Supplemental Data

DNA Methylation Analysis Identifies Loci

for Blood Pressure Regulation

Melissa A. Richard, Tianxiao Huan, Symen Ligthart, Rahul Gondalia, Min A. Jhun, Jennifer A. Brody, Marguerite R. Irvin, Riccardo Marioni, Jincheng Shen, Pei-Chien Tsai, May E. Montasser, Yucheng Jia, Catriona Syme, Elias L. Salfati, Eric Boerwinkle, Weihua Guan, Thomas H. Mosley, Jr., Jan Bressler, Alanna C. Morrison, Chunyu Liu, Michael M. Mendelson, André G. Uitterlinden, Joyce B. van Meurs, BIOS Consortium, Oscar H. Franco, Guosheng Zhang, Yun Li, James D. Stewart, Joshua C. Bis, Bruce M. Psaty, Yii-Der Ida Chen, Sharon L.R. Kardia, Wei Zhao, Stephen T. Turner, Devin Absher, Stella Aslibekyan, John M. Starr, Allan F. McRae, Lifang Hou, Allan C. Just, Joel D. Schwartz, Pantel S. Vokonas, Cristina Menni, Tim D. Spector, Alan Shuldiner, Coleen M. Damcott, Jerome I. Rotter, Walter Palmas, Yongmei Liu, Tomáš Paus, Steve Horvath, Jeffrey R. O'Connell, Xiuqing Guo, Zdenka Pausova, Themistocles L. Assimes, Nona Sotoodehnia, Jennifer A. Smith, Donna K. Arnett, Ian J. Deary, Andrea A. Baccarelli, Jordana T. Bell, Eric Whitsel, Abbas Dehghan, Daniel Levy, and Myriam Fornage
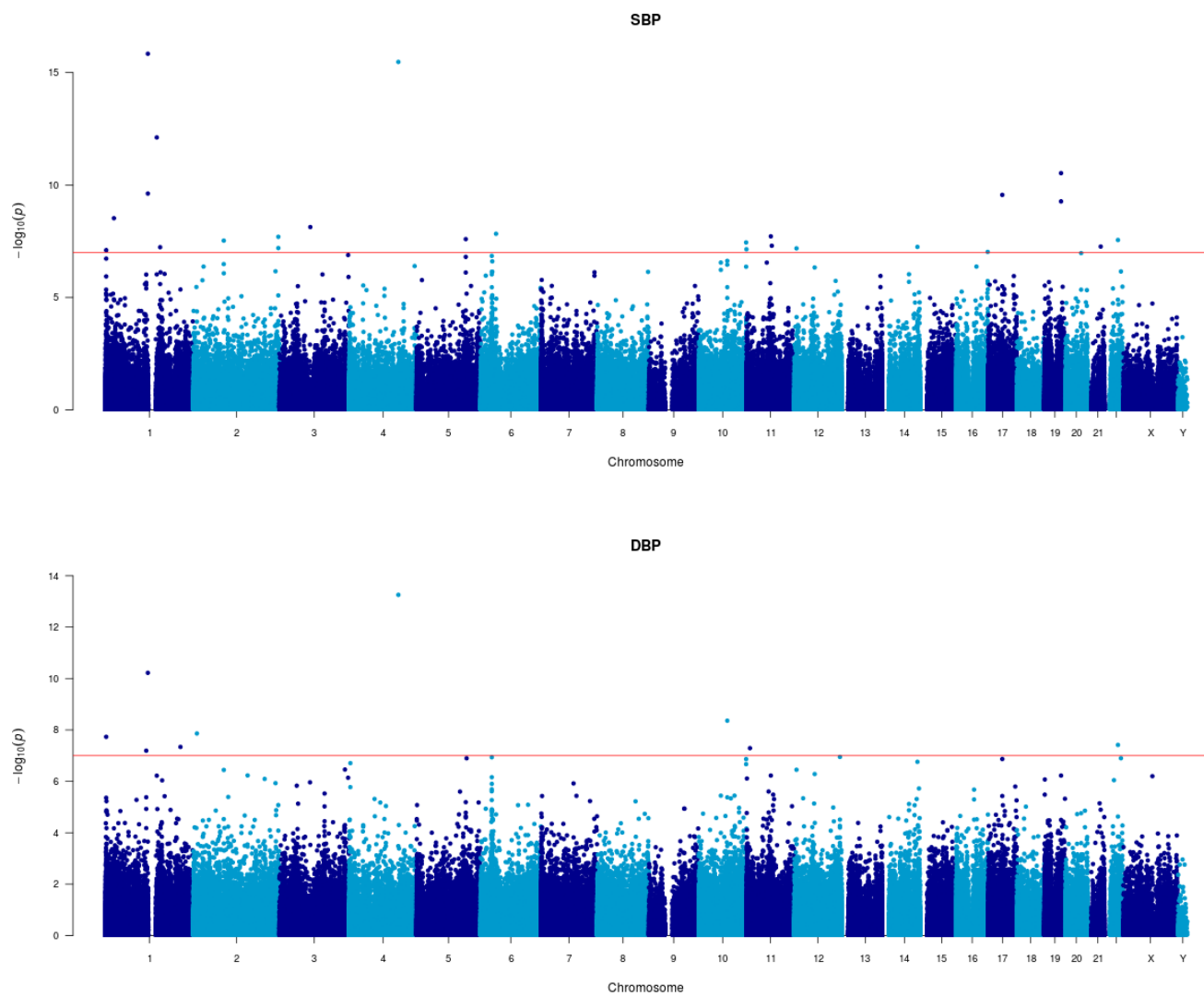
Figure S1 Manhattan plots showing systolic and diastolic BP associations at each CpG site in the discovery meta-analyses. The CpG chromosomal locations are plotted against the –log10(p value). The red line indicates the Bonferroni threshold (1×10^{-7}) for epigenome-wide significance.

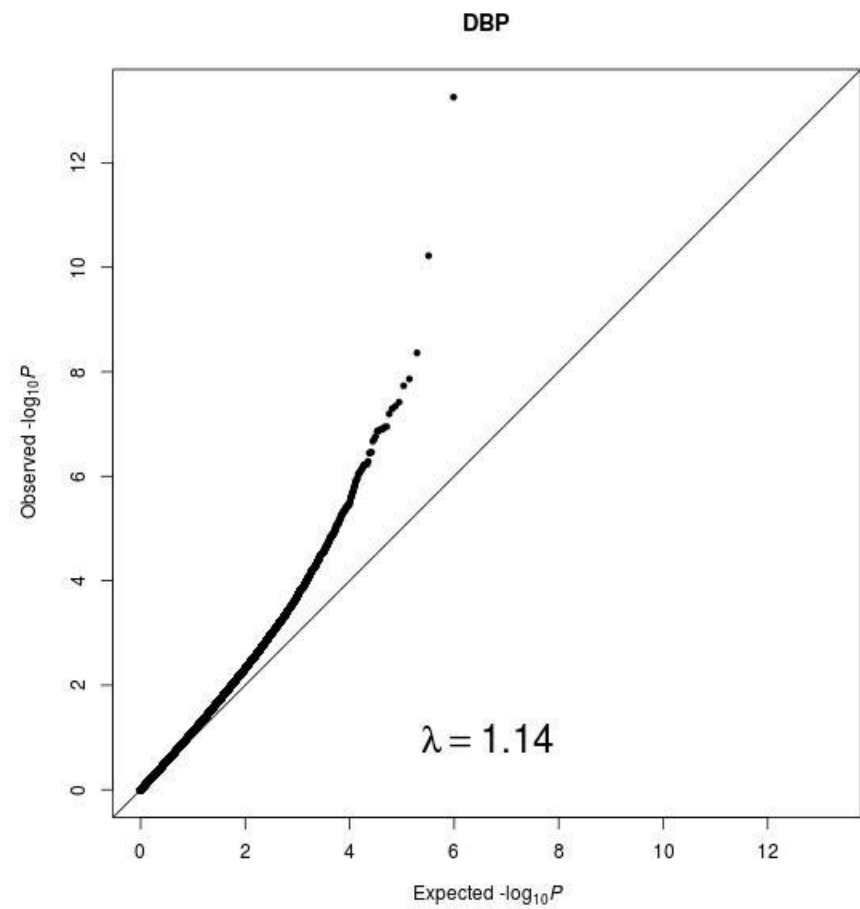Abbreviations: DBP, diastolic blood pressure; SBP, systolic blood pressure.

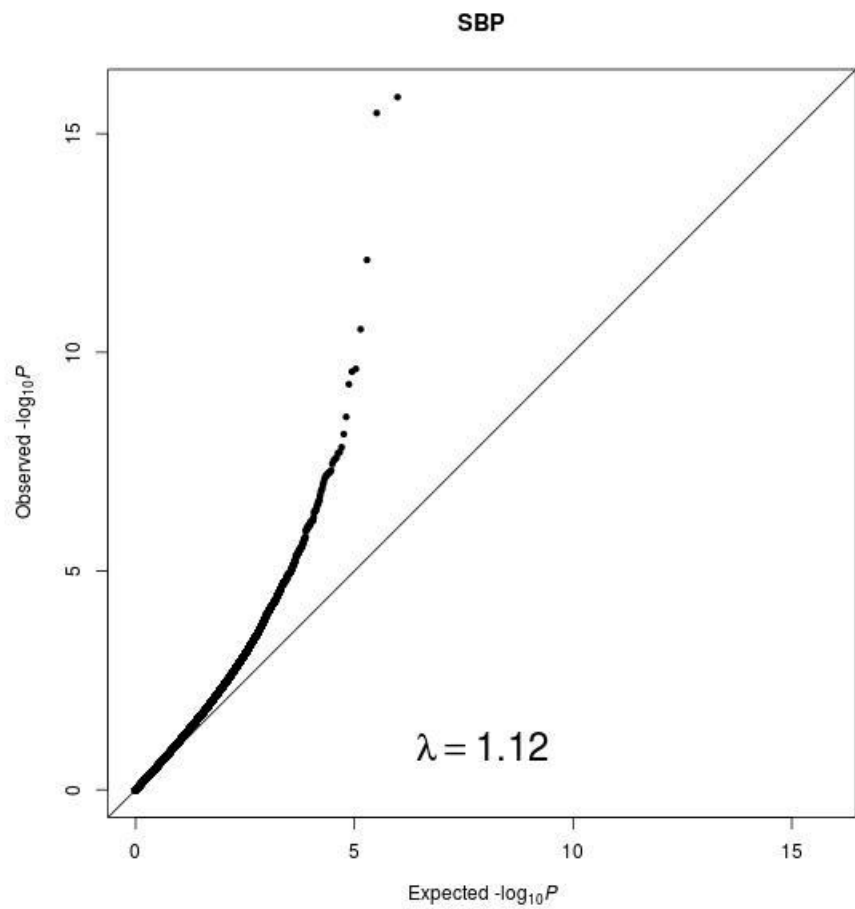**Figure S2 QQ plots for observed vs expected –log10(p value) at each CpG site in the discovery meta-analyses**.

Abbreviations: DBP, diastolic blood pressure; SBP, systolic blood pressure.
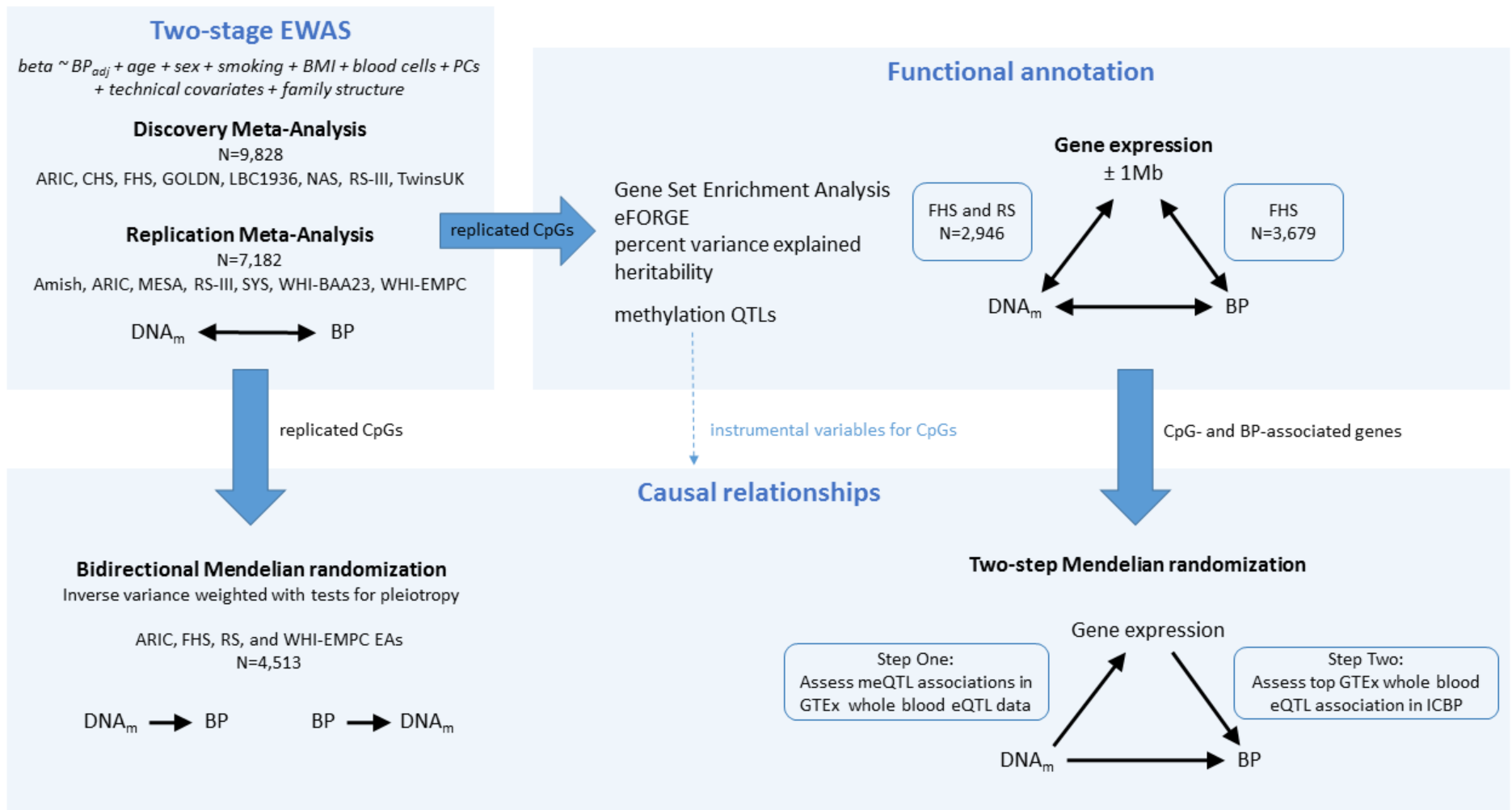
Figure S3 Schematic of the overall study design.

