Blood leukocyte DNA methylation predicts risk of future myocardial infarction and coronary heart disease: A longitudinal study of 11,461 participants from population-based cohorts

Golareh Agha†[1,2], Michael M. Mendelson†[3], Cavin K. Ward-Caviness†[4], Roby Joehanes†[5], TianXiao Huan[6], Rahul Gondalia[7], Elias Salfati[8], Jennifer A. Brody[9], Giovanni Fiorito[10], Jan Bressler[11], Brian H. Chen[12], Symen Ligthart[13], Simonetta Guarrera[10], Elena Colicino[14], Allan C. Just[15], Simone Wahl[16], Christian Gieger[16], Amy R. Vandiver[17], Toshiko Tanaka[12], Dena G. Hernandez[18], Luke C. Pilling[19], Andrew B. Singleton[18], Carlotta Sacerdote[20], Vittorio Krogh[21], Salvatore Panico[22], Rosario Tumino[23], Yun Li[24], Guosheng Zhang[25], James D. Stewart[26], Floyd JS[9], Kerri L. Wiggins[9], Jerome I. Rotter[27], Michael Multhaup[17] , Kelly Bakulski[28], Steven Horvath[29], Philip S.Tsao[8], Devin M. Absher[30], Pantel Vokonas[31], Joel Hirschhorn[32] , M Daniele Fallin[33], Chunyu Liu[34], Stefania Bandinelli[35], Eric Boerwinkle[36], Abbas Dehghan[37], Joel D. Schwartz[38], Bruce M. Psaty[39], Andrew P. Feinberg[40], Lifang Hou*[41], Luigi Ferrucci*[42], Nona Sotoodehnia*[43], Giuseppe Matullo*[10], Annette Peters*[44], Myriam Fornage*[45], Themistocles L. Assimes*[46], Eric A. Whitsel*[47], Daniel Levy*[48] , Andrea A. Baccarelli*[1]

† These authors contributed equally
* =last authors

1. Department of Environmental Health Sciences, Columbia University Mailman School of Public Health, NY, NY 10032, USA.
2. Columbia Aging Center, Columbia University Mailman School of Public Health, NY, NY 10032, USA.
3. Population Sciences Branch, National Heart, Lung, and Blood Institute, NIH, Bethesda, MD 20892, USA; Framingham Heart Study, Framingham, MA 01702, USA; Department of Cardiology, Boston Children's Hospital, Boston, MA 02115, USA.
4. National Health and Environmental Effects Research Laboratory, Environmental Public Health Division, Chapel Hill NC 27514, USA; Institute of Epidemiology II, Helmholtz Institute, Ingolstaedter Landstrasse 1, Neuherberg, Germany 85764.
5. National Heart, Lung and Blood Institute, Bethesda, MD 20824-0105, USA; Hebrew SeniorLife, Harvard Medical School, Boston, MA 02115, USA.
6. The Population Sciences Branch, Division of Intramural Research, National Heart, Lung and Blood Institute, Bethesda, MD 20892, USA.
7. Department of Epidemiology, University of North Carolina, Chapel Hill, NC 27514, USA.
8. Department of Medicine, Stanford University School of Medicine, Stanford, CA 94305, USA.
9. Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, WA 98195, USA.
10. Italian Institute for Genomic Medicine (IIGM/HuGeF) and Department of Medical Sciences, University of Turin, Turin, Italy.
11. Human Genetics Center, School of Public Health, University of Texas Health Science Center at Houston, Houston, TX 77030, USA.
12. Translational Gerontology Branch, National Institute on Aging, National Institutes of Health, Baltimore, MD 21250, USA.
13. Department of Epidemiology, Erasmus MC University Medical Center, Rotterdam, the Netherlands.
14. Icahn School of Medicine at Mount Sinai, NY 10029, NY.
15. Department of Environmental Medicine and Public Health, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA.
16. Research Unit Molecualr Epidemiology, Helmholtz Zentrum München, 1 InglastaedterLandstrasse 1 85764, München, Germany.
17. Center for Epigenetics, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA.
18. Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD 20892, USA.
19. Epidemiology and Public Health Group, University of Exeter Medical School, Barrack Road, Exeter, EX2 5DW, UK.
20. Unit of Cancer Epidemiology, Città della Salute e della Scienza University-Hospital and Center for Cancer Prevention (CPO), Turin, Italy.
21. Epidemiology and Prevention Unit, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy.
22. Dipartimento di Medicina Clinica e Chirurgia, Federico II University, Naples, Italy.
23. Cancer Registry And Histopathology Department, "Civic- M.P. Arezzo2 Hospital, Asp Ragusa, Italy.

