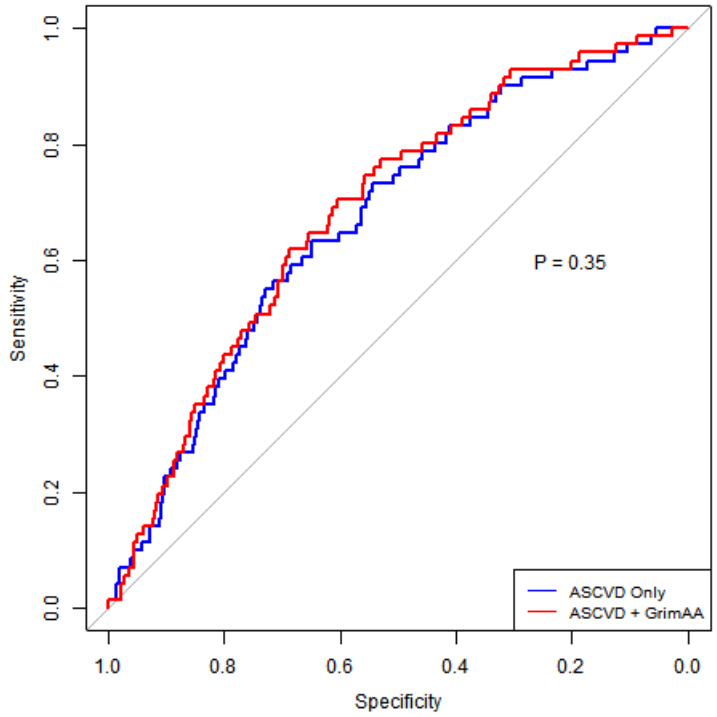


Figure SM 3-3: Receiver operator characteristic (ROC) curves for incident CVD in GENOA African Americans

(A) FRS vs. FRS + GrimAA



(B) ASCVD vs. ASCVD + GrimAA



(A) Receiver operator curves for time to CVD for models of Framingham risk score (blue line) vs. Framingham risk score + GrimAA (red line) (N = 945).
(B) Receiver operator curves for time to CVD for models of ASCVD score (blue line) vs. ASCVD + GrimAA (red line) (N = 988).
Models are adjusted for age, sex and familial relatedness.
Cardiovascular disease (CVD) was defined as self-reported myocardial infarction, coronary artery revascularization, cerebrovascular events, or surgical carotid artery revascularization.
CVD, cardiovascular disease; FRS, Framingham risk score; ASCVD, atherosclerotic cardiovascular disease.

Figure SM 3-4: Improvement in reclassification of incident CVD in GENOA African Americans

(A) FRS + GrimAA

FRS model		Standard CVD Model + GrimAA		
		≤ 7.5%	> 7.5%	Total no. (%) of participants
CVD Events	≤ 7.5%	7	0	**7**
	> 7.5%	0	62	**62**
	Total no. (%) of participants	**7 (10.1)**	**62 (89.9)**	**69**
CVD Nonevents	≤ 7.5%	337	1	**338**
	> 7.5%	49	489	**538**
	Total no. (%) of participants	**386 (44.1)**	**490 (55.9)**	**876**

(B) ASCVD + GrimAA

ASCVD model		Standard CVD Model + GrimAA		
		≤ 7.5%	> 7.5%	Total no. (%) of participants
CVD Events	≤ 7.5%	14	1	**15**
	> 7.5%	0	56	**56**
	Total no. (%) of participants	**14 (19.7)**	**57 (80.3)**	**71**
CVD Nonevents	≤ 7.5%	442	9	**451**
	> 7.5%	23	443	**466**
	Total no. (%) of participants	**465 (49.7)**	**452 (50.3)**	**917**

Columns and rows refer to categories of predicted risk using (A) FRS and GrimAA and (B) ASCVD and GrimAA. The counts in the cells represent the number of individuals assigned to the indicated risk category.
Blue shaded cells represent correctly reclassified individuals, and orange shaded cells represent incorrectly classified individuals.
NRI = P(up|case) – P(down|case) – P(up|noncase) + P(down|noncase)
In the model with FRS and GrimAA (A), 1 individual was incorrectly up-classified and 49 individuals were correctly down-classified (NRI = 49/876 – 1/876)
In the model with ASCVD and GrimAA (B), 1 individual was correctly-up classified. 9 individuals were incorrectly up-classified, and 23 individuals were correctly down-classified (NRI = 1/71 – 9/917 + 23/917)
Abbreviations: CVD, cardiovascular disease; FRS, Framingham risk score; ASCVD, atherosclerotic cardiovascular disease
Models are adjusted for age, sex and familial relatedness.
Cardiovascular disease (CVD) was defined as self-reported myocardial infarction, coronary artery revascularization, cerebrovascular events, or surgical carotid artery revascularization

Chapter 4. Epigenetics of Single and Multisite Atherosclerosis in African Americans from the Genetic Epidemiology Network of Arteriopathy (GENOA)

4.1 Abstract

DNA methylation, an epigenetic mechanism modulated by lifestyle and environmental factors, may be an important biomarker of complex diseases including cardiovascular diseases (CVD) and subclinical atherosclerosis. Novel biomarkers of CVD may help to improve risk prediction and advance precision medicine. DNA methylation in blood samples from 391 African Americans from the Genetic Epidemiology Network of Arteriopathy (GENOA) was assessed at baseline, and atherosclerosis was assessed 5 and 12 years later. Using linear mixed models, we examined the association between previously-identified CpGs for coronary artery and carotid artery calcification, both individually and aggregated into methylation risk scores (MRS_{CAC} and $MRS_{cartoid}$), and four measures of atherosclerosis (coronary artery calcification (CAC), abdominal aorta calcification (AAC), ankle brachial index (ABI), and a multi-atherosclerosis score combining the three measures). We also examined the association between four epigenetic age acceleration measures (IEAA, EEAA, PhenoAge acceleration, and GrimAge acceleration) and the four atherosclerosis measures. Finally, we characterized the temporal stability of the potential epigenetic markers for atherosclerosis (epigenetic age acceleration and MRSs) using longitudinal measures of DNA methylation measured five years after baseline for a subset of 193 participants. One and six CpGs were associated with AAC and multisite atherosclerosis, respectively, after adjusting for age, sex, and traditional CVD risk factors at false discovery rate (FDR) < 0.1. A one-unit increase in $MRS_{carotid}$ was associated with an

approximately 1.6-fold increase in the Agatston score of CAC and AAC, and a 0.7 units increase in the multisite atherosclerosis score (score range 0-12) after adjusting for CVD risk factors (Bonferroni adjusted $P < 0.05$). The $MRS_{carotid}$ explained 5.3%, 2.7%, and 5.5% of the variability of CAC, AAC, and multisite atherosclerosis. A 5-year increase in GrimAA (~1 SD) was associated with a 1.6-fold (95% CI 1.11–2.25) increase in the Agatston score of AAC and 0.7 units (95% CI 0.33–1.07) increase in multisite atherosclerosis score after adjusting for CVD risk factors (Bonferroni adjusted $P < 0.05$). All epigenetic measures were relatively stable,with the highest intraclass correlation coefficients observed for $MRS_{carotid}$ and GrimAge acceleration (0.82 and 0.89, respectively). Our study found evidence of an association between a number of CpGs and two epigenetic measures, an atherosclerosis methylation risk score ($MRS_{carotid}$) and GrimAge acceleration, and atherosclerosis at multiple vascular sites in a sample of African Americans. These epigenetic measures were relatively stable over time. These findings deepen our understanding of the relationship between aging and atherosclerosis and suggest that further evaluation of these potential biomarkers is warranted.

4.2 Introduction

Cardiovascular diseases (CVD), including coronary heart disease, myocardial infarction, stroke, and peripheral artery disease, are the leading cause of death in the US.[1] In African Americans, the CVD mortality burden is 21% higher compared to whites despite modest decreases in racial disparities at the national level since 2005.[2] Genetic factors, along with non-genetic risk factors, such as age, smoking, and hypertension contribute to CVD. However, much of the variability in CVD, as well as the persistent causes of CVD disparities, remains unexplained.

Vascular calcification can occur in either the intimal or medial layers of the arterial wall.[3] The intima is the innermost layer consisting of a smooth endothelium layer covered by elastic tissue. The aorta is associated with atherosclerosis.[3] Atherosclerosis is a chronic inflammatory age-related condition that develops over several decades and is a precursor for CVD.[4-6] Coronary artery calcification (CAC) is a strong predictor of incident CVD and coronary heart disease beyond traditional CVD risk factors.[7-10] Medial calcification occurs in the tunica media which consists of smooth muscles cells and elastic fibers. It affects lower limb arteries, in addition to the aorta, and is typically associated with peripheral artery disease.[3] Medial lesions are thought to calcify earlier than intimal ones and result in vascular stiffness and reduced vessel compliance,[11,12] and they may be also associated with CVD.[8,13] Medial calcification increases with aging and is prevalent in individuals with chronic kidney disease and diabetes mellitus. In the peripheral arteries, it can be assessed the using ankle-brachial pressure index ratio.[11] Epigenetics, which captures both genetic influences as well as imprints of lifestyle and environmental exposures throughout the life course, may help identify biological mechanisms contributing to CVD pathogenesis and atherosclerosis. As such, epigenetic markers may have

potential as biomarkers of CVD risk. Similarly, epigenetic age acceleration measures, including first generation measures (Horvath[14] and Hannum[15]), which were trained on chronological age, and more recent measures trained on a number of biological and physiological markers and chronological age, such as PhenoAge[16] and GrimAge,[17] are DNA methylation-based markers of biological aging that are associated with late-life onset diseases and mortality.[14,15,17-19] Previous studies of genome-wide DNA methylation patterns or epigenetic age acceleration measures have reported significant associations between epigenetic markers and CVD,[20-26] however, a majority of these studies were in cohorts of European ancestry and/or were cross-sectional rather than longitudinal.

Only a few epidemiological studies have examined the association between DNA methylation and subclinical CVD or atherosclerosis.[27,28] A recent cross-sectional transcriptome and epigenome analyses of atherosclerosis in 1,208 participants from the Multi-Ethnic Study of Atherosclerosis (MESA) identified 82 differentially methylated CpGs associated with either CAC or carotid plaque score at false discovery rate (FDR) ≤ 0.1.[27] The sample was comprised of 45.9% Caucasians, 21.5% African Americans, and 32.6% Hispanics. Race-specific analyses showed that the directions of the methylation changes were generally consistent across ethnicities, although some sites were not significant for African Americans and Hispanics.[27] The most significant CpG associated with carotid plaque, cg05575921, is located in the *AHRR* gene body and is a well-documented smoking marker.[29-32]

In this study, we evaluated the association between potential epigenetic markers of atherosclerosis and single- or multi-site atherosclerosis in 391 African Americans from the Genetic Epidemiology Network of Arteriopathy (GENOA) during a mean follow-up of 12 years. Our definition of atherosclerosis included both intimal and medial calcification measures: CAC,

abdominal aorta calcification (AAC), and ABI. Epigenetic markers included the previously-identified CpGs for atherosclerosis,[27] methylation risk scores (MRSs) derived from these CpGs, and four epigenetic age acceleration measures (Horvath (IEAA),[14] Hannum (EEAA),[15] PhenoAge (PhenoAA),[16] and GrimAge (GrimAA)).[17] Finally, we characterized the temporal stability of the epigenetic age acceleration measures and the MRSs using longitudinal measures of DNA methylation for a subset of the sample (N=129).

4.3 Methods

Study sample

Genetic Epidemiology Network of Arteriopathy (GENOA) is a community-based study in Rochester, MN and Jackson, MS that was established to identify genes influencing blood pressure.[33] In the first phase of GENOA (Phase I: 1996 – 2001), sibships with at least two adults with clinically diagnosed essential hypertension before age 60 were recruited, and all siblings in the sibship were invited to participate regardless of hypertension status. Exclusion criteria included secondary hypertension, alcoholism or drug abuse, pregnancy, insulin-dependent diabetes mellitus, or active malignancy.

At baseline (Phase I), a total of 1,583 non-Hispanic whites (Rochester, MN) and 1,854 African Americans (Jackson, MS) were enrolled. In the second phase (Phase II: 2001 – 2005), all participants were invited for a second examination. Eighty percent of African Americans (N = 1,482) and 75% of non-Hispanic whites (N = 1,213) from Phase 1 returned. At Phase III (2009-2011), 752 African Americans returned for a third examination. Demographic information, medical history, clinical characteristics, lifestyle factors, and blood samples were collected in each phase. This study includes 391 African American participants from 277 sibships who had their DNA methylation measured in whole blood samples collected at Phase I and were seen in

Phase III. For a subset of 129 participants that were seen in Phase III, DNA methylation was measured at both Phase I and Phase II. Written informed consent was obtained from all participants and approval was granted by participating institutional review boards (University of Michigan, University of Mississippi Medical Center, and Mayo Clinic).

Study measurements

Height was measured by stadiometer and weight by electronic balance. Body mass index (BMI) was calculated as weight in kilograms divided by the square of height in meters. Smoking was categorized as current, former, or never. Resting systolic (SBP) and diastolic blood pressure (DBP) were measured by a random zero sphygmomanometer and a cuff appropriate for arm size. The second and third of three readings, after the participant sat for at least 5 minutes, were averaged for analysis.[34] Information on current anti-hypertensive medication and statin use were collected. Hypertension was defined having an average SBP ≥ 140 mmHg or DBP ≥ 90 mmHg, or current anti-hypertensive medication use. Type 2 diabetes status (T2D) was defined as having fasting blood glucose levels ≥ 126 mg/dL or self-reported physician diagnosed diabetes and current diabetes medications use. Serum total cholesterol (TC), HDL-C, and triglycerides (TGs) were measured by standard enzymatic methods on a Hitcahi 911 Chemistry Analyzer (Roche Diagnostics, Indianapolis, IN). TC was adjusted for statin use as TC/0.8. LDL-C was calculated using the Friedewald formula for individuals with triglycerides below 400 mg/dl.[35]

DNA methylation, epigenetic age acceleration and methylation risk scores

Genomic DNA from 1,106 African American participants from Phase I and 304 from Phase II was extracted from stored peripheral blood leukocytes using AutoGen FlexStar (AutoGen, Holliston, MA) and DNA methylation was measured using the Infinium MethylationEPIC BeadChip. DNA methylation processing procedures have been previously

described.[36] Briefly, Sex mismatches and outliers were excluded using the shinyMethyl R package.[37] Probes with detection *P*-value < 10^{-16} were considered to be successfully detected.[38] Samples and probes that failed a detection rate of at least 10% were removed. The Noob method was used for individual background and dye-bias normalization.[39] Regression on Correlated Probes method was used to adjust for the probe-type bias in the data.[40] White blood cell type proportions within the blood sample were estimated using Houseman's method.[41]A total of 1,100 samples from Phase I and 294 from Phase II were available after quality control.

