

OPEN

## Multi-ancestry epigenome-wide analyses identify methylated sites associated with aortic augmentation index in TOPMed MESA

Xiaowei Hu<sup>1,18</sup>, Jeongok G. Logan<sup>2,18</sup>, Younghoon Kwon<sup>3</sup>, Joao A. C. Lima<sup>4</sup>, David R. Jacobs<sup>5</sup>, Daniel Duprez<sup>6</sup>, Lyndia Brumback<sup>7</sup>, Kent D. Taylor<sup>8</sup>, Peter Durda<sup>9</sup>, W. Craig Johnson<sup>7</sup>, Elaine Cornell<sup>9</sup>, Xiuqing Guo<sup>8</sup>, Yongmei Liu<sup>10</sup>, Russell P. Tracy<sup>9</sup>, Thomas W. Blackwell<sup>11</sup>, George Papanicolaou<sup>12</sup>, Gary F. Mitchell<sup>13</sup>, Stephen S. Rich<sup>1</sup>, Jerome I. Rotter<sup>8</sup>, David J. Van Den Berg<sup>14</sup>, Julio A. Chirinos<sup>15</sup>, Timothy M. Hughes<sup>16</sup>, Francine E. Garrett-Bakelman<sup>17</sup>✉ & Ani Manichaikul<sup>1</sup>✉

Despite the prognostic value of arterial stiffness (AS) and pulsatile hemodynamics (PH) for cardiovascular morbidity and mortality, epigenetic modifications that contribute to AS/PH remain unknown. To gain a better understanding of the link between epigenetics (DNA methylation) and AS/PH, we examined the relationship of eight measures of AS/PH with CpG sites and co-methylated regions using multi-ancestry participants from Trans-Omics for Precision Medicine (TOPMed) Multi-Ethnic Study of Atherosclerosis (MESA) with sample sizes ranging from 438 to 874. Epigenome-wide association analysis identified one genome-wide significant CpG (cg20711926-*CYP1B1*) associated with aortic augmentation index (Aix). Follow-up analyses, including gene set enrichment analysis, expression quantitative trait methylation analysis, and functional enrichment analysis on differentially methylated positions and regions, further prioritized three CpGs and their annotated genes (cg23800023-*ETS1*, cg08426368-*TGFB3*, and cg17350632-*HLA-DPB1*) for Aix. Among these, *ETS1* and *TGFB3* have been previously prioritized as candidate genes. Furthermore, both *ETS1* and *HLA-DPB1* have significant tissue correlations between Whole Blood and Aorta in GTEx, which suggests *ETS1* and *HLA-DPB1* could be potential biomarkers in understanding pathophysiology of AS/

<sup>1</sup>Center for Public Health Genomics, University of Virginia, Charlottesville, VA 22908, USA. <sup>2</sup>School of Nursing, University of Virginia, Charlottesville, VA, USA. <sup>3</sup>Department of Medicine, University of Washington, Seattle, WA, USA. <sup>4</sup>Department of Internal Medicine, School of Medicine, Johns Hopkins University, Baltimore, MD, USA. <sup>5</sup>Division of Epidemiology, School of Public Health, University of Minnesota, Minneapolis, MN, USA. <sup>6</sup>Cardiovascular Division, University of Minnesota, Minneapolis, MN, USA. <sup>7</sup>Department of Biostatistics, University of Washington, Seattle, WA, USA. <sup>8</sup>The Institute for Translational Genomics and Population Sciences, Department of Pediatrics, The Lundquist Institute for Biomedical Innovation at Harbor-UCLA Medical Center, Torrance, CA, USA. <sup>9</sup>Laboratory for Clinical Biochemistry Research, University of Vermont, Burlington, VT, USA. <sup>10</sup>Division of Cardiology, Department of Medicine, Duke University, Durham, NC, USA. <sup>11</sup>Department of Biostatistics, School of Public Health, University of Michigan, Ann Arbor, MI, USA. <sup>12</sup>Epidemiology Branch, National Heart, Lung and Blood Institute, Bethesda, MD, USA. <sup>13</sup>Cardiovascular Engineering, Inc, Norwood, MA, USA. <sup>14</sup>Department of Preventive Medicine and Center for Genetic Epidemiology, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA. <sup>15</sup>Division of Cardiovascular Medicine, Department of Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. <sup>16</sup>Department of Internal Medicine - Section of Gerontology and Geriatric Medicine, and Department of Epidemiology and Prevention, Wake Forest School of Medicine, Winston-Salem, NC, USA. <sup>17</sup>Department of Biochemistry and Molecular Genetics, Department of Medicine, University of Virginia, 1340 Jefferson Park Ave., Pinn hall 6054, Charlottesville, VA 22908, USA. <sup>18</sup>These authors contributed equally: Xiaowei Hu and Jeongok G. Logan. ✉email: fg5q@UVAhealth.org; am3xa@virginia.edu

## PH. Overall, our findings support the possible role of epigenetic regulation via DNA methylation of specific genes associated with AIx as well as identifying potential targets for regulation of AS/PH.

Arterial stiffness (AS) serves as an independent predictor of cardiovascular diseases (CVD) morbidity and mortality<sup>1,2</sup>. Structurally, AS is characterized by degeneration of the media layer of arterial walls and involves decreased intact elastin and increased collagen fibers<sup>3</sup>. Functionally, AS impairs the cushioning capacity of arteries that transform pulsatile flow at the ascending aorta into steady flow through the arterioles<sup>3</sup>. Pulsatile hemodynamics (PH) resulting from AS involves the augmentation of systolic blood pressure (BP) in late systole<sup>4</sup> which partially results from wave reflection arriving at the aorta during ejection, increasing left ventricular load<sup>5</sup>.

Both genome-wide association studies (GWAS) and candidate gene approaches have identified genetic variants that are associated with AS/PH, including those in the renin–angiotensin–aldosterone (RAA) system, elastic fiber structural components, metalloproteinases, the nitric oxide (NO) pathway, and  $\beta$ -adrenergic receptors<sup>6–8</sup>. However, a large group of identified AS/PH-associated genes are not related to known pathophysiological mechanisms of AS/PH, reflecting the likely multifactorial etiology of AS/PH<sup>6</sup>. Since variation in AS/PH is mediated through gene expression, exploration of factors that are involved in transcriptional pathways may offer new insights to variation of AS/PH.

DNA methylation (DNAm), one of the most studied epigenetic modifications, plays a critical role in the transcriptional regulation<sup>9,10</sup>. DNAm occurs when a methyl group is added to a cytosine followed by guanine (CpG) in the genome<sup>10</sup>. Despite the emphasis on the role of AS/PH in the development of CVD, few human studies have examined DNAm in relation to AS/PH. In contrast, studies in aged mice have demonstrated that compound H, a potential activator of DNA demethylases, attenuates aging-related AS and hypertension<sup>11</sup>. In hyperhomocysteinemia mice, 5-aza-2'-deoxycytidine (Aza), a DNA methyltransferase (*DNMT1*) inhibitor, reduced high BP and vascular stiffening via reduced expression of matrix metalloproteinases 9 (*MMP9*), and tissue inhibitors of metalloproteinases (TIMPs)<sup>12</sup>. Using umbilical cord blood DNA from 470 participants, an epigenome-wide association study (EWAS) identified differentially methylated CpGs associated with increased BP and AS at an 8–9-year follow-up examination<sup>13</sup>. Although single CpG positions have been associated with many phenotypes<sup>14–16</sup>, various studies have reported that methylation levels are strongly correlated across the genome, and the reported functionally-relevant findings are generally associated with genomic region rather than single CpG position<sup>17–19</sup>. Hence, understanding DNAm associations with AS/PH at both single CpG position and genomic region levels are important to gain a broader understanding of DNAm changes in AS/PH.

Our study aimed to identify differentially methylated CpG positions and regions that are associated with AS/PH traits. First, we examined differently methylated positions (DMPs) and differentially methylated regions (DMRs) that are associated with eight AS/PH traits, using both epigenome-wide and literature-based candidate gene approaches<sup>7</sup>. Second, we conducted three follow-up analyses including gene set enrichment analysis, expression quantitative trait methylation analysis, and functional enrichment analysis for significantly associated DMPs and DMRs for each AS/PH trait. Third, we prioritized CpGs and their annotated genes for each AS/PH trait using the overlap of significant findings across three follow-up analyses.

