SUPPLEMENTAL MATERIAL

Table of contents

1. Methods

1.1 Study Population

The present study included 6,662 European ancestry (EA), 2,702 African ancestry (AA), and 360 Hispanic ancestry (HA) participants from five cohorts of the Cohorts for Heart and Aging Research in Genetic Epidemiology Consortium (CHARGE), including ARIC, FHS, GOLDN, MESA, and RS. Details of participating cohorts and exclusion criteria are described in Descriptions of participating cohorts. Participants' characteristics are presented in Supplemental Material Table 1. The protocol was approved by all participating institutions' Institutional Review Board. All participants provided written informed consent.

1.2. Diet Scores

All participating cohorts used food frequency questionnaires (FFQs) to assess dietary intake. Each FFQ has been validated in previous studies.[1-8] FFQs used in each cohort, methods of quality control, and diet score development are detailed herein and in Supplemental Material Table 1. We calculated two complementary diet scores (Box 1), Mediterranean-style diet score (MDS) and the Alternative Healthy Eating Index (AHEI), based on prior studies.[9-11]

Box 1.

$MDS = \text{Vegetable}_{QR} + \text{Fruit}_{QR} + \text{Nut}_{QR} + \text{Legume}_{QR} + \text{Whole-grain}_{QR} + \text{Fish}_{QR} + \text{Red-meat}_{inv\_QR} + \text{MSR}_{QR} + \text{Alcohol}$

$AHEI = \text{Vegetable}_{10} + \text{Fruit}_{10} + \text{Nuts \& Legume}_{10} + \text{Whole-grain}_{10} + \text{SSB \& Fruit Juice}_{inv\_10} + \text{Red-meat}_{inv\_10} + \text{EPA \& DHA}_{10} + \text{PUFA}_{10} + \textit{trans}\text{-Fat}_{inv\_10} + \text{Sodium}_{inv\_10} + \text{Alcohol}$

QR: sex-specific quartile rank
inv_QR: inversed sex-specific quartile rank
For alcohol in MDS: a score of 1 if consumption ≥10 grams/day and ≤25 grams/day for men or ≥5 grams/day and ≤15 grams/day for women
10: A score of 0 to 10 using thresholds that largely based on dietary guidelines[11]
inv_10: an inversed score of 0 to 10 using thresholds that largely based on dietary guidelines[11]

The MDS consists of nine components: vegetables, fruits, nuts, legumes, whole grains, fish, red meat, ratio of monounsaturated fatty acids (MUFA) to saturated fatty acids (SFA), and alcohol.[9,10] We categorized consumption of each food component into sex-specific quartiles for each participating study cohort. We assigned scores of 0, 1, 2, and 3 to all components except red meat and alcohol, from the

Supplemental Matrial Table 1. Participant characeristics and methods for measurements of DNA methylation and overall diet quality