We derived two methylation risk scores based on the regression coefficients for CpG sites associated with CAC and/or carotid plaque at FDR≤ 0.1 in MESA.[27] At FDR of ≤ 0.1, 16 CpG sites were associated with CAC in MESA, 15 of which were available in GENOA. For carotid plaque, 68 CpG sites were reported in the MESA study, 62 of which were available in GENOA, and two of which overlapped with those associated with CAC (cg07033253 and cg23661483). Methylation M-values were adjusted for batch effects (modeled as random effects of plate, row and column) and white blood cell counts modeled as fixed effects. The adjusted means plus residuals were then used to calculate two methylation risk scores, MRS_{CAC} and $MRS_{carotid}$. The MRSs were scored so that higher values correspond to greater risk of atherosclerosis.

Methylation beta values for Phase I and Phase II DNA methylation were uploaded to the online Horvath epigenetic age calculator to calculate DNAm Age.[42] Four measures of epigenetic age – HorvathAge, HannumAge, PhenoAge and GrimAge – were estimated. Intrinsic epigenetic age acceleration (IEAA) is based on the regression residuals from a model of chronological age and blood cell counts and Horvath age as the outcome.[14,15] Extrinsic epigenetic age acceleration (EEAA) is derived similarly based on the Hannum epigenetic age measure but incorporates

weighted averages of three white blood cell types (naïve cytotoxic T cells, exhausted cytotoxic T cells, and plasmablasts).[15,43] PhenoAA is a residual measure from a model of phenotypic age, calculated based on clinical measures such as albumin, creatinine, and white blood cell counts, as well as chronological age.[18] GrimAA is based on the residuals from a model of GrimAge and chronological age. GrimAge is a composite biomarker of smoking pack-years and 7 surrogate measures of plasma proteins selected for their significant association with time-to-death. The components are adrenomedullin (ADM), beta-2-microglobulin, cystatin C, GDF-15, leptin, plasminogen activator inhibitor antigen type 1 (PAI-1), and tissue inhibitor metalloproteinases 1 (TIMP-1).[17]

Atherosclerosis measurements

Computed tomography (CT) imaging was used to quantify calcification in the coronary arteries and the abdominal aorta. CT images were read by trained technologists and the amount of calcified plaque was calculated by multiplying each lesion area are by a weighted attenuation score (Hounsfield units) on a TeraRecon Aquarius Workstation (TeraRecon, San Mateo, CA). The amount of calcification was quantified using the Agatston score. Both CAC and abdominal aorta calcification (AAC) were natural log-transformed as ln (measure +1) when examined individually as outcomes.

The ankle-brachial index was used to quantify atherosclerosis in the peripheral arteries. Details about the ABI measurement has been previously described.[44] Briefly, a Doppler ultrasonic instrument (Medisonics, Minneapolis, MN) was used to detect the pulse at each arm and ankle using appropriately sized blood pressure cuffs. ABI was calculated as the systolic blood pressure at each ankle site divided by the higher of the two brachial pressures. The lower

of the average ABIs from both legs was used. Individuals with ABI > 1.50 were excluded as they may have non-compressible arteries and medial arterial calcification.

Multisite atherosclerosis score was defined similarly to that described by Zhao et al.[45] Both CAC and AAC were scored separately as follows: 0 if absent, or for those with calcification as a score between 1 and 4 according to gender-specific quartiles of each measure. ABI was scored between 1 and 4 for the highest to lowest gender-specific quartiles for ABI < 1.0; 0 for 1 ≤ ABI < 1.4, and 1 if ABI ≥ 1.4 and ≤ 1.50. The multisite atherosclerosis score was then calculated as the sum of the three measures (range: 0-12) and was modeled as a continuous outcome.

Statistical analysis

Outliers beyond 5 standard deviations from the mean of the outcome and epigenetic measures were removed. We calculated the Pearson correlation coefficients between the single site atherosclerosis measures, and among the epigenetic age acceleration measures and MRSs. We also calculated the correlations between the individual GrimAge components and the MRSs. We use linear mixed models that account for familial relatedness to assess the association between the previously identified atherosclerosis-associated CpGs in MESA and MRSs derived from these CpGs (predictors) and the single- or multi-site atherosclerosis measures (outcomes). The minimally adjusted model (Model 1) was adjusted for age at baseline, sex, first 4 genetic principal components (PCs), and time between the measures. The time covariate was calculated based on the age difference of the participants at the time of atherosclerosis assessment (Phase II or Phase III) and DNA methylation assessment (Phase I). For models of multisite atherosclerosis, we included two time covariates – the age difference between Phase III and Phase I, and an additional covariate for the age difference between Phase III and Phase II – to account for the differences in the assessment times of ABI (at Phase II) and CAC and AAC (at Phase III). Model

2 was additionally adjusted for smoking status, and Model 3 was further adjusted for cardiovascular risk factors (T2D status, hypertension status, BMI, and statin-adjusted total cholesterol levels).

We used a similar procedure as described above to assess the association between the four epigenetic age acceleration measures and atherosclerosis without including the PCs. This approach of not adjusting for genetic PCs is consistent with other studies and allow for comparability of findings. As sensitivity analysis, we also performed this analysis after adjusting for the first four genetic PCs. We additionally carried out sensitivity analyses adjusting for white blood cell counts for significant associations with PhenoAA and GrimAA to assess confounding by changes in blood cell composition.[16,17,46] For GrimAA, we investigated the association between the individual GrimAge components and atherosclerosis to identify components that may be driving the association between GrimAA and atherosclerosis or that outperform the overall GrimAA measure itself. For this analysis, GrimAge components were scaled and centered. Models were adjusted for age, sex, and white blood cell counts.

For the subset of individuals with repeated DNA methylation measurements, we used chi-square and t-tests as appropriate to compare the demographic and clinical characteristics of the subset to the full sample. We used linear mixed-effect models adjusted for age and sex to calculate the intraclass correlation coefficients (ICC) between Phases I and II for the methylation risk scores and the epigenetic age acceleration measures. Parametric bootstrapping (1000 iterations) was used to calculate the 95% confidence intervals of the ICC coefficients.[47,48] Additionally, for these participants, we assessed the associations between all of the epigenetic biomarkers at Phase II and atherosclerosis after adjusting for age at Phase II, sex, time between measures, and the first 4 PCs (Model 1). Finally, we used generalized estimating equation (GEE)

models to assess the intra-individual changes in the epigenetic predictors during follow-up in models adjusted for age and sex.

Statistical tests were two-sided. False discovery rate (FDR)[49] of 0.1 was considered significant for the associations between the individual CpGs and the atherosclerosis measures. A Bonferroni adjusted *P* value of < 0.025 (for MRS association analyses) and *P* < 0.0125 (for epigenetic age acceleration association analyses) were considered significant. Analyses were conducted in R (Version 3.4.1),[50] using the lme4,[51] rptR,[48] and geepack [52] packages.

4.4 Results

Sample characteristics

Baseline characteristics of the participants are shown in **Table 1**. The mean age of the participants was 56 years (SD = 9.0) at Phase I. Women comprised about 76% of the sample. Phase III was on average 12 years (SD = 1.2) after Phase I. About 64% of the participants had hypertension and 16% had T2D at Phase I. Distributions of the atherosclerosis measures and MRSs are shown in **Supplemental Figures 1 and 2**, respectively. For MRS_{CAC}, one outlier beyond 5 SD from the mean was removed from the analysis.

Correlations among the methylation risk scores, epigenetic age acceleration measures, and atherosclerosis measures

CAC and AAC were correlated at r = 0.57 ($P < 2.2 \times 10^{-16}$). ABI was negatively correlated with AAC (r = -0.14, $P = 0.004$), but only weakly correlated with CAC (r = -0.08, p = 0.09). **Supplemental Table 1** shows the correlation between the epigenetic predictors. The two MRSs were correlated at r = 0.39 ($P = 6 \times 10^{-16}$). The correlation between the acceleration measures was weak to moderate (r range: 0.21 – 0.46). The epigenetic age acceleration measures and MRSs were weakly correlated (r range: 0.08 – 0.25), with the exception of GrimAA and

$MRS_{carotid}$ which were correlated at r = 0.62 ($P = 2.2 \times 10^{-16}$). **Supplemental Figure 3** shows the scatterplots of the MRSs against the epigenetic age acceleration measures. Of the GrimAge components, smoking pack years slightly more correlated with $MRS_{carotid}$ than the overall GrimAA measure (r = 0.68, $P = 2.2 \times 10^{-16}$), with the second highest correlation between $MRS_{carotid}$ and PAI-1 (r = 0.24, $P = 1.1 \times 10^{-6}$).

Associations between previously identified atherosclerosis-associated CpGs and atherosclerosis

CAC, AAC, ABI, and multisite atherosclerosis were associated with 11, 17, 4, and 21 CpGs in Model 1, respectively (FDR < 0.1). The regression results are shown in **Supplemental Table 2**. One CpG, cg05246522, was a CAC-associated CpG in MESA and the remainder of the significant CpGs in GENOA were carotid plaque-associated CpGs in MESA. **Supplemental Table 3** shows the regression results for the CpGs that remained significantly associated with atherosclerosis after adjustment for cardiovascular risk factors (Model 3) in GENOA. After adjusting for cardiovascular risk factors, 6 CpGs (cg05575921 (*AHRR*), cg04761231 (*RPL35*), cg08958747 (*RAB26*), cg09935388 (*GFI1*), cg16661609 (*LILRB4*), and cg21161138 (*AHRR*)) remained significantly associated with multisite atherosclerosis, and cg05575921 was also associated with AAC (FDR < 0.1). All of these CpGs were carotid plaque-associated CpGs in MESA. Hypermethylation at all of the significant CpGs (FDR < 0.1 in Model 3) was associated with decreased multisite atherosclerosis. Significant CpG,s in Model 3 explained between 0.3% and 6.5% of the variability of multisite atherosclerosis, and cg05575921 explained 2.7% of the variability of AAC.

Given the previous reports of an association between three of the identified CpGs (cg05575921, cg21161138, and cg09935388) and smoking, we examined their methylation levels by smoking status in GENOA. As expected, the plots show a dose response effect where

current smokers had the lowest methylation levels and never smokers had the highest methylation levels (**Supplemental Figure 4**).

Associations between methylation risk scores and atherosclerosis

Table 2 shows the associations between MRS_{CAC} and $MRS_{carotid}$ and single or multisite atherosclerosis measures. The beta coefficients shown in the tables correspond to the change in the atherosclerosis measures associated with a 1-unit increase in the MRS. In Model 1, MRS_{CAC} was associated with log-transformed CAC (Beta = 0.421, 95%CI 0.080–0.762, $P = 0.016$) but the association was not significant after adjusting for CVD risk factors (Models 2 and 3). $MRS_{carotid}$ was associated with both log-transformed CAC (Beta = 0.778, 95%CI 0.433–1.12, $P = 1.26 \times 10^{-5}$) and log-transformed AAC (Beta = 1.062, 95%CI 0.688–1.44, $P = 5.13 \times 10^{-8}$). Associations remained associated after adjusting for smoking (Model 2) and other traditional CVD risk factors (Model 3), where a one unit increase in $MRS_{carotid}$ was associated with a 0.479 units increase in log transformed CAC (95%CI 0.083–0.875, $P = 0.018$) and a 0.551 units increase in log-transformed AAC (95%CI 0.118–0.984). This is equivalent to an approximately 1.6-fold increase in the Agatston score of CAC and AAC. A one unit increase in $MRS_{carotid}$ was associated with a 0.7 units (95%CI 0.21–1.13) increase in the multisite atherosclerosis score after adjusting for CVD risk factors (Model 3). The $MRS_{carotid}$ explained 5.3%, 2.7%, 0.29%, and 5.5% of the variability of CAC, AAC, ABI, and multisite atherosclerosis after adjusting for CVD risk factors.

Associations between epigenetic age acceleration and atherosclerosis

Associations between the epigenetic age acceleration measures and atherosclerosis are shown in **Table 3**. PhenoAA was associated with multisite atherosclerosis in Model 1, but the association attenuated after adjusting for CVD risk factors. GrimAA was positively associated

with all measures of atherosclerosis in the minimally adjusted model (Model 1) and remained significant after adjusting for CVD risk factors for AAC and multisite atherosclerosis. A 5-year increase in GrimAA (~1 SD) was associated with a 1.6-fold (95% CI 1.11–2.25) increase in the Agatston score of AAC and a 0.7 units (95% CI 0.33–1.07) increase in multisite atherosclerosis score in Model 3. Increased GrimAA was nominally associated with lower ABI (higher atherosclerosis), but the association was not significant after accounting for multiple testing. The effect estimates were unchanged after adjusting for the first 4 genetic PCs. **Supplemental Table 4** shows the associations for PhenoAA and GrimAA after adjusting for white blood cell counts. The associations between GrimAA and AAC and multisite atherosclerosis in Model 3 slightly attenuated after adjusting for white blood cells counts but remained significant at $P < 0.05$. The associations between the DNA methylation based surrogate measures comprising GrimAge and atherosclerosis are shown in **Supplemental Table 5.** Associations were adjusted for age, sex, time between measures, and white blood cell counts. All components were associated with at least one single or multi-site atherosclerosis measure, with the exception of leptin and TIMP-1 ($P < 0.05$). DNAm smoking pack years was associated with each of the atherosclerosis measures and was the most significant predictor of CAC, AAC and multisite atherosclerosis, with consistent effect directions across all of the measures. Compared to the corresponding model using GrimAA **(Supplemental Table 3; Model 1)**, DNAm smoking pack years was more significantly associated with single- and multi-site atherosclerosis, with greater magnitude of effects, than the overall GrimAA measure.

Longitudinal correlation of methylation risk scores and epigenetic age acceleration measures between Phases I and II

Compared to the remaining sample, the 129 individuals with longitudinal measures of DNA methylation were younger (mean age of 53.4 vs. 57.3 years, $P < 0.001$), had a lower

multisite atherosclerosis score (mean of 4.51 vs. 5.26, $P = 0.026$) and a lower PhenoAA (mean of -1.11 vs. 0.41, $P = 0.043$). All the remaining epigenetic and atherosclerosis measures were similar across both samples ($P > 0.05$). The mean time difference between the DNA methylation measurements was 5.5 years (standard deviation = 1.1 years). All of the epigenetic measures were relatively stable between Phases I and II, with both $MRS_{carotid}$ and GrimAA showing the highest stability (ICC > 0.8), and MRS_{CAC} showing the lowest stability (ICC = 0.519) (**Table 4**). $MRS_{carotid}$ derived at Phase II was associated with consistent effect direction but reduced effect magnitude for log-transformed AAC (Beta = 0.894, 95% CI 0.241–1.55, $P = 0.008$) and multisite atherosclerosis (Beta = 0.891, 95% CI 0.238–1.54, $P = 0.009$) compared to the full sample analysis using Phase I MRSs **(Table 2, Model 1).** Neither MRS was associated with CAC or ABI at Phase II **(Supplemental Table 6)**. For the epigenetic age acceleration measures from Phase II, only GrimAA was associated with multisite atherosclerosis score (Beta = 0.123, 95%CI 0.015–0.231, $P = 0.028$ in Model 1). Supplemental Figure 5 shows the change MRSs and epigenetic age acceleration measures by age. As expected, both MRSs were significantly and positively associated with age ($p<0.05$), while the epigenetic acceleration measures were not.