## Results

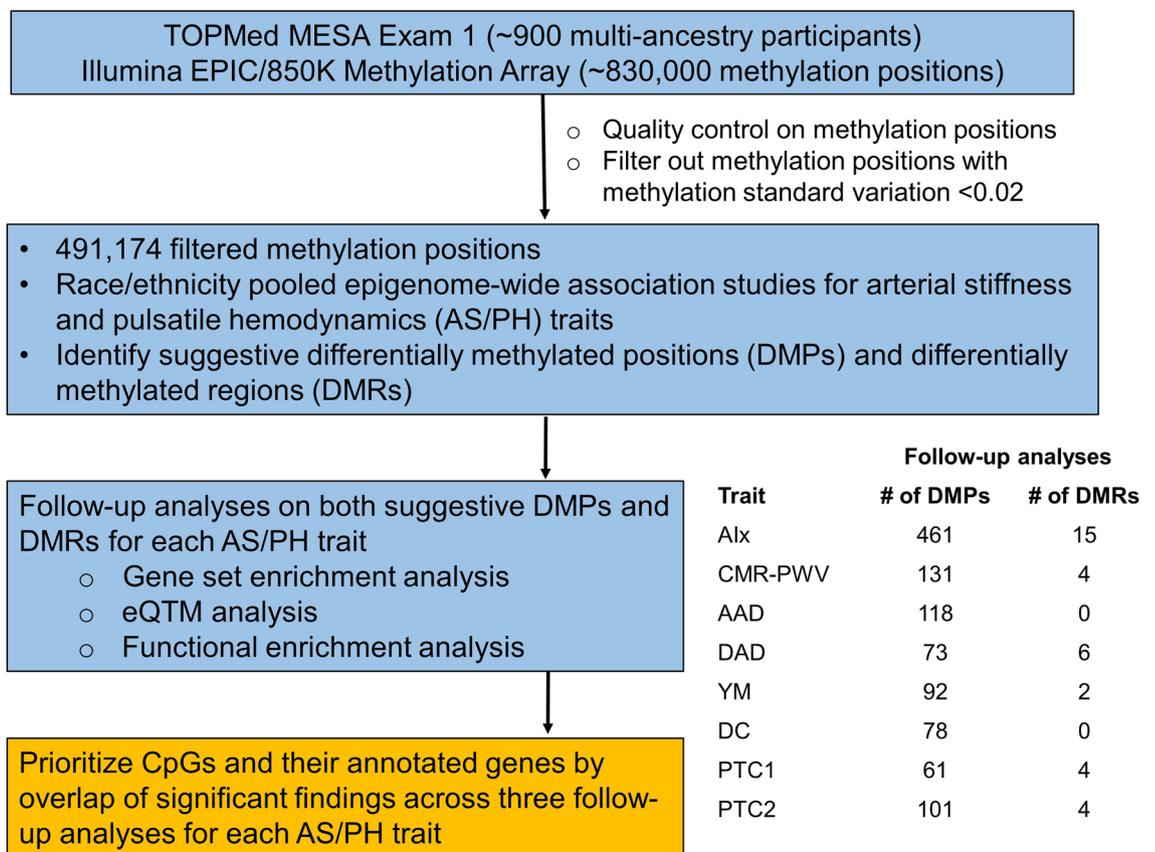
The overview of study design is shown in Fig. 1. After identifying suggestive differentially methylated positions (DMPs) and differentially methylated regions (DMRs) through race/ethnicity pooled epigenome-wide association studies, we performed three follow-up analyses on both suggestive DMPs and DMRs to prioritize CpGs and their annotated genes for each AS/PH trait. The assessment of AIx yielded prioritized genes from the subsequent three follow-up analyses.

### Participant characteristics

The demographic and clinical characteristics of study samples for AIx are summarized in Table 1, which includes 805 TOPMed MESA participants who have both AIx and DNAm measures from MESA Exam 1 (Methods). Based on participant self-reported race/ethnicity, 41% and 33% of the participants were categorized as Non-Hispanic White (NHW) and Hispanic/Latino (HIS) respectively. The remaining participants represented African American (AA) (18%) and Chinese (8%). The median value of AIx was 14.47 (25th percentile = 13.54 and 75th percentile = 15.83). After conducting linear regression to examine the association of log-transformed AIx with each clinical characteristic, we found that age, sex, AA group, mean arterial pressure, anti-hypertensive medication, and HDL cholesterol had nominally significant associations ( $P < 0.05$ ). The sample characteristics, and association of the other seven AS/PH traits with each clinical characteristic are summarized in Supplementary Table S1.

### Two FDR-significant CpGs associated with AIx

After applying standard methylation quality control and filtering out positions with low methylation variation (Methods), there were 491,174 CpGs identified for epigenetic association analyses in MESA. In the multi-ancestry epigenome-wide association studies (EWAS) of AS/PH traits, we first reported the Quantile–Quantile (QQ) plots with corresponding  $\lambda$ , a measure to quantify the inflation in the test statistics (Supplementary Fig. S1). The genomic inflation  $\lambda$  ranged from 0.97 (PTC1) to 1.05 (AIx). We observed two false discovery rate (FDR)-significant (FDR < 0.05) differentially methylated positions (DMPs), cg20711926 (*CYP11B1*) and cg25309493 (*NGEF*), associated with AIx, with cg20711926 ( $p$ -value =  $1.87 \times 10^{-9}$ ) also passing Bonferroni significance ( $0.05/491,174 = 1.02 \times 10^{-7}$ ) (Fig. 2). We further examined the association of these two FDR-significant DMPs with AIx in race/ethnic stratified analyses. The two DMPs were nominally significant in all ancestry groups except cg20711926 in AA group (Fig. 2), with greater statistical significance in HIS than for the other two ancestry



**Figure 1.** Study design. The suggestive DMP was defined as CpG with association  $p$ -value less than  $1 \times 10^{-4}$  and 0.05 respectively for epigenome-wide and candidate-gene approaches; the suggestive DMR was defined as co-methylated region with association  $p$ -value less than  $5 \times 10^{-4}$  and 0.05 respectively for epigenome-wide and candidate-gene approaches; eQTM, expression quantitative trait methylation; AIx, aortic augmentation index; CMR-PWV, aortic arch pulse-wave velocity measured by carotid magnetic resonance imaging; AAD, ascending aortic distensibility; DAD, descending aortic distensibility; YM, Young's Elastic Modulus; DC, distensibility coefficient; PTC1 and PTC2, radial artery pressure waveform index 1 and 2.

groups. Additionally, two FDR-significant DMPs were positively associated with AIx in both race/ethnic pooled and stratified analyses. However, there were no FDR-significant DMPs for the other seven AS/PH traits as shown in Supplementary Table S2. In analysis of CpGs annotated to genes in the candidate-gene list (Supplementary Table S3), there were no CpGs identified at FDR < 5% (Supplementary Table S2).

As there were few FDR-significant DMPs identified, we interrogated those CpGs that achieved suggestive significance based upon EWAS ( $p$ -value <  $1 \times 10^{-4}$ ) and candidate gene ( $p$ -value < 0.05) approaches to select CpGs for follow-up analyses. The suggestive thresholds for statistical significance were selected based on the distribution of EWAS  $p$ -values of all eight AS/PH traits (Supplementary Tables S4-S11). Accordingly, there were 461 unique suggestive DMPs included in the follow-up analyses for AIx (Supplementary Table S4). The suggestive DMPs included in the follow-up analyses for the other seven AS/PH traits are presented in Supplementary Tables S5-S11.

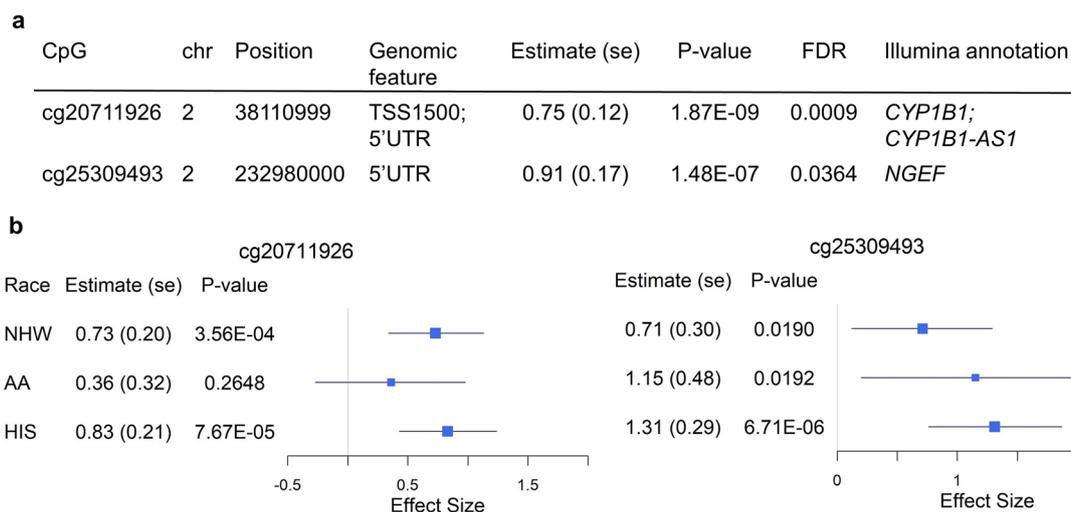
### EWAS at co-methylated genomic regions

We tested 58,222 pre-defined genomic regions and identified 8,165 co-methylated genomic regions for the identification of differentially methylated regions (DMRs) associated with AS/PH traits. There was no FDR-significant DMR identified by either the epigenome-wide or candidate-gene approach for any of the AS/PH traits (Supplementary Table S12). Similar to suggestive DMPs from EWAS, we reported and used co-methylated genomic regions reaching suggestive significance of association  $p$ -values from two approaches ( $p$ -value <  $5 \times 10^{-4}$  from epigenome-wide approach and  $p$ -value < 0.05 from candidate-gene approach) for follow-up analyses.