Cohort	N	Age (years)	Women (%)	Race/Ethnicity	DNA Methylation tissue	DNA Methylation array	Dietary assessment	Family Cohort	White blood cell counts	BMI (kg/m^2)	Current Smoking status	Physical activity score	Energy intake (kcal/day)	Alcohol (g/day)
ARIC	3,431	54 ± 6	63%	EA, AA	Whole blood	Illumina 450K	66 item, interviewer administered semi-quantitative modified Willett FFQ	No	Determined by automated particle Coulter Counters	28.6 ± 5.9	27%	2.3 ± 0.7	1572 ± 584	5 ± 13
FHS	3,359	58 ± 13	55%	EA, AA	Whole blood	Illumina 450K	126 item self-administered semi-quantitative Willett FFQ	Yes	Estimated using Houseman method	27.9 ± 5.5	8%	36 ± 6	1913 ± 635	9 ± 12
GOLDN	1,082	18 to 87	51.83%	EA	CD4+ T-cells	Illumina 450K	124 item, self-administered, (DHQ)	Yes	Cell sorted	16.56 to 52.66	8%	0 to 17 /day	471 to 5548	0 to 34
MESA	1,250 (971 with all covars)	70 ± 9	51.00%	AA, EA, HA	CD14+ Monocytes	Illumina 450K	120 item, self-administered, modified Block-FFQ	No	Purified monocytes, % residual cell types estimated from gene expression signatures	29.8 ± 5.5	8.9%	7.5 ± 4.6	1750 ± 883	6 ± 12
RSIII-1	440	58.3 ± 7.6	53.00%					No	Estimated using Houseman method	27.2 ± 4.8	25%	median MET hours per week: 37.7 95% range [2.3-201.4]	2380 ± 826	median: 6.8 95% range [0-35.4]
RSII-3	362	70.8 ± 3.2	57%					No	measured cell counts (monocytes, granulocytes, and lymphocytes)	27.5 ± 4.2	9%	83.9 [16-182]	1974 ± 635.7	6.2 [0-31.3]
RSIII-2	149	60.7 ± 4.5	67.80%	EA	Whole blood	Illumina 450K	389 item, self-administered FFQ for Dutch adults	No	measured cell counts (monocytes, granulocytes, and lymphocytes)	27.3 ± 4.0	9%	39.5 [1.5-250.5]	2257 ± 679.7	8.3 [0-29.1]

ARIC: Atherosclerosis Risk in Communities Study; FHS: The Framingham Heart Study; GOLDN: Genetics of Lipid Lowering Drugs and Diet Network; MESA: The Multi-Ethnic Study of Atherosclerosis; RS: The Rotterdam Study. AA: African ancestry; EA: European ancestry; HA: Hispanic ancestry. FFQ: food frequency questionnaire; DHQ: diet history questionnaire

lowest to highest quartiles. We reversed the order of the scores for red meat. For alcohol, we assigned a score of 1 if consumption was ≥ 10 grams/day and ≤ 25 grams/day for men or ≥ 5 grams/day and ≤ 15 grams/day for women and 0 for all other values. All the scores were summed which created the MDS, which ranged from 0 to 25, with higher scores indicating a healthier dietary pattern. The AHEI included 11 components including vegetables, fruits, nuts and legumes, sugar-containing beverages, whole grains, red meat, eicosapentaenoic acid (EPA) and docosahexaenoic acid (DHA), polyunsaturated fatty acids (PUFA without EPA and DHA), trans-fatty acids, sodium, and alcohol.[11] A score of 0 to 10 was assigned for each component, with a score of 10 indicating that the participant had met dietary recommendations based on the reported relations between each diet component and chronic diseases. The AHEI score ranges from possible values of 0 to 110. A higher AHEI score largely reflects a better adherence to the contemporaneous dietary guideline. Dietary recommendations for the 11 components of the AHEI were similar in the 2010 Dietary Guidelines for Americans compared to the 2015 Dutch food-based dietary guideline. We therefore used the same cut-off values for all AHEI components in our participating cohorts, including four US cohorts and the Dutch cohort.

1.3. DNA methylation assessment

Methylation profiles were measured using the Illumina 450K BeadChip using DNA derived from all leukocytes, CD4+ T-cells, or CD14+ monocytes in peripheral blood (Supplemental Material Table 1). DNA preparation, bisulfite conversion, methylation profiling, quality control procedures, and raw data normalization for participating cohorts are detailed in Descriptions of participating cohorts. We analyzed DNA methylation signal β values, i.e., methylated signals divided by the sum of methylated and unmethylated signals. We excluded probes on the X and Y chromosomes to avoid sex bias.