Abbreviations: BMI, body mass index; BP, blood pressure; CpG, cytosine-phosphate-guanine; DNA_m, DNA methylation; EA, European ancestry; eFORGE, **F**unctional element **O**verlap analysis of the **R**esults of **G**enome Wide Association Study **E**xperiments for EWAS studies; eQTL, expression quantitative trait locus; EWAS, epigenome-wide association study; GTEx, Genotype-Tissue Expression project; ICBP, International Consortium for Blood Pressure GWAS; Mb, megabase; PC, principal components; QTL, quantitative trait locus.
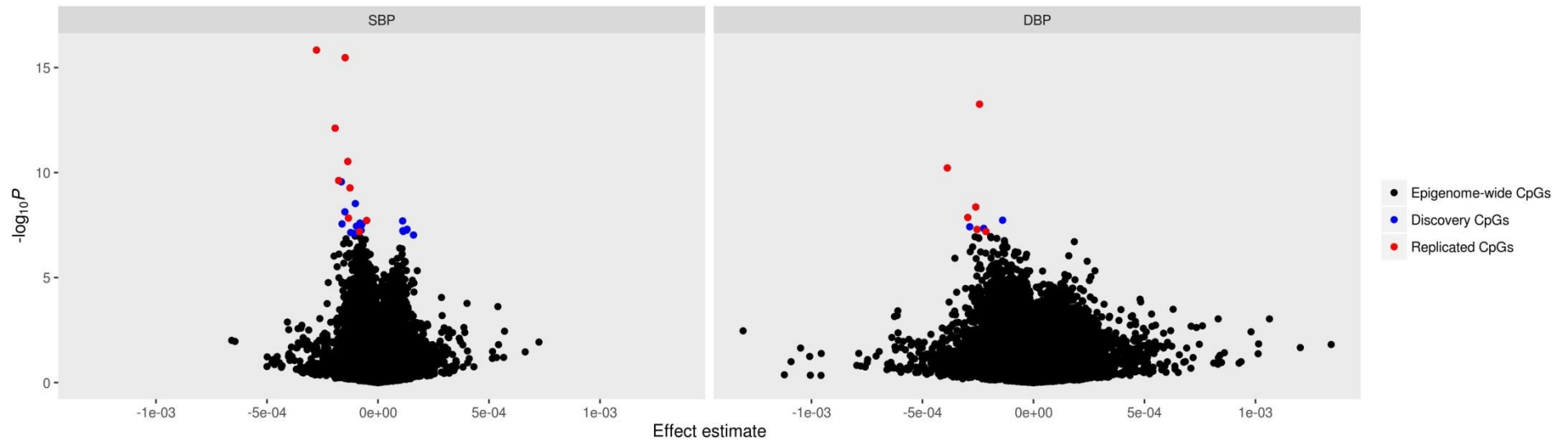
Figure S4 Volcano plots of the effect sizes plotted against –log10(p value) for epigenome-wide discovery meta-analysis associations with systolic and diastolic BP. Effect size units are percent change in DNA methylation per 1-unit change in blood pressure. CpG sites identified at the discovery stage are colored blue and replicated CpG sites are colored red; all other CpG sites are shown in black.

Abbreviations: CpG, cytosine-phosphate-guanine; DBP, diastolic blood pressure; SBP, systolic blood pressure.
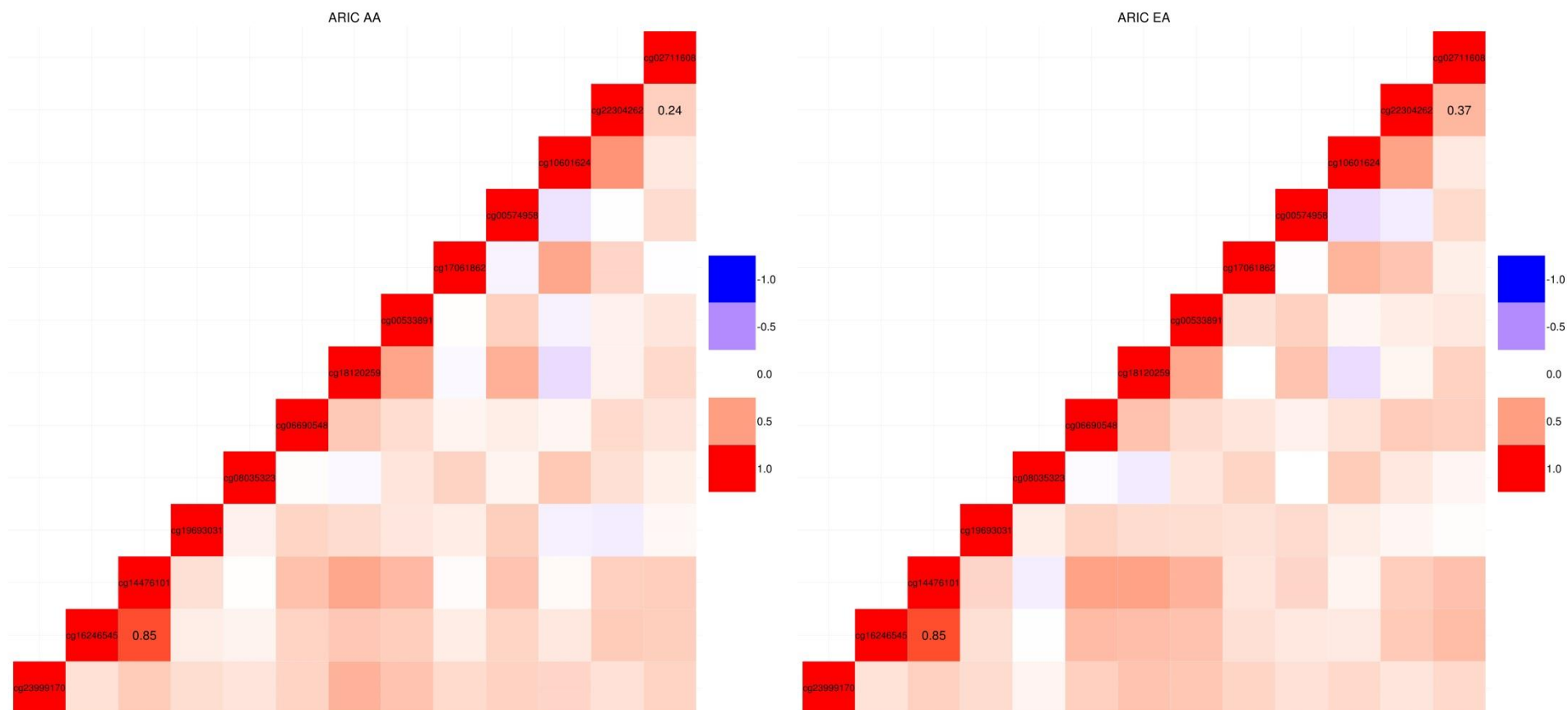
Figure S5 Spearman correlation of methylation levels at replicated CpG sites for BP among individuals of African American and European ancestry in the ARIC cohort. For CpG sites in the same gene (*PHGDH* cg16246545 and cg14476101; *SLC1A5* cg22304262 and cg02711608) the estimated correlation coefficients are shown. The color scale corresponds to strength of correlation, where inverse correlations are blue and direct correlations are red.

Abbreviations: AA, African American; ARIC, atherosclerosis Risk in Communities; EA, European Ancestry.
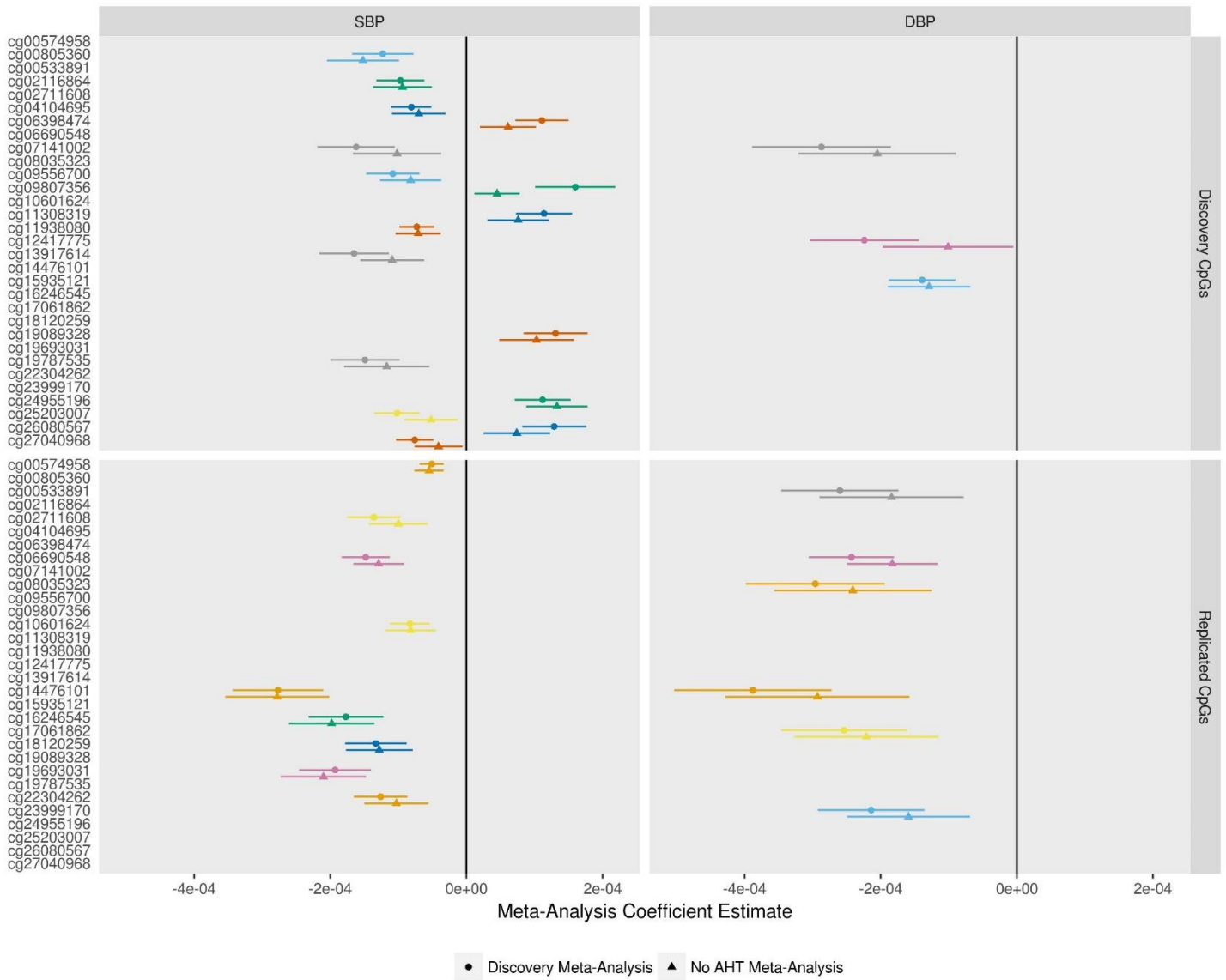
Figure S6 Forest plots showing coefficient point estimates and 95% confidence intervals for the 31 discovery CpG sites in the discovery meta-analysis (N=9,828; circle) and the meta-analysis among individuals not reporting antihypertensive treatment (N=9,894; triangle). Discovery CpG sites with attenuated coefficient estimates among non-medicated individuals tended to not replicate.