Overall, we identified suggestive DMRs for six out of eight AS/PH traits (no suggestive DMR for traits AAD and DC, Supplementary Table S12). The association results of suggestive DMRs and their inclusive CpGs for six AS/PH traits are presented in Supplementary Tables S13-S18. AIx had the most suggestive DMRs where 13 and 2 DMRs were from epigenome-wide and candidate-gene approaches, respectively (Table 2). Among these, 13 out of 15 suggestive DMRs were positively associated with AIx. Overall, the number of CpGs included in suggestive DMRs ranged from 3 to 7. Further, we checked the overlap between suggestive DMPs and suggestive DMRs. Supplementary Table S19 shows that methylation differences at individual CpGs and co-methylated

	N = 805		
	Summary statistics	Estimate (se)	P-value
Age, year	60 ± 10	0.0037 (0.0004)	6.89E-16
Male (%)	48	-0.0869 (0.0086)	1.04E-22
Non-Hispanic White (%)	41	Baseline	Baseline
African American (%)	18	0.0264 (0.0128)	0.04
Hispanic/Latino (%)	33	0.0087 (0.0107)	0.41
Chinese (%)	8	-0.0217 (0.0178)	0.22
BMI (kg/m <sup>2</sup> )	28.80 ± 5.23	-0.0007 (0.0009)	0.44
Never smoker (%)	49	Baseline	Baseline
Former smoker (%)	37	-0.0161 (0.0099)	0.10
Current smoker (%)	13	-0.0193(0.0142)	0.18
Smoking pack-years	9.68 ± 17.65	-0.0002 (0.0003)	0.49
Mean arterial pressure, mmHg	88.86 ± 12.12	0.0014 (0.0004)	2.42E-04
Anti-hypertensive medication (%)	34	0.0309 (0.0096)	1.29E-03
Anti-diabetic medication (%)	8	-0.0153 (0.0173)	0.37
Lipid-lowering medication (%)	16	-0.0049 (0.0124)	0.69
Fasting glucose (mg/dL)	95.24 ± 24.16	-0.0002 (0.0002)	0.30
HDL cholesterol (mg/dL)	50.52 ± 13.96	0.0015 (0.0003)	2.33E-06
LDL cholesterol (mg/dL)	118.13 ± 31.57	2.54E-05 (0.0001)	0.86
Triglycerides (mg/dL)	135.24 ± 79.88	-3.28E-05 (5.72E-05)	0.57
AIx (%)	14.47 (13.54, 15.83)		

**Table 1.** Clinical characteristics of samples in TOPMed MESA. Mean ± sd; median (25th percentile, 75th percentile); estimate (se) and *P*-value, estimated effect size (standard error) and *P*-value from linear regression model between log-transformed aortic augmentation index (AIx) and each clinical characteristic.



**Figure 2.** FDR-significant (FDR < 0.05) differentially methylated positions associated with aortic augmentation index. **(a)** Multi-ancestry epigenome-wide association studies (EWAS) summary statistics of two FDR-significant CpGs. **(b)** Forest plots of race-stratified EWAS of two FDR-significant CpGs. Position, based on GRCh38/hg38; Estimate (se) and *P*-value, effect size (standard error) and *P*-value of CpG respectively from EWAS; FDR, false discovery rate based on multiple testing correction on 491,174 CpGs; NHW, Non-Hispanic White; AA, African American; HIS, Hispanic/Latino.

regions partially overlap, demonstrating the value of analyzing both individual CpGs and co-methylated regions provided a more complete view of DNAm association with AS/PH traits.

### Gene set enrichment analysis

We conducted gene set enrichment analysis (Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis) for two sets of CpGs, the significant EWAS set and the significant WGCNA-module set, to identify FDR-significant (FDR < 0.05) GO terms and KEGG pathways for each trait of interest (Methods). By using 518 CpGs that were included in the significant EWAS set associated with AIx, we identified

Analysis approach	Suggestive significance of P-value	DMR (GRCh38/hg38)	No. CpGs	Estimate (se)	P-value	Illumina annotation
Epigenome-wide	$5 \times 10^{-4}$	chr11:2,300,571–2,300,821	7	0.72 (0.18)	9.05E-05	<i>C11orf21; TSPAN32</i>
		chr14:22,819,785–22,819,906	3	0.51 (0.13)	1.50E-04	<i>SLC7A7</i>
		chr22:44,676,732–44,677,170	4	0.49 (0.13)	1.61E-04	<i>PRR5</i>
		chr4:7,646,814–7,646,901	3	0.73 (0.19)	1.86E-04	<i>SORCS2</i>
		chr1:153,775,067–153,775,183	5	0.43 (0.12)	2.07E-04	<i>SLC27A3</i>
		chr13:40,188,014–40,188,333	6	–0.62 (0.17)	3.32E-04	<i>LINC00332</i>
		chr13:25,095,914–25,096,182	7	1.21 (0.34)	3.36E-04	<i>PABPC3</i>
		chr19:48,720,589–48,720,909	3	1.18 (0.33)	3.75E-04	<i>MAMSTR; RASIP1</i>
		chr17:35,241,580–35,241,948	3	–0.57 (0.16)	3.88E-04	<i>SLFN5</i>
		chr12:6,533,036–6,533,189	3	0.55 (0.16)	4.07E-04	<i>GAPDH</i>
		chr17:30,970,363–30,970,587	6	1.61 (0.46)	4.49E-04	<i>RNF135</i>
		chr8:133,576,001–133,576,067	3	1.02 (0.29)	4.82E-04	N/A
Candidate-gene	0.05	chr9:35,905,820–35,906,152	4	0.50 (0.14)	4.84E-04	<i>HRCT1</i>
		chr20:46,008,193–46,008,343	3	0.40 (0.15)	6.42E-03	<i>MMP9</i>
		chr7:74,062,328–74,062,346	3	0.28 (0.13)	0.0298	<i>ELN</i>

**Table 2.** Suggestive differentially methylated regions associated with aortic augmentation index. DMR, differentially methylated region; Estimate (se) and P-value, effect size (standard error) and P-value of co-methylated regions respectively from co-methylated region association test by *coMethDMR*; No. CpGs, number of CpGs included in the DMR.

22 and 36 FDR-significant GO terms and KEGG pathways respectively. The summary and details of FDR-significant enrichment results using CpGs included in the significant EWAS set for AS/PH traits is presented in Supplementary Table S20 and Supplementary Tables S21–S28 correspondingly. From WGCNA analysis, we identified five AIx-associated modules with nominal significance ( $p < 0.05$ ), while no significant modules were identified for PTC1 and PTC2 (Supplementary Table S29). The turquoise module had the most significant enrichment results for AIx, 448 and 48 FDR-significant GO terms and KEGG pathways were identified respectively (Supplementary Table S29). The two FDR-significant CpG sites for AIx, cg20711926 (*CYP1B1*) and cg25309493 (*NGEF*), were also included in the turquoise module. The summary of FDR-significant enrichment results using CpGs included in the significant WGCNA-module set for AS/PH traits is presented in Supplementary Table S29.

We further examined the overlap of significant enrichment results from the two sets of CpGs. There were seven overlapping significant GO terms and eleven overlapping significant KEGG pathways from the turquoise module for AIx, while there was only one overlapping significant GO term from the purple module (Supplementary Table S29). The top overlapping significant GO term for AIx from the turquoise module was a biological process regulating systemic arterial BP (GO:0,001,990), and the only overlapping significant GO term from the purple module is also a biological process regulating inflammatory response (GO:0,050,727) (Supplementary Table S30). Furthermore, 13 candidate genes (*ACE*, *AGER*, *AGTR1*, *ECE1*, *EDNRB*, *ESR1*, *ETS1*, *IL6*, *MMP9*, *NFKB1*, *NOS3*, *REN*, *TNF*) were involved in these two overlapping significant GO terms for AIx (Supplementary Table S30). For the other seven AS/PH traits, the overlap of significant enrichment results was only observed for AAD and DC (Supplementary Tables S29, S31, S32).

### Significant associations between gene expression and suggestive DNAm differences

To identify genes whose expression is associated with suggestive DNAm differences, we conducted expression quantitative trait methylation (eQTM) analysis for suggestive DMPs and DMRs with their annotated genes (Methods). We identified 60 FDR-significant eQTMs at DMPs among 343 expression-DMP association tests for AIx (Supplementary Table S33). Fewer FDR-significant eQTMs at DMPs were observed for the other seven AS/PH traits (range from 2 to 13, Supplementary Table S34). Additionally, there were five FDR-significant eQTMs at DMRs for AIx, and no more than two significant eQTMs at DMRs identified for each of DAD, YM, PTC1, and PTC2 (Supplementary Table S35).