1.4. Statistical Analyses

1.4.1. Epigenome-wide association analysis. We first conducted a two-step analysis in EA participants, i.e., discovery in the FHS (n=3,266) and replication in the meta-analysis of the ARIC, GOLDN, MESA,

and RS (n=3,396). In the discovery analysis, we considered CpGs with false discovery rate (FDR) < 0.05 significant and forwarded these CpGs for replication analysis. We conducted meta-analysis to aggregate results generated by each replication cohort. We used random-effects models to account for heterogeneity between studies. In meta-analysis, each study was given a weight in proportion to its precision by using the inverse of the variance (i.e., inverse variance weighted random-effects models). We considered those with p-value < 0.05/number of CpGs tested (i.e., Bonferroni correction) replicated CpGs. We used linear regression models with adjustment for sex, age, and energy intake in the discovery and replication analyses. Linear mixed models were used when adjustment for kinship was needed. Models were also adjusted for estimated leukocyte composition and technical variables in a cohort-specific manner (See Descriptions of participating cohorts).

We also conducted a one-step analysis, i.e., meta-analyzed all EA cohorts using random effects models with internal validation. We reported CpGs with FDR < 0.05 in the meta-analysis and had p-value < 0.05 with concordant direction of effect in the discovery and replication cohorts (i.e., internally validated). We adjusted for the same covariates in the two-step analysis. In both analyses, to select independent CpGs, we excluded those close to (within ± 500kb) and correlated with (Pearson $|r| > 0.7$) a CpG with a lower p-value. We also examined if the association remained significant with additional adjustment for smoking status, physical activity, and BMI.

We conducted similar analyses in AA and HA participants and performed trans-ethnic meta-analysis in all participants (including EA, AA, and HA participants). Because of the limited sample size of AA and HA participants, we considered significant CpGs (at FDR < 0.05) in the trans-ethnic analysis as secondary findings.

1.4.2. Association with CVD risk factors. We evaluated the association of diet score-associated CpGs with CVD risk factors reported in the MRC-IEU EWAS catalog [http://www.ewascatalog.org].

1.4.3. Mendelian Randomization (MR) analysis. We conducted bidirectional MR analysis to infer if there were causal relations between diet-associated CpGs and cardiometabolic risk factors including BMI, waist-to-hip ratio adjusted for BMI (WHRadjBMI), triglyceride and high-density lipoprotein cholesterol (HDL) concentrations, and type 2 diabetes. We performed the two-sample MR analysis using R package *TwoSampleMR*.[12] To test if differential DNA methylation may trigger cardiometabolic risk factor changes (i.e., CpG → trait), we used independent *cis*-meQTL variants (pair-wise linkage disequilibrium (LD) R-squared < 0.1, minor allele frequency > 0.01, imputation R-squared > 0.5, p-value for association with CpGs < 5×10^{-8}) as instrumental variables (IVs). The *cis*-meQTL variants were identified in 4,170 FHS participants (See section for Methylation Quantitative Trait Loci).[13] The number of IVs utilized in these analyses varied for different CpGs. We obtained effect size and corresponding standard error of these IVs in relation to cardiometabolic risk factors from the published genome-wide association studies (GWAS) with largest sample size for BMI,[14] WHRadjBMI,[15] triglycerides and HDL concentrations,[16] and type 2 diabetes.[17] We also tested if the observed differential DNA methylation was caused by cardiometabolic risk factor (i.e., trait → CpG). To test this hypothesis, we used single nucleotide polymorphisms (SNPs) significantly associated with cardiometabolic risk factors identified by the above mentioned GWAS.[14-17] We excluded SNPs with low imputation quality (imputation R-squared ≤ 0.5), minor allele frequency ≤ 0.01 in the FHS, and LD R-squared ≥ 0.1 based on the 1000 Genomes Project CEU reference population.[18] The number of SNPs utilized in the final MR analyses included 801 SNPs for BMI, 43 SNPs for WHR, 19 SNPs for triglycerides, 53 SNPs for HDL, and 294 SNPs for type 2 diabetes. The effect size and corresponding standard error of these SNPs with CpGs were obtained in the FHS.[13] We performed the MR analysis primarily using the inverse variance weighted (IVW) method. To test if the assumption of no horizontal pleiotropy effect, we evaluated the p-value for the intercept generated using the MR-Egger method.