Abbreviations: AHT, anti-hypertensive treatment; CpG, cytosine-phosphate-guanine; DBP, diastolic blood pressure; SBP, systolic blood pressure.
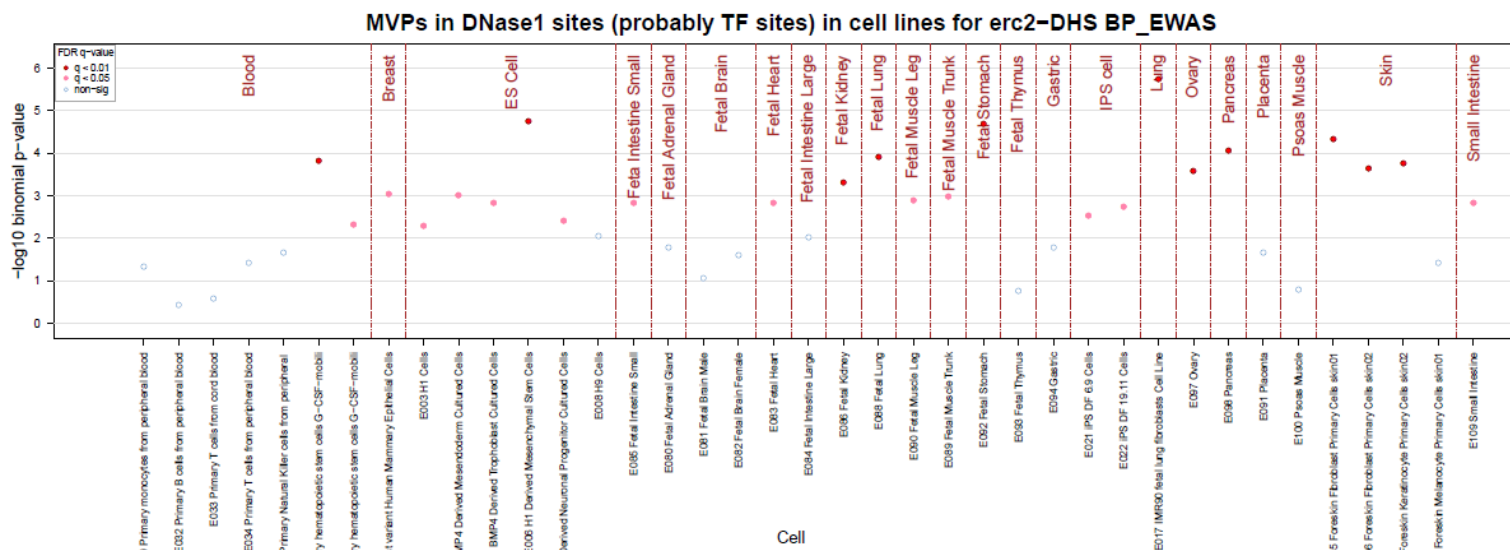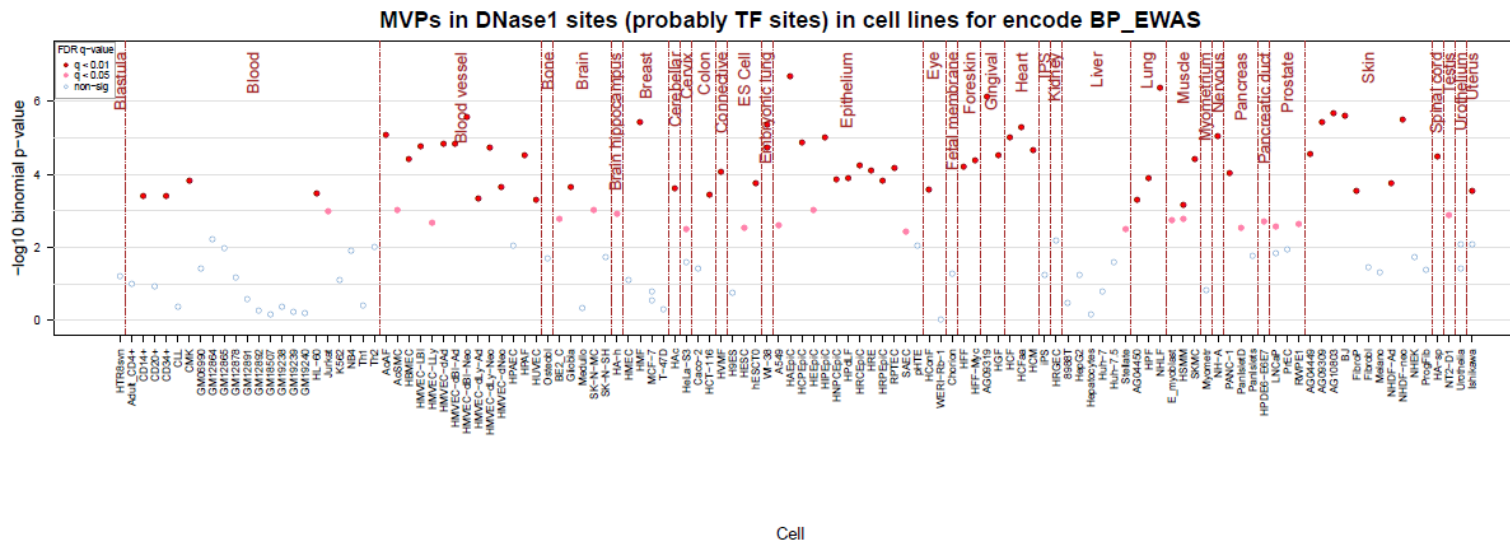
Figure S7 Enrichment of replicated CpG sites for DNase I hypersensitive sites across various tissues in the ENCODE (upper panel) and Roadmap Epigenomics (lower panel) projects. Cell types showing enrichment at FDR-adjusted p value <0.01 are shown in red and FDR-adjusted p value <0.05 are shown in pink.

Abbreviations: BP, blood pressure; EWAS, epigenome-wide association study; FDR, false discovery rate; DHS, DNase I hypersensitive site; MVPs, methylation-varying probe; TF, transcription factor.
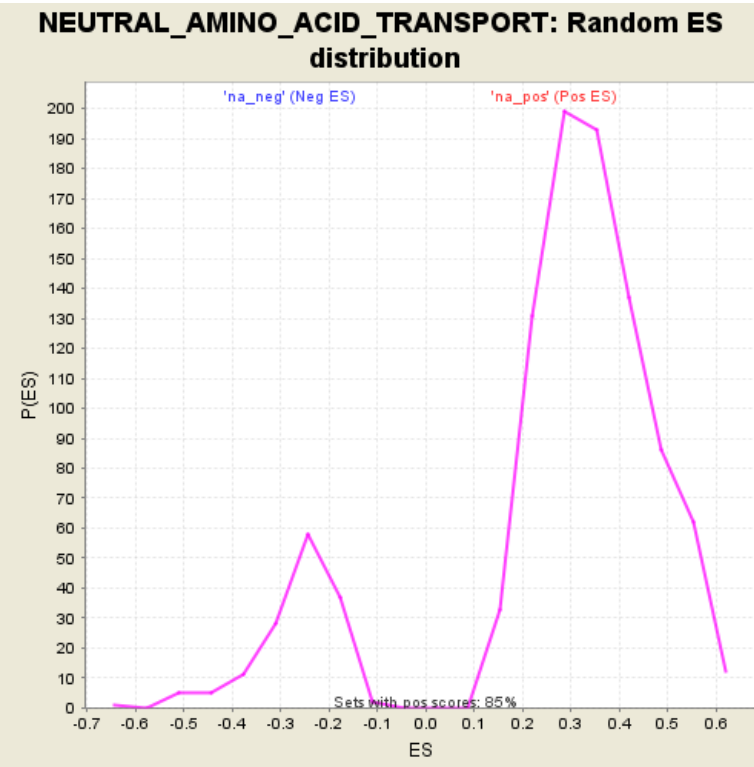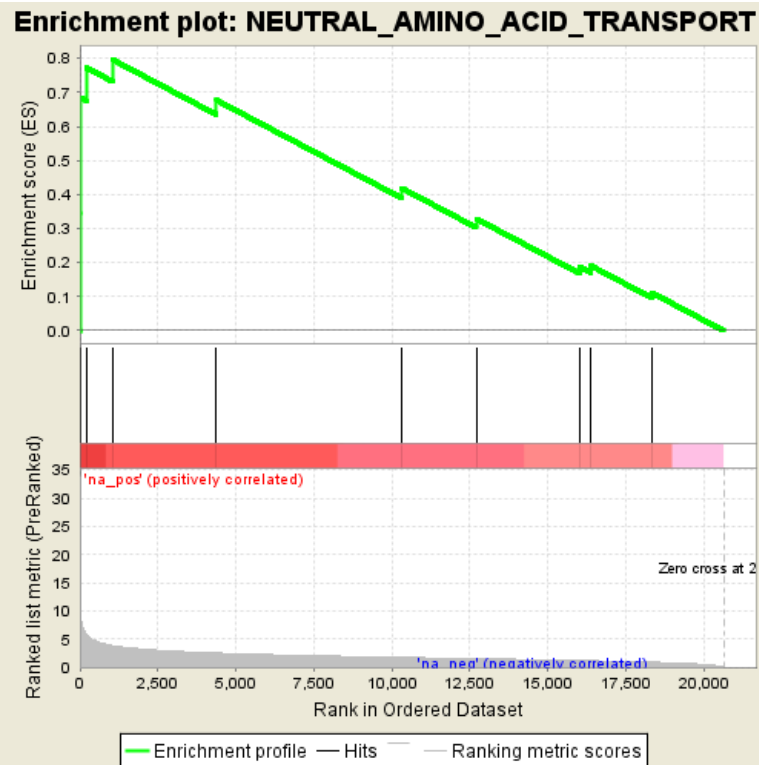
Figure S8 DNA methylation at genes involved in neutral amino acid transport were identified by Gene Set Enrichment Analysis of the overall meta-analyses of systolic and diastolic BP. Genes belonging to this category are primarily solute carrier (*SLC*) genes. The plots show the distribution of the ranked genes in the overall meta-analysis.

Abbreviations: ES, enrichment score.

| CpG site	Chr	Position	UCSC Gene	Systolic BP				Diastolic BP			
				Discovery		Replication		Discovery		Replication	
				Coeff	p value	Coeff	p value	Coeff	p value	Coeff	p value
cg15935121	1	2230601	*SKI*	-7.2E-05	1.9E-07	-1.0E-05	5.3E-01	-0.0001	1.9E-08	-6.8E-05	2.2E-02
cg09556700	1	2230668	*SKI*	-0.0001	7.9E-08	-4.1E-05	5.8E-02	-0.0001	7.9E-04	-0.0001	2.4E-03
cg25203007	1	24126017	*GALE*	-0.0001	3.0E-09	-3.2E-05	1.3E-01	-0.0001	6.7E-05	-2.6E-05	5.0E-01
cg23999170	1	115628111	*TSPAN2*	-0.0001	2.7E-06	-0.0001	1.6E-05	-0.0002	6.4E-08	-0.0002	3.4E-07
cg16246545	1	120255941	*PHGDH*	-0.0002	2.4E-10	-0.0002	3.3E-14	-0.0002	2.2E-04	-0.0003	4.3E-07
cg14476101	1	120255992	*PHGDH*	-0.0003	1.5E-16	-0.0004	7.0E-21	-0.0004	6.0E-11	-0.0005	1.9E-12
cg19693031	1	145441552	*TXNIP*	-0.0002	7.7E-13	-0.0003	3.8E-19	-0.0002	6.0E-07	-0.0004	7.5E-10
cg24955196	1	154982621	*ZBTB7B*	0.0001	5.8E-08	5.0E-05	1.5E-02	0.0001	2.6E-03	0.0001	7.5E-04
cg12417775	1	212463238	*PPP2R5A*	-0.0001	4.4E-06	-4.8E-05	6.7E-02	-0.0002	4.6E-08	-0.0001	1.9E-03
cg08035323	2	9843525		-0.0001	4.2E-05	-8.7E-05	4.1E-03	-0.0003	1.4E-08	-0.0002	2.6E-04
cg11938080	2	85924685	*GNLY*	-7.3E-05	3.0E-08	-2.8E-05	1.3E-01	-0.0001	3.6E-07	-5.9E-05	7.7E-02
cg11308319	2	240291426	*HDAC4*	0.0001	6.4E-08	9.0E-06	6.6E-01	0.0001	8.3E-05	1.4E-05	7.0E-01
cg06398474	2	240291509	*HDAC4*	0.0001	2.0E-08	1.6E-05	3.8E-01	0.0002	8.4E-06	2.2E-05	5.0E-01
cg19787535	3	87843755		-0.0001	7.4E-09	-1.4E-05	6.7E-01	-0.0002	3.0E-04	6.3E-05	2.8E-01
cg06690548	4	139162808	*SLC7A11*	-0.0001	3.4E-16	-0.0002	8.3E-20	-0.0002	5.5E-14	-0.0003	9.9E-14
cg04104695	5	139058749	*CXXC5*	-8.1E-05	2.5E-08	-3.7E-05	1.1E-01	-0.0001	1.7E-05	-1.0E-04	1.7E-02
cg18120259	6	43894639	*LOC100132354*	-0.0001	1.5E-08	-0.0002	9.4E-15	-0.0002	1.9E-05	-0.0003	6.9E-10
cg00533891	10	80919242	*ZMIZ1*	-0.0001	2.4E-07	-9.2E-05	3.7E-03	-0.0003	4.4E-09	-0.0002	8.9E-04
cg02116864	10	134222453	*PWWP2B*	-9.7E-05	3.6E-08	-5.5E-05	1.1E-02	-0.0002	1.4E-07	-0.0001	1.1E-02
cg00805360	10	135091210	*ADAM8*	-0.0001	7.2E-08	-6.0E-05	2.3E-02	-0.0002	6.9E-05	-8.7E-05	7.1E-02
cg17061862	11	9590431		-0.0001	6.9E-05	-0.0002	6.6E-09	-0.0003	5.1E-08	-0.0003	1.2E-06
cg00574958	11	68607622	*CPT1A*	-5.1E-05	1.9E-08	-4.8E-05	1.4E-06	-8.2E-05	5.9E-07	-6.7E-05	2.5E-04
cg19089328	11	71210210	*NADSYN1*	0.0001	5.0E-08	1.4E-05	6.4E-01	0.0002	3.2E-05	3.6E-05	5.2E-01
cg10601624	12	6404377		-8.3E-05	6.6E-08	-0.0001	1.6E-10	-0.0001	3.5E-07	-0.0002	1.7E-07
cg27040968	14	99787363		-7.6E-05	5.6E-08	-1.0E-06	9.4E-01	-0.0001	1.7E-07	-5.9E-05	6.9E-02
cg09807356	16	88905700	*GALNS*	0.0002	9.4E-08	6.0E-06	5.8E-01	0.0002	4.2E-03	1.3E-05	5.6E-01
cg13917614	17	40125660	*CNP*	-0.0002	2.7E-10	-4.0E-05	3.3E-02	-0.0002	1.4E-07	-7.1E-05	3.9E-02
cg22304262	19	47287778	*SLC1A5*	-0.0001	5.4E-10	-0.0001	8.7E-09	-0.0002	6.0E-07	-0.0002	4.9E-05
cg02711608	19	47287964	*SLC1A5*	-0.0001	3.0E-11	-0.0001	1.1E-11	-0.0002	3.2E-05	-0.0002	3.0E-06
cg26080567	21	37536137	*DOPEY2*	0.0001	5.5E-08	2.7E-05	2.0E-01	0.0002	1.3E-05	-4.0E-05	3.1E-01
cg07141002	22	38201690	*H1F0*	-0.0002	2.8E-08	-6.3E-05	3.8E-02	-0.0003	3.8E-08	-0.0002	1.6E-03

Table S1 Discovery and replication meta-analyses results for the thirty-one CpG sites with p value <1.0 X 10^{-7} at the discovery stage. Position is Hg19. Coefficients give the percent change in DNA methylation for every 1-unit change in blood pressure.