4.5 Discussion

In this study of African Americans, we examined the association between whole blood DNA methylation patterns, measured at baseline, and multisite atherosclerosis assessed 5 and 12 years later. After adjusting for CVD risk factors, six CpGs were associated with multisite atherosclerosis and an aggregate risk score ($MRS_{carotid}$) was associated with CAC, AAC, and multisite atherosclerosis. The mean of $MRS_{carotid}$ was negative, indicative of a general trend of hypomethylation at the individual CpGs associated with carotid plaque. GrimAA was the only epigenetic age acceleration measure associated with atherosclerosis. GrimAA and $MRS_{carotid}$

were moderately correlated with each other, and both measures were relatively stable over 5 years in a small subset of participants with repeated DNA methylation measurement. Although our findings cannot establish causality, the relatively long time period between the DNA methylation and atherosclerosis assessments help establish temporal patterning between changes in DNA methylation and the development of subclinical markers of CVD. Our findings additionally highlight the role of age-related methylation changes, as measured by the epigenetic clocks, and vascular calcification.

After adjusting for CVD risk factors, 1 CpG was associated with AAC, and another 5 were also associated with multisite atherosclerosis at FDR < 0.1. The direction of effects between the individual CpG sites and atherosclerosis in our study were consistent with MESA, where hypomethylation was associated with increased carotid plaque. Three of the significant CpGs (cg05575921, cg21161138, and cg09935388) were previously found to be associated with smoking.[29-32] Both cg05575921 and cg21161138 are located in the *AHRR* gene body. *AHRR* is an aryl hydrocarbon receptor repressor, which among other roles, inhibits the metabolism of polycyclic aromatic hydrocarbons and dioxins by competing with AHR.[53,54] The significant association between these CpGs and atherosclerosis after adjusting for smoking status in our study could potentially be related to residual confounding, errors in self-reporting of smoking, and/or interindividual sensitivities to smoking with lasting biological effects. In a previous analysis in MESA that used a candidate gene approach to assess the association between CpG sites in *AHRR* and atherosclerosis, hypomethylation at cg05575921 ($P = 3.08\times10^{-10}$) and cg21161138 ($P = 7.73\times10^{-8}$) was significantly associated with carotid plaque score.[55] Similar to our findings, the association with cg05575921 remained significant after adjusting for self-reported smoking exposure, urinary cotinine, and other CVD risk factors, and remained

significant in stratified analysis of former smokers and current smokers but not never smokers. Other studies have also reported evidence of cg05575921 differential methylation by air pollution in adults,[56] by maternal smoking in neonates,[57,58] and by smoking in atherosclerotic plaque specimens.[59]

Similarly, cg09935388 located in the *GFI1* gene body has been found to be associated with smoking [32,60,61] and exposure to maternal smoking in fetuses.[62,63] The growth factor independent 1 transcriptional repressor gene, *GFI1*, encodes a nuclear zinc finger protein that plays a role in hematopoiesis, oncogenesis, and in controlling histone modification as part of a complex with other cofactors.[64,65] cg04761231 is located in the gene body of *RPL35*, which encodes a ribosomal protein, and was also associated with smoking.[66,67] cg08958747 is located in the gene body region of *RAB26*, a member of the RAB protein family which are important regulators of vesicular fusion and trafficking.[68] cg16661609 is located upstream of a transcription start site of the *LILRB4* gene. *LILRB4* (leukocyte immunoglobulin like receptor B4) is a member of the leukocyte immunoglobulin-like receptor family. The receptor is expressed on immune cells where it transduces a negative signal that inhibits stimulation of an immune response and controls inflammatory responses and cytotoxicity to help focus the immune responses. One notable finding in MESA was that both the transcriptome signature of AT-rich interaction domain 5B (*ARID5B*) and a cg25953130 site in the gene were associated with CAC and carotid plaque atherosclerosis measures.[27] In GENOA, cg25953130 was only nominally associated with multisite atherosclerosis after adjusting for CVD risk factors (Beta = -0.625, 95%CI 0.17–1.23, $P = 0.044$), but the effect direction was consistent across studies.

In our study, MRS_{CAC} was associated only with CAC, and the association became non-significant after adjusting for CVD risk factors. $MRS_{carotid}$ was associated with CAC, AAC, and

the derived multisite atherosclerosis score. Two of the significant CpGs in $MRS_{carotid}$ explained a slightly higher percent of the variability of multisite atherosclerosis compared to that explained by $MRS_{carotid,}$ potentially due to added noise generated by including CpGs not as strongly associated with atherosclerosis in African Americans. It is not clear why in our sample, carotid plaque-associated CpGs were better predictors of atherosclerosis than CAC-associated ones. In MESA, out of 7 CpGs associated with CAC (FDR < 0.05) in the full sample, only 3 were significant in the African American sub-sample, although the magnitude and direction of effect were consistent across ancestry groups. Also, despite having similar age distributions, MESA multi-ethnic participants had a higher median CAC score than African Americans in GENOA, and MESA also had a lower percentage of females. The prevalence of CAC in GENOA was low (median: 18.7 and IQR: 0 – 195.7) with about 40% of the sample having an Agatston score of zero while in MESA, participants had a median CAC score of 46 (IQR: 0 – 305), and the proportion with no CAC was not reported. This is consistent with evidence from epidemiological studies showing that African Americans tend to have lower calcification in the coronary arteries compared to whites.[69-71] Hence, the CpGs included in the MRS_{CAC} could be related to a more extreme form of the trait versus the lower atherosclerosis burden seen in GENOA. Another difference between MESA and GENOA is that the methylation signature in MESA was measured in monocytes, which have a well-established role in atherogenesis,[72,73] while the DNA methylation in GENOA was measured in all white blood cells. This could have the effect of diluting the associations observed in MESA if some cell types have different methylation patterns compared to monocytes.[27] Additionally, the ICC of MRS_{CAC} was low, which could indicate a lower stability of methylation at these sites. This could be particularly relevant to

GENOA, as we assessed the association with CAC 12 years after DNA methylation measurement, while in MESA methylation and atherosclerosis were assessed concurrently.

GrimAA was the only epigenetic age acceleration measure associated with atherosclerosis in GENOA. Increased GrimAA, indicative of increased biological aging, was associated with higher atherosclerosis. Very few other studies have examined epigenetic age associations with subclinical measures of CVD. In a cross-sectional analysis of 2,500 African Americans from the Atherosclerosis Risk in Communities (ARIC) study, a 5-year increase in both Horvath and the Hannum acceleration measures was associated with an approximately 0.01 mm increase in carotid intima thickness.[28] We did not find associations between the Horvath and Hannum measures we evaluated (IEAA and EEAA) and atherosclerosis, and in general our findings were varied across our four epigenetic age acceleration measures. These measures include different CpG sites, differ in how they were trained, and are hypothesized to capture different biological processes and aspects of aging.[18,46] Both GrimAA and PhenoAA were trained using longitudinal data, making them better predictors of aging-related outcomes.[74,75] Two recent studies of participants of European ancestry have additionally shown that GrimAA outperforms the other acceleration measures in its association with incident CVD [25,26] and all-cause mortality.[25,76] Our reported associations attenuated slightly after adjusting for white blood cell counts but remained significant, suggesting that the effects were not mediated by blood cell composition. In our study, GrimAA was correlated with $MRS_{carotid}$, with the smoking pack-years component of GrimAge being slightly more correlated with $MRS_{carotid}$ than the overall GrimAA measure. While we do not know individual the CpGs comprising GrimAge, we know of one CpG overlap, cg05575921,[17] between GrimAge and $MRS_{carotid}$, and potentially other smoking related CpGs, which partially explains the observed correlations.

Our results show that the epigenetic measures most strongly associated with atherosclerosis, GrimAA and $MRS_{carotid}$, had moderate stability (ICC: 0.822 and 0.888, respectively) across repeated samples taken approximately 5 years apart. The characteristic of reliability over time is an important consideration for biomarkers that may be used for risk prediction. Further studies are needed to fully characterize the longitudinal patterns of DNA methylation, especially in response to known drivers of DNA methylation changes, such as smoking. In GENOA, smoking status was unchanged for the majority of the sample between Phases I and III, with only 27 current smokers becoming former smokers. One study that looked at the longitudinal changes of fetal DNA methylation in response to maternal smoking using serial samples at birth, age 7, and age 17 found evidence of reversible methylation changes at cg09935388 (*GFI1*) and persistent methylation changes at cg05575921 (*AHRR*).[63] An epigenome-wide study of adult smoking reported that out of approximately 2,600 CpGs that were differentially methylated between current versus never smokers, 185 CpGs showed patterns of persistent methylation changes between former versus never smokers, including cg05575921, cg09935388, and cg21161138 CpGs.[66] Most recently, Dugue et al. reported a reversibility coefficient (ratio of regression coefficients comparing former to current smokers and never to current smokers) between 69% and 75% at cg05575921, cg09935388, and cg21161138.[77] Little is known regarding the longitudinal trends of GrimAA; however, a longitudinal trend of increased GrimAA with increasing age has been observed in one study.[78]

One strength of our study is that we examined atherosclerosis at multiple vascular sites which reflect both intimal and medial vascular changes that may manifest differentially over time. CAC has been more extensively studied because it appears in a more clinically relevant vascular site [79] and is more strongly associated with coronary disease compared to carotid intima

thickness.[80] Abdominal aorta atherosclerosis has been less extensively studied, yet recent evidence suggests that the associations of CAC and AAC with CVD are independent and additive.[8,13] AAC starts earlier in life and is more prevalent than CAC and may be associated with an increased risk of onset and progression of CAC and/or lower ABI.[81,82] Given the low prevalence of CAC and previous findings that it may not carry the same pathobiologic significance in African Americans,[83] incorporating extracoronary calcification may be more informative and useful for risk assessment. A limitation of our study is the attrition of participants between Phases I and III. In a previous work, we have noted that participants who were lost to follow-up had higher epigenetic age acceleration and higher CVD risk.[84] Additionally, we did not have measures of carotid plaque which was included in the multisite atherosclerosis score that we used as a model for our score in GENOA,[45] and the GENOA atherosclerosis measures were not all assessed concurrently. Furthermore, seven significant CpGs from MESA were not available in GENOA because of the different arrays used.

In conclusion, our study found evidence of associations between DNA methylation and atherosclerosis at multiple vascular sites after accounting for traditional CVD risk factors. DNA methylation changes were at CpGs with inflammatory and smoking-related regulatory functions, which further highlights the relevance of inflammation and smoking in atherosclerosis. Despite being derived from CpGs associated with carotid plaque, $MRS_{carotid}$ was associated with atherosclerosis in the coronary arteries and abdominal aorta suggesting common pathobiological mechanisms of atherosclerosis on a systemic level. GrimAA was also associated with multisite atherosclerosis beyond traditional CVD risk factors. This is one of the very few studies to examine the DNA methylation signature of multiple atherosclerosis measures in a population-based cohort. These results further our understanding of the relationship between aging and

atherosclerosis. Further work in this area may lead to better prediction of those at increased risk for atherosclerosis.

4.6 References

1. Virani SS, Alonso A, Aparicio HJ, et al. Heart Disease and Stroke Statistics-2021 Update: A Report From the American Heart Association. *Circulation.* 2021;143(8):e254-e743.
2. Van Dyke M, Greer S, Odom E, et al. Heart disease death rates among blacks and whites aged≥ 35 years—United States, 1968–2015. *MMWR Surveillance Summaries.* 2018;67(5):1.
3. Mackey RH, Venkitachalam L, Sutton-Tyrrell K. Calcifications, arterial stiffness and atherosclerosis. *Adv Cardiol.* 2007;44:234-244.
4. Ross R. Atherosclerosis--an inflammatory disease. *N Engl J Med.* 1999;340(2):115-126.
5. Libby P, Buring JE, Badimon L, et al. Atherosclerosis. *Nat Rev Dis Primers.* 2019;5(1):56.
6. Frostegard J. Immunity, atherosclerosis and cardiovascular disease. *BMC Med.* 2013;11:117.
7. Criqui MH, Langer RD, Fronek A, et al. Mortality over a period of 10 years in patients with peripheral arterial disease. *N Engl J Med.* 1992;326(6):381-386.
8. Criqui MH, Denenberg JO, McClelland RL, et al. Abdominal aortic calcium, coronary artery calcium, and cardiovascular morbidity and mortality in the Multi-Ethnic Study of Atherosclerosis. *Arterioscler Thromb Vasc Biol.* 2014;34(7):1574-1579.
9. Shaw LJ, Raggi P, Schisterman E, Berman DS, Callister TQ. Prognostic value of cardiac risk factors and coronary artery calcium screening for all-cause mortality. *Radiology.* 2003;228(3):826-833.
10. Kondos GT, Hoff JA, Sevrukov A, et al. Electron-beam tomography coronary artery calcium and cardiac events: a 37-month follow-up of 5635 initially asymptomatic low- to intermediate-risk adults. *Circulation.* 2003;107(20):2571-2576.
11. Ho CY, Shanahan CM. Medial Arterial Calcification: An Overlooked Player in Peripheral Arterial Disease. *Arterioscler Thromb Vasc Biol.* 2016;36(8):1475-1482.
12. Lanzer P, Boehm M, Sorribas V, et al. Medial vascular calcification revisited: review and perspectives. *Eur Heart J.* 2014;35(23):1515-1525.
13. Bastos Goncalves F, Voute MT, Hoeks SE, et al. Calcification of the abdominal aorta as an independent predictor of cardiovascular events: a meta-analysis. *Heart.* 2012;98(13):988-994.
14. Horvath S. DNA methylation age of human tissues and cell types. *Genome Biol.* 2013;14(10):R115.
15. Hannum G, Guinney J, Zhao L, et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol Cell.* 2013;49(2):359-367.
16. Levine ME, Lu AT, Quach A, et al. An epigenetic biomarker of aging for lifespan and healthspan. *Aging (Albany NY).* 2018;10(4):573-591.

17. Lu AT, Quach A, Wilson JG, et al. DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging (Albany NY).* 2019;11(2):303-327.

18. Liu Z, Leung D, Levine M. Comparative analysis of epigenetic aging clocks from CpG characteristics to functional associations. *bioRxiv.* 2019:512483.

19. Fransquet PD, Wrigglesworth J, Woods RL, Ernst ME, Ryan J. The epigenetic clock as a predictor of disease and mortality risk: a systematic review and meta-analysis. *Clin Epigenetics.* 2019;11(1):62.

20. Rask-Andersen M, Martinsson D, Ahsan M, et al. Epigenome-wide association study reveals differential DNA methylation in individuals with a history of myocardial infarction. *Hum Mol Genet.* 2016;25(21):4739-4748.

21. Fernández-Sanlés A, Sayols-Baixeras S, Subirana I, et al. DNA Methylation Biomarkers Of Myocardial Infarction And Cardiovascular Disease. *bioRxiv.* 2020:707315.

22. Nakatochi M, Ichihara S, Yamamoto K, et al. Epigenome-wide association of myocardial infarction with DNA methylation sites at loci related to cardiovascular disease. *Clin Epigenetics.* 2017;9:54.

23. Agha G, Mendelson MM, Ward-Caviness CK, et al. Blood Leukocyte DNA Methylation Predicts Risk of Future Myocardial Infarction and Coronary Heart Disease. *Circulation.* 2019;140(8):645-657.