1.4.4. Association with all-cause mortality. In parallel with the present study, another project of the CHARGE Epigenetic Working Group analyzed the epigenome-wide association between DNA methylation and all-cause mortality using Cox proportional hazard models. The mortality study meta-analyzed up to 10,083 samples from 10 EA cohorts, including 2,536 death events documented with median follow-up 6.4 to 20 years (Supplemental Material Table 2). We reported their meta-analysis results generated using random-effects models for diet score-associated CpGs identified in the present study.

1.4.5. Functional and regulatory annotation. We performed hypergeometric tests using the phyper function in R with default values to examine the genomic characteristics of the diet score-associated CpGs. We calculated the Spearman ranked correlation coefficients for mean DNA methylation levels of these CpGs measured in peripheral blood samples with those measured in other tissues using a dataset in the Gene Expression Omnibus (GEO Series accession number GSE48472).[19] We queried *cis*-meQTL variants in the NHGRI-EBI Catalogue of Published GWAS to test if these variants are associated with CVD and risk factors.[20] We used the platform of the Functional Mapping and Annotation of Genome-Wide Association Studies[21] to examine expression of genes annotated to diet-associated CpGs in 30 general tissues included in the Genotype-Tissue Expression (GTEx v6) database22 and to perform gene-set enrichment analysis using default values. Bonferroni corrected p-values were reported for all enrichment analyses.

2. Descriptions of participating cohorts

2.1. Framingham Heart Study (FHS)

2.1.1. Cohort and Study Sample. The Framingham Heart Study (FHS) is a population-based, prospective study.[23,24] For the present study, we included a sub-cohort of participants who attended the FHS Offspring examination cycle 8 (2005-2008; n=3,021) or the FHS Third Generation cohort examination cycle 2 (2008-2011; n=3,411). Participants were included if they had had a history of myocardial infarction or

Supplemental Material Table 2. Cohorts for analysis of diet-associated CpGs and all-cause mortality in EA participants

Cohort	N	N.Death	Age	Women, %	Median follow-up years	DNA methylation tissue
Atherosclerosis Risk in Communities (ARIC) Study	969	331	59.8 (5.5)	59	20.0 (5.2)	Whole blood
Cardiovascular Health Study (CHS)	419	373	75.0 (4.9)	60	12.7 (6.1)	Whole blood
ESTHER	1000	265	62.1 (6.5)	50	13.7 (3.5)	Whole blood
Framingham Heart Study (FHS)	2427	403	66.3 (9.0)	55	9.1 (2.2)	Whole blood
InChianti Study	488	104	10.0 (1.6)	62	10.0 (1.6)	Whole blood
KORA F4 Study	1727	89	61.0 (8.9)	51	6.4 (0.9)	Whole blood
Lothian Birth Cohort (LBC) Study - 1921	418	366	79.1 (0.6)	60	9.8 (4.7)	Whole blood
Lothian Birth Cohort (LBC) Study - 1936	900	192	69.6 (0.8)	50	10.2 (2.4)	Whole blood
Normative Aging Study (NAS)	640	221	72.8 (6.8)	0	10.5 (3.3)	Whole blood
Women's Health Initiative (WHI)*	1095	192	62 (6.9)	100	11.5 (3.5)	Whole blood

EA: European ancestry

* European ancestry participants in The Epigenetic Mechanisms of PM-Mediated CVD Risk (WHI-EMPC)

stroke, cancer (other than basal cell carcinoma), or bariatric surgery, as well as had no data on DNA methylation, diet, and covariates. After exclusion, data collected from 3,266 FHS participants were analyzed. All participants provided written informed consent and the Framingham Heart Study protocols and procedures were approved by the Institutional Review Board for Human Research at Boston University Medical Center.