24. Department of Genetics, Department of Biostatistics, Department of Computer Science, University of North Carolina, Chapel Hill, NC 27599, USA.
25. Curriculum in Bioinformatics and Computational Biology, and Department of Genetics, and Department of Statistics, University of North Carolina, Chapel Hill, NC 27514, USA.
26. Carolina Population Center and Department of Epidemiology, University of North Carolina at Chapel Hill, NC 27514, USA.
27. The Institute for Translational Genomics and Population Sciences, Departments of Pediatrics and Medicine, LABioMed at Harbor-UCLA Medical Center, Torrance, CA 90502, USA.
28. Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, MI 48109, USA.
29. Department of Human Genetics, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA 90095, USA.
30. HudsonAlpha institute of Biotechnology, Huntsville, AL, USA.
31. VA Normative Aging Study, VA Boston Healthcare System, Department of Medicine, Boston University School of Medicine, Boston, MA 02118, USA.
32. Department of Medicine, Division of Endocrinology, Boston Children's Hospital, Boston, MA 02115, USA; Departments of Medicine and Pediatrics, Harvard Medical School, Boston, MA 02115, USA.
33. Department of Mental Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 21205, USA.
34. Department of Biostatistics, Boston University School of Public Health, Boston, MA 02118, USA. .
35. Azienda Sanitaria, USL Centro FIrenze, Firenze Italy.
36. Human Genetics Center, School of Public Health, The University of Texas Health Science Center at Houston, Houston, Texas, USA; Human Genome Sequencing Center, Baylor College of Medicine, Houston, Texas, USA.
37. Department of Epidemiology and Biostatistics, MRC–PHE Centre for Environment & Health, School of 346 Public Health, Imperial College London, UK.
38. Department of Epidemiology, and Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, MA 02115, United States.
39. Cardiovascular Health Research Unit, Departments of Medicine, Epidemiology, and Health Services, University of Washington, Seattle, WA 98195, USA; Kaier Permanente Washington Health Research Institute, Seattle, WA 98195, USA.
40. Departments of Medicine, Biomedical Engineering and Mental Health, Johns Hopkins University, Baltimore, MD 21205, USA.
41. Center for Population Epigenetics, Robert H. Lurie Comprehensive Cancer Center and Department of Preventive Medicine, Northwestern University , Feinberg School of Medicine, Chicago, IL 60611, USA.
42. Intramural Research Program, National Institute on Aging, National Institutes of Health, Baltimore, MD 21224, USA.
43. Division of Cardiology, Departments of Medicine and Epidemiology, Cardiovascular Health Research Unit, University of Washington, Seattle, WA 98101, USA.
44. Helmholtz Zentrum München, Institute of Epidemiology, Neuherberg, Germany; German Research Center for Cardiovascular Disease (DzHK e.V. - partner site Munich), Munich, Germany; Ludwig-Maximilians University, Institute for Biometry, Medical Information Science and Epidemiology, Munich, Germany.
45. Brown Foundation Institute of Molecular Medicine McGovern Medical School, and Human Genetics Center, School of Public Health, The University of Texas Health Science Center at Houston, Houston, TX 77030, USA.
46. Department of Medicine (Cardiovascular Medicine), and Department of Health Research & Policy, Stanford University School of Medicine, Stanford, CA 94305, USA.
47. Department of Epidemiology, Gillings School of Global Public Health, and Department of Medicine, School of Medicine, University of North Carolina, Chapel Hill, NC 27514, USA.
48. Framingham Heart Study, Framingham, MA 01702, USA; Population Sciences Branch, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD 20892, USA.

Corresponding author:
Golareh Agha, PhD
1) Laboratory of Environmental Precision Biosciences; 2) Columbia Aging Center
Columbia University Mailman School of Public Health
Columbia University, NY, NY, 10032
Golareh.agha@columbia.edu; Golareh.agha@gmail.com
401-261-3365

Manuscript word count (not including title, abstract, acknowledgment, references, tables, and figure legends): 3,093; (limit: 3000)

KEY POINTS

Question:

Does blood leukocyte DNA methylation predict risk of coronary heart disease (CHD) and myocardial infarction (MI)?

Finding:

In this large-scale study of more than 11,000 participants, we assessed blood DNA methylation at more than 400,000 cytosine-phosphate-guanine (CpG) dinucleotide across the genome and identified 52 CpG sites at which DNA methylation levels are associated with risk of CHD and MI. Several of these CpGs localize to genes that implicate biological function related to calcium regulation, among others. The addition of CHD-associated CpGs to traditional CHD risk factors improved CHD risk prediction, and two CpGs showed evidence of a causal effect of DNA methylation on incident CHD.

Meaning

Our findings demonstrate a robust link between blood-derived DNA methylation markers and risk of future CHD, supporting the role of DNA methylation as a potential molecular biomarker to inform CHD risk.

ABSTRACT (word count 347, limit 350)

IMPORTANCE: Epigenetic modifications such as DNA methylation have been implicated in coronary heart disease (CHD), but prior evidence in humans is based on small, cross-sectional studies. Whether DNA methylation is associated with risk of future CHD remains undetermined.

OBJECTIVE: We examined whether blood DNA methylation is associated with incident CHD across large observational studies, and whether DNA methylation improves CHD risk prediction beyond traditional CHD risk factors.

DESIGN/SETTING: Nine population-based cohorts from the United States and Europe obtained genome-wide DNA methylation data via the Illumina Infinium 450k microarray, and prospectively ascertained new-onset CHD events. Each cohort conducted race-specific analyses adjusted for age, sex, smoking status, education, body mass index, differential blood cell counts, and technical variables. We then conducted fixed-effect meta-analyses across cohorts.

PARTICIPANTS: A total of 11,461 individuals (mean age 64 years, 67% women, 35% African-American) who were free of CHD at baseline

EXPOSURE: Blood leukocyte DNA methylation levels across 442,192 cytosine-phosphate-guanine (CpG) sites.

OUTCOME(S): Incident CHD including coronary insufficiency/unstable angina, recognized myocardial infarction (MI), coronary revascularization, and coronary death. Recognized nonfatal and fatal MI was examined as a secondary outcome.

RESULTS: During a mean follow-up of 11.2 years, 1,895 individuals developed new-onset CHD. Methylation levels at 52 CpG sites were associated with incident CHD or MI (false discovery rate<0.05); these sites were not previously identified in genomic or DNA methylation studies of CHD.
Several of these CpGs map to genes involved in serum calcium regulation (*ATP2B2*, *CASR, GUCA1B, HPCAL1)* or kidney function decline (*CDH23*, *HPCAL1).* Mendelian randomization analyses revealed that DNA methylation at two CpGs is causally associated with incident CHD; these CpGs map to active regulatory regions proximal to long non-coding RNA transcripts.

Analyses restricted to participants of African-American or European ancestry demonstrated concordant results with analysis of the combined group. In an independent discovery and test-set design, adding CHD-associated CpGs to established CHD risk factors improved the risk prediction model performance (c-statistic increased from 0.735 to 0.766 [p=0.04]).

CONCLUSION/RELEVANCE: Methylation of blood-derived DNA is associated with risk of future CHD across diverse populations, and may serve as an informative biomarker for prediction of CHD.

INTRODUCTION

Coronary heart disease (CHD) is a major contributor to global morbidity and mortality.[1] Despite substantial progress in CHD prevention, more-targeted approaches are needed to further reduce the persistent burden due to high incidence of CHD. Methylation of DNA at cytosine-phosphate-guanine (CpG) dinucleotides is a widely-characterized epigenetic regulatory mechanism that is stable, but is modifiable and thus can be environmentally responsive. DNA methylation marks at a CpG site can thus reflect both underlying genetic variation as well as exposures to external environmental factors.[2] In vitro and animal-based studies provide evidence that DNA methylation changes are involved in the development of CHD,[3] and large-scale population-based studies have more recently shown that risk factors for CHD including smoking,[4] obesity,[5] and hypertension[6,7] are linked to persistent DNA methylation marks in blood. Hence DNA methylation, as a molecular bio-archive integrating genetic predisposition and exposure to risk factor exposures, may serve as a predictor of CHD and identify potentially modifiable pathways related to CHD. Prior studies of DNA methylation and CHD in humans[8-12] have generally been small in sample size ($n < 300$), focused on repetitive elements[8,10] or selective genomic regions,[12] or have been cross-sectional or case-control in design.[8,10-12] However, whether blood DNA methylation predicts future CHD has not been comprehensively investigated to date.