24. Horvath S, Gurven M, Levine ME, et al. An epigenetic clock analysis of race/ethnicity, sex, and coronary heart disease. *Genome Biol.* 2016;17(1):171.

25. Hillary RF, Stevenson AJ, McCartney DL, et al. Epigenetic measures of ageing predict the prevalence and incidence of leading causes of death and disease burden. *Clin Epigenetics.* 2020;12(1):115.

26. Wang C, Ni W, Yao Y, et al. DNA methylation-based biomarkers of age acceleration and all-cause death, myocardial infarction, stroke, and cancer in two cohorts: The NAS, and KORA F4. *EBioMedicine.* 2020;63:103151.

27. Liu Y, Reynolds LM, Ding J, et al. Blood monocyte transcriptome and epigenome analyses reveal loci associated with human atherosclerosis. *Nat Commun.* 2017;8(1):393.

28. Roetker NS, Pankow JS, Bressler J, Morrison AC, Boerwinkle E. Prospective Study of Epigenetic Age Acceleration and Incidence of Cardiovascular Disease Outcomes in the ARIC Study (Atherosclerosis Risk in Communities). *Circ Genom Precis Med.* 2018;11(3):e001937.

29. Zhang Y, Wilson R, Heiss J, et al. DNA methylation signatures in peripheral blood strongly predict all-cause mortality. *Nat Commun.* 2017;8:14617.

30. Gao X, Jia M, Zhang Y, Breitling LP, Brenner H. DNA methylation changes of whole blood cells in response to active smoking exposure in adults: a systematic review of DNA methylation studies. *Clin Epigenetics.* 2015;7:113.

31. Zhang Y, Schottker B, Florath I, et al. Smoking-Associated DNA Methylation Biomarkers and Their Predictive Value for All-Cause and Cardiovascular Mortality. *Environ Health Perspect.* 2016;124(1):67-74.

32. Zeilinger S, Kuhnel B, Klopp N, et al. Tobacco smoking leads to extensive genome-wide changes in DNA methylation. *PLoS One.* 2013;8(5):e63812.

33. Daniels PR, Kardia SL, Hanis CL, et al. Familial aggregation of hypertension treatment and control in the Genetic Epidemiology Network of Arteriopathy (GENOA) study. *Am J Med.* 2004;116(10):676-681.

34. Turner ST, Kardia SL, Mosley TH, Rule AD, Boerwinkle E, de Andrade M. Influence of genomic loci on measures of chronic kidney disease in hypertensive sibships. *J Am Soc Nephrol.* 2006;17(7):2048-2055.

35. Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. *Clin Chem.* 1972;18(6):499-502.

36. Ammous F, Zhao W, Ratliff SM, et al. Epigenome-wide association study identifies DNA methylation sites associated with target organ damage in older African Americans. *Epigenetics.* 2020:1-14.

37. Fortin JP, Fertig E, Hansen K. shinyMethyl: interactive quality control of Illumina 450k DNA methylation arrays in R. *F1000Res.* 2014;3:175.

38. Lehne B, Drong AW, Loh M, et al. A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies. *Genome Biol.* 2015;16:37.

39. Fortin JP, Triche TJ, Jr., Hansen KD. Preprocessing, normalization and integration of the Illumina HumanMethylationEPIC array with minfi. *Bioinformatics.* 2017;33(4):558-560.

40. Niu L, Xu Z, Taylor JA. RCP: a novel probe design bias correction method for Illumina Methylation BeadChip. *Bioinformatics.* 2016;32(17):2659-2663.

41. Houseman EA, Accomando WP, Koestler DC, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics.* 2012;13:86.

42. DNA Methylation Age Calculator. https://dnamage.genetics.ucla.edu/. Accessed 06 Nov, 2020.

43. Chen BH, Marioni RE, Colicino E, et al. DNA methylation-based measures of biological age: meta-analysis predicting time to death. *Aging (Albany NY).* 2016;8(9):1844-1865.

44. Kullo IJ, Turner ST, Kardia SL, Mosley TH, Jr., Boerwinkle E, de Andrade M. A genome-wide linkage scan for ankle-brachial index in African American and non-Hispanic white subjects participating in the GENOA study. *Atherosclerosis.* 2006;187(2):433-438.

45. Zhao Y, Evans MA, Allison MA, et al. Multisite atherosclerosis in subjects with metabolic syndrome and diabetes and relation to cardiovascular events: The Multi-Ethnic Study of Atherosclerosis. *Atherosclerosis.* 2019;282:202-209.

46. Horvath S, Raj K. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. *Nat Rev Genet.* 2018;19(6):371-384.

47. Nakagawa S, Schielzeth H. Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. *Biol Rev Camb Philos Soc.* 2010;85(4):935-956.

48. Stoffel MA, Nakagawa S, Schielzeth H. rptR: Repeatability estimation and variance decomposition by generalized linear mixed-effects models. *Methods in Ecology and Evolution.* 2017;8(11):1639-1644.

49. Storey JD, Tibshirani R. Statistical significance for genomewide studies. *Proc Natl Acad Sci U S A.* 2003;100(16):9440-9445.

50. Team RC. R: A language and environment for statistical computing. In: Vienna, Austria; 2019.

51. Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using lme4. *2015.* 2015;67(1):48.

52. Højsgaard SH, Ulrich; Yan, Jun. The R Package geepack for Generalized Estimating Equations. *Journal of Statistical Software.* 2005;15(i02).

53. Esser C. Biology and function of the aryl hydrocarbon receptor: report of an international and interdisciplinary conference. *Arch Toxicol.* 2012;86(8):1323-1329.

54. Vogel CFA, Haarmann-Stemmann T. The aryl hydrocarbon receptor repressor - More than a simple feedback inhibitor of AhR signaling: Clues for its role in inflammation and cancer. *Curr Opin Toxicol.* 2017;2:109-119.

55. Reynolds LM, Wan M, Ding J, et al. DNA Methylation of the Aryl Hydrocarbon Receptor Repressor Associations With Cigarette Smoking and Subclinical Atherosclerosis. *Circ Cardiovasc Genet.* 2015;8(5):707-716.

56. Tantoh DM, Wu MC, Chuang CC, et al. AHRR cg05575921 methylation in relation to smoking and PM2.5 exposure among Taiwanese men and women. *Clin Epigenetics.* 2020;12(1):117.

57. Novakovic B, Ryan J, Pereira N, Boughton B, Craig JM, Saffery R. Postnatal stability, tissue, and time specific effects of AHRR methylation change in response to maternal smoking in pregnancy. *Epigenetics.* 2014;9(3):377-386.

58. Wiklund P, Karhunen V, Richmond RC, et al. DNA methylation links prenatal smoking exposure to later life health outcomes in offspring. *Clin Epigenetics.* 2019;11(1):97.

59. Siemelink MA, van der Laan SW, Haitjema S, et al. Smoking is Associated to DNA Methylation in Atherosclerotic Carotid Lesions. *Circ Genom Precis Med.* 2018;11(9):e002030.

60. Besingi W, Johansson A. Smoke-related DNA methylation changes in the etiology of human disease. *Hum Mol Genet.* 2014;23(9):2290-2297.

61. Parmar P, Lowry E, Cugliari G, et al. Association of maternal prenatal smoking GFI1-locus and cardio-metabolic phenotypes in 18,212 adults. *EBioMedicine.* 2018;38:206-216.

62. Joubert BR, Haberg SE, Nilsen RM, et al. 450K epigenome-wide scan identifies differential DNA methylation in newborns related to maternal smoking during pregnancy. *Environ Health Perspect.* 2012;120(10):1425-1431.

63. Richmond RC, Simpkin AJ, Woodward G, et al. Prenatal exposure to maternal smoking and offspring DNA methylation across the lifecourse: findings from the Avon

Longitudinal Study of Parents and Children (ALSPAC). *Hum Mol Genet.* 2015;24(8):2201-2217.

64. van der Meer LT, Jansen JH, van der Reijden BA. Gfi1 and Gfi1b: key regulators of hematopoiesis. *Leukemia.* 2010;24(11):1834-1843.

65. Duan Z, Zarebski A, Montoya-Durango D, Grimes HL, Horwitz M. Gfi1 coordinates epigenetic repression of p21Cip/WAF1 by recruitment of histone lysine methyltransferase G9a and histone deacetylase 1. *Mol Cell Biol.* 2005;25(23):10338-10351.

66. Joehanes R, Just AC, Marioni RE, et al. Epigenetic Signatures of Cigarette Smoking. *Circ Cardiovasc Genet.* 2016;9(5):436-447.

67. Sikdar S, Joehanes R, Joubert BR, et al. Comparison of smoking-related DNA methylation between newborns from prenatal exposure and adults from personal smoking. *Epigenomics.* 2019;11(13):1487-1500.

68. Seki N, Yoshikawa T, Hattori A, Miyajima N, Muramatsu M, Saito T. cDNA cloning of a human RAB26-related gene encoding a Ras-like GTP-binding protein on chromosome 16p13.3 region. *J Hum Genet.* 2000;45(5):309-314.

69. Wassel CL, Pankow JS, Peralta CA, Choudhry S, Seldin MF, Arnett DK. Genetic ancestry is associated with subclinical cardiovascular disease in African-Americans and Hispanics from the multi-ethnic study of atherosclerosis. *Circ Cardiovasc Genet.* 2009;2(6):629-636.

70. Kawakubo M, LaBree L, Xiang M, et al. Race-ethnic differences in the extent, prevalence, and progression of coronary calcium. *Ethn Dis.* 2005;15(2):198-204.

71. Manolio TA, Arnold AM, Post W, et al. Ethnic differences in the relationship of carotid atherosclerosis to coronary calcification: the Multi-Ethnic Study of Atherosclerosis. *Atherosclerosis.* 2008;197(1):132-138.

72. Pamukcu B, Lip GY, Devitt A, Griffiths H, Shantsila E. The role of monocytes in atherosclerotic coronary artery disease. *Ann Med.* 2010;42(6):394-403.

73. Ghattas A, Griffiths HR, Devitt A, Lip GY, Shantsila E. Monocytes in coronary artery disease and atherosclerosis: where are we now? *J Am Coll Cardiol.* 2013;62(17):1541-1551.

74. Levine ME. Assessment of Epigenetic Clocks as Biomarkers of Aging in Basic and Population Research. *J Gerontol A Biol Sci Med Sci.* 2020;75(3):463-465.

75. Nelson PG, Promislow DEL, Masel J. Biomarkers for Aging Identified in Cross-sectional Studies Tend to Be Non-causative. *J Gerontol A Biol Sci Med Sci.* 2020;75(3):466-472.

76. McCrory C, Fiorito G, Hernandez B, et al. GrimAge outperforms other epigenetic clocks in the prediction of age-related clinical phenotypes and all-cause mortality. *J Gerontol A Biol Sci Med Sci.* 2020.

77. Dugue PA, Jung CH, Joo JE, et al. Smoking and blood DNA methylation: an epigenome-wide association study and assessment of reversibility. *Epigenetics.* 2019:1-11.

78. Li X, Ploner A, Wang Y, et al. Longitudinal trajectories, correlations and mortality associations of nine biological ages across 20-years follow-up. *Elife.* 2020;9.

79. Bielak LF, Peyser PA. Genetics of Subclinical Coronary Atherosclerosis. *Curr Genet Med Rep.* 2018;6(3):116-123.

80. Gepner AD, Young R, Delaney JA, et al. Comparison of Carotid Plaque Score and Coronary Artery Calcium Score for Predicting Cardiovascular Disease Events: The Multi-Ethnic Study of Atherosclerosis. *J Am Heart Assoc.* 2017;6(2).

81. Wong ND, Lopez VA, Allison M, et al. Abdominal aortic calcium and multi-site atherosclerosis: the Multiethnic Study of Atherosclerosis. *Atherosclerosis.* 2011;214(2):436-441.

82. Onuma OK, Pencina K, Qazi S, et al. Relation of Risk Factors and Abdominal Aortic Calcium to Progression of Coronary Artery Calcium (from the Framingham Heart Study). *Am J Cardiol.* 2017;119(10):1584-1589.

83. Doherty TM, Tang W, Detrano RC. Racial differences in the significance of coronary calcium in asymptomatic black and white subjects with coronary risk factors. *J Am Coll Cardiol.* 1999;34(3):787-794.

84. Ammous F, Zhao W, Ratliff SM, et al. Epigenetic age acceleration is associated with cardiometabolic risk factors and clinical cardiovascular disease risk scores in African Americans. *Clin Epigenetics.* 2021;13(1):55.

4.7 Tables

Table 4-1: Descriptive characteristics of GENOA African Americans

Characteristics[a]	**Overall (N = 391)**
Females (%)	297 (76.0%)
Age at Phase I (years)	56.0 ± 9.0
Age at Phase III (years)	68.0 ± 8.4
Cardiovascular risk factors at Phase I	
Smoking status	
Never (%)	241 (61.6%)
Former (%)	93 (23.8%)
Current (%)	57 (14.6%)
Body mass index (kg/m^2)	31.4 ± 6.2
Hypertension	249 (63.7%)
Type 2 diabetes	63 (16.1%)
Total cholesterol[b] (mg/dl)	208.56 ± 50.4
Low density lipoprotein[c] (mg/dl)	121.8 (42.8)
Statin use (%)	18 (4.6%)
Atherosclerosis measures at Phase III	
Coronary artery calcium score, median (IQR)	18.7 (0 – 195.7)
Abdominal aorta calcium score, median (IQR)	500.8 (18.9 – 1800.5)
Ankle-brachial index, median (range)[d]	0.985 (0.474 – 1.274)
Multisite atherosclerosis score, median (IQR) (range: 0 – 12)	5 (2 – 7)
Methylation risk scores at Phase I	
MRS_{CAC} [e]	0.86 ± 0.71
$MRS_{carotid}$	-3.60 ± 0.71
Epigenetic age acceleration at Phase I	
Horvath (years)	52.8 ± 8.9
Hannum (years)	46.2 ± 9.5
PhenoAge (years)	42.7 ± 11.5
GrimAge (years)	52.8 ± 7.9

, intrinsic epigenetic age acceleration; EEAA, extrinsic epigenetic age acceleration; IQR, interquartile range

[a] Means ± standard deviation

[b] Adjusted for statin use as total cholesterol/0.8

[c] Calculated using the Freidwald formula for participants with triglycerides < 400 mg/dl (N=386)

[d] Measured at Phase II (mean: 5.2 ± 1.3 years from Phase I)

[e] N = 390, after removing one outlier beyond 5 standard deviation units from the mean

Table 4-2: Association between methylation risk scores and atherosclerosis measures in GENOA African Americans