### Significant enrichment in Heart enhancers for AIx-associated DNAm differences

We conducted functional enrichment analysis to better understand the complex interpretation of suggestive DNAm differences (Methods). Specifically, we checked the enrichment results of two AS/PH-relevant tissues (Blood and Heart) for suggestive DMPs and DMRs. The analysis for AIx highlighted enrichment in both Blood and Heart enhancers from 15 chromatin states for CpGs included in AIx-associated DMPs and DMRs (Supplementary Fig. S2). Furthermore, the significant enrichment in Heart enhancers was observed in two AS/PH-relevant cell-types: E104 Right Atrium (Q-value =  $8.6 \times 10^{-3}$ ) and E65 Aorta (Q-value = 0.04) (Supplementary Table S36). For the other seven AS/PH traits, we did not observe any significant enrichment in the 15 chromatin states for either Blood or Heart tissues (Supplementary Fig. S3–S9).

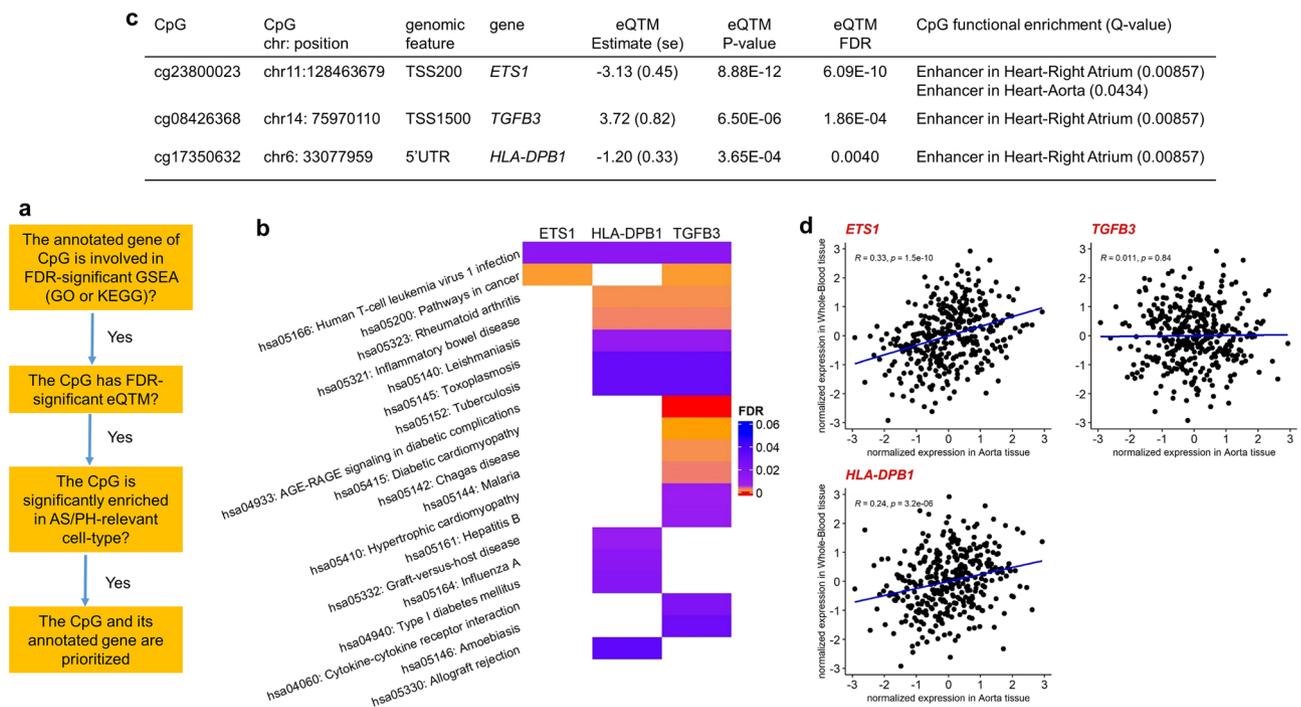
### Three genes prioritized for AIx

After conducting follow-up analyses for CpGs included in suggestive DMPs and DMRs, we checked CpGs included in the overlap of significant findings across three follow-up analyses (gene set enrichment analysis, expression quantitative trait methylation analysis, and functional enrichment analysis). The work flow of prioritization is shown in Fig. 3a. We first summarized results from the three follow-up analyses for the two FDR-significant DMPs associated with AIx in Supplementary Table S37. No significant results were observed in any of the three follow-up analyses for these two DMPs. However, there were three other CpGs (cg23800023-*ETS1*, cg17350632-*HLA-DPB1*, and cg08426368-*TGFB3*) prioritized by follow-up analyses for AIx.

Among the prioritized genes, both *ETS1* and *TGFB3* are on our candidate gene list, although the candidate gene list itself was not used in our overall approach for prioritization of genes. All three genes were involved in KEGG pathway hsa05166: Human T-cell leukemia virus 1 infection pathway and both *ETS1* and *TGFB3* were in KEGG pathway hsa05200: Pathways in cancer (Fig. 3b). Additionally, all three CpGs were FDR-significant eQTM and both CpGs cg23800023 and cg17350632 were negatively associated with their annotated genes, *ETS1* and *HLA-DPB1*, respectively (Fig. 3c). Furthermore, the CpG cg23800023 was enhancer-enriched in both AS/PH-relevant cell-types, Right Atrium (enrichment Q-value =  $8.57 \times 10^{-3}$ ) and Aorta (enrichment Q-value = 0.0434); the other two CpGs were enhancer-enriched in Right Atrium (enrichment Q-value =  $8.57 \times 10^{-3}$ ) (Supplementary Fig. S2 and Fig. 3c). Finally, we checked the correlation of gene expression levels between GTEx Whole-Blood and Aorta tissues respectively for these three genes (Methods). Both *ETS1* and *HLA-DPB1* have significant tissue correlation (*ETS1*, Pearson's correlation  $R = 0.33$  with p-value =  $1.50 \times 10^{-10}$ ; *HLA-DPB1*, Pearson's correlation  $R = 0.24$  with p-value =  $3.20 \times 10^{-6}$ ; Fig. 3d). However, *TGFB3* was not significantly correlated in those two tissues in GTEx. We also checked the overlap of significant findings across the three follow-up analyses for the other seven AS/PH traits, but no overlap was observed.

### Discussion

Arterial stiffness (AS) is a subclinical condition that has a significant prognostic value for future development of CVD events and end organ damage<sup>20</sup>. AS results in excess pulsatile hemodynamics (PH), which can be measured by aortic augmentation index<sup>21</sup> (AIx). While CVD is often identified late requiring interventions to manage disease progress, targeting aortic AS and excessive PH present an early opportunity to identify individuals at greater risk as well as means to monitor effects of preventative interventions. Although the pathophysiology and prognostic significance of AS/PH have been well described, epigenetic influences that impact transcript level expression of AS/PH remain poorly known. The current study contributes to the field by investigating



**Figure 3.** Three prioritized genes (*ETS1*, *HLA-DPB1*, *TGFB3*) for aortic augmentation index. **(a)** Workflow of prioritization; **(b)** FDR-significant KEGG pathways that three prioritized genes were involved; **(c)** Summary table of follow-up analyses results of three prioritized CpGs associated with aortic augmentation index; **(d)** Pearson's correlation ( $R$ ) of inverse-normalized gene expressions between GTEx v8 Whole-Blood and Aorta tissues for three prioritized genes. Position, based on (GRCh38/hg38); FDR, false discovery rate; functional enrichment, eFORGE 15 chromatin state enrichment analysis; GSEA, gene set enrichment analysis; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; eQTM, expression quantitative trait methylation; AS/PH, arterial stiffness and pulsatile hemodynamics.

individual CpG positions and co-methylated regions that are associated with multiple AS/PH traits using both epigenome-wide and candidate-gene approaches. Among a diverse cohort of individuals free from clinical CVD, we identified two significant differentially methylated positions (DMPs), cg20711926 (Bonferroni-significant) and cg25309493 (FDR-significant) that were associated with AIx independent of potential confounders. These two DMPs were positively associated with AIx across all groups in race/ethnic stratified analysis.

AIx is a commonly used measure of late systolic pressure augmentation obtained from pulse wave analysis (PWA)<sup>22</sup>. PWA is a noninvasive method to generate the ascending aorta pressure wave from the pressure waveform measured in the radial artery<sup>21</sup>. The pressure waveform is a composite of the forward pressure wave (incident wave) and a reflected wave. In elastic vessels, because pulse wave velocity (PWV) is low, the reflected wave tends to arrive back at the aortic root during diastole. In stiff vessels, PWV is high and the reflected wave arrives back at the central arteries earlier, augmenting the systolic pressure. This augmented pressure (AP) is calibrated by PWA, and AIx is defined as AP which is expressed as a percentage of pulse pressure<sup>23</sup>. While PWV is the gold standard of AS<sup>24</sup>, AIx can be affected by multiple factors (e.g., left ventricular ejection, PWV, timing of reflection, arterial tone, structure at peripheral reflecting sites, BP, and heart rate)<sup>25</sup>. Nevertheless, the current study demonstrated that AIx was the only AS/PH trait that identified the most significant associated DNAm differences among eight AS/PH traits, including CMR-PWV, AAD, DAD along with all other derives of PWA including YM, DC, PTC 1 & 2.