Abbreviations: BP, blood pressure; Chr, chromosome; Coeff, coefficient; CpG, cytosine-phosphate-guanine; UCSC, University of California Santa Cruz.

CpG site	Chr	Position	UCSC Gene	Gene Location	CpG Island	DHS	Enhancer
cg23999170	1	115628111	*TSPAN2*	Body	N Shelf		
cg16246545	1	120255941	*PHGDH*	Body	S Shore		
cg14476101	1	120255992	*PHGDH*	Body	S Shore		
cg19693031	1	145441552	*TXNIP*	3'UTR			
cg08035323	2	9843525				DHS	Enhancer
cg06690548	4	139162808	*SLC7A11*	Body			
cg18120259	6	43894639	LOC100132354	Body		DHS	Enhancer
cg00533891	10	80919242	*ZMIZ1*	5'UTR		DHS	Enhancer
cg17061862	11	9590431			N Shelf		
cg00574958	11	68607622	*CPT1A*	5'UTR	N Shore		
cg10601624	12	6404377			S Shelf	DHS	Enhancer
cg22304262	19	47287778	*SLC1A5*	Body;5'UTR	N Shelf		
cg02711608	19	47287964	*SLC1A5*	1stExon;5'UTR;Body	N Shelf		

Table S2 Illumina annotation for replicated blood pressure CpG sites. Position is Hg19.

Abbreviations: Chr, chromosome; CpG, cytosine-phosphate-guanine; DHS, DNase I hypersensitive site; UCSC, University of California Santa Cruz; UTR, untranslated region.

CpG site	Chr	Position	Systolic BP					Diastolic BP				
			Discovery			Replication	Overall	Discovery			Replication	Overall
			Het p value	Race p value	Race and cell type p	Het p value	Het p value	Het p value	Race p value	Race and cell type p	Het p value	Het p value
cg23999170	1	115628111	0.8268	0.6674	0.5586	0.0634	0.9521	0.4965	0.8864	0.7823	0.2852	0.7553
cg16246545	1	120255941	0.6613	0.4293	0.1304	0.5039	0.1098	0.1588	0.8459	0.2307	0.6393	0.1271
cg14476101	1	120255992	0.2434	0.5686	0.0744	0.4511	0.0487	0.0947	0.5398	0.0626	0.8208	0.1481
cg19693031	1	145441552	0.2155	0.9776	0.5260	0.2938	0.0081	0.1174	0.4529	0.3516	0.3328	0.0599
cg08035323	2	9843525	0.3762	0.2123	0.1343	0.2910	0.4229	0.1455	0.0245	0.0037	0.6021	0.2192
cg06690548	**4**	**139162808**	**3.15E-07**	**0.0008**	**4.45E-06**	**0.1621**	**0.0149**	**2.20E-05**	**0.0001**	**4.39E-06**	**0.0047**	**0.0921**
cg18120259	6	43894639	0.1549	0.7064	0.4899	0.2389	0.1862	0.3308	0.9610	0.7416	0.2316	0.1954
cg00533891	10	80919242	0.1800	0.5954	0.6065	0.1649	0.3622	0.6205	0.4702	0.8903	0.7653	0.3503
cg17061862	11	9590431	0.0133	0.2222	0.0077	0.1004	0.0779	0.4931	0.6978	0.4094	0.5484	0.8683
cg00574958	11	68607622	0.1491	0.3041	0.1967	0.0047	0.8235	0.0267	0.9703	0.7466	0.3659	0.5334
cg10601624	12	6404377	0.7570	0.0853	0.3723	0.4898	0.0854	0.6368	0.1311	0.5958	0.3886	0.2762
cg22304262	19	47287778	0.3399	0.3279	0.5860	0.3930	0.5529	0.5831	0.2701	0.1851	0.8915	0.9046
cg02711608	19	47287964	0.0763	0.4415	0.8385	0.7613	0.9155	0.0357	0.6652	0.8892	0.8832	0.7833

Table S3 Tests for heterogeneity in epigenome-wide meta-analyses for systolic and diastolic BP among the replicated CpG sites. Position is Hg19. Cochran's *Q* statistic was used to assess variation in the discovery, replication, and overall meta-analyses. Differences in the discovery meta-analyses that may be explained by race (EA compared to AA) or cell type (whole blood compared to T cells) were tested using a 1 degree of freedom Chi-square test for differences between strata.

Abbreviations: BP, blood pressure; Chr, chromosome; CpG, cytosine-phosphate-guanine; Het, Cochran's heterogeneity statistic's (*Q*) *P*-value.

Supplemental Excel file provided.

Table S4 Results of overall epigenome-wide meta-analyses for systolic and diastolic BP (p value <1.0 x 10⁻⁷). Position is Hg19. Coefficients give the percent change in DNA methylation for every 1-unit change in blood pressure.

Abbreviations: BP, blood pressure; Chr, chromosome; Coeff, coefficient; CpG, cytosine-phosphate-guanine; UCSC, University of California Santa Cruz.

CpG site	Chr	Position	UCSC Gene	Systolic BP				Diastolic BP			
				No AHT				No AHT			
				Coeff	p value	Discovery p	Replication p	Coeff	p value	Discovery p	Replication p
cg16246545	1	120,255,941	*PHGDH*	-0.0002	**8.0E-10**	2.4E-10	3.3E-14	-0.0002	7.2E-03	2.2E-04	4.3E-07
cg14476101	1	120,255,992	*PHGDH*	-0.0003	**5.8E-13**	1.5E-16	7.0E-21	-0.0003	1.9E-05	6.0E-11	1.9E-12
cg19693031	1	145,441,552	*TXNIP*	-0.0002	**5.9E-11**	7.7E-13	3.8E-19	-0.0002	8.1E-05	6.0E-07	7.5E-10
cg24955196	1	154,982,621	*ZBTB7B*	0.0001	**8.9E-09**	5.8E-08	1.5E-02	0.0003	**1.8E-09**	2.6E-03	7.5E-04
cg06690548	4	139,162,808	*SLC7A11*	-0.0001	**1.6E-11**	3.4E-16	8.3E-20	-0.0002	1.1E-07	5.5E-14	9.9E-14
cg05014727	10	6,214,016	*PFKFB3*	-0.0002	**2.3E-08**	3.0E-04	6.8E-04	-0.0001	1.1E-02	5.7E-03	7.0E-02
cg08994060	10	6,214,026	*PFKFB3*	-0.0002	**6.3E-08**	2.8E-03	1.7E-03	-0.0002	2.4E-03	4.2E-03	8.0E-02
cg26262157	10	6,214,079	*PFKFB3*	-0.0002	**5.2E-08**	3.7E-03	6.1E-04	-0.0001	7.9E-03	7.6E-03	5.9E-02
cg00805360	10	135,091,210	*ADAM8*	-0.0002	**2.4E-08**	7.2E-08	2.3E-02	-0.0003	**2.6E-08**	6.9E-05	7.1E-02

Table S5 Results of an overall epigenome-wide meta-analyses for systolic and diastolic BP among individuals not reporting the use of antihypertensive medication (p value <1.0 x 10^{-7}). Position is Hg19. Coefficients give the percent change in DNA methylation for every 1 mmHg change in blood pressure.

Abbreviations: AHT, antihypertensive treatment; BP, blood pressure; Chr, chromosome; Coeff, coefficient; CpG, cytosine-phosphate-guanine; UCSC, University of California Santa Cruz.

CpG site	top BP-associated meQTL	meQTL-CpG dist, bp	Trait	ICBP 1000G SBP p value	ICBP 1000G DBP p value
cg00533891	rs737012	13,056	SBP + DBP	0.1967	0.1572
cg02711608	rs56050670	17,333	SBP + DBP	0.0410	0.0813
cg08035323	rs17592239	15,359	SBP	0.0241	0.4423
cg08035323	rs4669419	-12,904	DBP	0.0998	0.0936
cg10601624	rs4764572	13,621	SBP	0.5050	0.6137
cg10601624	rs984337	-20,852	DBP	0.8157	0.3748
cg14476101	rs561931	-1,486	SBP + DBP	0.0074	0.0129
cg16246545	rs561931	-1,435	SBP + DBP	0.0074	0.0129
cg17061862	rs6486429	-2,878	SBP	0.1288	0.8563
cg17061862	rs11042429	9,727	DBP	0.4218	0.0239
cg18120259	rs7745517	456	SBP + DBP	0.7853	0.7433
cg22304262	rs4578775	12,736	SBP + DBP	0.0408	0.1099
cg23999170	rs10858064	-2,415	SBP + DBP	0.0525	0.0195

Table S6 Association of meQTLs with BP in ICBP 1000Genomes data. The meQTLs for each CpG were pruned to be independent ($r^2 < 0.2$) and the meQTL most associated with systolic and/or diastolic BP is shown. Distance between SNPs and meQTL is from position in Hg19.

Abbreviations: bp, base pairs; CpG, cytosine-phosphate-guanine; DBP, diastolic blood pressure; dist, distance; meQTL, methylation quantitative trait locus; SBP, systolic blood pressure.

CpG site	Trait	IV SNPs, n	Mean estimate	SE	p value	Pleiotropy p value
cg00533891	*SBP*	4	-1.5	2.7	0.4972	0.3643
	DBP	4	-0.8	1.8	0.6948	0.1208
cg02711608	*SBP*	3	-7.2	5.4	0.3156	0.3135
	DBP	3	-6.1	3.0	0.1107	0.5290
cg08035323	*SBP*	6	6.2	2.1	0.0067	0.0762
	DBP	6	0.9	1.3	0.5016	0.0720
cg14476101	*SBP*	7	-3.7	2.3	0.1567	0.6649
	DBP	7	-1.8	1.3	0.2291	0.9083
cg16246545	*SBP*	6	-3.5	3.0	0.2922	0.8693
	DBP	6	-1.5	1.7	0.3964	0.7974
cg17061862	*SBP*	10	-1.0	1.6	0.4559	0.4309
	DBP	10	-0.5	1.0	0.5896	0.6703
cg22304262	*SBP*	7	-2.2	2.5	0.4186	0.3419
	DBP	7	-1.9	1.3	0.1465	0.2795
cg23999170	*SBP*	5	-1.1	4.7	0.8320	0.7656
	DBP	5	1.9	3.0	0.5631	0.8470

Table S7 Two-sample inverse variance-weighted Mendelian randomization tests for causal relationships of CpG sites with BP (forward causality). Estimates of meQTL association with CpGs were derived from meta-analyses of ARIC, FHS, RS, and WHI-EMPC. Estimates of meQTL association with BP are from 1000Genomes analysis in the International Consortium for Blood Pressure. All samples are of European ancestry.

Abbreviations: CpG, cytosine-phosphate-guanine; DBP, diastolic blood pressure; IV, instrumental variable; meQTL, methylation quantitative trait locus; SBP, systolic blood pressure; SNP, single nucleotide polymorphism.

| | | Forward Mendelian Randomization: CpG → BP | | | | | | Reverse Mendelian Randomization: BP → CpG | | | | | |
| | | Inverse-Variance Weighted Test for Causality | | | | Egger Test | | Inverse-Variance Weighted Test for Causality | | | | Egger Test | |
CpG	Trait	IV SNPs, n	Mean estimate	(SE)	p value	Pleiotropy p	Causal p	IV SNPs, n	Mean estimate	(SE)	p value	Pleiotropy p	Causal p
cg00533891	SBP	6	-10.3	(13.5)	0.3130	0.0318	0.1911	29	-8.3E-04	(3.8E-04)	0.0388	0.3590	0.0358
	DBP	6	-14.9	(7.8)	0.0405	0.1299	0.1135	29	-2.0E-03	(5.7E-04)	0.0013	0.3785	0.0065
cg00574958	SBP	-						29	1.4E-04	(1.2E-04)	0.2301	0.9356	0.4243
	DBP							29	1.4E-04	(1.9E-04)	0.4346	0.0254	0.0409
cg02711608	SBP	3	-31.1	(30.2)	0.2953	0.3424	0.7072	29	-5.5E-04	(2.3E-04)	0.0204	0.4964	0.0361
	DBP	3	-30.2	(17.3)	0.1879	2.4E-05	0.5167	29	-7.7E-04	(3.7E-04)	0.0495	0.6032	0.1057
cg06690548	SBP	-						29	-4.3E-04	(3.5E-04)	0.2267	0.4193	-
	DBP	-						29	-1.7E-04	(5.7E-04)	0.7724	0.8800	-
cg08035323	SBP	7	20.9	(11.1)	0.0091	0.5892	-	29	-4.1E-04	(3.5E-04)	0.2206	0.6500	-
	DBP	7	15.1	(6.4)	0.0111	0.5803	-	29	-1.2E-03	(5.7E-04)	0.0226	0.8343	-
cg10601624	SBP	-						29	-3.6E-04	(2.2E-04)	0.1069	0.1527	0.9929
	DBP	-						29	-1.0E-03	(3.5E-04)	0.0033	0.0164	0.9232
cg14476101	SBP	7	-2.5	(14.3)	0.8669	0.5364	-	29	-1.3E-04	(5.0E-04)	0.7977	0.2059	-
	DBP	7	1.0	(5.6)	0.8623	0.7469	-	29	-2.4E-04	(8.2E-04)	0.7757	0.5550	-
cg16246545	SBP	6	-10.7	(17.5)	0.5684	0.2205	-	29	-1.3E-04	(4.1E-04)	0.7618	0.0758	-
	DBP	6	-1.8	(6.7)	0.8042	0.4186	-	29	-5.1E-05	(6.7E-04)	0.9404	0.1424	-
cg17061862	SBP	10	-8.6	(10.2)	0.4224	0.0959	-	29	1.6E-04	(3.5E-04)	0.6574	0.8119	-
	DBP	10	4.8	(5.1)	0.1112	0.9740	-	29	1.1E-05	(5.7E-04)	0.9844	0.4935	-
cg18120259	SBP	-						29	6.8E-05	(3.1E-04)	0.8084	0.6888	-
	DBP	-						29	-1.8E-04	(5.1E-04)	0.6968	0.4091	-
cg19693031	SBP	-						29	3.3E-04	(3.8E-04)	0.3889	0.6004	-
	DBP	-						29	5.9E-04	(6.2E-04)	0.3509	0.6332	-
cg22304262	SBP	9	0.4	(11.4)	0.9696	0.8178	-	29	-5.4E-04	(3.0E-04)	0.0516	0.4670	-
	DBP	9	-8.5	(6.5)	0.2219	0.9145	-	29	-9.3E-04	(4.9E-04)	0.0367	0.2711	-
cg23999170	SBP	5	5.9	(18.4)	0.7547	0.2738	0.3934	29	4.1E-04	(3.1E-04)	0.1954	0.8319	-
	DBP	5	-1.2	(10.6)	0.9151	0.0363	0.1370	29	2.7E-04	(5.1E-04)	0.6080	0.9852	-

Table S8 Multi-instrument bidirectional Mendelian randomization tests for causality between DNA methylation and BP using inverse-variance weighted effects and Egger regression tests. Tests for causality based on Egger regression have substantially reduced power compared to inverse-variance weighted tests. Egger causal p values are provided only when pleiotropic effects of IVs are significant for either BP trait (p < 0.05) or when both the inverse-variance weighted and Egger tests both support reverse causation.