		Model 1			**Model 2**			**Model 3**		
Outcome	**Methylation risk score**	**Beta**	**SE**	***P***	**Beta**	**SE**	***P***	**Beta**	**SE**	***P***
Coronary artery calcium score (CAC)[a]	MRS_{CAC}	**0.421**	**0.174**	**0.016**	0.330	0.173	0.057	0.290	0.165	0.079
	$MRS_{carotid}$	**0.778**	**0.176**	**1.26×10^{-5}**	**0.484**	**0.211**	**0.022**	**0.479**	**0.202**	**0.018**
Abdominal aorta calcification score (AAC)[a]	MRS_{CAC}	**0.448**	**0.191**	**0.020**	0.268	0.184	0.145	0.267	0.180	0.138
	$MRS_{carotid}$	**1.062**	**0.191**	**5.13×10^{-8}**	**0.527**	**0.224**	**0.019**	**0.551**	**0.221**	**0.013**
Ankle-brachial index (ABI)	MRS_{CAC}	0.012	0.007	0.074	0.014	0.007	0.041	0.014	0.007	0.042
	$MRS_{carotid}$	-0.012	0.007	0.085	-0.005	0.008	0.586	-0.003	0.008	0.717
Multisite atherosclerosis score	MRS_{CAC}	0.443	0.204	0.030	0.295	0.196	0.134	0.283	0.193	0.143
	$MRS_{carotid}$	**1.196**	**0.202**	**7.58×10^{-9}**	**0.679**	**0.238**	**0.005**	**0.669**	**0.236**	**0.005**

Model 1 is adjusted for age, sex, time between measures, and 4 genetic principal components
Model 2 is adjusted for Model 1 covariates and smoking status
Model 3 is adjusted for Model 2 covariates, hypertension status, diabetes status, body mass index, and total cholesterol levels adjusted for statin use
Beta is the change in the atherosclerosis measure associated with a 1 unit increase in the MRS.
P-values significant after Bonferroni correction ($P < 0.025$) are shown in bold font
[a] Coronary artery calcium score and abdominal aorta calcification score were transformed as ln[(CAC+1)] and ln[(AAC+1)]

Table 4-3: Associations between epigenetic age acceleration and atherosclerosis measures in GENOA African Americans

		Model 1			**Model 2**			**Model 3**		
Outcome	**Epigenetic age acceleration**	Beta	SE	*P*	Beta	SE	*P*	Beta	SE	*P*
Coronary artery calcium score (CAC)[a]	IEAA	0.016	0.027	0.555	0.008	0.026	0.750	-0.008	0.025	0.741
	EEAA	0.018	0.023	0.436	0.012	0.022	0.589	0.003	0.021	0.901
	PhenoAA	0.034	0.018	0.056	0.023	0.017	0.191	0.010	0.017	0.532
	GrimAA	**0.149**	**0.028**	**2.04×10^{-7}**	**0.105**	**0.034**	**0.002**	0.074	0.033	0.025
Abdominal aorta calcification score (AAC)[a]	IEAA	-0.016	0.029	0.591	-0.025	0.028	0.366	-0.040	0.027	0.139
	EEAA	0.027	0.025	0.285	0.024	0.024	0.311	0.018	0.024	0.437
	PhenoAA	0.046	0.019	0.018	0.033	0.018	0.073	0.028	0.018	0.133
	GrimAA	**0.188**	**0.031**	**2.94×10^{-9}**	**0.106**	**0.036**	**0.004**	**0.092**	**0.036**	**0.012**
Ankle-brachial index (ABI)	IEAA	-0.001	0.001	0.562	0	0.001	0.644	-0.001	0.001	0.580
	EEAA	0	0.001	0.649	0	0.001	0.755	0	0.001	0.795
	PhenoAA	-0.001	0.001	0.076	-0.001	0.001	0.133	-0.001	0.001	0.149
	GrimAA	**-0.004**	**0.001**	**0.002**	-0.003	0.001	0.024	-0.003	0.001	0.026
Multisite atherosclerosis score	IEAA	0.021	0.031	0.489	0.011	0.029	0.713	-0.002	0.029	0.942
	EEAA	0.027	0.027	0.318	0.021	0.026	0.417	0.013	0.025	0.613
	PhenoAA	**0.067**	**0.020**	**0.001**	**0.052**	**0.020**	**0.008**	0.044	0.019	0.023
	GrimAA	**0.231**	**0.032**	**4.59×10^{-12}**	**0.158**	**0.038**	**4.58×10^{-5}**	**0.140**	**0.038**	**3.01×10^{-4}**

IEAA: intrinsic epigenetic age acceleration; EEAA: extrinsic epigenetic age acceleration; PhenoAgeAccel: PhenoAge acceleration; GrimAgeAccel: GrimAge acceleration

Model 1 is adjusted for age, sex, and time between measures

Model 2 is adjusted for Model 1 covariates and smoking status

Model 3 is adjusted for Model 2 covariates, hypertension status, diabetes status, body mass index, and total cholesterol levels adjusted for statin use

Beta is the change in the atherosclerosis measure associated with a 1 unit increase in the epigenetic age acceleration measure

P-values significant after Bonferroni correction ($P < 0.0125$) are shown in bold font

[a] Coronary artery calcium score and abdominal aorta calcification score were transformed as ln[(CAC+1)] and ln[(AAC+1)]

Table 4-4: Inter-individual correlations for methylation risk scores and epigenetic age acceleration measures between Phases I and II in GENOA African Americans (N=129)

Measures	Intraclass correlation coefficient (95%CI)
MRS_{CAC}	0.519 (0.385 – 0.645)
$MRS_{Carotid}$	0.822 (0.764 – 0.874)
IEAA	0.726 (0.635 – 0.802)
EEAA	0.799 (0.726 – 0.854)
PhenoAA	0.676 (0.569 – 0.758)
GrimAA	0.888 (0.848 – 0.921)

IEAA: intrinsic epigenetic age acceleration; EEAA: extrinsic epigenetic age acceleration; PhenoAgeAccel: PhenoAge acceleration; GrimAgeAccel: GrimAge acceleration

5.1 Supplementary material

Table SM 4-1: Pearson correlations between epigenetic age acceleration measures and methylation risk scores in GENOA African Americans

	IEAA	**EEAA**	**AgeAccelPheno**	**AgeAccelGrim**	**MRS$_{CAC}$**
IEAA					
EEAA	0.39****				
AgeAccelPheno	0.44****	0.46****			
AgeAccelGrim	0.21****	0.27****	0.31****		
MRS$_{CAC}$	0.08	0.20***	0.15**	0.19****	
MRS$_{carotid}$	0.19***	0.25****	0.21****	0.62****	0.39****

*** p < .0001; ** p < .01; * p < .05

IEAA, intrinsic epigenetic age acceleration; EEAA, extrinsic epigenetic age acceleration

Table SM 4-2: Association between previously-identified atherosclerosis-associated CpGs and atherosclerosis measures in GENOA African Americans [a]

	Multisite atherosclerosis				Coronary artery calcium score (CAC)[b]				Abdominal aorta calcification score (AAC)[b]				Ankle-brachial index (ABI)			
CpGs	**Beta**	**SE**	**P**	**FDR**	**Beta**	**SE**	**P**	**FDR**	**Beta**	**SE**	**P**	**FDR**	**Beta**	**SE**	**P**	**FDR**
cg05575921	-1.109	0.160	1.64E-11	8.75E-10†	-0.570	0.141	6.32E-05	0.005	-1.050	0.150	1.16E-11	8.32E-10†	0.019	0.006	0.001	0.039
cg21161138	-2.495	0.435	1.93E-08	5.14E-07†	-1.382	0.381	3.26E-04	0.008	-2.168	0.412	2.34E-07	8.37E-06	0.047	0.015	0.002	0.039
cg09935388	-1.158	0.212	8.67E-08	1.54E-06†	-0.676	0.184	2.81E-04	0.008	-1.013	0.201	6.83E-07	1.63E-05	0.023	0.007	0.001	0.039
cg21566642	-1.525	0.353	1.97E-05	2.63E-04	-0.759	0.306	0.013	0.092	-1.520	0.331	6.05E-06	1.08E-04	0.017	0.012	0.161	
cg01940273	-1.530	0.398	1.41E-04	0.001	-0.614	0.343	0.075		-1.519	0.374	5.79E-05	0.001	0.017	0.014	0.204	
cg08958747	-1.869	0.488	1.49E-04	0.001†	-1.310	0.420	0.002	0.036	-0.818	0.466	0.080		0.032	0.017	0.054	
cg24859433	-1.224	0.340	3.69E-04	0.003	-0.860	0.295	0.004	0.057	-1.064	0.322	0.001	0.009	0.015	0.011	0.186	
cg14753356	-1.511	0.447	0.001	0.005	-1.069	0.384	0.006	0.063	-1.143	0.422	0.007	0.046	0.022	0.015	0.138	
cg19572487	-1.325	0.390	0.001	0.005	-0.707	0.340	0.038		-1.479	0.364	0.000	0.001	0.021	0.013	0.108	
cg04761231	-2.169	0.676	0.001	0.007†	-0.945	0.581	0.104		-1.420	0.640	0.027		0.030	0.023	0.200	
cg25953130	-1.019	0.319	0.001	0.007	-0.509	0.274	0.064		-0.759	0.301	0.012	0.061	0.021	0.011	0.057	
cg03636183	-0.841	0.273	0.002	0.010	-0.597	0.232	0.011	0.079	-0.767	0.258	0.003	0.025	0.010	0.009	0.286	
cg18168448	-1.076	0.369	0.004	0.016	-0.872	0.316	0.006	0.063	-0.627	0.349	0.073		0.012	0.012	0.348	
cg15342087	-1.070	0.375	0.005	0.017	-0.865	0.322	0.008	0.063	-0.926	0.354	0.009	0.055	0.018	0.013	0.151	
cg18446336	-0.658	0.275	0.017	0.061	-0.462	0.235	0.050		-0.903	0.257	0.000	0.005	0.000	0.009	0.961	
cg16661609	-1.386	0.603	0.022	0.070†	-0.250	0.519	0.631		-1.312	0.568	0.021	0.090	0.010	0.020	0.619	
cg19979108	-0.988	0.429	0.022	0.070	-0.683	0.369	0.065		-0.497	0.406	0.222		0.023	0.014	0.119	
cg09646173	-1.115	0.497	0.025	0.075	-0.502	0.430	0.244		-0.211	0.470	0.653		0.023	0.017	0.177	
cg05246522	-0.838	0.386	0.030	0.085	-0.658	0.334	0.050		-0.612	0.363	0.093		-0.007	0.013	0.599	
cg15501219	-1.820	0.878	0.039	0.099	-1.090	0.753	0.148		0.157	0.831	0.850		0.056	0.029	0.056	
cg21271420	-1.345	0.650	0.039	0.099	-1.008	0.560	0.073		-0.612	0.614	0.320		0.025	0.022	0.262	
cg03295554	-0.649	0.328	0.049		-0.439	0.281	0.119		-0.763	0.308	0.014	0.061	-0.002	0.011	0.878	
cg03738331	0.979	0.492	0.047		0.165	0.424	0.697		1.144	0.461	0.014	0.061	-0.034	0.016	0.040	
cg12547807	-1.191	0.598	0.047		-0.375	0.514	0.466		-0.792	0.565	0.162		0.057	0.020	0.005	0.070
cg15344028	-0.480	0.308	0.120		-0.371	0.261	0.156		-0.832	0.287	0.004	0.029	-0.005	0.010	0.619	
cg20507228	0.303	0.351	0.389		0.805	0.299	0.007	0.063	0.844	0.327	0.010	0.057	0.028	0.012	0.016	

Model 1 is adjusted for age, sex, time between measures, and 4 genetic principal components

Beta is the change in the atherosclerosis measure associated with a 1 unit increase in the CpG methylation.

FDR > 0.10 are left blank

[a] Only associations with FDR<0.1 in Model 1 for any of the four measures of atherosclerosis are shown.

[b] Coronary artery calcium score and abdominal aorta calcification score were transformed as ln[(CAC+1)] and ln[(AAC+1)]

†Associations that remained significant in Model 3 (adjusted for Model 1 plus smoking status, type 2 diabetes status, hypertension status, BMI, and statin-adjusted total cholesterol levels) at FDR< 0.1.

Table SM 4-3: Association between previously-identified atherosclerosis-associated CpGs and atherosclerosis measures in GENOA African Americans (FDR < 0.1 in Model 3)[a]

	Multisite atherosclerosis				Coronary artery calcium score (CAC)[b]				Abdominal aorta calcification score (AAC)[b]				Ankle-brachial index (ABI)			
CpGs	**Beta**	**SE**	**P**	**FDR**	**Beta**	**SE**	**P**	**FDR**	**Beta**	**SE**	**P**	**FDR**	**Beta**	**SE**	**P**	**FDR**
cg05575921	-0.806	0.223	3.38E-04	0.020	-0.390	0.193	0.044		-0.760	0.208	2.88E-04	0.019	0.016	0.008	0.040	
cg04761231	-1.710	0.646	0.008	0.091	-0.812	0.550	0.141		-1.169	0.604	0.054		0.017	0.023	0.469	
cg08958747	-1.267	0.474	0.008	0.091	-0.841	0.404	0.038		-0.338	0.446	0.449		0.022	0.017	0.199	
cg09935388	-0.609	0.232	0.009	0.091	-0.323	0.198	0.104		-0.482	0.217	0.027		0.018	0.008	0.027	
cg16661609	-1.583	0.570	0.006	0.091	-0.496	0.487	0.309		-1.541	0.531	0.004		0.011	0.021	0.577	
cg21161138	-1.415	0.497	0.005	0.091	-0.721	0.427	0.092		-1.226	0.464	0.009		0.035	0.018	0.048	

Model 3 is adjusted for age, sex, time between measures, 4 genetic principal components, smoking, T2D status, hypertension status, BMI, and statin-adjusted total cholesterol levels.
Beta is the change in the atherosclerosis measure associated with a 1 unit increase in the CpG methylation.
FDR > 0.10 are left blank
[a] Only associations with FDR<0.1 in Model 3 for any of the four measures of atherosclerosis are shown.
[b] Coronary artery calcium score and abdominal aorta calcification score were transformed as ln[(CAC+1)] and ln[(AAC+1)

Table SM 4-4: Association between epigenetic age acceleration measures and single and multisite atherosclerosis after adjusting for white blood cell counts in GENOA African Americans

Outcome	Epigenetic age acceleration	Model 1			Model 2			Model 3		
		Beta	**SE**	***P***	**Beta**	**SE**	***P***	**Beta**	**SE**	***P***
Coronary artery calcium score (CAC)[a]	PhenoAA	0.041	0.019	0.030	0.024	0.019	0.202	0.010	0.018	0.595
	GrimAA	**0.149**	**0.029**	**2.85×10^{-7}**	**0.108**	**0.036**	**0.003**	0.073	0.035	0.036
Abdominal aorta calcification score (AAC)[a]	PhenoAA	**0.052**	**0.020**	**0.012**	0.029	0.020	0.149	0.020	0.019	0.300
	GrimAA	**0.194**	**0.031**	**1.07×10^{-9}**	**0.106**	**0.038**	**0.005**	**0.087**	**0.038**	**0.023**
Ankle-brachial index (ABI)	PhenoAA	-0.001	0.001	0.069	-0.001	0.001	0.135	-0.001	0.001	0.135
	GrimAA	**-0.003**	**0.001**	**0.002**	-0.003	0.001	0.016	-0.004	0.001	0.016
Multisite atherosclerosis score	PhenoAA	**0.074**	**0.022**	**0.001**	0.050	0.021	0.019	0.040	0.021	0.054
	GrimAA	**0.231**	**0.033**	**7.25×10^{-11}**	**0.157**	**0.040**	**1.08×10^{-4}**	**0.137**	**0.040**	**7.87×10^{-4}**