We conducted a longitudinal, large-scale, multi-cohort, epigenome-wide investigation of incident CHD among 11,461 participants in the Cohorts for Heart and Aging Genetic Epidemiology (CHARGE) consortium,[13] an international consortium formed to conduct large-scale genome-wide research across population-based cohort studies worldwide. We first assessed whether leukocyte DNA methylation was associated with risk of CHD, and further combined knowledge of CHD-associated methylation changes with genetic sequence variation to assess causality between DNA methylation and incident CHD. We further assessed whether DNA methylation could improve CHD risk prediction beyond traditional CHD risk factors.

METHODS

Study design and population

We selected cohorts participating in the CHARGE Consortium who assayed genome-wide leukocyte DNA methylation using an Infinium 450k microarray, and prospectively ascertained CHD events thereafter. Nine population-based cohorts comprising a total of 11,461 participants from the United States and Europe were included: the Atherosclerosis Risk in Communities Study (ARIC), Cardiovascular Health Study (CHS), long-tErm follow-up of antithrombotic management Patterns In acute CORonary syndrome patients (EPICOR), the Framingham Heart Study (FHS), the Invecchiare in Chianti study (InCHIANTI), the Kooperative Gesundheitsforschung in der Region Augsburg study (KORA), the Normative Aging Study (NAS), the Women's Health Initiative "Epigenetic Mechanisms of Particulate Matter-Mediated CVD" (WHI-EMPC) ancillary study, and the "Integrative genomics and risk of CHD and related phenotypes in the Women's Health Initiative" (WHI-BAA23) Ancillary study (Detailed information on each cohort's methods in **Cohort Supplementary info**). Each cohort study obtained informed consent from participants and ethics approval from its respective institutional review board and ethics committee. Cohorts comprising participants of both African-American and European Ancestry were separated into race-specific samples for analyses. Accordingly, we performed an epigenome-wide analysis for each of 12 study samples, and then meta-analyzed the resulting summary statistics from the 12 analyses. We also examined the association between DNA methylation and *cis*-genetic variants (±500 KB) in a subset of the cohorts (with available genomic data) and conducted Mendelian Randomization to evaluate potential causal relations between DNA methylation and incident CHD. We further applied an independent discovery and test set design to evaluate the CHD prediction model performance with the addition of DNA methylation information to clinical CHD risk factors (complete study workflow depicted in **eFigure 1**).

Measurement of DNA methylation

For all cohorts, DNA was extracted from whole blood samples and bisulfite-converted using a Zymo EZ DNA methylation kit. The Illumina Infinium Human Methylation450K BeadChip (Illumina Inc, San Diego, CA, USA) was used to measure DNA methylation. Following standard practice,[4,6] quality control, filtering, and normalization of the methylation data were independently conducted for each cohort according to standard criteria and other diagnostics unique to the cohort (details in **Cohort Supplementary info**). For each CpG, methylation = M/(M+U+ε), where M and U are the average fluorescence intensity from the probe (i.e., the oligonucleotide that hybridizes to the target CpG) corresponding to the methylated (M) and unmethylated (U) target CpG, respectively, and ε=100 to protect against division by zero. Therefore, the methylation at each CpG is contained in the interval 0-1, with 0 indicating no methylation and 1 indicating 100% methylation across DNA from blood cells in the sample.

Definition of coronary heart disease (CHD) and myocardial infarction (MI) events

Our primary outcome of interest was incident CHD, defined as any of the following: recognized nonfatal or fatal MI (hospitalization with diagnostic electrocardiographic (ECG) changes and/or biomarkers of MI), coronary insufficiency/unstable angina, coronary revascularization, or coronary death. We also conducted a secondary meta-analyses restricted to incident MI-only (recognized nonfatal or fatal MI), in order to evaluate whether analysis with this more homogenous outcome measure altered the results.

Individual study epigenome-wide analyses

Baseline was defined as the time of blood sampling for DNA methylation assays, and all cohorts excluded individuals with prevalent CHD at baseline. Seven cohorts conducted time-to-event analyses using Cox proportional hazard models, and three of these adapted Firth's penalized Cox regression[14] due to a low number of CHD events. Two prospective cohorts, EPICOR and WHI-BAA23, employed a nested case-control design with incident CHD events and performed logistic regression analyses, which—under

specific assumptions—provide risk estimates that are unbiased in relation to the estimates derived from Cox regression. We conducted a sensitivity meta-analysis that excluded these two cohorts and evaluated the concordance with the meta-analysis that included all cohorts. All analyses were race-specific, and adjusted for age, sex, body mass index (kg/m2), smoking status (current, former, never), education (as years of education or categorical levels of school degrees completed), differential cell counts,[15] family structure (if present), and batch-related technical variables (**Cohort Supplementary info**).

Meta-analysis

We performed an inverse-variance-weighted fixed-effects meta-analysis using the *metafor* package in R. We accounted for multiple-testing by controlling the false discovery rate (FDR) at 5%. Of the CpGs that exceeded this *a priori* multiple-testing threshold, we excluded CpGs harboring a single nucleotide polymorphism [SNP] assayed in the 1000 Genomes Project with an ancestry-specific minor allele frequency >0.01, and CpGs that had high inter-study heterogeneity assessed using Cochran's Q test (Q <0.05/ number of significant CpGs).