*CYP1B1* (cytochrome P450 family 1 subfamily B member 1) is the gene annotated to cg20711926, the Bonferroni-significant DMP associated with AIx. *CYP1B1* is a member of the CYP1 subfamily and encodes *CYP1B1* enzyme which is involved in drug metabolism and synthesis of cholesterol, steroids, and other lipids<sup>26</sup>. While many studies reported the role of *CYP1B1* in glaucoma and cancer<sup>27,28</sup>, recent studies have indicated that *CYP1B1* is involved in cardiac pathophysiological changes. For example, *CYP1B1* has been found to mediate angiotensin II-induced aortic smooth muscle cell migration, proliferation, and protein synthesis in rats<sup>29</sup>, as well as to contribute to cardiac hypertrophy induced by uremic toxin in mice<sup>30</sup>. Our finding further suggests that methylation of *CYP1B1* may play a crucial role in determining AIx, although additional research is required to determine whether the association we observed is causal. Despite the need for further investigation, *CYP1B1* may have noteworthy therapeutic implications for cardiac hypertrophy and AS/PH that are the important underlying mechanisms of CVD<sup>31</sup>.

In addition to *CYP1B1*, our study also prioritized three CpGs and their annotated genes (cg23800023-*ETS1*, cg08426368-*TGFB3*, and cg17350632-*HLA-DPB1*) for AIx, based on the overlap of significant results from three follow-up analyses of DMPs, including gene set enrichment analysis, expression quantitative trait methylation (eQTM) analysis, and functional enrichment analysis. Both *ETS1* and *TGFB3* have been previously prioritized as candidate genes<sup>7</sup>, although the candidate gene information was not used in our overall approach for the prioritization. Thus, our approach brings new independent evidence to support the roles of these two genes in processes regulating AS/PH.

First, the cg23800023 annotated gene *ETS1* encodes the founding member of the family of ETS transcription factors. ETS family proteins have a conserved ETS DNA-binding domain that recognizes the GGAA/T in target genes and function as transcriptional activators or repressors of many genes<sup>32</sup>. In endothelial cells, vascular smooth muscle cells, and epithelial cancer cells, ETS is involved in regulating expressions of matrix metalloproteinase (MMP)1, 2, 3, and 9 and vascular endothelial growth factor<sup>33–36</sup> that are the well-known regulators of AS/PH<sup>37</sup>. Our study adds to the existing evidence supporting the link between *ETS1* and AIx via the regulation of DNAm. However, further research is needed to establish the causality of this association.

Second, the cg17350632 annotated gene *HLA-DPB1* (histocompatibility complex, class II, DP beta 1) belongs to a family of human leukocyte antigen (HLA) complex genes that help the immune system to distinguish the body's own proteins from proteins from foreign bodies. Among three subregions (DP, DQ, and DR) of HLA-D, HLA-DP molecule has not been extensively studied<sup>38</sup>, but available studies have demonstrated that HLA-DPB1 is associated with autoimmune disorder such as rheumatoid arthritis<sup>39</sup> and Behcet's disease<sup>40</sup>. The role of *HLA-DPB1* in the vascular system has not been revealed yet. Our study presents novel information on the connection between *HLA-DPB1* and AIx via the regulation of DNAm. Furthermore, both *ETS1* and *HLA-DPB1* have significant tissue correlations between Whole Blood and Aorta in GTEX, which suggests *ETS1* and *HLA-DPB1* could be potential biomarkers in understanding pathophysiology of AS/PH.

Lastly, the cg08426368 annotated gene *TGFB3* (transforming growth factor beta 3) encodes a *TGFB3* protein which is part of a large family of cytokines called TGFB superfamily<sup>41</sup>. *TGFB3* regulates molecules involved in cell proliferation, cell differentiation, and apoptosis, and plays a critical role in the formation of blood vessels and wound healing<sup>41,42</sup>. Previous studies have indicated that TGFB pathways regulate the expressions of elastin, collagen, and *MMP2* & *9*<sup>43</sup>, that are the major determinants of mechanical properties of large arteries<sup>44</sup>. Although *TGFB3* did not show a significant tissue correlation between Whole Blood and Aorta in GTEX, the present study further confirms the potential relationship between *TGFB3* and AIx mediated by DNAm.

It is noteworthy that all these three prioritized genes were involved in KEGG pathway hsa05166, Human T-cell leukemia virus 1 (HTLV-1) infection pathway. HTLV-1 infection can cause adult T-cell leukemia/lymphoma and HTLV-1-associated myelopathy<sup>45,46</sup>, as well as an inflammatory disease such as arthritis<sup>47</sup>. HTLV-1 is known to disturb the regulation of cytokines including interferon gamma (IFN- $\gamma$ ), tumor necrosis factor alpha (TNF- $\alpha$ ), transforming growth factor beta (TGF- $\beta$ ), and IL-10<sup>48</sup>. Given the previous evidence that HTLV-1 triggers chronic inflammatory cascade, which is an important risk factor of atherosclerosis and CVD<sup>49</sup>, potential roles of HTLV infection in AS/PH are considered as an important gap that needs to be filled by future studies.

In summary, our study provides valuable information on epigenetic modifications associated with AS/PH. The major strengths of our study include being one of the very few studies that examined the epigenome-wide association with AS/PH and the use of population-based multi-ancestry data that include surrogate measures of AS and related hemodynamic parameters. In addition, we conducted a series of follow-up analyses to prioritize

genes that are potentially relevant to AS/PH. A few limitations of our study should be noted. First, the study didn't use AS gold standard measure, carotid–femoral PWV<sup>50</sup>, due to data unavailability. Secondly, our study has limited statistical power of identifying significant DNAm differences associated with AS/PH traits due to the sample size. Thirdly, our study lacks a replication study due to the discrepancies in AS/PH measures across public data. Lastly, the study sample included relatively old adults with mean age of 60 ( $\pm$  10) years; thus, the study findings may be confounded with other CVD risks associated with aging. Future studies should aim to replicate the study findings in younger and healthy populations and to investigate the biological processes in which the identified genes contribute to AS/PH. Such studies will help better understand pathophysiology of AS/PH and identify potential therapeutic targets of AS/PH.

## Methods

### Overview of approach

The overview of study design is shown in Fig. 1. We conducted analyses using Illumina's Methylation EPIC Bead-Chip (850 K) and NHLBI Trans-Omics for Precision Medicine (TOPMed) Multi-Ethnic Study of Atherosclerosis (MESA) Exam 1 arterial stiffness and pulsatile hemodynamics (AS/PH) traits with about 900 multi-ancestry participants. We first applied quality control (QC) on DNA methylation (DNAm) data and filtered out methylation positions with low methylation variation (i.e., the standard deviation of methylation Beta-values  $<$  0.02). The epigenetic analyses were then performed on the filtered methylation positions. First of all, we performed epigenome-wide association studies (EWAS) between AS/PH traits and DNAm at individual CpG position and co-methylated genomic region in race/ethnicity pooled analyses respectively. Then the CpGs included in both differentially methylated positions (DMPs) and differentially methylated regions (DMRs) were examined further in three follow-up analyses (i.e., gene set enrichment analysis, expression quantitative trait methylation analysis, and functional enrichment analysis) for each AS/PH trait. Finally, we used the overlap of significant findings across three follow-up analyses to prioritize CpGs and their annotated genes for each AS/PH trait.

### Study participants

The study population consisted of participants free from clinical CVD enrolled in MESA, a longitudinal study of subclinical CVD and risk factors that predict progression to clinically overt cardiovascular disease or progression of the subclinical disease<sup>51</sup>. Between 2000 and 2002, MESA recruited 6,814 men and women 45 to 84 years of age from Forsyth County, North Carolina; New York City; Baltimore; St. Paul, Minnesota; Chicago; and Los Angeles. Exclusion criteria were clinical CVD, weight exceeding 136 kg (300 lb.), pregnancy, and impediment to long-term participation. Approximately 38 percent of the recruited participants are white, 28 percent African American, 22 percent Hispanic/Latino, and 12 percent Asian, predominantly of Chinese descent. Among the specific exclusions related to CVD were the following physician-diagnosed criteria: heart attack, angina or taking nitroglycerin, stroke or TIA, heart failure, current atrial fibrillation, or having undergone procedures related to cardiovascular disease (<https://www.mesa-nhlbi.org/eligibility.aspx>). All research was performed in accordance with relevant guidelines and regulations. All participants provided informed consent and the protocols of MESA were approved by the IRBs of collaborating institutions and the National Heart, Lung and Blood Institute and informed consent was obtained from all participants. Research involving human research participants was performed in accordance with the Declaration of Helsinki. The participants in this study are a subset of MESA individuals who have both AS/PH measures (described in detail previously<sup>52–56</sup>) and DNAm measures<sup>57</sup> from MESA Exam 1 (2000–2002).