PhenoAgeAccel: PhenoAge acceleration; GrimAgeAccel: GrimAge acceleration
Model 1 is adjusted for age, sex, time between measures, and 5 white blood cell counts
Model 2 is adjusted for Model 1 covariates and smoking status
Model 3 is adjusted for Model 2 covariates, hypertension status, diabetes status, body mass index, and total cholesterol levels adjusted for lipid lowering medications
Beta is the change in the atherosclerosis measure associated with a 1 year increase in the epigenetic age acceleration measure.
Association with $P < 0.05$ are shown in bold font
[a] Coronary artery calcium score and abdominal aorta calcification score were transformed as ln[(CAC+1)] and ln[(AAC+1)]

Table SM 4-5: Association between components of GrimAge and single and multisite atherosclerosis in GENOA African Americans

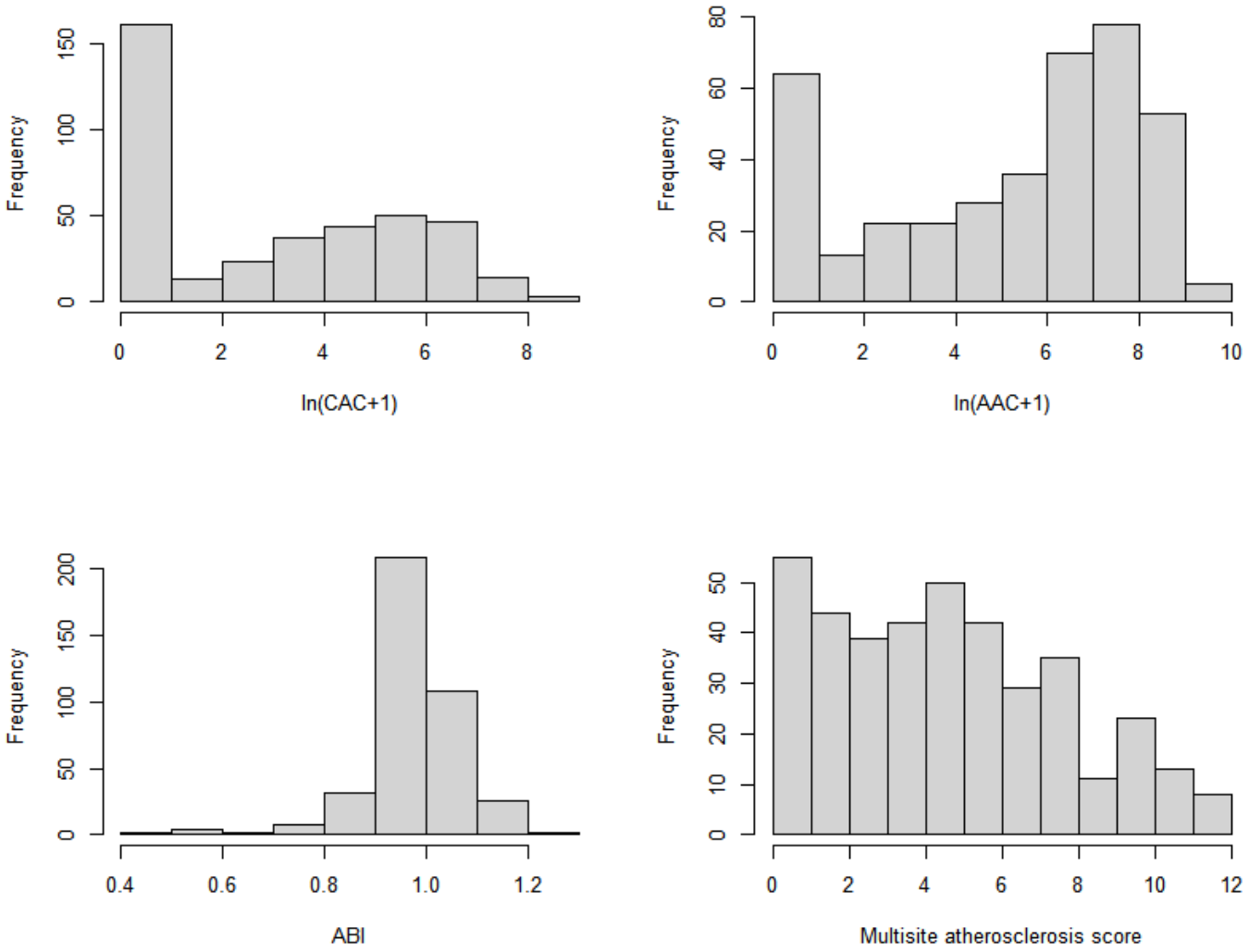
	Coronary artery calcium score (CAC)[a]			Abdominal aorta calcification score (AAC)[a]			Ankle brachial index (ABI)			Multisite atherosclerosis score		
GrimAge component (Predictor)	**Beta**	**SE**	***P***	**Beta**	**SE**	***P***	**Beta**	**SE**	***P***	**Beta**	**SE**	***P***
DNAm Adrenomedullin (ADM)	0.333	0.182	0.068	0.037	0.2	0.853	-0.012	0.007	0.104	**0.472**	**0.211**	**0.026**
DNAm Beta-2-microglobulim (B2M)	**0.372**	**0.167**	**0.026**	**0.582**	**0.184**	**0.002**	-0.013	0.007	0.052	**0.557**	**0.195**	**0.005**
DNAm Cystatin C	**0.558**	**0.23**	**0.016**	0.393	0.256	0.126	**-0.035**	**0.009**	**1.12E-04**	**0.859**	**0.269**	**0.002**
DNAm Growth Differentiation Factor 15 (GDF15)	**0.354**	**0.158**	**0.026**	**0.512**	**0.174**	**0.003**	-0.003	0.006	0.579	**0.431**	**0.185**	**0.021**
DNAm Leptin	0.424	0.238	0.076	0.175	0.26	0.502	4.68E-04	0.009	0.959	0.281	0.276	0.31
DNAm Smoking pack-years	**0.569**	**0.132**	**2.21E-05**	**1.05**	**0.139**	**3.19E-13**	**-0.012**	**0.005**	**0.024**	**1.021**	**0.149**	**3.23E-11**
DNAm Plasminogen activator inhibitor antigen type 1 (PAI1)	**0.457**	**0.127**	**3.48E-04**	**0.282**	**0.141**	**0.046**	0.001	0.005	0.892	**0.447**	**0.148**	**0.003**
DNAm Tissue inhibitor metalloproteinases 1 (TIMP1)	0.413	0.216	0.058	-0.013	0.241	0.958	-0.009	0.009	0.288	0.345	0.256	0.178

Table SM 4-6: Association between methylation risk scores at Phase II and atherosclerosis measures in GENOA African Americans (N = 129)

Outcome	Epigenetic measure	Beta	SE	*P*
Coronary artery calcium score (CAC) [a]	MRS_{CAC}	0.156	0.315	0.621
	$MRS_{Carotid}$	0.297	0.306	0.333
Abdominal aorta calcification score (AAC) [a]	MRS_{CAC}	0.154	0.353	0.663
	$MRS_{Carotid}$	**0.894**	**0.333**	**0.008**
Ankle Brachial Index (ABI)	MRS_{CAC}	0	0.011	0.949
	$MRS_{Carotid}$	-0.010	0.011	0.596
Multisite atherosclerosis score	MRS_{CAC}	0.133	0.354	0.707
	$MRS_{Carotid}$	**0.891**	**0.333**	**0.009**

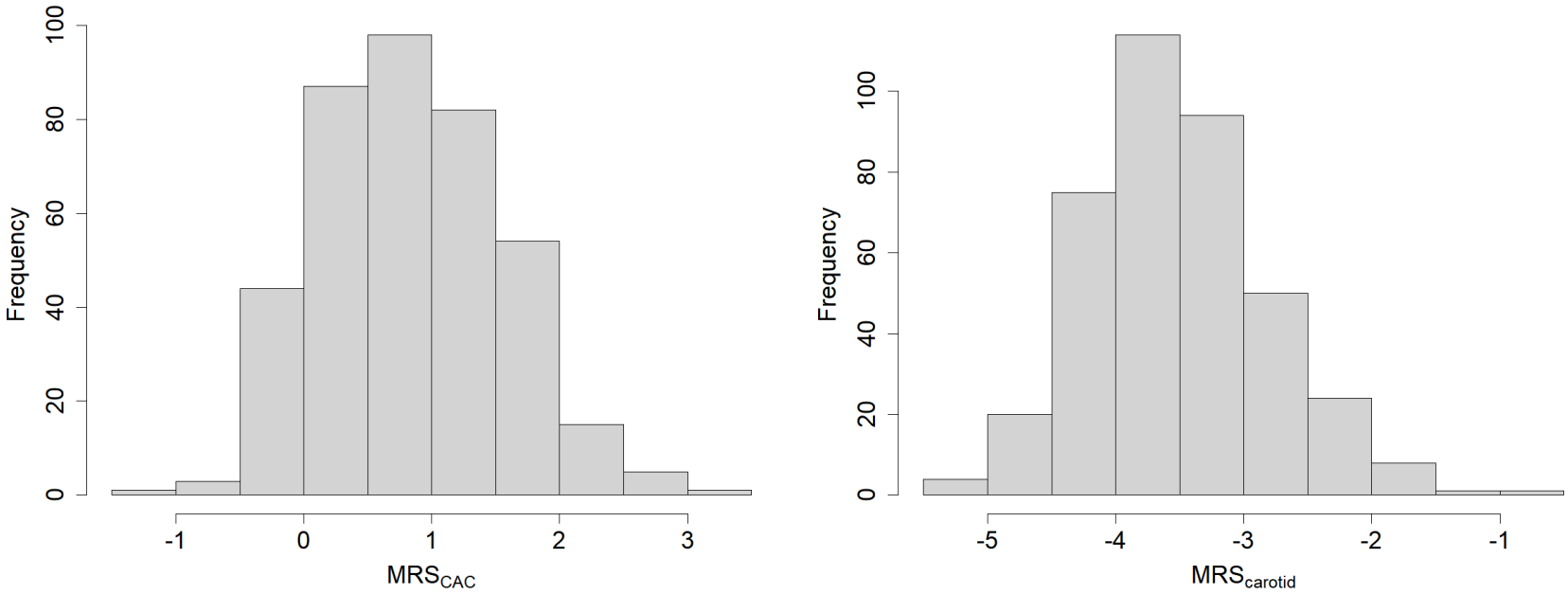
Model is adjusted for age, sex, time between measures, and 4 genetic principal components
Beta is the change in the atherosclerosis measure associated with a 1 unit increase in the MRS.
P-values significant after Bonferroni correction ($P < 0.025$) are shown in bold font
[a] Coronary artery calcium score and abdominal aorta calcification score were transformed as ln[(CAC+1)] and ln[(AAC+1)]

Figure SM 4-1: Distribution of single- and multi-site atherosclerosis measures in GENOA African Americans



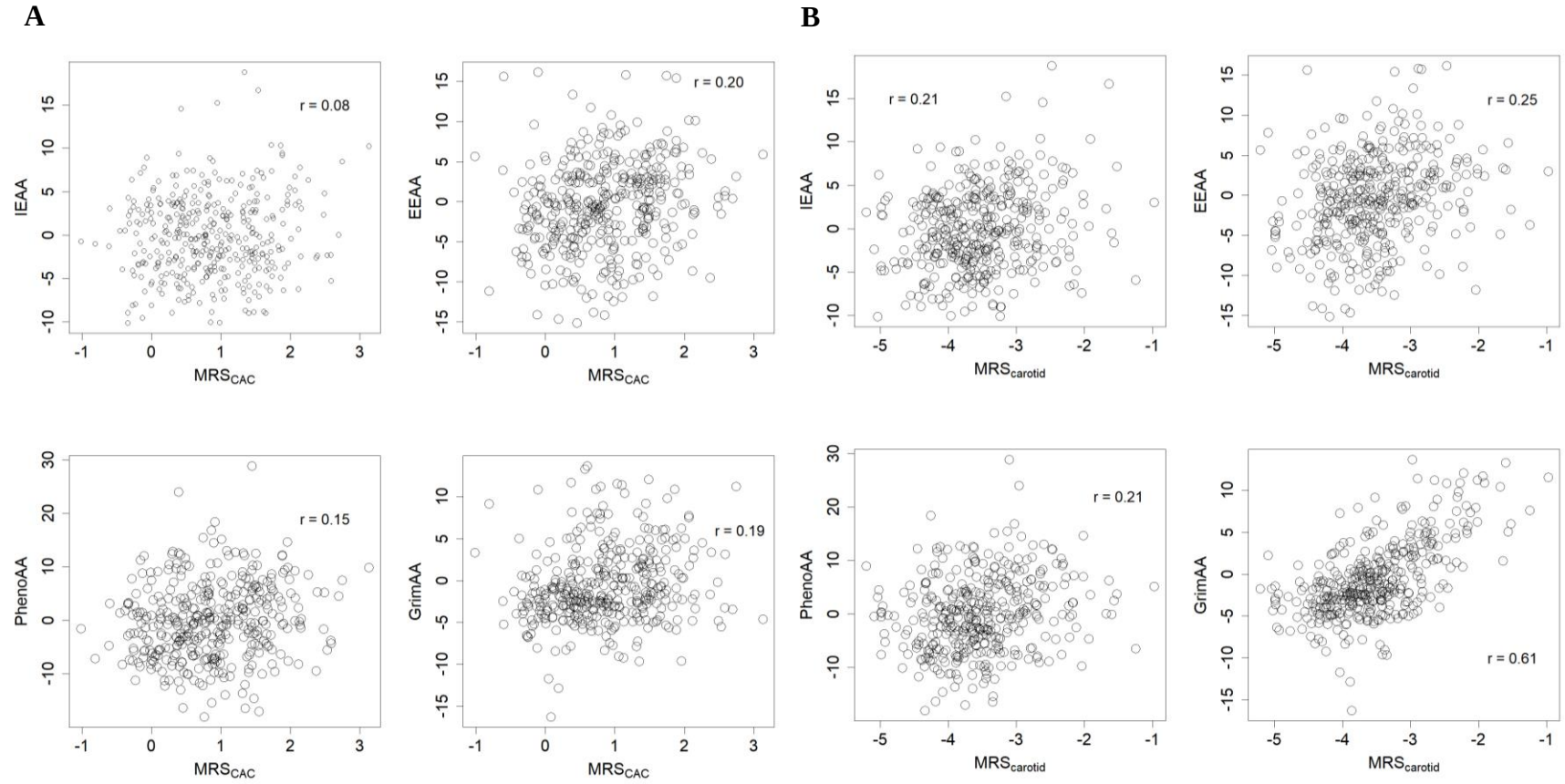
CAC, coronary artery calcification; AAC, abdominal aorta calcification; ABI, ankle brachial index

Figure SM 4-2: Distribution of MRS_{CAC} and $MRS_{carotid}$ in GENOA African Americans



CAC, coronary artery calcification; MRS, methylation risk score

Figure SM 4-3: Scatterplots and Pearson correlation coefficients for MRS_{CAC} (A) and $MRS_{carotid}$ (B) and epigenetic age acceleration (IEAA, EEAA, PhenoAA, and GrimAA) in GENOA African Americans.