Identification of associated genetic variants and Mendelian Randomization analyses

We investigated whether genetic variants within ±500 kb (*cis*) of the incident CHD- and MI-associated CpGs contributed to variation in methylation levels, i.e., were methylation-quantitative trait loci (meQTLs). The discovery analysis was conducted on 3868 individuals from the FHS, followed by replication in KORA. Genotyping was conducted with the Affymetrix 500K and MIPs 50K platforms in FHS, and the Affymetrix Axiom array in KORA, and imputation was performed using the 1000 Genomes reference panel in both cohorts. meQTLs with P $<1x10^{-4}$ at the discovery stage, followed by P<bonferroni threshold (i.e. P<0.05/number of significant discovery stage meQTLs) at the replication stage were selected. Using these selected meQTLs, we then implemented a two-sample instrumental variable approach as implemented in MRbase (http://www.mrbase.org/) to infer causal relations between DNA

methylation and incident CHD. Genotype associations for CHD and MI were obtained from the CARDIoGRAMplusC4D 2015 GWAS (n= 60,801 cases and n= 123,504 controls).[16]

Improvement in risk prediction metrics

We conducted an additional discovery meta-analysis that included all cohorts except the FHS. We used CpGs associated with incident CHD at an FDR < 0.05 to test for improvement in the Cox proportional hazard model performance in the FHS. We tested for improvement in the c-statistic compared to a base model of clinical CHD variables in the American College of Cardiology / American Heart Association (ACC/AHA) Pooled Cohort Atherosclerotic Cardiovascular Disease (ASCVD) Risk Calculator: age, sex, total cholesterol, high density lipoprotein cholesterol, systolic blood pressure, diastolic blood pressure, hypertension treatment, diabetes, and current cigarette smoking.

RESULTS

Participant characteristics

Among the 11,461 participants included, mean age at baseline was 64 years, 67% were female, and 35% were of African-American ancestry (**Table 1**). During a mean follow-up of 11.2 years, a total of 1,895 CHD events occurred, and 1,183 of these were MI events.

Association of DNA methylation with risk of Coronary heart disease (CHD) and myocardial infarction (MI)

Among 442,192 CpGs analyzed, methylation levels at 30 CpGs were associated with incident CHD (multiple testing-adjusted FDR p-value <0.05; **Table 2** and **eFigure 2).** Methylation levels at 29 CpGs were associated with our secondary outcome of incident MI at a FDR p-value <0.05 (**Table 3** and **eFigure 3)**. Among these 30 and 29 CpGs identified in the incident CHD and incident MI-only meta-analyses, respectively, seven CpGs met the FDR<0.05 threshold in both analyses and thus overlapped,

meaning that there were 52 unique CpGs identified across the two meta-analyses. When we compared results for these 52 CpGs, we found that the direction, magnitude, and precision of estimated effects were highly concordant when comparing results from these two meta-analyses (**Figure 1**). Manhattan plots indicated that significant associations were distributed across the genome (**Figure 2**). Neither meta-analysis was strongly influenced by inflation from technical or batch effects. In addition, both had a uniform distribution of p-values and symmetry in the coefficient direction of effect (**Figure 2**). Therefore, we concluded that results obtained from the secondary, incident MI-only meta-analysis did not materially differ from the primary, CHD meta-analysis, and henceforth combined the results from the two meta-analyses and simply refer to all 52 CpGs as CHD-associated CpGs.

Additional sensitivity-analyses and race-specific meta-analyses

Our findings for the 52 CHD-associated CpGs were highly consistent when we compared results obtained from the meta-analysis of all nine cohorts as described above (n=11,461) with results from a meta-analysis on only the 7 cohorts that performed Cox regression (n=9,255) (**eFigure 4**). Similarly, we performed four additional meta-analyses, each time excluding one of the four largest cohorts (FHS, NAS, ARIC, KORA), and found similar results across these meta-analyses for the 52 CHD-associated CpGs (**eFigure 5**). In race-specific meta-analyses, the effect size and direction of effects for the majority of the CHD-associated CpGs were similar when comparing those of European ancestry with African-American ancestry (**eFigure 6**). However, 11 of the 52 CpGs showed race-specific differences in the association of DNA methylation with incident CHD (p-value <0.05 for difference in t-statistic; **eTable 1**).

Mendelian Randomization for identifying causal associations between DNA methylation and incident CHD

We identified and replicated an association between DNA methylation and cis genetic variation (i.e. meQTLs) for 10 of the 52 CHD-associated CpGs, and used these meQTLs in a two-step Mendelian

Randomization analysis (**eTable 2**). For two of the 10 CpGs with replicated meQTLs, we observed a causal effect of DNA methylation on incident CHD: cg26470101 (β [95% CI] for 1% increase in DNA methylation = 0.042 [0.002, 0.08]; P = 0.037) and cg07289306 (β [95% CI] for 1% = -0.148 [-0.288, -0.009]; P = 0.04) on CHD (**Table 4).** Both CpGs map to regulatory active intergenic regions within CpG islands (i.e. regions important for transcription regulation), and cg07289306 is located proximal to two long non-coding RNA transcripts (**eFigure 7**).

Improvement in CHD risk prediction using DNA methylation

DNA methylation at 30 CpGs was associated (FDR <0.05) with incident CHD in the discovery meta-analyses that excluded the FHS cohort. We then tested for improvement in model performance in the FHS cohort: adding these 30 CpGs to a baseline model that included the same covariates as the ACC/AHA pooled cohort ASVCD risk calculator (except for race) increased the c-statistic from 0.735 to 0.766 (p=0.04).

DISCUSSION

We conducted a large-scale analysis of DNA methylation in relation to incident CHD and MI in 11,461 adults across multiple North American and European cohorts. Methylation levels at 52 CpGs across the genome were associated with risk of CHD and MI. A 5% increase in DNA methylation was related to differences in CHD risk of a clinically relevant magnitude, ranging from a 46% decrease in the risk of CHD (cg12766383) to a 65% increase in risk (cg05820312). Mendelian randomization analyses revealed that methylation at two CpGs had a causal effect on incident CHD. Further, the performance of a CHD risk model with variables from a gold-standard clinical risk calculators was improved upon addition of CHD-associated CpGs – and the improvements in the c-statistic were of a similar magnitude to those obtained when adding coronary artery calcium scores.[17,18]