### DNA methylation profiling

DNAm profiles were generated using the Infinium Methylation EPIC BeadChip (850 K) (Illumina, Inc., San Diego, CA). As part of the TOPMed MESA Multi-Omics project, 900 participants were selected for MESA Exam 1 DNAm profiling based on the following criteria: (1) restrict to those already included in the TOPMed Whole Genome Sequencing effort<sup>57</sup>, (2) preserve the race/ethnic distribution of participants in the parent MESA cohort, (3) maximize the amount of overlapping 'omics data (with the other 'omics included in the TOPMed MESA Multi-Omics pilot requiring availability of plasma samples for proteomics/metabolomics and RNA from PBMCs, monocytes or T cells for RNA-seq). The QC was applied to DNAm data prior to analysis, including color bias correction, median background adjustment, standard quantile adjustment, batch effect correction (using *ComBat* function in *sva*<sup>58</sup> R package), sex or race mismatch check, and outlier detection (using *Gaphunter* function in *minifi*<sup>59</sup> R package). We further filtered out low methylation variation positions that the standard deviation of their methylation Beta-values is less than 0.02. Hence, 491,174 CpGs were left for epigenetic analyses. To preserve better statistical properties (i.e., homoscedasticity), the M-values (i.e.,  $M = \log(\text{Beta}/(1-\text{Beta}))$ ) were used in the analyses.

### Arterial stiffness and pulsatile hemodynamics traits

The study examined eight AS/PH traits, including: aortic augmentation index (AIx, %) measured by PWA; aortic arch pulse-wave velocity measured by cardiac magnetic resonance imaging (CMR-PWV, m/sec), ascending aortic distensibility (AAD, mmHg), descending aortic distensibility (DAD, mmHg); Young's Elastic Modulus (YM, mmHg) and distensibility coefficient (DC, mmHg), measured by carotid ultrasound; PTC1 (milliseconds) and PTC2 (milliseconds) obtained from radial artery pressure waveforms. The definition and measurement of all eight AS/PH traits in MESA are described in previous studies<sup>21,54,56</sup>. Due to the skewness of raw phenotype data of eight AS/PH traits, the log-transformation was applied to AS/PH traits except for DC where the square-root transformation was applied.

### EWAS at individual CpG

The linear regression adjusted for covariates was performed to test association between each AS/PH trait and M-value of individual CpG at epigenome-wide scale excluding sex chromosomes for 491,174 CpGs. The covariates included age, sex, race/ethnicity, BMI, smoking status (never, former, or current), smoking pack-years, mean arterial pressure, anti-hypertensive medication usage (yes or no), anti-diabetic medication usage (yes or no), lipid-lowering medication usage (yes or no), fasting glucose, high-density lipoprotein (HDL) cholesterol, low-density lipoprotein (LDL) cholesterol, triglycerides, estimated cell type proportions (Lymph, Mono, and Neu), and the first four genomic principal components (PCs) of ancestry. We carried out pooled analyses across self-reported race/ethnic groups for EWAS. The Illumina<sup>60</sup> was used for CpG annotation. We then generated a list of 59 candidate genes from literature-based approach<sup>7</sup>. The EWAS results for CpGs annotated to 59 candidate genes were referred to as candidate-gene approach results. All association tests were adjusted for multiple comparisons using false discovery rate (FDR) correction (Benjamini-Hochberg) at 5%. Due to the limited statistical power of this study, we considered CpGs with EWAS p-values passed suggestive significance as suggestive DMPs. The suggestive significance cut-offs for epigenome-wide approach and candidate-gene approach were  $1 \times 10^{-4}$  and 0.05, respectively. All suggestive DMPs were then used in the follow-up analyses.

### EWAS at co-methylated genomic region

We applied coMethDMR<sup>61</sup> on 491,174 CpGs to identify DMRs. First, coMethDMR identified co-methylated sub-regions with closely located and co-methylated CpGs. We first extracted clusters of CpGs located closely within genomic regions, i.e., the CpG cluster has at least three CpGs and the maximum separation between any two consecutive CpGs within the cluster is 200 base pairs. This step helps to ensure the sub-regions with similar CpG densities. Then we used the correlation between methylation levels among CpGs (i.e., *rdrop* statistics  $> 0.4$  in coMethDMR) in a sub-region to identify co-methylated CpGs. Next, the median of M-values of CpGs within a co-methylated region was used to test association with AS trait in a random coefficient mixed model in coMethDMR that allowed us to model both variations between CpGs within the region and differential methylation simultaneously. The mixed model was also adjusted for the same covariates as for EWAS at individual CpG. We used the *AnnotateResults* function in coMethDMR to annotate co-methylated regions. The association results for co-methylated regions annotated to 59 candidate genes were referred to as candidate-gene approach results. Similar to the identification of suggestive DMP, the co-methylated region with association p-values passed suggestive significance was defined as suggestive DMR. The suggestive significance cut-offs for DMR were  $5 \times 10^{-4}$  and 0.05 for epigenome-wide approach and candidate-gene approach, respectively. All suggestive DMRs were then used in the follow-up analyses as well.

### Weighted gene co-expression network analysis (WGCNA)

WGCNA<sup>62</sup> is a systems biology method that can be used to find modules (clusters) with highly correlated methylation levels and to relate modules to clinical traits. We applied the WGCNA R package<sup>63</sup> on 491,174 CpGs to identify modules significantly associated with AS/PH traits. First, an unsigned co-methylation network was constructed by using *blockwiseModules* function (soft thresholding power = 6, merge cut height = 0.25, and minimum module size = 30). 41 modules were identified from the WGCNA network. DNAm levels of CpGs within a module were summarized by the module eigengene (ME) value which represents the overall methylation level of CpGs clustering in a module. Next, the linear regression model adjusted for covariates was performed between ME value and AS/PH traits for each module to identify significantly associated modules with AS/PH traits. The covariates were the same as EWAS analysis. We considered the module with association p-value  $< 0.05$  as a significant module. Finally, we used CpGs in the significant modules to carry out gene set enrichment analysis.

### Gene set enrichment analysis (GSEA)

We conducted both Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis for each AS/PH trait on two sets of CpGs respectively. The first set of CpGs, the significant EWAS set, was composed of CpGs from the union of both suggestive DMPs and DMRs identified based on suggestive significance of EWAS for each trait of interest. The second set of CpGs, the significant WGCNA-module set, contained CpGs in trait-associated significant modules ( $P < 0.05$ ) from the WGCNA network. We used the *gsameth* function in the missMethyl<sup>64</sup> R package that was developed for genome-wide DNAm data to conduct GSEA. The FDR at 5% was applied to GSEA results to correct for multiple testing comparisons.

### Expression quantitative trait methylation (eQTM) analysis

To identify genes whose expression is associated with significant DNAm differences, we conducted eQTM analysis for both suggestive DMPs and DMRs with their annotated genes (Illumina reference table was used to annotate genes for CpGs and DMR annotation was done by coMethyDMR). We used 587 multi-ancestry samples with both MESA Exam 1 RNA-seq normalized gene expression and DNA methylation for association analysis. First, we removed confounding effects in DNA methylation by fitting the linear regression model  $M \text{ value} \sim \text{age} + \text{gender} + \text{race} + \text{first 4 genomic PCs of ancestry} + \text{estimated cell type proportions (Lymph, Mono, and Neu)}$  and extracting DNA methylation residuals from the model. We used the median M value for the DMR. Similarly, we removed potential confounding effects in RNA-seq by fitting model  $\text{normalized gene expression} \sim \text{age} + \text{gender} + \text{race} + \text{first 4 genomic PCs of ancestry} + \text{PEER factors 1-10}$  and extracting gene expression residuals from the model. Next, for each gene expression and significant DNAm difference pair, we tested association between gene expression residuals (outcome) and DNA methylation residuals via a simple linear regression to quantify eQTM. The FDR at 5% was applied to eQTM results for suggestive DMPs and DMRs respectively to correct for multiple testing comparisons.

## Functional enrichment analysis

To understand the complex interpretation of significant DNAm differences better, we applied eFORGE<sup>65,66</sup> to test whether our AS/PH trait-associated DNAm differences were enriched in regulatory elements from the Roadmap Epigenomics Consortium<sup>67</sup> across 20 tissues and cell types. The CpGs included in both suggestive DMPs and DMRs for each AS/PH trait were used for eFORGE 15 chromatin states enrichment analysis. eFORGE selects a background of 1,000 random CpGs with matching properties based on gene-centric categories (first Exon, 3' untranslated region or UTR, 5'UTR, Body, intergenic region or IGR, TSS1500 and TSS200) and CpG island-centric categories (CpG island, CpG island shore/shelf, N/A or “open sea”). Then the eFORGE uses a 1 kb proximity to filter out highly correlated input CpGs. Finally, eFORGE compares the number of CpGs overlapping regulatory elements from the reference panel with those obtained randomly to calculate enrichment scores for each of the selected cell types. eFORGE performs the Benjamini–Yekutieli approach to account for multiple testing corrections for a cell-type level significance.

## Tissue correlation look-up in GTEx

We used GTEx v8 gene expression data<sup>68</sup> of both Whole-Blood and Aorta tissue (AS/PH-relevant tissue) to check their tissue correlation for the prioritized genes after follow-up analyses. The GTEx v8 gene expression data was downloaded from the GTEx portal. The inverse normalization was first applied to 360 tissue-overlapped GTEx samples. Then both Pearson's correlation and its p-value were reported to measure tissue-correlation level.