Abbreviations: BP, blood pressure; CpG, cytosine-phosphate-guanine; DBP, diastolic blood pressure; IV, instrumental variable; SBP, systolic blood pressure; SE, standard error; SNP, single nucleotide polymorphism.

CpG site	Chr	Position	1) Genes tested for GE-Methylation, n	2) Genes tested for GE-BP significant in 1), n	Significant genes in triangular GE analysis, n
cg23999170	1	115628111	22	5	2
cg16246545	1	120255941	20	1	0
cg14476101	1	120255992	21	1	0
cg19693031	1	145441552	32	0	-
cg08035323	2	9843525	22	0	-
cg06690548	4	139162808	8	1	1
cg18120259	6	43894639	50	0	-
cg00533891	10	80919242	14	0	-
cg17061862	11	9590431	26	3	0
cg00574958	11	68607622	30	4	2
cg10601624	12	6404377	61	10	2
cg22304262	19	47287778	52	8	0
cg02711608	19	47287964	52	7	0

Table S9 Number of gene expression analyses per replicated CpG site used for FDR-correction of p values. Position is in Hg19.

Abbreviations: BP, blood pressure; Chr, chromosome; CpG, cytosine-phosphate-guanine; GE, gene expression.

Supplemental Excel file provided.

Table S10 Genes in a *cis*-region (+/- 1Mb) of replicated CpG sites 1) suggestively associated with methylation in meta-analyses of FHS and RS at p value <0.05, and 2) associated with blood pressure traits with at least one p value <0.05. Start and stop positions are in Hg19.

Abbreviations: Chr, chromosome; Coeff, coefficient; Corr, correlation coefficient for continuous traits; CpG, cytosine-phosphate-guanine; DBP, diastolic blood pressure; HTN, hypertension; SBP, systolic blood pressure; UCSC, University of California Santa Cruz.

CpG site	Chr	Position	UCSC Gene	Phenotype(s)
cg23999170	1	115628111	*TSPAN2*	
cg16246545	1	120255941	*PHGDH*	Alcohol intake[1], BMI[2]
cg14476101	1	120255992	*PHGDH*	Lipids[3], A-diol[4], BMI[2,5], waist circumference[5]
cg19693031	1	145441552	*TXNIP*	Lipids[3,4], type 2 diabetes[6–9]
cg08035323	2	9843525		
cg06690548	4	139162808	*SLC7A11*	Lipids[3], A-diol[4], BMI[2], IgE[10]
cg18120259	6	43894639	LOC100132354	A-diol [4], BMI[2]
cg00533891	10	80919242	*ZMIZ1*	
cg17061862	11	9590431		
cg00574958	11	68607622	*CPT1A*	Lipids[3,4,11], BMI[2,5,8,12], waist circumference[5], metabolic syndrome[13], hypertriglyceridemic waist[14], adiponectin[15]
cg10601624	12	6404377		
cg22304262	19	47287778	*SLC1A5*	A-diol[4]
cg02711608	19	47287964	*SLC1A5*	Lipids[3], BMI[2]

Table S11 Phenotypes previously associated with the 13 CpG sites replicated for association with BP. Position is in Hg19.

Abbreviations: A-diol, 4-androsten-3-beta,17-beta-diol disulfate; BMI, body mass index; Chr, chromosome; Coeff, coefficient; CpG, cytosine-phosphate-guanine; DBP, diastolic blood pressure; HTN, hypertension; IgE, immunoglobulin E; SBP, systolic blood pressure; UCSC, University of California Santa Cruz.

SNP	Chr	CpG	CpG position (bp)	SNP-CpG distance (bp)	SNP-CpG		CpG-BP EWAS		Additional GWAS loci in high LD ($r^2>0.8$) with Kato SNP (ARIC SNP-CpG p value)
					ARIC p[a] N=790	Kato p[b] N=6,684	SBP p N=17,010	DBP p N=17,010	
rs880315	1	cg02903756	10,750,680	46,186	NR	4.69E-12	0.7525	0.7593	
rs12567136	1	cg05228408	11,865,352	18,379	NR	3.01E-71	0.1157	0.4805	
		cg20946054	11,761,766	121,965	**8.09E-05**	NR	0.0318	0.0015	rs17367504 (**1.69E-04**)
rs1344653	2	cg13996430	19,741,587	-10,742	NR	5.09E-17	0.6825	0.7047	
		cg17314700	19,548,291	182,554	8.61E-03	NR	0.8128	0.4091	
rs1275988	2	cg02952978	27,165,719	-251,355	**1.32E-03**	NR	0.2293	0.2035	
		cg19115882	26,919,145	-4,781	NR	2.09E-32	0.0041	0.0174	
rs7629767	3	cg03022575	42,003,672	39,837	2.43E-03	NR	0.4034	0.2647	rs3774372 (**5.21E-04**), rs9815354 (**6.73E-04**)
		cg02108620	42,002,230	41,279	NR	2.54E-152	0.9744	0.1441	
rs13149993	4	cg05974274	81,123,369	35,176	8.65E-03	NR	0.1271	0.4145	
		cg05452645	81,117,647	40,898	NR	2.38E-29	0.0010	0.0074	
rs2014912	4	cg01368160	87,187,259	-471,589	9.89E-03	NR	0.0323	0.4454	
		cg20784207	86,597,598	118,072	NR	3.57E-37	0.1274	0.8120	
rs7733331	5	cg24363955	32,788,467	40,379	NR	1.89E-28	0.7480	0.9826	
		cg15198736	32,709,396	119,450	**5.82E-04**	NR	0.8623	0.1140	rs1173771 (**6.49E-04**)
rs13359291	5	cg07849972	122,759,670	-283,213	3.03E-03	NR	0.5951	0.8575	
		cg23290100	122,435,626	40,831	NR	3.34E-189	0.0337	0.7627	
rs9687065	5	cg12302647	148,533,875	-142,735	2.65E-02	NR	0.8404	0.5647	
		cg18129178	148,520,854	-129,714	NR	4.05E-50	0.5191	0.8933	
rs11960210	5	cg22790839	157,883,933	-66,299	NR	6.09E-27	0.8453	0.7847	
		cg12515908	157,413,937	403,697	7.43E-03	NR	0.7405	0.5791	
rs1563788	6	cg12585005	43,655,766	-347,403	**1.36E-03**	NR	0.4349	0.4140	
		cg00084398	43,249,983	58,380	NR	3.76E-48	0.4597	0.1026	
rs17080102	6	cg01354656	151,411,710	-406,940	9.02E-03	NR	0.9834	0.9821	
		cg02784464	151,121,916	-117,146	NR	2.35E-17	0.5820	0.8370	
rs10260816	7	cg12244052	45,961,469	48,631	NR	1.24E-08	0.5673	0.8391	
		cg13070193	45,613,752	396,348	**6.41E-04**	NR	0.4437	0.9674	
rs731141	10	cg10751070	96,143,568	-244,887	NR	1.43E-08	0.8070	0.9242	
		cg04143348	95,656,819	241,862	**9.18E-04**	NR	0.8809	0.9756	rs9663362 (**6.44E-04**), rs932764 (4.18E-03)
rs11191375	10	cg03275084	104,617,031	-152,374	3.56E-02	1.75E-06	0.0016	0.0066	
		cg07119830	104,412,306	52,351	NR	3.E-746	0.6838	0.4193	

SNP	Chr	CpG	Position	Distance	p-value	p-value			
		cg12331743	104,153,933	310,724	5.94E-03	NR	0.0484	0.0096	
rs2484294	10	cg20663200	116,163,392	-371,330	**1.68E-03**	NR	0.8672	0.2613	
		cg05575054	115,804,968	-12,906	NR	4.83E-27	0.0163	0.0283	
rs751984	11	cg00009053	61,283,865	-5,619	NR	2.76E-58	0.7452	0.3231	
		cg26786382	61,277,328	918	5.65E-03	1.69E-06	0.0628	0.2037	
rs2055450	11	cg11637980	101,000,717	-450,300	3.10E-03	NR	0.3830	0.9702	rs633185 (**1.73E-03**)
		cg05925497	100,734,094	-183,677	2.07E-02	1.09E-12	0.2339	0.6334	
rs10894192	11	cg03927812	130,271,903	-5,786	NR	1.31E-55	0.1087	0.4554	
		cg04713042	129,912,454	353,663	**1.49E-03**	NR	0.8072	0.6620	rs11222084 (3.94E-03)
rs11105354	12	cg00757033	89,920,650	105,873	NR	1.91E-194	0.7233	0.8567	
		cg06869160	89,919,815	106,708	2.58E-02	NR	0.5176	0.4349	rs2681472 (2.29E-02), rs2681492 (2.06E-02), rs17249754 (2.24E-02)
rs3184504	12	cg16423624	112,220,997	-336,389	6.35E-03	NR	0.4666	0.2621	rs653178 (2.29E-03)
		cg10833066	111,807,467	96,904	NR	1.29E-48	0.2043	0.4457	
rs1378942	15	cg16377819	75,575,161	-497,794	**1.44E-03**	NR	0.4735	0.7943	
		cg02696790	75,250,997	-173,630	NR	1.42E-86	0.0262	0.0234	
		cg20668952	75,082,885	-5,518	2.32E-02	6.54E-09	0.0521	0.0665	
rs8032315	15	cg06330618	91,428,456	-10,159	NR	2.01E-167	0.1058	0.4800	
		cg02864248	91,411,838	6,459	1.18E-02	2.91E-16	0.4764	0.4012	
		cg01994513	91,260,603	157,694	2.22E-03	NR	0.7373	0.2635	rs2521501 (**1.58E-03**)
rs2301597	17	cg24180402	43,221,464	-48,191	1.96E-02	4.39E-30	0.0039	0.0999	
		cg19407385	43,099,144	74,129	NR	3.51E-283	0.3725	0.6644	
		cg07822074	43,098,904	74,369	8.13E-03	8.36E-27	0.5166	0.7629	
		cg06471905	42,835,688	337,585	**6.28E-04**	NR	0.9597	0.4451	
rs7405452	17	cg06602723	46,693,336	-18,666	2.00E-02	2.34E-16	0.8131	0.7141	
		cg22053945	46,651,360	23,310	NR	1.05E-151	0.1920	0.3117	
		cg04171235	46,604,393	70,277	4.45E-02	1.75E-14	0.8655	0.4466	
		cg08036188	46,522,560	152,110	3.01E-03	NR	0.3476	0.1459	
rs2240736	17	cg06762332	59,573,001	-87,608	2.25E-03	NR	0.9046	0.5538	
		cg00730441	59,483,863	1,530	NR	6.66E-131	0.8128	0.9765	
rs740406	19	cg14103263	2,727,148	-494,927	**1.24E-03**	NR	0.4348	0.7746	
		cg04052466	2,251,061	-18,840	NR	2.33E-45	0.7375	0.7853	

Table S12 Top associations of methylation +/-500 kb of methylation-mediated GWAS SNPs identified by Kato et al. The meQTL analyses were conducted among ARIC whites (N=790) using methylation data +/-500 kb of each sentinel SNP reported by Kato et al. Genotypes were directly measured or imputed to

1000Genomes. meQTL p values are shown for lead CpGs reported either by Kato et al or in ARIC analyses and any non-lead CpGs (in the online supplement) with p value <0.05 in ARIC. ARIC meQTL analyses were Bonferroni-corrected for the 28 GWAS loci reported by Kato et al (p value <0.0018), with significant p values shown in bold. Lookups in the overall meta-analysis for association of each CpG with systolic and diastolic BP is shown (N=17,058). Position is in Hg19. The 1000Genomes CEU reference population was used to calculate r^2.

[a]meQTL p values not reported (NR) in ARIC are >0.05. [b]meQTL p values not reported (NR) in Kato et al are >3.8E-6 in discovery (N=1,904). Combined p values for discovery and replication are shown (N=6,684).

Abbreviations: bp, base pairs; BP, blood pressure; Chr, chromosome; Coeff, coefficient; CpG, cytosine-phosphate-guanine; DBP, diastolic blood pressure; GWAS, genome-wide association studies; LD, linkage disequilibrium; NR, not reported; SBP, systolic blood pressure; SNP, single nucleotide polymorphism.

| SNP | Chr | CpG | SNP-CpG distance (bp) | Kato CpG? | Systolic BP | | | | | Diastolic BP | | | | |
					Coeff	p value	CpG Adj Coeff	p value	Coeff Percent Change	Coeff	p value	CpG Adj Coeff	p value	Coeff Percent Change
rs12567136	1	cg05228408	18,379	Yes	2.45	0.0366	0.75	0.5817	69%	1.90	0.0039	1.81	0.0178	4%
rs12567136	1	cg20946054	121,965		2.45	0.0366	1.34	0.2908	45%	1.90	0.0039	1.74	0.0145	8%

Table S13 Adjustment of the association of a GWAS locus (*CLCN6* rs12567136) for DNA methylation at two CpG sites among ARIC participants of European ancestry (N=790). Position is in Hg19.

Abbreviations: Adj, adjusted; bp, base pairs; BP, blood pressure; Chr, chromosome; Coeff, coefficient; CpG, cytosine-phosphate-guanine; GWAS, genome-wide association studies; SNP, single nucleotide polymorphism.