Biological relevance and clinical implications

The identified 52 CHD-associated CpGs in our study implicate calcium regulation, kidney function, and TGF-beta pathways in CHD risk. Specifically, CpG cg2261787 maps to the ATPase plasma membrane calcium transporter 2 (*ATP2B2*) gene, which encodes a protein in the plasma membrane calcium transporter (PMCA) family with critical roles in intracellular calcium homeostasis. GWAS [19]have linked genes encoding PMCAs to blood pressure and hypertension, and there is evidence that PMCAs have therapeutic potential in novel anti-hypertensive treatments.[20] Similarly, CpG cg06582394 maps to the calcium sensing receptor (*CASR*) gene, which encodes a protein with a key role in calcium homeostasis. In a recent Mendelian randomization analysis of 184,305 individuals, Larsson *et al*. reported that a genetic variant at the *CASR* locus showed strong associations with coronary artery disease and MI.[21] CpGs cg14010194 and cg03467256 map to guanylate cyclase activator 1B (*GUCA1B*) and hippocalcin-like 1 (*HPCAL1*), respectively, which encode calcium-binding proteins with roles in calcium-dependent regulation.[22,23] Observational studies and calcium supplementation randomized clinical trials provide evidence of associations between serum calcium levels and increased risk of CHD and MI.[24,25] Our results provide the first evidence that epigenetic regulation may be involved in calcium-related CHD risk, an underdeveloped area of therapeutics.

CpGs cg19227382, cg03467256, and cg25497530 map to gene loci that are potentially relevant to kidney function, including cadherin-related 23 (*CDH23*), *HPCAL1*, and protein tyrosine phosphatase receptor type N2 (*PTPRN2*). Both *CDH23 and HPCAL1* were identified in a GWAS of kidney function in approximately 64,000 participants of European decent.[26] An epigenome-wide study of 400 individuals showed differential blood DNA methylation at the *PTPRN2* locus in chronic kidney disease cases relative to controls.[27] Kidney function is a well-recognized risk factor for CVD, with a recent AHA report highlighting that individual with an estimated glomerular filtration rate (eGFR) of 15 to 30 mL/min per 1.73 m^2 have the highest adjusted relative risk of CVD mortality.[28] In a study of 1.27 million individuals with a median follow-up for 48 months, the rate of incident MI in people with diabetes was substantially lower than for those with chronic kidney disease.[29]

Other gene loci identified include the insulin growth factor 1 receptor (*IGF1R*), transforming growth factor beta receptor 1 (*TGFBR1*), and integrin subunit beta 2 (*ITGB2)*. The roles of IGF1R and the TGF-beta signaling in cardiac remodeling and function are well recognized,[30,31] and recently *TGFBR1* gene expression levels in blood samples from acute MI patients strongly predicted left-ventricular dysfunction.[32] Finally, *ITGB2* encodes a leukocyte cell-surface adhesion molecule that directly facilitates leukocyte transendothelial migration, a key step in formation of atherosclerosis.[33]

Both CpGs with evidence of a causal effect on CHD, cg20545941 and cg07289306, are located within active CpG islands in intergenic regions, indicating that they may be involved in important regulatory processes with respect to CHD. Additionally, cg07289306 is proximal to two long non-coding RNAs (lncRNAs), and increasing evidence indicates that lncRNAs are key components of transcriptional regulatory pathways that govern cardiac development and cardiovascular pathophysiology.[34,35]

When we added the CHD-associated DNA methylation CpGs to traditional CVD risk factors, we observed a significant improvement in CHD risk prediction (according to the c-statistic) equal in magnitude to risk prediction improvements observed when coronary artery calcification scores, an established indicator of CHD risk, are added to traditional risk prediction models.[17,18] Additional studies may thus benefit from comprehensively evaluating DNA methylation biomarkers for improvement in prediction, discrimination, and reclassification of CHD.

Findings in the context of prior evidence

Our findings did not overlap with those of previous studies, which may be because prior studies were individual, smaller, single-cohort studies and were often composed of select populations geographically and ethnically distinct from the populations in our meta-analysis. For example, Sharma *et al*. identified differentially methylated regions (DMRs) near or within genes *C1QL4*, *CCDC47*, and *TGFBR3* in a study of 36 men (18 CAD, 18 controls) from India.[12] Nakatochi et al. compared 192 MI cases with 192 controls

in an epigenome-wide whole-blood analysis on elderly Japanese individuals, and reported DNA methylation at two CpGs, located in the *ZFHX3* and *SMARCA4* genes, to be associated with MI.[11] In the prospective Italian EPICOR cohort, Guarrera *et al*. compared 292 MI cases with 292 matched controls ascertained prospectively during follow-up, and reported that a differentially-methylated region (DMR), within the Zinc Finger And BTB Domain Containing 2 (*ZBTB12)* gene body was associated with MI.[9]

Study limitations

We used well-established statistical procedures to remove the effect cell-type heterogeneity, which can serve as a major source of confounding in epigenome-wide studies and has previously been well-described.[15,36] However, residual confounding is still possible. Additionally, our Mendelian randomization analyses provide evidence supporting a causal association but do not prove causality, and thus, follow-up experimental work is warranted. Another limitation is the relatively large contribution from cohorts based in primarily Western countries in Europe and the United States due to limited availability of DNA methylation and incident CHD data in more ethnically diverse cohorts. Thus our findings have not been evaluated for generalizability to other populations.

Study strengths

Our study is by far the largest on the topic to date, with nearly 12,000 participants. We also made use of incident cases that were stringently adjudicated over a long-term follow-up, another novelty with respect to epigenome-wide studies. Our analysis allows for detection of effects at individual epigenetic loci that were potentially missed in prior smaller studies. Our longitudinal epigenome-wide approach provides evidence for the role of DNA methylation in determining risk of future CHD, which was not possible in previous studies due to their cross-sectional design. Furthermore, we used Mendelian randomization to build evidence regarding causal effects of DNA methylation on incident CHD, and evaluated CHD risk prediction improvement as compared to traditional risk factors.

CONCLUSION

We present novel findings on associations of leukocyte DNA methylation with risk of CHD, with effect sizes and risk prediction metrics that are of a clinically relevant magnitude. In addition, our findings highlight known as well as under-recognized pathways to CHD, including calcium regulation, kidney function, and gene regulation mechanism involving non-coding RNA. Overall, the findings provide a deeper understanding of the molecular landscape of incident CHD and may present novel avenues for the development of therapeutic interventions and improvements in risk prediction.