## Data availability

The arterial stiffness and pulsatile hemodynamics data for the Multi-Ethnic Study of Atherosclerosis (MESA) are available by application through dbGaP. The dbGaP accession number for MESA is phs000209. The Infinium Methylation EPIC BeadChip (850 K) array data for MESA/TOPMed participants is available through dbGaP via accession phs001416.v3.p1.

Received: 29 June 2023; Accepted: 12 October 2023

Published online: 17 October 2023

## References

1. Laurent, S. *et al.* Expert consensus document on arterial stiffness: Methodological issues and clinical applications. *Eur. Heart J.* **27**, 2588–2605 (2006).
2. Redheuil, A. *et al.* Proximal aortic distensibility is an independent predictor of all-cause mortality and incident CV events: The MESA study. *J. Am. Coll. Cardiol.* **64**, 2619–2629 (2014).
3. Ziemann, S. J., Melenovsky, V. & Kass, D. A. Mechanisms, pathophysiology, and therapy of arterial stiffness. *Arterioscler. Thromb. Vasc. Biol.* **25**, 932–943 (2005).
4. Dernellis, J. & Panaretou, M. Aortic stiffness is an independent predictor of progression to hypertension in nonhypertensive subjects. *Hypertension* **45**, 426–431 (2005).
5. Chung, C.-M. *et al.* Arterial stiffness is the independent factor of left ventricular hypertrophy determined by electrocardiogram. *Am. J. Med. Sci.* **344**, 190–193 (2012).
6. Lacolley, P., Challande, P., Osborne-Pellegrin, M. & Regnault, V. Genetics and pathophysiology of arterial stiffness. *Cardiovasc. Res.* **81**, 637–648 (2009).
7. Logan, J. G., Engler, M. B. & Kim, H. Genetic determinants of arterial stiffness. *J. Cardiovasc. Tran. Res.* **8**, 23–43 (2015).
8. Yasmin, & O'Shaughnessy, K. M. Genetics of arterial structure and function: Towards new biomarkers for aortic stiffness?. *Clin. Sci. (London)* **114**, 661–677 (2008).
9. The role of methylation in gene expression | Learn Science at Scitable. <https://www.nature.com/scitable/topicpage/the-role-of-methylation-in-gene-expression-1070/>.
10. Gibney, E. R. & Nolan, C. M. Epigenetics and gene expression. *Heredity* **105**, 4–13 (2010).
11. Chen, K. & Sun, Z. Activation of DNA demethylases attenuates aging-associated arterial stiffening and hypertension. *Aging. Cell* **17**, e12762 (2018).
12. Narayanan, N. *et al.* Epigenetic regulation of aortic remodeling in hyperhomocysteinemia. *FASEB J.* **28**, 3411–3422 (2014).
13. Murray, R. *et al.* Influence of maternal lifestyle and diet on perinatal dna methylation signatures associated with childhood arterial stiffness at 8 to 9 years. *Hypertension* **78**, 787–800 (2021).
14. Richard, M. A. *et al.* DNA methylation analysis identifies loci for blood pressure regulation. *Am. J. Hum. Genet.* **101**, 888–902 (2017).
15. Barfield, R. *et al.* Epigenome-wide association analysis of daytime sleepiness in the multi-ethnic study of atherosclerosis reveals African-American-specific associations. *Sleep* **42**, zsz1101 (2019).
16. Lee, M. *et al.* Pulmonary function and blood DNA methylation: A multiethnic epigenome-wide association meta-analysis. *Am. J. Respir. Crit. Care Med.* **206**, 321–336 (2022).
17. Irizarry, R. A. *et al.* Comprehensive high-throughput arrays for relative methylation (CHARM). *Genom. Res.* **18**, 780–790 (2008).
18. Ventham, N. T. *et al.* Integrative epigenome-wide analysis demonstrates that DNA methylation may mediate genetic risk in inflammatory bowel disease. *Nat. Commun.* **7**, 13507 (2016).
19. Zhang, L. *et al.* Epigenome-wide meta-analysis of DNA methylation differences in prefrontal cortex implicates the immune processes in Alzheimer's disease. *Nat. Commun.* **11**, 6114 (2020).
20. Boutouyrie, P., Chowienczyk, P., Humphrey, J. D. & Mitchell, G. F. Arterial stiffness and cardiovascular risk in hypertension. *Circ. Res.* **128**, 864–886 (2021).
21. Chen, C. H. *et al.* Estimation of central aortic pressure waveform by mathematical transformation of radial tonometry pressure. Validation of generalized transfer function. *Circulation* **95**, 1827–1836 (1997).
22. Oliver, J. J. & Webb, D. J. Noninvasive assessment of arterial stiffness and risk of atherosclerotic events. *Arterioscler. Thromb. Vasc. Biol.* **23**, 554–566 (2003).
23. O'Rourke, M. F. Steady and pulsatile energy losses in the systemic circulation under normal conditions and in simulated arterial disease. *Cardiovasc. Res.* **1**, 313–326 (1967).
24. O'Rourke, M. F., Nichols, W. W. & Safar, M. E. Pulse waveform analysis and arterial stiffness: realism can replace evangelism and scepticism. *J. Hypertens.* **22**, 1633–1634 (2004).
25. Hamilton, P. K., Lockhart, C. J., Quinn, C. E. & McVeigh, G. E. Arterial stiffness: Clinical relevance, measurement and treatment. *Clin. Sci. (London)* **113**, 157–170 (2007).