1. Cohort Descriptions and Acknowledgements

The Amish Complex Disease Research Studies (Amish)

The Old Order Amish (OOA) subjects included in this study were participants of several studies of cardiovascular health in relatively healthy volunteers from the OOA community of Lancaster County, PA and their family members. The studies were carried out at the University of Maryland as part of the Amish Complex Disease Research Program (ACDRP); see http://medschool.umaryland.edu/endocrinology/amish/research-program.asp. The OOA population of Lancaster County, PA immigrated to the Colonies from Western Europe in the early 1700's. There are now over 30,000 OOA individuals in the Lancaster area, nearly all of whom can trace their ancestry back 12-14 generations to approximately 750 founders. Investigators at the University of Maryland, School of Medicine have been studying the genetic determinants of cardiometabolic health in this population since 1993. To date, over 7,000 Amish adults have participated in one or more of our studies. The subjects on whom the methylation chip was used were participants of the Heredity and Phenotype Interaction (HAPI) heart study[16], the Pharmacogenomics of Anti-Platelet Intervention (PAPI) study[17], or the Amish Family Diabetes Study (AFDS)[18]. These studies collected large numbers of variables including demographic and anthropometric information, medical history, clinical characteristics, lifestyle factors, and study specific variables, as well as blood and urine samples. Blood pressure was measured in triplicate using a standard sphygmomanometer in the sitting position after 5 minutes rest, and the average of the 3 measures was used for analysis. All study protocols were approved by the institutional review board at the University of Maryland and participating institutions. Informed consent was obtained from each of the study participants.

Genomic DNA was extracted from whole blood of 432 samples, and quantitated using picogreen. Bisulfite conversion of 500 ng genomic DNA was performed using the EZ DNA Methylation Kit (Zymo Research; Irvine, CA, USA) according to the manufacturer's instructions, and the Illumina Infinium Human Methylation 450K Bead chip array (HM450K)[19] was used for methylation profiling. Sample preparation and hybridization were performed by the same technician, and the arrays were processed using the same scanner to reduce batch effect. Genome Studio (GS) (V2011.1) was used for processing the raw intensities and performing QC for samples and probes. Examining internal control probes identified one array that did not perform well and all 12 samples in it were missing 25% of the data, so we completely excluded this array.GS normalization and background correction were applied according to Illumina's recommendation for the remaining 420 samples. By testing Y chromosome probes we identified 19 potentially mismatched or contaminated samples that were removed. We also removed one from each pair of 18 sets of duplicates leaving 384 samples for analysis all with >95% of probes detected with average p value <0.01. Blood pressure measures were available for only 192 individuals at the same time of blood draw. Probes with average detection p value >0.01 or call rate <95% were removed. Also sex chromosome probes and previously identified cross-reactive probes[20–22] were excluded leaving 323,747 probes for analysis. Blood cell subtype composition were estimated using the Housman method[23,24].

Atherosclerosis Risk in Communities (ARIC)

The Atherosclerosis Risk in Communities (ARIC) study is a prospective cohort study of cardiovascular disease risk in four U.S. communities[25]. Between 1987 and 1989, men and women aged 45–64 years were recruited from four sites: Forsyth County, North Carolina; Jackson, Mississippi (African Americans only); suburban Minneapolis, Minnesota; and Washington County, Maryland. The ARIC study protocol was approved by the institutional review board of each participating university. After written informed consent was obtained, including consent for genetic studies, participants underwent a baseline clinical examination and four subsequent follow-up clinical examinations (visits 1-5). Systolic and diastolic blood pressures were measured three times using a random zero sphygmomanometer with subjects in a seated position. The average of the second and third readings was used in analyses. DNA methylation data are available for

African American members of the cohort from two study sites (Forsyth County and Jackson) and were included in the discovery meta-analysis. Measurements of DNA methylation for white members of the cohort later became available and were included in the replication meta-analysis. Cross-sectional analyses within each race were conducted using blood pressure, methylation, and covariates measured concurrently at visit 2 or visit 3.

Genomic DNA was extracted from peripheral blood leukocyte samples using the Gentra Puregene Blood Kit (Qiagen; Valencia, CA, USA). DNA samples were bisulfite-converted using the EZ-96 DNA Methylation Kit (Deep Well Format) (Zymo Research; Irvine, CA, USA) and hybridized to the Illumina HumanMethylation 450K beadarray (Illumina, Inc., San Diego, CA, USA). Probe intensities were extracted using Illumina GenomeStudio 2011.1, Methylation module 1.9.0 software. Samples were removed based on pass rate <95%, gender mismatch, SNP discordance with previous genotyping, and outliers in principal component analysis (N = 107). Additionally, CpG sites missing in ≥5% of samples were excluded. Methylation values were normalized using the Beta MIxture Quantile dilation (BMIQ) method[26]. Blood cell proportions were imputed using the Houseman method[23] among African Americans and Houseman cell types from the Horvath method[27] among white individuals. Surrogate variables were estimated among African Americans after removing the effects of age, sex, BMI, smoking, and the first four principal components of ancestry.

Cardiovascular Health Study (CHS)

The Cardiovascular Health Study (CHS) is a population-based cohort study of risk factors for coronary heart disease and stroke in adults ≥65 years conducted across four field centers. The original predominantly European ancestry cohort of 5,201 persons was recruited in 1989-1990 from random samples of the Medicare eligibility lists; subsequently, an additional predominantly African-American cohort of 687 persons was enrolled for a total sample of 5,888. DNA methylation was measured on 200 European ancestry and 200 African-American ancestry participants. The samples were randomly selected among participants without presence of coronary heart disease, congestive heart failure, peripheral vascular disease, valvular heart disease, stroke or transient ischemic attack at study baseline or lack of available DNA at study year 5.

CHS was approved by institutional review committees at each field center and individuals in the present analysis had available DNA and gave informed consent including consent to use of genetic information for the study of cardiovascular disease. Methylation measurements were performed at the Institute for Translational Genomics and Population Sciences at the Harbor-UCLA Medical Center Institute for Translational Genomics and Population Sciences using the Infinium HumanMethylation450 BeadChip (Illumina Inc, San Diego, CA). Quality control was performed in in the minfi R package (version 1.12.0, http://www.bioconductor.org/packages/release/bioc/html/minfi.html). Samples with low median intensities of below 10.5 (log2) across the methylated and unmethylated channels, samples with a proportion of probes falling detection of greater than 0.5%, samples with QC probes falling greater than 3 standard deviation from the mean, sex-check mismatches, failed concordance with prior genotyping or > 0.5% of probes with a detection p-value > 0.01 were removed. Methylation values were normalized using the SWAN[28] quantile normalization method. White blood cell proportions were not directly measured in CHS and were estimated from the methylation data using the Houseman

method[23].

Framingham Heart Study (FHS)

The Framingham Heart Study Offspring cohort (FHS-Offspring) was initially recruited in 1971 and included 5,124 offspring of the FHS Original cohort. From 2002 to 2005, the adult children (third generation cohort, N=4,095) of the offspring cohort participants were recruited and examined (FHS-3rd Gen). Detailed descriptions of cohorts have been published[29,30]. A total of 2,836 FHS-Offspring participants who attended the eighth exam cycle from 2005-2008 were included in the DNA methylation meta-analysis study. The 1,549 FHS-3rd Gen participants who attended the second exam cycle from 2005-2008 were used as the validation set for testing the proportion of SBP / DBP phenotypic variation explained by the top BP associated DNA methylation loci identified in this study. All participants provided written consent for genetic research.

DNA methylation measurements of the FHS Offspring were conducted in two laboratories: lab 1 samples included an ongoing case-control study of cardiovascular disease (n=576), and lab 2 included the remainder of the offspring cohort (n=2,270). DNA methylation measurements of the FHS 3rd Generation participants (n=1,549) were conducted in one laboratory. Buffy coat preparations were obtained from the whole blood samples and genomic DNA was extracted using the Gentra Puregene Blood Kit (Qiagen, Venlo, Netherlands) according to the manufacturer's instructions. Subsequently, bisulfite conversion of 1 ug genomic DNA was performed (Zymo Research, Irvine, CA, USA) and bisulfite conversion efficiency was determined by PCR amplification of the converted DNA using Universal Methylated Human DNA Standard and Control Primers (Zymo Research, Irvine, CA, USA). The Infinium HumanMethylation450 array (Illumina, San Diego, CA) was used to quantify genome-wide DNA methylation and Illumina GenomeStudio software was used to estimate β scores. β scores underwent normalization within the two laboratory batches using the DASEN methodology implemented in the wateRmelon R package[31] which includes background adjustment of the methylated and unmethylated intensities and quantile normalization of the methylated and unmethylated probes within the two types of probes technologies separately. Quality control filters excluded samples with missing rate > 1% at detection P-value <0.01, poor matching to the 65 single nucleotide polymorphism (SNP) control probe locations, and outliers by multi-dimensional scaling techniques. For quality control at the probe level, those with missing rate >20% at detection P-value <0.01, as well as probes previously identified to map to multiple locations or have an underlying SNP (minor allele frequency >5% in European ancestry 1000 genomes project data) at the CpG site or <=10 bp of the single base extension (n=42,251) were excluded, as were probes on the sex chromosomes.

These procedures left 2,377 FHS-Offspring (n=442 for batch 1 and n=1,935 for batch 2) and 1,522 FHS-3rd Gen participants with phenotype data, and 443,252 probes for analyses. Surrogate variable analysis (SVA) was used to adjust for unmeasured technical and batch effects, including cell count composition[23] SVAs were created within laboratory batches and cohorts after removing the signals of interest from BMI, age, sex, smoking status, systolic blood pressure, low-density lipoprotein cholesterol, high-density lipoprotein cholesterol, triglycerides, fasting plasma glucose, and exam date. SVAs that were associated with SBP / DBP with a P-value < 0.05 were included in subsequent analysis of testing associations of DNA methylation and BP.

Association testing was performed using linear mixed effect regression models, with DNA methylation β score as the dependent variable, SBP or DBP as the independent variables, age, sex, BMI, and SVAs as fixed effects, and family structure as a random effect in the FHS-Offspring cohort . For individuals receiving antihypertensive treatment, treatment-adjusted SBP and DBP were calculated by adding 10 mm Hg and 5 mm Hg to the measured SBP and DBP values, respectively. Regression models were fit for the two laboratory batches separately. An inverse variances weighted meta-analysis of the two batches were conducted by the *metagen()* function in the R package Meta (http://cran.r-project.org/web/packages/meta/index.html). These meta-analysis p values were used to represent the best estimate of the associations of the replication probes in the FHS-Offspring cohort.

Genetic Epidemiology Network of Atherosclerosis (GENOA)

The Genetic Epidemiology Network of Arteriopathy (GENOA) study is a community-based study of hypertensive sibships that was designed to investigate the genetics of hypertension and target organ damage in African Americans from Jackson, Mississippi and non-Hispanic whites from Rochester, Minnesota[32]. In the initial phase of the GENOA study (Phase I: 1996-2001), all members of sibships containing ≥ 2 individuals with essential hypertension clinically diagnosed before age 60 were invited to participate, including both hypertensive and normotensive siblings. Exclusion criteria of the GENOA study were secondary hypertension, alcoholism or drug abuse, pregnancy, insulin-dependent diabetes mellitus, or active malignancy. Eighty percent of African Americans (1,482 subjects) and 75% of non-Hispanic whites (1,213 subjects) from the initial study population returned for the second examination (Phase II: 2001-2005). Study visits were made in the morning after an overnight fast of at least eight hours. Demographic information, medical history, clinical characteristics, lifestyle factors, and blood samples were collected in each phase. Written informed consent was obtained from all subjects and approval was granted by participating institutional review boards. DNA methylation levels were measured only in African Americans participants, so only African Americans were included in the current analysis. Participants were excluded from this analysis if they were also participants in the ARIC study.

Genomic DNA of 422 participants was extracted from stored peripheral blood leukocytes, bisulfite converted, and then measured for DNA methylation using the Illumina Infinium HumanMethylation450 BeadChip. DNA was extracted from stored blood samples collected during the Phase I examination. The Minfi R package[33] was used to preprocess, normalize (SWAN[28]), and calculate beta values. The proportions of each cell type were estimated using Houseman's method[23]. Detection p-values were calculated for each site, and beta values were set to missing if a site had detection p-value>0.01. In all samples, > 95% of probes had a detection P-value<0.01; thus, no samples were excluded from analysis. A total of 4,070 probes were removed due to detection P-value>0.01 in >1% of samples.

Genetics of Lipid Lowering Drugs and Diet Network (GOLDN)

The National Heart, Lung, and Blood Institute GOLDN study was designed to identify genetic determinants of lipid response to two interventions (a high-fat meal challenge and fenofibrate treatment for 3 weeks). Briefly, the study ascertained and recruited families from the Family Heart Study at two centers, Minneapolis, MN and Salt Lake City, UT, who self-reported to be white. The study protocol was approved by Institutional Review Boards at the University of Minnesota, University of Utah, and Tufts University/New England Medical Center. For the current study we evaluated fasting systolic and diastolic blood pressure among 991 participants for whom epigenetic data were available. Resting blood pressure was measured twice between 7:00 and 11:00 am in a sitting position. The average of the two measurements was used in this analysis.