Role of the funding source

ARIC has been funded in whole or in part with Federal funds from the National Heart, Lung, and Blood Institute, National Institutes of Health, Department of Health and Human Services (contract numbers HHSN268201700001I, HHSN268201700002I, HHSN268201700003I, HHSN268201700004I and HSN268201700005I). The authors thank the staff and participants of the ARIC study for their important contributions. Funding was also supported by 5RC2HL102419 and R01NS087541. **CHS** was supported by NHLBI contracts HHSN268201200036C, HHSN268200800007C, HHSN268201800001C, N01HC55222, N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, N01HC85086; and NHLBI grants U01HL080295, R01HL087652, R01HL105756, R01HL103612, R01HL120393, and U01HL130114 with additional contribution from the National Institute of Neurological Disorders and Stroke (NINDS). Additional support was provided through R01AG023629 from the National Institute on Aging (NIA). A full list of principal CHS investigators and institutions can be found at CHS-NHLBI.org. The provision of genotyping data was supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR000124, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The work in **EPICOR** is supported by the Compagnia di San Paolo for the EPIC-Italy and EPICOR projects, the Human Genetics Foundation-Turin (HuGeF) / Italian Institute for Genomic Medicine (IIGM), and by Ministero dell'Istruzione, dell'Università e della Ricerca – MIUR project "Dipartimenti di Eccellenza 2018 – 2022" to Department of Medical Sciences, University of Turin. EPIC-Italy is further supported by a grant from the "Associazione Italiana per la Ricerca sul Cancro" (AIRC, Milan). **FHS** is funded by National Institutes of Health contract N01-HC-25195. The laboratory work for this investigation was funded by the Division of Intramural Research, National Heart, Lung, and Blood Institute, National Institutes of Health. The analytical component of this project was funded by the Division of Intramural Research, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD. The **InCHIANTI** data are available on request at http://inchiantistudy.net/wp/inchianti-dataset/;**KORA** was initiated and financed by the Helmholtz Zentrum München-Germany Research Center for Environmental Health, which is funded by the German Federal Ministry of Education and Research (BMBF) and by the State of Bavaria. Furthermore, KORA research was supported within the Munich Center of health Sciences, Ludwig Maximilians-Universität, as part of LMUinnovativ. This work was supported in part by the BMBF within the framework of the e:Med research and funding concept (e:AtheroSysMed, grant 01ZX1313A-2014). **NAS** is supported by grants from the National Institute of Environmental Health Sciences (NIEHS) (R01ES021733, R01ES025225, R01ES027747). Other support comes from NIEHS grants ES015172, ES014663 and ES020010, Environmental Protection Agency (EPA) grant

RD832416, as well as grant P30ES009089. It is supported by the Cooperative Studies Program/ERIC, US Department of Veterans Affairs, and is a research component of the Massachusetts Veterans Epidemiology Research and Information Center (MAVERIC). The views expressed in this paper are those of the authors and do not necessarily represent the views of the US Department of Veterans Affairs. Additional support was provided by the US Department of Agriculture, Agricultural Research Service (contract 53-K06-510). The views expressed in this paper are those of the authors and do not necessarily represent the views of the US Department of Veterans Affairs. **WHI-EMPC** is funded by the National Institute of Environmental Health Sciences (R01-ES020836; Whitsel, Baccarelli, Hou). The WHI program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, U.S. Department of Health and Human Services through contracts HHSN268201100046C, HHSN268201100001C, HHSN268201100002C, HHSN268201100003C, HHSN268201100004C, and HHSN271201100004C. LH was funded by AHA grant "14SFRN20790000". The authors thank the WHI investigators and staff for their dedication, and the study participants for making the program possible. A listing of WHI investigators can be found at http://www.whi.org/researchers/Documents%20%20Write%20a%20Paper/WHI%20Investigator%20Short%20List.pdf. **WHI-BAA23** was supported by NHLBI's Broad Agency Announcement contract HHSN268201300006C. The WHI program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, U.S. Department of Health and Human Services through contracts HHSN268201100046C, HHSN268201100001C, HHSN268201100002C, HHSN268201100003C, HHSN268201100004C, and HHSN271201100004C.

Acknowledgments

The authors would like to thank the KORA participants for their participation in the KORA cohort. The authors wish to thank all who participated in, or collaborated with, EPIC; the Italian AVIS blood donor organization, AIRE-ONLUS di Ragusa, and the Sicilian Government.

Data and Materials relevant to this study

The DNA methylation dataset from **ARIC** is available on request at https://www2.cscc.unc.edu/aric/distribution-agreements. The **CHS** data can be requested at https://chs-nhlbi.org/node/6222. **EPICOR** data are available upon request from HuGeF (http://www.hugef-torino.org/) via a project agreement. Requests should be sent to info@hugef-torino.org. The **FHS** DNA methylation datasets are available from the dbGAP repository: phs000724. The genotype datasets are available from the dbGAP repository: phs000007. The **InCHIANTI** data are available on request at http://inchiantistudy.net/wp/inchianti-dataset/. The **KORA** data can be requested at KORA PASST: https://epi.helmholtz-muenchen.de/. The **NAS** DNA methylation datasets are available at the dbGAP repository: phs000853. The **WHI-BAA24** DNA methylation dataset is available at dbGAP repository phs001335. **WHI-EMPC** data are available on request: http://www.whi.org/.

Author Contributions

GA, AAB, MMM, CWC, RJ, and DLwere involved in designing the study and preparing the analysis plan. AAB, DL, EAW, TLA, MF, AP, GM, NS, LF, LH were involved in the acquisition of DNA methylation data. GA and RJ were involved in preparing the analytic programming code. GA, CWC, TH, RG, ES, JAB, GF, JB, BHC were involved in conducting the individual cohort epigenome-wide analyses. CWC, MMM, RJ, TH, GA were involved in additional analyses of gene variants and Mendelian randomization in FHS and KORA. GA, AAB, MMM, CWC, RJ were involved in drafting of the manuscript.
All authors listed contributed to interpretation of the data and results, enhancement of the analyses, multiple revisions of the manuscript, and final approval of the submitted manuscript.

Disclosures

Bruce M. Psaty serves on the DSMB of a clinical trial funded by Zoll LifeCor and on the steering committee of the Yale Open Data Access Project funded by Johnson & Johson. Brian H. Chen was an employee of the US National

Institutes of Health during this study, but is currently employed by Life Epigenetics, Inc,, which had no influence on the analysis, interpretation, or reporting of this publication.