IEAA, intrinsic epigenetic age acceleration; EEAA, extrinsic epigenetic age acceleration.

Figure SM 4-4: Methylation at the cg05575921 (A), cg21161138 (B) and cg09935388 (C) by smoking status at Phase I in GENOA African Americans

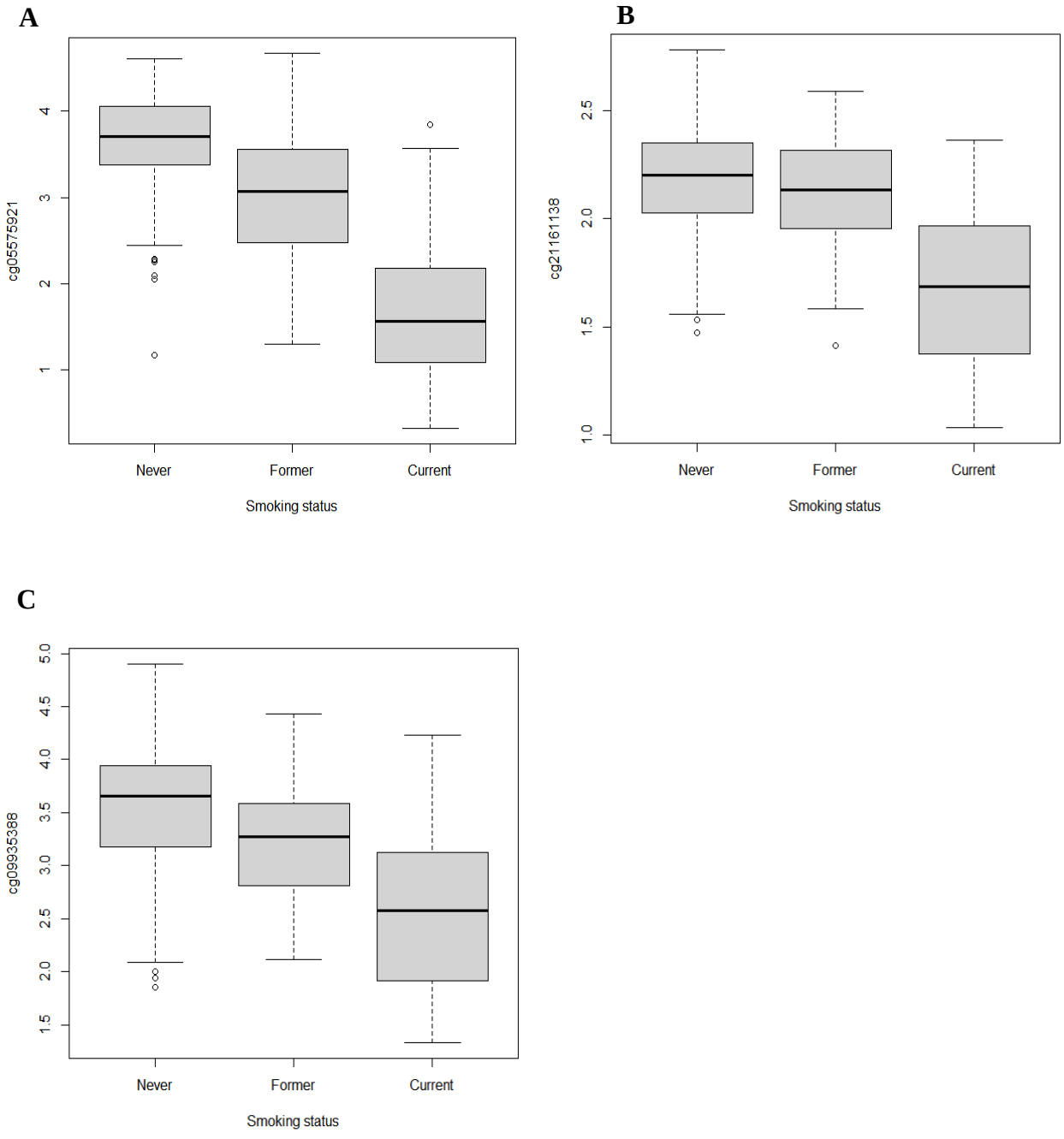
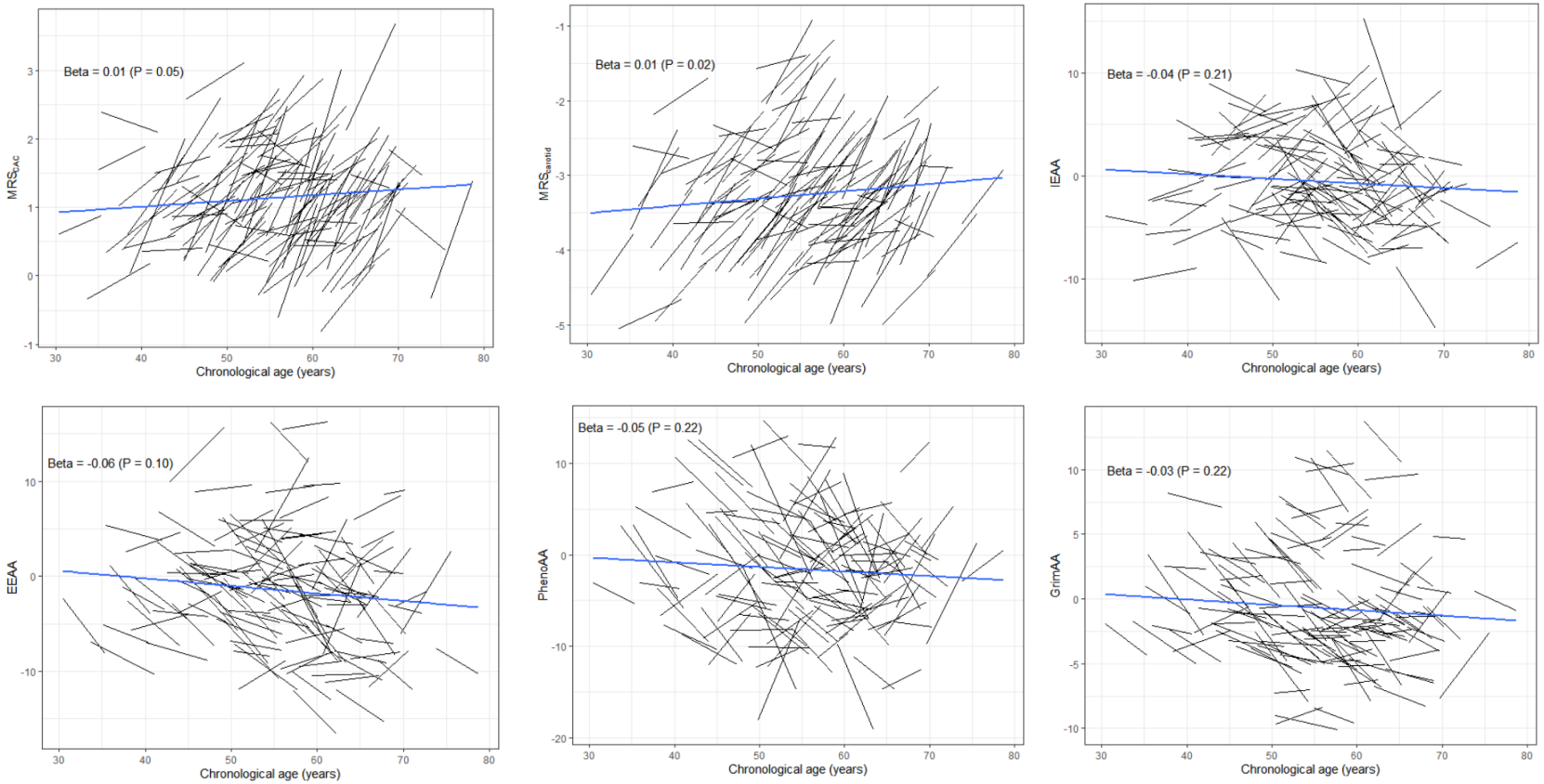


Figure SM 4-5: Spaghetti plots for the change in participant methylation risk scores (MRS_{CAC} and $MRS_{carotid}$) and epigenetic age acceleration (IEAA, EEAA, PhenoAA, and GrimAA) between Phase I and Phase II in GENOA African Americans



IEAA, intrinsic epigenetic age acceleration; EEAA, extrinsic epigenetic age acceleration.
Blue lines indicate the mean change of the methylation risk scores or epigenetic age acceleration measures by age using generalized estimating equation (GEE) models adjusted for sex.

Chapter 5. Conclusions

5.1 Summary and Implications of Main Findings

The primary goals of this dissertation were to: (1) examine the epigenome wide association between DNA methylation and 5 TOD measures of functional and structural changes in the heart, kidneys and brain using univariate and multivariate models; (2) evaluate whether epigenetic age acceleration measures are associated with cardiometabolic markers and CVD incidence; and (3) examine the association between epigenetic age acceleration measures and previously-identified CpGs for atherosclerosis, both individually and aggregated into MRSs, and single- or multi-site atherosclerosis. Findings from this work advance our knowledge about the relationship between DNA methylation, aging, and CVD, highlighting the important role of epigenetics in subclinical and clinical CVD and TOD.

In Chapter 2, we addressed whether DNA methylation sites are associated with eGFR, albuminuria, left ventricular mass index, relative wall thickness, and white matter hyperintensity in GENOA African Americans. For TOD traits related to the heart and kidneys, we used a multivariate (pleiotropy) model that leverages the correlation between those traits to identify differentially methylated CpGs. Using this model, we successfully identified seven CpGs: cg21134922, cg04816311 near *C7orf50*, cg09155024, cg10254690 near *OAT*, cg07660512, cg12661888 near *IFT43*, and cg02264946 near *CATSPERD* at FDR < 0.1. Three CpGs remained significant after adjusting for hypertension, BMI, and type 2 diabetes. DNA methylation was associated with proximal gene expression for some CpGs, but we did not find evidence of

mediation by gene expression. Mendelian randomization analyses suggested causality between three CpGs and eGFR (cg04816311, cg10254690, and cg07660512). In a replication analysis in an independent cohort of 614 African Americans from the Hypertension Genetic Epidemiology Network (HyperGEN), One site, cg04816311, was significantly associated with eGFR (P = 0.0003), LVMI (P = 0.0003), and RWT (P = 0.002).

Our study findings are important because they fill in the gap in the current literature on the role of epigenetics in TOD. To date, very few studies have examined the association between genome-wide DNA methylation and TOD traits.[1] Our sample which was comprised of largely hypertensive African Americans is representative of the population group most burdened by hypertension and downstream TOD. We have additionally showed the utility of using a pleiotropy informed analysis to increase statistical power and improve detection of significant DNA methylation sites. Finally, our conclusions were further supported by our ability to adjust for important comorbidities to delineate potentially important biological TOD pathways, the use of Mendelian randomization to assess causality, and the replication of some of our findings in an independent cohort of African Americans.

In Chapter 3, we investigated the association between four DNAm-based epigenetic age acceleration measure and ten cardiometabolic risk factors and CVD incidence in GENOA African Americans. Overall, increased epigenetic age acceleration, indicative of faster biological aging was associated with a worse cardiometabolic risk profile, although the associations with specific cardiometabolic risk factors varied across the age acceleration measures. All of the epigenetic acceleration measures were correlated with risk of CVD onset as modeled by clinical CVD risk scores (FRS and ASCVD equation). The direction of association was as expected (i.e. increased biological aging was associated with increased risk of CVD as predicted by FRS

and/or ASCVD). GrimAA outperformed IEAA, EEAA, and PhenoAA in predicting CVD incidence, and the association remained significant after adjusting for traditional CVD risk factors. A 5-year increase in GrimAA, was associated with a hazard ratio of 1.47 (95% CI: 1.05 – 2.01, $P = 0.024$) for incident CVD. Finally, we examined whether GrimAA could be used to improve the predictive accuracy of clinical CVD risk scores. GrimAA improved model fit over clinical risk scores using likelihood ratio tests ($P = 0.013$ for FRS, $P = 0.008$ for ASCVD), did not improve C statistics ($P > 0.05$), and marginally improved the net reclassification index (NRI = 0.055, 95% CI: 0.040 – 0.071 for FRS; 0.029, 95% CI: 0.006 – 0.064 for ASCVD).

The results of this study are important because they add to the growing literature examining the associations between epigenetic age acceleration measures, CVD, and its risk factors. Our study included two "first generation" and two "second generations" measures. The observation that GrimAA, a "second generation" measure, outperforms the other 3 measures in predicting incident CVD has also been observed in two large cohorts with participants of Europeans/Caucasian ancestry.[2,3] Our findings suggest that the association between GrimAA and CVD incidence is shared across these two racial groups, with similar effect sizes across the three studies. Our study was the first to go beyond evaluating statistical associations by assessing the predictive utility of epigenetic age acceleration measures to improve CVD risk prediction. Our findings highlight the potential relevance of epigenetic markers for the improvement of risk stratification of CVD beyond traditional risk factors and CVD risk scores currently used clinically.

In Chapter 4, we build upon our findings in Chapter 3 where we found evidence of association between epigenetic age acceleration measures and CVD. In this study, we examined the association between previously identified atherosclerosis-associated CpGs (individually and

aggregated into MRSs), epigenetic age acceleration measures, and subclinical CVD. To quantify subclinical CVD, we used three measures of atherosclerosis in the coronary arteries (CAC), abdominal aorta (AAC), and peripheral arteries (ABI). We additionally derived a multisite atherosclerosis score combining these three measures. When we examined the association with previously-identified CpGs, one and six CpGs were associated with AAC and multi-site atherosclerosis, respectively (FDR < 0.1). A one unit increase in the MRS for carotid artery calcification was associated with a 1.6-fold increase in AAC and 0.7 units increase in multi-site atherosclerosis (score range: 0-12) after adjusting for CVD risk factors (Bonferroni-adjusted $P < 0.05$). As we expected based on our findings in Chapter 3, GrimAA was the only measure associated with AAC and multisite atherosclerosis after adjustment for CVD risk factors (Bonferroni-adjusted $P < 0.05$). We also observed that the epigenetic measures most strongly associated with atherosclerosis, GrimAA and $MRS_{carotid}$, had moderate stability (ICC: 0.822 and 0.888, respectively) across repeated samples taken approximately 5 years apart.

To our knowledge, this is one of the few studies to examine the whole blood DNA methylation signature of multiple atherosclerosis measures in a population-based cohort. It is also one of the few studies to use a prospective design where DNA methylation was measured approximately 5 and 12 years prior to the assessment of atherosclerosis. Our finding that GrimAA, but not other measures of epigenetic age, is associated with subclinical CVD supports our previously reported association between GrimAA and CVD incidence. We also found that some associations between DNA methylation and atherosclerosis are consistent across African Americans in GENOA and white participants in other studies. Finally, our results support the notion that associations between epigenetic biomarkers and atherosclerosis have similarities

across vascular sites and suggests common underlying biological mechanisms of atherosclerosis at these sites.