DNA was extracted from CD4+ T-cells harvested from stored buffy coats using antibody-linked Invitrogen Dynabeads. We lysed cells captured on the beads and extracted DNA using DNeasy kits (Qiagen, Venlo, Netherlands). We used the Illumina Infinium Human Methylation450 Beadchip (Illumina Inc, San Diego, CA) to interrogate ~470,000 autosomal CpG sites across the genome.[10] A description of the array as well as CpG site nomenclature conventions can be found at http://www.illumina.com/products/methylation_450_beadchip_kits.ilmn. For each assay, 500ng of DNA was treated with sodium bisulfite (EZ DNA, Zymo Research, Irvine, CA) prior to standard Illumina amplification, hybridization, and imaging steps. The resulting intensity files were analyzed with Illumina's GenomeStudio which generated beta scores (ie, the proportion of total signal from the methylation specific probe or color channel) and "detection *P*-values" (defined as the 1-p-value computed from the background model characterizing the chance that the target sequence signal was distinguishable from the negative controls). Beta scores with an associated detection *P*-value greater than 0.01 were removed and samples with more than 1.5% missing data points were eliminated from further analysis. Furthermore, any CpG probes where more than 10% of samples failed to yield adequate intensity were removed. A total of 58 samples were removed. The filtered beta scores were then subjected to batch normalization with the ComBat package for R software in non-parametric mode (http://www.bu.edu/jlab/wp-assets/ComBat/Abstract.html). To correct for probe chemistry we separately normalized probes from the Infinium I and II chemistries and subsequently adjusted the β scores for Infinium II probes using the equation derived from fitting a second order polynomial to the observed methylation values across all pairs of probes located <50bp apart (within-chemistry correlations >0.99), where one probe was Infinium I and one was Infinium II. Principal components (PCs) based on the beta scores of all autosomal CpGs passing QC were generated using the *prcomp* function in R (V 2.12.1) and used to adjust for cell purity in association analysis.

Lothian Birth Cohort 1936 (LBC1936)

The Lothian Birth Cohort of 1936 is a follow-up study of the Scottish Mental Survey of 1947. The survey had tested the intelligence of almost every child born in 1936 and attending school in Scotland in the month of June 1947. DNA methylation and blood pressure concurrently measured at a mean age of 69.5 years was used for analyses. Blood pressure was estimated as the average of three sitting systolic and diastolic readings from an Omron 705IT monitor. DNA methylation was assessed using the Illumina Human Methylation 450k BeadChip from whole blood of consenting

participants (n=1,005). Background correction and quality control were performed to remove probes with a low detection rate, low quality based on manual inspection, and samples with a mismatch between genotypes and SNP control probes, and incorrectly predicted sex. Full details are provided in Shah et al.[34]

Multi-Ethnic Study of Atherosclerosis (MESA)

The Multi-Ethnic Study of Atherosclerosis (MESA) was designed to investigate the prevalence, correlates, and progression of subclinical cardiovascular disease in a population cohort of 6,814 participants. Since its inception in 2000, five clinic visits collected extensive clinical, socio-demographic, lifestyle, behavior, laboratory, nutrition, and medication data[35]. DNA methylation and gene expression were measured in purified (CD14+) monocyte samples from the April 2010 – February 2012 examination (exam 5) of 1,264 randomly selected MESA participants from four MESA field centers (Baltimore, MD; Forsyth County, NC; New York, NY; and St. Paul, MN) as previously described[36]. The study protocol was approved by the Institutional Review Board at each site. All participants signed informed consent.

As previously described[36], blood was initially collected in sodium heparin-containing Vacutainer CPT™ cell separation tubes (Becton Dickinson, Rutherford, NJ, USA) to separate peripheral blood mononuclear cells from other elements within 2 h from blood draw. Subsequently, monocytes were isolated with the anti-CD14-coated magnetic beads, using AutoMACs automated magnetic separation unit (Miltenyi Biotec, Bergisch Gladbach, Germany). Based on flow cytometry analysis of 18 specimens, monocyte samples were consistently >90% pure. DNA and RNA were isolated from samples simultaneously using the AllPrep DNA/RNA Mini Kit (Qiagen, Inc., Hilden, Germany). DNA and RNA QC metrics included optical density measurements, using a NanoDrop spectrophotometer and evaluation of the integrity of 18s and 28s ribosomal RNA. Illumina HumanMethylation450 BeadChips and HiScan reader were used to perform the epigenome-wide methylation analysis. Bead-level methylation data were summarized in GenomeStudio. Because a two-channel system and both Infinium I and II assays were used, normalization was performed in several steps using the lumi package. "Smooth quantile normalization" was used to adjust for color bias. Next, the data were background adjusted by subtracting the median intensity value of the negative control probes. Lastly, data were normalized across all samples by standard quantile normalization applied to the bead-type intensities and combined across Infinium I and II assays and both colors. QC measures included checks for sex and race/ethnicity mismatches, and outlier identification by multidimensional scaling plots. To estimate residual sample contamination for data analysis, we generated separate enrichment scores for neutrophils, B cells, T cells, monocytes, and natural killer cells. We implemented a Gene Set Enrichment Analysis[37] as previously described[36] to calculate the enrichment scores using the gene signature of each blood cell type from previously defined lists[38]. To remove technical error in methylation levels associated with batch effects across the multiple chips, positional effects of the sample on the chip, and residual sample contamination with non-monocyte cell types, we adjusted methylation values for chip, sample position on the chip, and estimated residual sample contamination with neutrophils, B cells, T cells, monocytes, and natural killer cells. The final methylation value for each methylation probe was computed as the beta-value, essentially the proportion of the methylated to the total intensity.

The Normative Aging Study (NAS)

The Normative Aging Study (NAS) is an ongoing longitudinal study of aging established in 1963 by the US Department of Veterans Affairs (VA). It enrolled community-dwelling men living in the Greater Boston area, 21-80 years old and free of known chronic medical conditions at entry[39]. Subsequently participants were invited to medical examinations every three to five years. DNA samples were collected from active participants between 1999-2013[39]. In total, 668 individuals (650 white, 12 black, 5 Hispanic white, 1 Hispanic black; 279 without medication) with cross-sectional data are included in this analysis. The study was reviewed and approved by the Institutional Review Boards (IRBs) of all participating institutions. All participants have provided written informed consent at each visit. At each in-person visit, participants completed questionnaires regarding demography, life-style and medical history. Chronological age and smoking status (never, former, current) were self-reported. High blood pressure was defined as antihypertensive medication use or SBP ≥140 mmHg or DBP ≥90 mmHg at study visit. Body Mass Index (BMI) was computed from anthropometric measures, performed with participants in undershorts and socks[40].

DNA was extracted from buffy coat using the QIAamp DNA Blood Kit (QIAGEN, Valencia, CA). 500 ng of DNA was used to perform bisulfite conversion using the EZ-96 DNA Methylation Kit (Zymo Research, Orange, CA). To reduce the chip and plate effects, we used a two-stage age-stratified algorithm to randomize samples and ensure similar age distributions across chips and plates; 12 samples – which were sampled across all age quartiles – were randomized to each chip, then chips were randomized to plates (each housing eight chips). Quality control analysis was performed to remove samples with a detection p-value > 0.05 in more than 1% probes, and probes with a detection p-value > 0.05 in more than 1% samples, respectively. The remaining samples were preprocessed using noob background correction[41] and dye bias adjustment[42]. Quantile normalization was then performed for methylated and unmethylated intensities of both type I and type II probes with nasen method[31], followed by BMIQ method[26] for probe types on beta methylation values. After preprocessing there were 484,613 CpG probes in the final working set. We adjusted for patient characteristics including age, BMI, smoking status and technical covariates for possible batch effects. Besides, due to the influence of blood cell count on methylation levels, we additionally adjusted each model for white blood cell (WBC) counts[23] and measured cell proportions. In addition, we used first four PCs calculated from the probes within 50bp of SNPs to account for potential population stratification[43].

Rotterdam Study (RS)

The Rotterdam Study is a prospective population based cohort study in a well-defined area of Rotterdam, the Netherlands. The design of the Rotterdam Study has been detailed elsewhere[44]. For the current analysis we used data from individuals aged 45 years and older that participated in the third cohort of the Rotterdam Study. In the first visit of

the third cohort, 3,934 participants were examined between February 2006 and December 2008. Whole blood DNA methylation was quantified in a random subset of 750 individuals with genotyping and RNA expression data available.

DNA was extracted from whole peripheral blood (stored in EDTA tubes) by standardized salting out methods. Genome-wide DNA-methylation levels in 750 subjects from the Rotterdam Study-III were determined using the Illumina HumanMethylation 450K beadarray (Illumina, Inc., San Diego, CA, USA). In short, samples (500ng of DNA per sample) were first bisulfite treated using the Zymo EZ-96 DNA-methylation kit (Zymo Research, Irvine, CA, USA). Next, they were hybridized to the arrays according to the manufacturer's protocol. During quality control samples showing incomplete bisulfite treatment were excluded (n=5) as were samples with a low detection rate (<99%) (n=7), or gender swaps (n=4). Probes with a detection p-value>0.01 in >1% samples, were filtered out. A total number of 474,528 probes passed the quality control and the filtered β values were normalized with DASEN implemented in the *wateRmelon* package in R statistical software[31].

The generation and management of the Illumina 450 K methylation array data (EWAS data) for the Rotterdam Study was executed by the Human Genotyping Facility of the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, the Netherlands. The EWAS data were funded by the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, and by the Netherlands Organization for Scientific Research (NWO; project number 184021007). The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam, Netherlands Organization for the Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam.

Saguenay Youth Study (SYS)

The Saguenay Youth Study (SYS) is a population-based study of adolescents and their middle-aged parents investigating the etiology, early stages and trans-generational trajectories of common cardiometabolic and brain diseases[45]. Half of the adolescents were exposed prenatally to maternal cigarette smoking. Sibships of adolescents (adolescents with one or more siblings aged 12-18 years) and their biological parents were recruited from Saguenay-Lac-St. Jean region of Quebec, Canada. All participants were of French-Canadian origin (i.e., both maternal and paternal grandparents of the adolescents were of French-Canadian ancestry and born in the region). Data collection was conducted in two waves. In Wave 1 (2003 – 2012) adolescents (n=1,029, aged 15.0 ±1.8 years) were recruited and underwent a 'complete assessment' (cardiovascular and other phenotyping, and genotyping) and parents (n=962, aged 43.3 ± 4.6 years) underwent a 'partial assessment' (questionnaires and genotyping). In Wave 2 (2012 – 2015), a subset of the parents (n=664, aged 49.2 ±5.0 years) underwent 'complete assessment'. The research Ethics Committee of the Chicoutimi Hospital and the Hospital for Sick Children in Toronto approved the study protocol. Written informed consent was obtained from all participants. In the present study, only data from the parents was considered[45]. A total of 105 parent participants with complete and quality controlled blood-pressure (BP) and DNA-methylation data have been studied here. The cardiovascular assessment involved a 52-minute protocol during which beat-by-beat BP was monitored, at rest and in response to physical and mental challenges, using a Finometer™ (FNS Finapres, Amsterdam, The Netherlands). Systolic and diastolic BP values analyzed were 1-minute averages obtained after 5 minutes at rest in a sitting position. Average time interval between BP and DNA-methylation measurements for the participants studied was 3.11 (1.07) years.

Genomic DNA was extracted from peripheral blood cells using the Human610-Quad and HumanOmniExpress BeadChips (Illumina, San Diego, CA, USA), bisulfite converted, and then epityped. Epityping was conducted using the Infinium HumanMethylation450K BeadChip (Illumina, San Diego, CA, USA) at the Montreal Genome Centre (Montreal, Quebec,

Canada). DNA-methylation data was available in 288 parents. A total of 6673 probes were removed due to detection P-value>0.01 in >1% of samples. One sample was removed based on pass rate <95%. Methylation values were normalized using the SWAN method[28]. Blood-cell proportions were imputed using the Houseman method[23].

TwinsUK

The TwinsUK cohort was established in 1992 as a collection of healthy research volunteers who were monozygotic and dizygotic same-sex twins. The participants are mostly adult female Caucasians and in total there are more than 13,000 participants across the UK[46]. TwinsUK blood methylation profiles were available from 690 female subjects who also had records of blood pressure and blood cell counts (eosinophils, lymphoctyes, monocytes, and neutrophils). These subjects are aged from 29 to 82 years old and were not affected with type 2 diabetes or cancer. The fasting and resting blood pressure have been measured three times in a sitting position, and the average of the latter two measurements was used in this analysis.

DNA was extracted from whole blood samples stored in EDTA tubes at 4 degree. Infinium HumanMethylation450 BeadChips (Illumina Inc, San Diego, CA) were used to measure DNA methylation. Details of the experimental approach have been previously described[47]. To correct for technical issues including two Illumina probe types, a normalization using the beta mixture quantile dilation (BMIQ) approach was applied[26]. The DNA methylation probes that mapped incorrectly or to multiple locations in the reference sequence were removed. Probes with detection P-value > 0.01 among > 10% of samples were also removed. All remaining probes were with non-missing values. After quality control, 452,785 probes were used for subsequent analyses. The epigenome-wide association models were adjusted for cell counts and technical covariates that included plate, position on the plate, family structure, and zygosity structure. A linear mixed effect regression model was applied as the data contained twin pairs, and all covariates apart from family and zygosity were taken as fixed-effect term.

Women's Health Initiative – Broad Agency Award 23 (WHI-BAA23)

Women were selected from one of two large Women's Health Initiative (WHI) sub cohorts that had previously undergone genome-wide genotyping as well as profiling for seven cardiovascular disease related biomarkers including total cholesterol, HDL, LDL, triglycerides, CRP, creatinine, insulin, and glucose through two core WHI ancillary studies. The first cohort is the WHI SNP Health Association Resource (SHARe) cohort of minorities that includes >8000 African American (AA) women and >3500 Hispanic women. These women were genotyped through WHI core study M5-SHARe (www.whi.org/researchers/data/WHIStudies/StudySites/M5) and underwent biomarker profile through WHI core study W54-SHARe (www.whi.org/researchers/data/WHIStudies/StudySites/W54). The second cohort consists of a combination of European Americans (EA) from the two Hormonal Therapy (HT) trials selected for GWAS and biomarkers in core studies W58 (www.whi.org/researchers/data/WHIStudies/StudySites/W58) and W63 (www.whi.org/researchers/data/WHIStudies/StudySites/W63). From these two cohorts, two sample sets were formed. The first (sample set 1) is a sample set of 637 CHD cases and 631 non-CHD cases as of Sept 30, 2010. The second sample set (sample set 2) is a non-overlapping sample of 432 cases of coronary heart disease and 472 non-cases as of September 17, 2012. All women with measures of inflammation that passed QC were included in this analysis. Written informed consent was obtained from all participants. All participants with complete and quality controlled blood pressure (BP) and DNA-methylation data have been studied here. Blood pressure was measured in all WHI participants

at baseline and annual clinic visits by certified staff in the WHI clinic using standardized procedures. The average of two baseline readings taken at the same clinic visit will be used for our analyses. Systolic and diastolic BP values analyzed were one-minute averages obtained after 5 minutes at rest in a sitting position.