REFERENCES

1. Butler D. UN targets top killers. *Nature* 2011; **477**(7364): 260-1.
2. Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nature genetics* 2003; **33 Suppl**: 245-54.
3. Joehanes R, Just AC, Marioni RE, et al. Epigenetic Signatures of Cigarette Smoking. *Circulation Cardiovascular genetics* 2016; **9**(5): 436-47.
4. Dick KJ, Nelson CP, Tsaprouni L, et al. DNA methylation and body-mass index: a genome-wide analysis. *Lancet* 2014.
5. Richard MA, Huan T, Ligthart S, et al. DNA Methylation Analysis Identifies Loci for Blood Pressure Regulation. *American journal of human genetics* 2017; **101**(6): 888-902.
6. Ligthart S, Marzi C, Aslibekyan S, et al. DNA methylation signatures of chronic low-grade inflammation are associated with complex diseases. *Genome biology* 2016; **17**(1): 255.
7. Psaty BM, Sitlani C. The Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium as a model of collaborative science. *Epidemiology* 2013; **24**(3): 346-8.
8. Heinze G, Schemper M. A solution to the problem of monotone likelihood in Cox regression. *Biometrics* 2001; **57**(1): 114-9.
9. Nikpay M, Goel A, Won HH, et al. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nature genetics* 2015; **47**(10): 1121-30.
10. Pierce BL, Burgess S. Efficient design for Mendelian randomization studies: subsample and 2-sample instrumental variable estimators. *American journal of epidemiology* 2013; **178**(7): 1177-84.
11. 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).
12. Levy D, Ehret GB, Rice K, et al. Genome-wide association study of blood pressure and hypertension. *Nature genetics* 2009; **41**(6): 677-87.
13. Little R, Cartwright EJ, Neyses L, Austin C. Plasma membrane calcium ATPases (PMCAs) as potential targets for the treatment of essential hypertension. *Pharmacology & therapeutics* 2016; **159**: 23-34.
14. Larsson SC, Burgess S, Michaelsson K. Association of Genetic Variants Related to Serum Calcium Levels With Coronary Artery Disease and Myocardial Infarction. *Jama* 2017; **318**(4): 371-80.
15. Lanfear DE, Yang JJ, Mishra S, Sabbah HN. Genome-wide approach to identify novel candidate genes for beta blocker response in heart failure using an experimental model. *Discovery medicine* 2011; **11**(59): 359-66.
16. Korkmaz S, Radovits T, Barnucz E, et al. Pharmacological activation of soluble guanylate cyclase protects the heart against ischemic injury. *Circulation* 2009; **120**(8): 677-86.
17. Rohrmann S, Garmo H, Malmstrom H, et al. Association between serum calcium concentration and risk of incident and fatal cardiovascular disease in the prospective AMORIS study. *Atherosclerosis* 2016; **251**: 85-93.
18. Bolland MJ, Grey A, Avenell A, Gamble GD, Reid IR. Calcium supplements with or without vitamin D and risk of cardiovascular events: reanalysis of the Women's Health Initiative limited access dataset and meta-analysis. *Bmj* 2011; **342**: d2040.
19. Gorski M, Tin A, Garnaas M, et al. Genome-wide association study of kidney function decline in individuals of European descent. *Kidney international* 2015; **87**(5): 1017-29.
20. Smyth LJ, McKay GJ, Maxwell AP, McKnight AJ. DNA hypermethylation and DNA hypomethylation is present at different loci in chronic kidney disease. *Epigenetics : official journal of the DNA Methylation Society* 2014; **9**(3): 366-76.

21. Laustsen PG, Russell SJ, Cui L, et al. Essential role of insulin and insulin-like growth factor 1 receptor signaling in cardiac development and function. *Molecular and cellular biology* 2007; **27**(5): 1649-64.
22. Dobaczewski M, Chen W, Frangogiannis NG. Transforming growth factor (TGF)-beta signaling in cardiac remodeling. *Journal of molecular and cellular cardiology* 2011; **51**(4): 600-6.
23. Devaux Y, Bousquenaud M, Rodius S, et al. Transforming growth factor beta receptor 1 is a new candidate prognostic biomarker after acute myocardial infarction. *BMC medical genomics* 2011; **4**: 83.
24. Klarin D, Zhu QM, Emdin CA, et al. Genetic analysis in UK Biobank links insulin resistance and transendothelial migration pathways to coronary artery disease. *Nature genetics* 2017; **49**(9): 1392-7.
25. Thum T, Condorelli G. Long noncoding RNAs and microRNAs in cardiovascular pathophysiology. *Circulation research* 2015; **116**(4): 751-62.
26. Baccarelli A, Rienstra M, Benjamin EJ. Cardiovascular epigenetics: basic concepts and results from animal and human studies. *Circulation Cardiovascular genetics* 2010; **3**(6): 567-73.
27. Baccarelli A, Wright R, Bollati V, et al. Ischemic heart disease and stroke in relation to blood DNA methylation. *Epidemiology* 2010; **21**(6): 819-28.
28. Guarrera S, Fiorito G, Onland-Moret NC, et al. Gene-specific DNA methylation profiles and LINE-1 hypomethylation are associated with myocardial infarction risk. *Clinical epigenetics* 2015; **7**: 133.
29. Kim M, Long TI, Arakawa K, Wang R, Yu MC, Laird PW. DNA methylation as a biomarker for cardiovascular disease risk. *PloS one* 2010; **5**(3): e9692.
30. Sharma P, Garg G, Kumar A, et al. Genome wide DNA methylation profiling for epigenetic alteration in coronary artery disease patients. *Gene* 2014; **541**(1): 31-40.
31. Nakatochi M, Ichihara S, Yamamoto K, et al. Epigenome-wide association of myocardial infarction with DNA methylation sites at loci related to cardiovascular disease. *Clinical epigenetics* 2017; **9**: 54.