More studies are needed to better elucidate the biological pathways implicated in the associations described in this dissertation and fully explore regulatory and downstream gene expression changes. In Chapter 2, our multivariate model and subsequent univariate models confirm our hypothesis that common biological mechanisms underly the TOD measures of multiple organ systems. For the identified CpGs, we only investigated proximal gene expression changes and lacked the power to detect mediating effects. However, our findings shed light on the potential role of pathways related to nitric oxide (NO) as an underlying mechanism of TOD. NO is an important messenger molecule that regulates blood vessel dilation and has other thrombotic and inflammatory effects.[4] Further, one identified CpG, cg04816311 near *C7orf50*, has been previously associated with type 2 diabetes and BMI. This suggests the possibility of common biological mechanisms underlying comorbidities in multiple organ systems.

In Chapter 3, we found a significant association between GrimAA and CVD incidence. Since information about the individual CpGs comprising GrimAge are not available, we explored the association between each component of GrimAge and CVD incidence to understand potential biological pathways leading to the observed association. Of the GrimAge components, we found that adrenomedullin, plasminogen activator inhibitor antigen type 1, and smoking pack-years are associated with CVD incidence. Adrenomedullin is a free-circulating peptide that is an important vasodilator and has other biological functions, especially involving the heart.[5] Studies suggest that adrenomedullin levels are increased in acute myocardial infarction and in hypertrophied and failing hearts as a protective mechanism against cardiac dysfunction, myocardial remodeling, or both.[5] Plasminogen activator inhibitor antigen type I (PAI-1) is a

biomarker of fibrinolysis, the degradation of fibrin as a result of complex interaction among multiple plasminogen activators and inhibitors. PAI-1 is one of the most important inhibitors of plasma fibrinolytic activity. *In vivo*, PAI-1 expression suppresses fibrinolysis leading to pathological fibrin deposition resulting in atherothrombosis and tissue damage.[6-8] PAI-1 expression has been found to play an important role in subclinical and clinical conditions related to aging including increased inflammation, atherosclerosis, and obesity.[6]

Smoking is a well-established risk factor for CVD. Evidence from animal and clinical studies show that a number of mechanisms precipitate the effects of smoking on CVD. Smoking has been reported to reduce flow-mediated dilatation in the systemic arteries which is an early marker for endothelial dysfunction.[9] It has also been associated with quantitative and qualitative modulation of lipids leading to increased levels of total cholesterol, LDL-C, and triglycerides in addition to increased lipid oxidation.[9] Furthermore, smoking is associated with increased activation of the immune system, both systematically and locally, which is an important constituent of atherogenesis. The increased inflammation results in increased white blood cell counts, changes in the vascular wall leading to increased expression of matrix metalloproteinases, and increased endothelial adhesion molecules leading to leukocyte recruitment.[9] Hence, our findings lend further evidence to the potential role of these components in shaping cardiovascular health and risk. Finally, it is important to note that the DNAm-based surrogate measures of these components may be more strongly associated than the measures themselves.[10] In the case of smoking, this may be due to inaccurate reporting or exposure to secondhand smoking, but it can also be due to DNAm capturing the intrinsic variations associated with long-term biological changes related to the plasma proteins and/or smoking.

In Chapter 4, we found similar evidence of the role of smoking in atherosclerosis, which was evident in the hypomethylation signature at a number of CpGs in the *AHRR* and *GFI1* genes. Additionally, we found evidence of an association between hypomethylation at cg16661609, located upstream of a transcription start site of the *LILRB4*, and atherosclerosis. This gene encodes a protein belonging to the subfamily B class of the leukocyte immunoglobulin-like receptors and is expressed on immune cells. This receptor binds to antigen-presenting cells and transduces a negative signal that inhibits stimulation of an immune response. We additionally found evidence of association between increased biological aging, as measured by GrimAge, and atherosclerosis. Each of the GrimAge components was associated with multisite atherosclerosis with the exception of Leptin and Tissue inhibitor metalloproteinases 1 (TIMP1). Smoking pack-years was the most strongly associated component, and it was associated with all the atherosclerosis measures in GENOA.

While it is well known that atherosclerosis is an aging disease, it is not clearly understood how the association between aging and atherosclerosis unfolds. DNA methylation, under both genetic and environmental control, may be an important link by which increased aging leads to atherosclerosis. Epigenetic changes with chronological age are currently recognized to be either age-related methylation changes or epigenetic drift.[11,12] Age-related methylation refers to the predictable, direction-specific changes in DNA methylation levels that occur with chronological age.[13,14] This concept is linked to epigenetic age acceleration measures or clocks. Epigenetic drift, on the other hand, is a reflection of deficient maintenance of epigenetic marks leading to the loss of hypomethylation in hypermethylated regions and gains in hypomethylated ones. These changes in the epigenome are thought to be stochastic, bidirectional, and to have varying contribution to the epigenome across the life course.[14] Recently, Kochmanski and colleagues

coined the term “environmental deflection” to refer to a process by which an environment can increase or decrease age-related methylation and epigenetic drift.[14] Our study found evidence of association between age-related DNA methylation changes (GrimAA), other CpGs combined into MRSs, and atherosclerosis. It is possible that the epigenetic biomarkers associated with atherosclerosis in our study could be showing evidence of environmental deflection, potentially due to traditional CVD risk factors including smoking. Evidence of epigenetic deflection due to smoking has been observed, for example, in our most significant CpG, cg05575921 in the *AHRR* gene region.[15]

Collective findings from this dissertation advance our knowledge about the associations between DNA methylation changes and DNA methylation based aging measures and CVD risk factors, subclinical and clinical disease in African Americans. Replication studies are needed to characterize whether our findings generalize to other cohorts of African Americans, as well as other African and non-African ancestries. Our GENOA cohort was recruited from Jackson, Mississippi, and most African ancestry individuals in this area are from coastal Western African and Nigerian origins.[16,17] Ancestry analysis of GENOA African Americans using genetic principal components confirms that this is a population with relatively homogenous ancestry. Both genetic variation and environmental factors impact DNA methylation.[18] Hence, our findings may not generalize to other African American cohorts of similar genetic ancestry due to differences in environmental exposures which can influence the epigenome during the early stages of embryo development and later in life.[17-20] Likewise, it remains unclear whether our findings are generalizable to other cohorts of different ancestries. Future studies with multiethnic participants and repeated DNA methylation measurements, including early in life, may elucidate

the differences in DNA methylation between different groups and improve our understanding of the contribution of differential DNA methylation to human variation and phenotypes.

5.2 Strengths and Limitations

While this dissertation contributes to our understanding of the role of DNA methylation in CVD, TOD, and their risk factors, it has several important limitations. First, our sample is comprised of individuals from sibships with a strong family history of hypertension. This may limit the generalizability of our findings. To address these concerns, we have adjusted our findings for hypertension or blood pressure measurements when appropriate. Second, we note the attrition of participants between Phase I and the follow-up visits (Phases II and III). We addressed concerns regarding selection bias by comparing the baseline characteristics of those lost to follow up and those who returned for later Phases. We noted that loss to follow-up may have biased the associations towards the null. Third, CVD events were defined based on participant reports, and were not validated using medical records examination. Given this, our findings regarding CVD risk prediction should be validated in studies with access to medical records. Last, although we used longitudinal data and Mendelian randomization for some of the reported associations, it remains difficult to establish causality between changes in the epigenome and CVD or TOD.

This dissertation has several strengths. First, we used a prospective design where DNA methylation was assessed prior to our outcomes of interest. For CVD and TOD, it is difficult to determine the timeline of true onset. However, our study design partially addresses concerns related to reverse causation and offers further validation of previous cross-sectional reports. Second, our rich phenotyping, which included a number of lifestyle and clinical measures in addition to genetic and transcriptomic data, allowed us to control for important confounders,

explore the causality of our associations using Mendelian randomization analysis, and characterize potential functional effects by assessing gene expression changes. Third, our findings further our understanding of the role of DNAm in CVD in African Americans who are underrepresented in previous literature. Finally, we employed multiple methods to assess the associations between DNA methylation and related outcomes. In Aim 1, we used multivariate modelling that leverages the correlation between TOD measures and increases statistical power and decreases the burden of multiple testing. In Aim 3, we took a different approach by creating a multisite atherosclerosis score based on three single site atherosclerosis measures. This approach allowed us to model the associations with the different atherosclerosis measures and draw conclusions on the presence of multisite atherosclerosis. Such conclusions would not have been possible using multivariate approaches which would have allowed us to only test whether DNA methylation is associated with at least one of the single site atherosclerosis measures.

5.3 Future Directions

This dissertation builds upon previous work that identified DNAm to be an important mechanism that captures the imprints of lifestyle and environmental exposures throughout the life course. The collective results generated from this dissertation have important implications that may be leveraged to improve risk stratification and prediction of CVD and TOD. Our findings emphasize that the etiology of these complex disease is intricate and is a result of interaction between different "omic" layers and environmental factors. Future studies that are able to more fully integrate "omic" data, including transcriptomics, proteomics, and metabolmics, can answer important questions regarding underlying biological mechanisms and processes for the associations identified. Furthermore, closer examination of downstream gene

expression changes due to DNAm changes can answer important questions on the proximal and distal regulatory effects of DNAm on gene expression in relation to CVD and TOD.

Additionally, beyond identification of significant CpGs and/or biomarkers, future studies with the capacity to delineate pathways of associations are of utmost importance. This could be achieved by using Mendelian randomization to identify causal effects of DNAm on CVD and its risk factors white controlling for confounding and reverse causation. DNAm associations with CVD and TOD could be due to causal pathways that can be targeted for potential interventions. Mendelian randomization could be used to refute causality where the DNAm patterns in the white blood cells are reflective of early signs of disease onset and elevated disease risk. Hence, DNAm offers the potential for identifying and quantifying key epigenetic modifications that may help to improve risk stratification and prediction of CVD and TOD. Combined with other precision medicine efforts, this may translate into tailored interventions based on genomic and molecular profiles, in addition to environmental exposures, of individuals.

Finally, larger studies with longer follow-up periods and repeated DNA methylation measurement would be better powered to investigate the role of DNA methylation in these age-related outcomes. Such studies would be less biased by reverse causality and are better powered to detect associations with health conditions that have long onset durations. In this dissertation, our outcome measures were assessed after DNA methylation, which helped establish temporal patterning for our associations. We additionally used Mendelian randomization to help further investigate the causality of our findings when possible. Repeated DNA methylation measurement was also essential to help further our understanding of DNA methylation changes over time. In Aim 3, we used a small sample of about 130 participants with repeated DNA methylation measurements to investigate the temporal stability of our epigenetic predictors. GEE models

allow us to model the methylation as repeated measures while incorporating within-subject and between-subject variations into model fitting. However, a larger sample size would have allowed us to further investigate trends and increased our statistical power. Hence, future adequately powered studies with repeated measures are essential to further our understanding of the role of DNA methylation in CVD and TOD.

5.5 Conclusion

Together, these studies support the premise that DNAm plays an important role in CVD and TOD and is a promising biomarker that may improve risk assessment in African Americans. Chapter 2 identified seven CpGs associated with TOD with suggestive evidence of causality. Findings from Chapters 3 and 4 shed important insights on the role of DNAm in atherosclerosis, CVD risk factors, and subsequent CVD incidence. Our findings contribute to the rapidly evolving picture of the role of epigenetics in CVD and TOD in African Americans. Insights from this dissertation can chart a new path forward for improved risk prediction, advances in precision medicine and ultimately inform efforts to reduce disease burden, especially among highly burdened populations.

5.4 References

1. Chu AY, Tin A, Schlosser P, et al. Epigenome-wide association studies identify DNA methylation associated with kidney function. Nat Commun. 2017;8(1):1286.
2. Hillary RF, Stevenson AJ, McCartney DL, et al. Epigenetic measures of ageing predict the prevalence and incidence of leading causes of death and disease burden. Clin Epigenetics. 2020;12(1):115.
3. Wang C, Ni W, Yao Y, et al. DNA methylation-based biomarkers of age acceleration and all-cause death, myocardial infarction, stroke, and cancer in two cohorts: The NAS, and KORA F4. EBioMedicine. 2020;63:103151.
4. Huynh NN, Chin-Dusting J. Amino acids, arginase and nitric oxide in vascular health. Clin Exp Pharmacol Physiol. 2006;33(1-2):1-8
5. Nishikimi T, Kuwahara K, Nakagawa Y, Kangawa K, Nakao K. Chapter 2 - Adrenomedullin. In: Schisler JC, Lang CH, Willis MS, eds. Endocrinology of the Heart in Health and Disease. Academic Press; 2017:41-58
6. Cesari M, Pahor M, Incalzi RA. Plasminogen activator inhibitor-1 (PAI-1): a key factor linking fibrinolysis and age-related subclinical and clinical conditions. *Cardiovasc Ther.* 2010;28(5):e72-91.
7. Weisberg AD, Albornoz F, Griffin JP, et al. Pharmacological inhibition and genetic deficiency of plasminogen activator inhibitor-1 attenuates angiotensin II/salt-induced aortic remodeling. *Arterioscler Thromb Vasc Biol.* 2005;25(2):365-371.
8. Aso Y. Plasminogen activator inhibitor (PAI)-1 in vascular inflammation and thrombosis. *Front Biosci.* 2007;12:2957-2966.
9. Messner B, Bernhard D. Smoking and cardiovascular disease: mechanisms of endothelial dysfunction and early atherogenesis. *Arterioscler Thromb Vasc Biol.* 2014;34(3):509-515.
10. Lu AT, Quach A, Wilson JG, et al. DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging (Albany NY).* 2019;11(2):303-327.
11. Teschendorff AE, West J, Beck S. Age-associated epigenetic drift: implications, and a case of epigenetic thrift? *Hum Mol Genet.* 2013;22(R1):R7-R15.
12. Li Y, Tollefsbol TO. Age-related epigenetic drift and phenotypic plasticity loss: implications in prevention of age-related human diseases. *Epigenomics.* 2016;8(12):1637-1651.
13. Jung M, Pfeifer GP. Aging and DNA methylation. *BMC Biol.* 2015;13:7.
14. Kochmanski J, Montrose L, Goodrich JM, Dolinoy DC. Environmental Deflection: The Impact of Toxicant Exposures on the Aging Epigenome. *Toxicol Sci.* 2017;156(2):325-335.
15. Shah S, McRae AF, Marioni RE, et al. Genetic and environmental exposures constrain epigenetic drift over the human life course. *Genome Res.* 2014;24(11):1725-1733.15.

16. Salas A, Richards M, Lareu MV, et al. The African diaspora: mitochondrial DNA and the Atlantic slave trade. *Am J Hum Genet.* 2004;74(3):454-465.

17. Micheletti SJ, Bryc K, Ancona Esselmann SG, et al. Genetic Consequences of the Transatlantic Slave Trade in the Americas. *Am J Hum Genet.* 2020;107(2):265-277.

18. Kader F, Ghai M. DNA methylation-based variation between human populations. *Mol Genet Genomics.* 2017;292(1):5-35.

19. Adkins RM, Krushkal J, Tylavsky FA, Thomas F. Racial differences in gene-specific DNA methylation levels are present at birth. *Birth Defects Res A Clin Mol Teratol.* 2011;91(8):728-736.

20. King K, Murphy S, Hoyo C. Epigenetic regulation of Newborns' imprinted genes related to gestational growth: patterning by parental race/ethnicity and maternal socioeconomic status. *J Epidemiol Community Health.* 2015;69(7):639-647.