DNA methylation analysis of blood was performed at HudsonAlpha Institute of Biotechnology using the Illumina Infinium Human- Methylation450 BeadChip. The Illumina BeadChips measures bisulfite-conversion-based, single-CpG resolution DNA methylation levels at 485,577 different CpG sites in the human genome. These data were generated by following the standard protocol of Illumina methylation assays, which quantifies methylation levels by the β value using the ratio of intensities between methylated and un-methylated alleles. Specifically, the β value is calculated from the intensity of the methylated (M corresponding to signal A) and un-methylated (U corresponding to signal B) alleles, as the ratio of fluorescent signals β = Max(M,0) / [Max(M,0) + Max(U,0) + 100]. Thus, β values range from 0 (completely un-methylated) to 1 (completely methylated) (Dunning, 2008).

Women's Health Initiative - Epigenetic Mechanisms of PM-Mediated CVD (WHI-EMPC)

The Women's Health Initiative – Epigenetic Mechanisms of Particulate Matter-Mediated Cardiovascular Disease (WHI-EMPC) is an ancillary study of epigenetic mechanisms underlying associations between ambient particulate matter air pollution and cardiovascular disease in the Women's Health Initiative clinical trials (CT) cohort, funded by the National Institute of Environmental Health Sciences (R01-ES020836)[48]. The WHI-EMPC study population is a stratified, random sample of 2,200 WHI CT participants who were examined between 1993 and 2001 and had available buffy coat, core analytes, electrocardiograms, and ambient concentrations of PM but were not taking anti-arrhythmic medications at the time. As such, WHI-EMPC is representative of the larger, multiethnic WHI CT population from which it was sampled: n=68,132 participants aged 50-79 years who were randomized to hormone therapy, calcium/vitamin D supplementation, and/or dietary modification in 40 U.S. clinical centers at the baseline exam (1993-1998) and re-examined in the fasting state one, three, six, and nine years later[49,50]. Current analyses involved information collected at the first visit with DNA methylation data available and were stratified by race/ethnicity (black, Hispanic/Latino, and white). On the day of DNA methylation sample collection, BP was measured in the seated position after a five-minute resting period with a conventional mercury sphygmomanometer and an appropriately sized cuff. The mean of two sequential recordings, taken at least 30 seconds apart, of systolic and diastolic BP was used for analyses.

Genome-wide DNA methylation at CpG sites was measured using the Illumina 450K Infinium Methylation BeadChip, quantitatively represented by beta (the percentage of methylated cytosines over the sum of methylated and unmethylated cytosines), and quality controlled using the following filters: detection p-values > 0.01 in > 10% of samples, detection p-values > 0.01 or missing in > 1% of probes, and probes with a coefficient of variation < 5%, yielding values of beta at 293,171 sites. DNAm data was normalized using BMIQ[26], then stage-adjusted using ComBat[51]. Epigenome-wide association models were adjusted for Houseman-estimated[23] cell subtype proportions (CD8-T, CD4-T, B cell, natural killer, monocyte, and granulocyte) and technical covariates including chip, row and column.

2. Gene Expression Measurements

Rotterdam Study Gene Expression

Whole-blood was collected (PAXGene Tubes – Becton Dickinson) and total RNA was isolated (PAXGene Blood RNA kits - Qiagen). To ensure a constant high quality of the RNA preparations, all RNA samples were analysed using the Labchip GX (Calliper) according to the manufacturer's instructions. Samples with an RNA Quality Score more than 7 were amplified and labelled (Ambion TotalPrep RNA), and hybridized to the Illumina HumanHT12v4 Expression Beadchips as described by the manufacturer's protocol. Processing of the Rotterdam Study RNA samples was performed at the Genetic Laboratory of Internal Medicine, Erasmus University Medical Centre Rotterdam. The RS-III expression dataset is available at GEO (Gene Expression Omnibus) public repository under the accession GSE33828. Illumina gene expression data was quantile-normalized to the median distribution and subsequently log2-transformed. The probe and sample means were centered to zero. Genes were declared significantly expressed when the detection p-values calculated by GenomeStudio were less than 0.05 in more than 10% of all discovery samples, which added to a total number of 21,238 probes[52]. Quality control was done using the eQTL-mapping pipeline. We only analyzed probes that uniquely mapped to the human genome build 37[53].

Framingham Heart Study Gene Expression

Affymetrix Human Exon Array ST 1.0 (Affymetrix, Inc., Santa Clara, CA) was utilized to measure mRNA expression levels genome wide (N=~18,000 transcripts). Details of the design, sampling, RNA isolation, and mRNA measurement have been described previously[54].

Supplementary References

1. Liu, C. *et al.* A DNA methylation biomarker of alcohol consumption. *Mol. Psychiatry* (2016). doi:10.1038/mp.2016.192

2. Wahl, S. *et al.* Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. *Nature* **541,** 81–86 (2017).

3. Dekkers, K. F. *et al.* Blood lipids influence DNA methylation in circulating cells. *Genome Biol.* **17,** 138 (2016).

4. Petersen, A. K. *et al.* Epigenetics meets metabolomics: An epigenome-wide association study with blood serum metabolic traits. *Hum. Mol. Genet.* **23,** 534–545 (2014).

5. Aslibekyan, S. *et al.* Epigenome-wide study identifies novel methylation loci associated with body mass index and waist circumference. *Obesity (Silver Spring).* **23,** 1493–501 (2015).

6. Soriano-Tárraga, C. *et al.* Epigenome-wide association study identifies TXNIP gene associated with type 2 diabetes mellitus and sustained hyperglycemia. *Hum. Mol. Genet.* **25,** 1–11 (2015).

7. Chambers, J. C. *et al.* Epigenome-wide association of DNA methylation markers in peripheral blood from Indian Asians and Europeans with incident type 2 diabetes: A nested case-control study. *Lancet Diabetes Endocrinol.* **3,** 526–534 (2015).

8. Al Muftah, W. A. *et al.* Epigenetic associations of type 2 diabetes and BMI in an Arab population. *Clin. Epigenetics* **8,** 13 (2016).

9. Florath, I. *et al.* Type 2 diabetes and leucocyte DNA methylation: an epigenome-wide association study in over 1,500 older adults. *Diabetologia* **59,** 130–138 (2016).

10. Liang, L. *et al.* An epigenome-wide association study of total serum immunoglobulin E concentration. *Nature* **520,** 670–674 (2015).

11. Irvin, M. R. *et al.* Epigenome-wide association study of fasting blood lipids in the genetics of lipid-lowering drugs and diet network study. *Circulation* **130,** 565–572 (2014).

12. Demerath, E. W. *et al.* Epigenome-wide association study (EWAS) of BMI, BMI change and waist circumference in African American adults identifies multiple replicated loci. *Hum. Mol. Genet.* **24,** 4464–79 (2015).

13. Das, M. *et al.* Association of DNA Methylation at CPT1A Locus with Metabolic Syndrome in the Genetics of Lipid Lowering Drugs and Diet Network (GOLDN) Study. *PLoS One* **11,** e0145789 (2016).

14. Mamtani, M. *et al.* Genome- and epigenome-wide association study of hypertriglyceridemic waist in Mexican American families. *Clin. Epigenetics* **8,** 6 (2016).

15. Aslibekyan, S. *et al.* CPT1A methylation is associated with plasma adiponectin. *Nutr. Metab. Cardiovasc. Dis.* **27,** 225–233 (2017).

16. Mitchell, B. D. *et al.* The genetic response to short-term interventions affecting cardiovascular function: rationale and design of the Heredity and Phenotype Intervention (HAPI) Heart Study. *Am. Heart J.* **155,** 823–8 (2008).

17. Shuldiner, A. R. *et al.* Association of cytochrome P450 2C19 genotype with the antiplatelet effect and clinical efficacy of clopidogrel therapy. *JAMA* **302,** 849–57 (2009).

18. Hsueh, W. C. *et al.* Diabetes in the Old Order Amish: characterization and heritability analysis of the Amish Family Diabetes Study. *Diabetes Care* **23,** 595–601 (2000).

19.   Bibikova, M. *et al.* High density DNA methylation array with single CpG site resolution. *Genomics* **98,** 288–95 (2011).

20.   Dedeurwaerder, S. *et al.* A comprehensive overview of Infinium HumanMethylation450 data processing. *Brief. Bioinform.* **15,** 929–41 (2014).

21.   Chen, Y. *et al.* Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics* **8,** 203–9 (2013).

22.   Moen, E. L. *et al.* Genome-wide variation of cytosine modifications between European and African populations and the implications for complex traits. *Genetics* **194,** 987–96 (2013).

23.   Houseman, E. A. *et al.* DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics* **13,** 86 (2012).

24.   Jaffe, A. E. & Irizarry, R. A. Accounting for cellular heterogeneity is critical in epigenome-wide association studies. *Genome Biol.* **15,** R31 (2014).

25.   The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. *Am. J. Epidemiol.* **129,** 687–702 (1989).

26.   Teschendorff, A. E. *et al.* A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. *Bioinformatics* **29,** 189–96 (2013).

27.   Horvath, S. DNA methylation age of human tissues and cell types. *Genome Biol.* **14,** R115 (2013).

28.   Maksimovic, J., Gordon, L. & Oshlack, A. SWAN: Subset-quantile within array normalization for illumina infinium HumanMethylation450 BeadChips. *Genome Biol.* **13,** R44 (2012).

29.   Feinleib, M., Kannel, W. B., Garrison, R. J., McNamara, P. M. & Castelli, W. P. The framingham offspring study. Design and preliminary data. *Prev. Med. (Baltim).* **4,** 518–525 (1975).

30.   Splansky, G. L. *et al.* The Third Generation Cohort of the National Heart, Lung, and Blood Institute's Framingham Heart Study: Design, Recruitment, and Initial Examination. *Am. J. Epidemiol.* **165,** 1328–1335 (2007).

31.   Pidsley, R. *et al.* A data-driven approach to preprocessing Illumina 450K methylation array data. *BMC Genomics* **14,** 293 (2013).

32.   Daniels, P. R. *et al.* Familial aggregation of hypertension treatment and control in the Genetic Epidemiology Network of Arteriopathy (GENOA) study. *Am. J. Med.* **116,** 676–81 (2004).

33.   Aryee, M. J. *et al.* Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* **30,** 1363–9 (2014).

34.   Shah, S. *et al.* Genetic and environmental exposures constrain epigenetic drift over the human life course. *Genome Res.* **24,** 1725–33 (2014).

35.   Bild, D. E. *et al.* Multi-Ethnic Study of Atherosclerosis: objectives and design. *Am. J. Epidemiol.* **156,** 871–81 (2002).

36.   Liu, Y. *et al.* Methylomics of gene expression in human monocytes. *Hum. Mol. Genet.* **22,** 5065–74 (2013).

37.   Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. …* **102,** 15545–15550 (2005).

38.   Abbas, A. R. *et al.* Immune response in silico (IRIS): immune-specific genes identified from a compendium of

microarray expression data. *Genes Immun.* **6,** 319–31 (2005).

39.   Bell, B., Rose, C. L. & Damon, A. The Veterans Administration longitudinal study of healthy aging. *Gerontologist* **6,** 179–84 (1966).

40.   Troisi, R. J., Heinold, J. W., Vokonas, P. S. & Weiss, S. T. Cigarette smoking, dietary intake, and physical activity: effects on body fat distribution--the Normative Aging Study. *Am. J. Clin. Nutr.* **53,** 1104–11 (1991).

41.   Triche, T. J., Weisenberger, D. J., Van Den Berg, D., Laird, P. W. & Siegmund, K. D. Low-level processing of Illumina Infinium DNA Methylation BeadArrays. *Nucleic Acids Res.* **41,** e90 (2013).

42.   Davis, S., Du, P., Bilke, S., Triche, T. J. & Bootwalla, M. methylumi: Handle Illumina methylation data. *R Packag. version 2.14.0* (2015).

43.   Barfield, R. T. *et al.* Accounting for population stratification in DNA methylation studies. *Genet. Epidemiol.* **38,** 231–41 (2014).

44.   Hofman, A. *et al.* The Rotterdam Study: 2016 objectives and design update. *Eur. J. Epidemiol.* **30,** 661–708 (2015).

45.   Pausova, Z. *et al.* Cohort Profile: The Saguenay Youth Study (SYS). *Int. J. Epidemiol.* (2016). doi:10.1093/ije/dyw023

46.   Moayyeri, A., Hammond, C. J., Valdes, A. M. & Spector, T. D. Cohort Profile: TwinsUK and healthy ageing twin study. *Int. J. Epidemiol.* **42,** 76–85 (2013).

47.   Tsaprouni, L. G. *et al.* Cigarette smoking reduces DNA methylation levels at multiple genomic loci but the effect is partially reversible upon cessation. *Epigenetics* **9,** 1382–96 (2014).

48.   U.S. Department of Health and Human Services. National Institutes of Health. Epigenetic Mechanisms of PM-Mediated CVD. *Research Portfolio Online Reporting Tools* Available at: https://projectreporter.nih.gov/project_info_description.cfm?aid=8369149&icde=27485446.

49.   Design of the Women's Health Initiative clinical trial and observational study. The Women's Health Initiative Study Group. *Control. Clin. Trials* **19,** 61–109 (1998).

50.   Anderson, G. L. *et al.* Implementation of the Women's Health Initiative study design. *Ann. Epidemiol.* **13,** S5-17 (2003).

51.   Johnson, W. E., Li, C. & Rabinovic, A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **8,** 118–127 (2007).

52.   Westra, H.-J. *et al.* Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat. Genet.* **45,** 1238–43 (2013).

53.   Schurmann, C. *et al.* Analyzing illumina gene expression microarray data from different tissues: methodological aspects of data analysis in the metaxpress consortium. *PLoS One* **7,** e50938 (2012).

54.   Joehanes, R. *et al.* Gene Expression Signatures of Coronary Heart Disease. *Arterioscler. Thromb. Vasc. Biol.* **33,** 1418–1426 (2013).