1. Butler D. UN targets top killers. *Nature.* 2011;477(7364):260-261.
2. Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nature genetics.* 2003;33 Suppl:245-254.
3. Baccarelli A, Rienstra M, Benjamin EJ. Cardiovascular epigenetics: basic concepts and results from animal and human studies. *Circulation Cardiovascular genetics.* 2010;3(6):567-573.
4. Joehanes R, Just AC, Marioni RE, et al. Epigenetic Signatures of Cigarette Smoking. *Circulation Cardiovascular genetics.* 2016;9(5):436-447.
5. Dick KJ, Nelson CP, Tsaprouni L, et al. DNA methylation and body-mass index: a genome-wide analysis. *Lancet.* 2014.
6. Ligthart S, Marzi C, Aslibekyan S, et al. DNA methylation signatures of chronic low-grade inflammation are associated with complex diseases. *Genome biology.* 2016;17(1):255.
7. Richard MA, Huan T, Ligthart S, et al. DNA Methylation Analysis Identifies Loci for Blood Pressure Regulation. *American journal of human genetics.* 2017;101(6):888-902.
8. Baccarelli A, Wright R, Bollati V, et al. Ischemic heart disease and stroke in relation to blood DNA methylation. *Epidemiology.* 2010;21(6):819-828.
9. Guarrera S, Fiorito G, Onland-Moret NC, et al. Gene-specific DNA methylation profiles and LINE-1 hypomethylation are associated with myocardial infarction risk. *Clinical epigenetics.* 2015;7:133.
10. Kim M, Long TI, Arakawa K, Wang R, Yu MC, Laird PW. DNA methylation as a biomarker for cardiovascular disease risk. *PloS one.* 2010;5(3):e9692.

11. Nakatochi M, Ichihara S, Yamamoto K, et al. Epigenome-wide association of myocardial infarction with DNA methylation sites at loci related to cardiovascular disease. *Clinical epigenetics.* 2017;9:54.
12. Sharma P, Garg G, Kumar A, et al. Genome wide DNA methylation profiling for epigenetic alteration in coronary artery disease patients. *Gene.* 2014;541(1):31-40.
13. Psaty BM, Sitlani C. The Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium as a model of collaborative science. *Epidemiology.* 2013;24(3):346-348.
14. Heinze G, Schemper M. A solution to the problem of monotone likelihood in Cox regression. *Biometrics.* 2001;57(1):114-119.
15. Houseman EA, Accomando WP, Koestler DC, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC bioinformatics.* 2012;13:86.
16. Nikpay M, Goel A, Won HH, et al. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nature genetics.* 2015;47(10):1121-1130.
17. Folsom AR, Kronmal RA, Detrano RC, et al. Coronary artery calcification compared with carotid intima-media thickness in the prediction of cardiovascular disease incidence: the Multi-Ethnic Study of Atherosclerosis (MESA). *Archives of internal medicine.* 2008;168(12):1333-1339.
18. 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).
19. Levy D, Ehret GB, Rice K, et al. Genome-wide association study of blood pressure and hypertension. *Nature genetics.* 2009;41(6):677-687.
20. Little R, Cartwright EJ, Neyses L, Austin C. Plasma membrane calcium ATPases (PMCAs) as potential targets for the treatment of essential hypertension. *Pharmacology & therapeutics.* 2016;159:23-34.
21. Larsson SC, Burgess S, Michaelsson K. Association of Genetic Variants Related to Serum Calcium Levels With Coronary Artery Disease and Myocardial Infarction. *Jama.* 2017;318(4):371-380.
22. Korkmaz S, Radovits T, Barnucz E, et al. Pharmacological activation of soluble guanylate cyclase protects the heart against ischemic injury. *Circulation.* 2009;120(8):677-686.
23. Lanfear DE, Yang JJ, Mishra S, Sabbah HN. Genome-wide approach to identify novel candidate genes for beta blocker response in heart failure using an experimental model. *Discovery medicine.* 2011;11(59):359-366.
24. Bolland MJ, Grey A, Avenell A, Gamble GD, Reid IR. Calcium supplements with or without vitamin D and risk of cardiovascular events: reanalysis of the Women's Health Initiative limited access dataset and meta-analysis. *Bmj.* 2011;342:d2040.
25. Rohrmann S, Garmo H, Malmstrom H, et al. Association between serum calcium concentration and risk of incident and fatal cardiovascular disease in the prospective AMORIS study. *Atherosclerosis.* 2016;251:85-93.
26. Gorski M, Tin A, Garnaas M, et al. Genome-wide association study of kidney function decline in individuals of European descent. *Kidney international.* 2015;87(5):1017-1029.
27. Smyth LJ, McKay GJ, Maxwell AP, McKnight AJ. DNA hypermethylation and DNA hypomethylation is present at different loci in chronic kidney disease. *Epigenetics : official journal of the DNA Methylation Society.* 2014;9(3):366-376.
28. Benjamin EJ, Blaha MJ, Chiuve SE, et al. Heart Disease and Stroke Statistics-2017 Update: A Report From the American Heart Association. *Circulation.* 2017;135(10):e146-e603.
29. Tonelli M, Muntner P, Lloyd A, et al. Risk of coronary events in people with chronic kidney disease compared with those with diabetes: a population-level cohort study. *Lancet.* 2012;380(9844):807-814.

30. Dobaczewski M, Chen W, Frangogiannis NG. Transforming growth factor (TGF)-beta signaling in cardiac remodeling. *Journal of molecular and cellular cardiology.* 2011;51(4):600-606.
31. Laustsen PG, Russell SJ, Cui L, et al. Essential role of insulin and insulin-like growth factor 1 receptor signaling in cardiac development and function. *Molecular and cellular biology.* 2007;27(5):1649-1664.
32. Devaux Y, Bousquenaud M, Rodius S, et al. Transforming growth factor beta receptor 1 is a new candidate prognostic biomarker after acute myocardial infarction. *BMC medical genomics.* 2011;4:83.
33. Klarin D, Zhu QM, Emdin CA, et al. Genetic analysis in UK Biobank links insulin resistance and transendothelial migration pathways to coronary artery disease. *Nature genetics.* 2017;49(9):1392-1397.
34. Ono K, Kuwabara Y, Horie T, Kimura T. Long Non-Coding RNAs as Key Regulators of Cardiovascular Diseases. *Circ J.* 2018;82(5):1231-1236.
35. Thum T, Condorelli G. Long noncoding RNAs and microRNAs in cardiovascular pathophysiology. *Circulation research.* 2015;116(4):751-762.
36. Jaffe AE, Irizarry RA. Accounting for cellular heterogeneity is critical in epigenome-wide association studies. *Genome biology.* 2014;15(2):R31.