26. Li, F., Zhu, W. & Gonzalez, F. J. Potential role of CYP1B1 in the development and treatment of metabolic diseases. *Pharmacol. Ther.* **178**, 18–30 (2017).
27. Bejjani, B. A. *et al.* Mutations in CYP1B1, the gene for cytochrome P4501B1, are the predominant cause of primary congenital glaucoma in Saudi Arabia. *Am. J. Hum. Genet.* **62**, 325–333 (1998).
28. Carrera, A. N., Grant, M. K. O. & Zordoky, B. N. CYP1B1 as a therapeutic target in cardio-oncology. *Clin. Sci. (London)* **134**, 2897–2927 (2020).
29. Yaghini, F. A. *et al.* Angiotensin II-induced vascular smooth muscle cell migration and growth are mediated by cytochrome P450 1B1-dependent superoxide generation. *Hypertension* **55**, 1461–1467 (2010).
30. Zhang, Y. *et al.* Inhibition of CYP1B1 ameliorates cardiac hypertrophy induced by uremic toxin. *Mol. Med. Rep.* **21**, 393–404 (2020).
31. Kass, D. A. Ventricular arterial stiffening: Integrating the pathophysiology. *Hypertension* **46**, 185–193 (2005).
32. Garrett-Sinha, L. A. Review of Ets1 structure, function, and roles in immunity. *Cell Mol. Life Sci.* **70**, 3375–3390 (2013).
33. Vandenbunder, B., Queva, C., Desbiens, X., Wernert, N. & Stehelin, D. Expression of the transcription factor c-Ets1 correlates with the occurrence of invasive processes during normal and pathological development. *Invasion Metastasis* **14**, 198–209 (1994).
34. Vandenbunder, B., Wernert, N., Queva, C., Desbiens, X. & Stehelin, D. Does the transcription factor c-ets1 take part in the regulation of angiogenesis and tumor invasion?. *Folia Biol. (Praha)* **40**, 301–313 (1994).
35. Dittmer, J. The biology of the Ets1 proto-oncogene. *Mol. Cancer* **2**, 29 (2003).
36. Puzovic, V., Brcic, I., Ranogajec, I. & Jakic-Razumovic, J. Prognostic values of ETS-1, MMP-2 and MMP-9 expression and co-expression in breast cancer patients. *Neoplasma* **61**, 439–446 (2014).
37. Yasmin, *et al.* Matrix metalloproteinase-9 (MMP-9), MMP-2, and serum elastase activity are associated with systolic hypertension and arterial stiffness. *Arterioscler. Thromb. Vasc. Biol.* **25**, 372 (2005).
38. McCanlies, E. C., Kreiss, K., Andrew, M. & Weston, A. HLA-DPB1 and chronic beryllium disease: A HuGE review. *Am. J. Epidemiol.* **157**, 388–398 (2003).
39. Jiang, L., Jiang, D., Han, Y., Shi, X. & Ren, C. Association of HLA-DPB1 polymorphisms with rheumatoid arthritis: A systemic review and meta-analysis. *Int. J. Surg.* **52**, 98–104 (2018).
40. Capittini, C. *et al.* Global meta-analysis on the association between Behcet syndrome and polymorphisms from the HLA class I (A, B, and C) and class II (DRB1, DQB1, and DPB1) genes. *Dis. Markers* **2021**, 9348697 (2021).
41. Haque, S. & Morris, J. C. Transforming growth factor- $\beta$ : A therapeutic target for cancer. *Hum. Vaccin. Immunother* **13**, 1741–1750 (2017).
42. Goumans, M.-J. & Ten Dijke, P. TGF- $\beta$  signaling in control of cardiovascular function. *Cold Spring Harb. Perspect. Biol.* **10**, a022210 (2018).
43. Scola, L. *et al.* Role of TGF- $\beta$  pathway polymorphisms in sporadic thoracic aortic aneurysm: rs900 TGF- $\beta$ 2 is a marker of differential gender susceptibility. *Mediators Inflamm* **2014**, 165758 (2014).
44. Wagenseil, J. E. & Mecham, R. P. Elastin in large artery stiffness and hypertension. *J. Cardiovasc. Tran. Res.* **5**, 264–273 (2012).
45. Eusebio-Ponce, E., Anguita, E., Paulino-Ramirez, R. & Candel, F. J. HTLV-1 infection: An emerging risk. Pathogenesis, epidemiology, diagnosis and associated diseases. *Rev. Esp. Quimioter* **32**, 485–496 (2019).
46. Miura, M., Naito, T. & Saito, M. Current perspectives in human T-cell Leukemia virus type 1 infection and its associated diseases. *Front. Med.* **9**, 867478 (2022).
47. Nishioka, K. *et al.* Chronic inflammatory arthropathy associated with HTLV-I. *Lancet* **1**, 441 (1989).
48. Quaresma, J. A. S. *et al.* HTLV-1, immune response and autoimmunity. *Viruses* **8**, 5 (2015).
49. Abolbashari, S., Ghayour-Mobarhan, M., Ebrahimi, M. & Meshkat, Z. The role of human T-lymphotropic virus (HTLV) in cardiovascular diseases: A review of literature. *ARYA Atheroscler* **14**, 183–187 (2018).
50. Laurent, S. *et al.* Carotid artery distensibility and distending pressure in hypertensive humans. *Hypertension* **23**, 878–883 (1994).
51. Bild, D. E. *et al.* Multi-ethnic study of atherosclerosis: Objectives and design. *Am. J. Epidemiol.* **156**, 871–881 (2002).
52. Chirinos, J. A. *et al.* Arterial wave reflections and incident cardiovascular events and heart failure: MESA (Multiethnic Study of Atherosclerosis). *J. Am. College Cardiol.* **21**, 2170–2177 (2012).
53. Ohyama, Y. *et al.* Ten-year longitudinal change in aortic stiffness assessed by cardiac MRI in the second half of the human lifespan: the multi-ethnic study of atherosclerosis. *Eur. Heart J. Cardiovasc. Imaging* **17**, 1044–1053 (2016).
54. Duprez, D. A. *et al.* Race/ethnic and sex differences in large and small artery elasticity—results of the multi-ethnic study of atherosclerosis (MESA). *Ethn. Dis.* **19**, 243–250 (2009).
55. Nechyporenko, A. *et al.* Association of statin therapy with progression of carotid arterial stiffness: The Multi-Ethnic Study of Atherosclerosis (MESA). *Hypertens. Res.* <https://doi.org/10.1038/s41440-022-01095-9> (2022).
56. Brumback, L. C., Jacobs, D. R. Jr. & Duprez, D. A. PTC1 and PTC2: New indices of blood pressure waveforms and cardiovascular disease. *Am. J. Epidemiol.* **189**, 726–734 (2020).
57. Taliun, D. *et al.* Sequencing of 53, 831 diverse genomes from the NHLBI TOPMed program. *Nature* **590**, 290–299 (2021).
58. Johnson, W. E., Li, C. & Rabinovic, A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **8**, 118–127 (2007).
59. Andrews, S. V., Ladd-Acosta, C., Feinberg, A. P., Hansen, K. D. & Fallin, M. D. “Gap hunting” to characterize clustered probe signals in Illumina methylation array data. *Epigenetics Chromatin* **9**, 56 (2016).
60. Infinium MethylationEPIC BeadChip Product Files. [https://support.illumina.com/array/array\\_kits/infinium-methylationepic-beadchip-kit/downloads.html](https://support.illumina.com/array/array_kits/infinium-methylationepic-beadchip-kit/downloads.html).
61. Gomez, L. *et al.* coMethDMR: accurate identification of co-methylated and differentially methylated regions in epigenome-wide association studies with continuous phenotypes. *Nucleic Acids Res.* **47**, e98 (2019).
62. Zhang, B. & Horvath, S. A general framework for weighted gene co-expression network analysis. *Stat. Appl. Genet. Mol. Biol.* **4**, 17 (2005).
63. Langfelder, P. & Horvath, S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinform.* **9**, 559 (2008).
64. Maksimovic, J., Oshlack, A. & Phipson, B. Gene set enrichment analysis for genome-wide DNA methylation data. *Genome Biol.* **22**, 173 (2021).
65. Breeze, C. E. *et al.* eFORGE: a tool for identifying cell type-specific signal in epigenomic data. *Cell Rep.* **17**, 2137–2150 (2016).
66. Breeze, C. E. *et al.* eFORGE v2.0: Updated analysis of cell type-specific signal in epigenomic data. *Bioinformatics* **35**, 4767–4769 (2019).
67. Kundaje, A. *et al.* Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317–330 (2015).
68. GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **369**, 1318–1330 (2020).

## Acknowledgements

This work is supported by a University of Virginia Pinn Scholar Award to Dr. Manichaikul and NHLBI R01-HL153248 (AM and XH). Dr. Chirinos is supported by NIH grants U01-TR003734, U01-TR003734-01S1, U01-HL160277, R33-HL-146390, R01-HL153646, K24-AG070459, R01-AG058969, R01-HL157108, R01-HL155599, R01-HL104106 and R01HL155764. The data used for the analyses described in this manuscript were obtained from the GTEx Portal (<https://gtexportal.org/home/>). The authors acknowledge Research

Computing at The University of Virginia for providing computational resources and technical support that have contributed to the results reported within this manuscript. URL: <https://rc.virginia.edu>. Whole genome sequencing (WGS) for the Trans-Omics in Precision Medicine (TOPMed) program was supported by the National Heart, Lung and Blood Institute (NHLBI). WGS for “NHLBI TOPMed: Multi-Ethnic Study of Atherosclerosis (MESA)” (phs001416.v1.p1) was performed at the Broad Institute of MIT and Harvard (3U54HG003067-13S1). Centralized read mapping and genotype calling, along with variant quality metrics and filtering were provided by the TOPMed Informatics Research Center (3R01HL-117626-02S1). Phenotype harmonization, data management, sample-identity QC, and general study coordination, were provided by the TOPMed Data Coordinating Center (3R01HL-120393-02S1), and TOPMed MESA Multi-Omics (HHSN2682015000031/HSN26800004). The MESA projects are conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with MESA investigators. Support for the Multi-Ethnic Study of Atherosclerosis (MESA) projects are conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with MESA investigators. Support for MESA is provided by contracts 75N92020D00001, HHSN268201500003I, N01-HC-95159, 75N92020D00005, N01-HC-95160, 75N92020D00002, N01-HC-95161, 75N92020D00003, N01-HC-95162, 75N92020D00006, N01-HC-95163, 75N92020D00004, N01-HC-95164, 75N92020D00007, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168, N01-HC-95169, UL1-TR-000040, UL1-TR-001079, UL1-TR-001420, UL1TR001881, DK063491, and R01HL105756. The authors thank the other investigators, the staff, and the participants of the MESA study for their valuable contributions. A full list of participating MESA investigators and institutes can be found at <http://www.mesa-nhlbi.org>. The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS.

### Author contributions

X.H., J.G.L., F.E.G., and A.M. contributed to the conception of ideas of the study, to the analysis of the data, and to the writing of the manuscript. J.G.L., Y.K., G.F.M., J.A.C. and T.M.H. contributed to selection of phenotypes for analyses. J.A.C.L., D.R.J., D.D., L.B., K.D.T., P.D., W.C.J., E.C., X.G., Y.L., R.P.T., T.W.B., G.P., S.S.R., J.I.R., and D.J.V. contributed to data collection, data management and distribution. All authors contributed to reviewing and editing the manuscript.

### Competing interests

Dr. Chirinos has recently consulted for Bayer, Fukuda-Denshi, Bristol-Myers Squibb, JNJ, Edwards Life Sciences, Merck, and NGM Biopharmaceuticals. He received University of Pennsylvania research grants from National Institutes of Health, Fukuda-Denshi, Bristol-Myers Squibb, Microsoft and Abbott. He is named as inventor in a University of Pennsylvania patent for the use of inorganic nitrates/nitrites for the treatment of Heart Failure and Preserved Ejection Fraction and a patent application for the use of plasma and urine biomarkers in heart failure with preserved ejection fraction. He has received payments for editorial roles from the American Heart Association, the American College of Cardiology, Elsevier and Wiley. He has received research device loans from Atcor Medical, Fukuda-Denshi, Unex, Uscom, NDD Medical Technologies, Microsoft and MicroVision Medical. Other authors declare no conflicts of interests.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-44806-z>.

**Correspondence** and requests for materials should be addressed to F.E.G.-B. or A.M.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023