2.1.2. Genome-wide DNA methylation profiling. Buffy coat fractions were obtained from peripheral whole blood samples and extracted genomic DNA using the Gentra Puregene DNA extraction kit (Qiagen, Venlo, Netherlands). We used EZ DNA Methylation Kit (Zymo Research, Irvine, CA) to perform bisulfite conversion to the extracted genomic DNA. After whole genome amplification, fragmentation, array hybridization, and single-base pair extension, we then quantified DNA methylation in three laboratories using the Illumina Infinium HumanMethylation450 (450K) BeadChip. We used the wateRmelon R package with the DASEN methodology to normalize laboratory-specific DNA methylation.[25] DNA methylation beta values were calculated as the ratio of methylated probe intensity to the overall intensity. We implemented rigorous quality control that has been described elsewhere.[26] Briefly, we excluded study participants with a probe missing rate >1%, poor SNP matching to the 65 SNP control probe locations, and outliers by multi-dimensional scaling techniques. At the probe level, we also excluded CpGs that: (1) have a missing rate of >20% at $p<0.01$, (2) have been previously identified to map to multiple locations, or (3) have an underlying SNP (minor allele frequency >5% in European ancestry (EUR) 1,000 genomes project data) at the CpG site or within 10 bp of the single base extension. After quality control, about 430,000 autosomal CpG sites were analyzed subsequently. The FHS methylation data are available at dbGaP under the accession number phs000724.v2.p9.

2.1.3. Diet assessment. A previously validated, self-administered semi-quantitative 126-item Harvard food frequency questionnaire (FFQ) was used to determine dietary intake for the previous year leading up to an examination.[27] The FFQ was mailed to participants to be completed at home and returned during the

study appointment. FFQ data were excluded if the reported energy intake was <2.5 MJ/d (600 kcal/d) for both men and women, ≥16.7 MJ/d (4000 kcal/d) for women, ≥17.5 MJ/d (4200 kcal/d) for men, and if ≥13 food items were left blank.

2.1.4. Covariates assessment. All covariates were assessed when participants visited the research clinic in accordance with standard protocols. Body mass index (BMI) was calculated as weight divided by height squared (kg/m^2). Current smokers were defined as participants who self-reported smoking at least one cigarette per day in the prior year. Energy intake was calculated using the Harvard FFQ. A physical activity score was generated using the intensity and time spent performing each type of activity assessed by the physical activity questionnaire.[28] White blood cell counts (B cells, granulocytes, monocytes, NK cells, CD4+ T-cells, CD8+ T-cells) was estimated using the Houseman method.[29]

2.1.5. Epigenome-wide association analysis. We first applied the sva R package to construct surrogate variables in each laboratory. To remove potential batch effects, we calculated the residuals of DNA methylation beta value with adjustment for surrogate variables that associated with diet scores at p-value <0.1. In the subsequent analysis, we used the sum of laboratory-specific residuals of DNA methylation beta values and median of DNA methylation beta values as the dependent variables. Diet scores were analyzed continuously as the independent variable. Three models were implemented. In the primary model, covariates included age, sex, energy intake, laboratories, and white blood cell counts. We conducted two additional models in sensitivity analysis. The first sensitivity model additionally adjusted for current smoking status and physical activity level. The second sensitivity model additionally adjusted for BMI. The relatedness in our study sample (kinship) was accounted for using linear mixed models.

2.1.6. *Methylation Quantitative Trait Loci (meQTLs).* We analyzed meQTL variants in 4,170 FHS participants. We genotyped our participants using Affymetrix 550K Array and imputed single nucleotide polymorphisms (SNPs) with the 1,000 Genomes Project reference panel.[18] We first calculated the

residuals for DNA methylation using linear mixed regression models with adjustment for sex, age, estimated leukocyte counts, technical covariates, and kinship and then regressed the residuals on SNPs. We defined cis-meQTL variants as SNPs associated with DNA methylation levels of nearby CpGs (±500 kilobases) with minor allele frequency > 0.01 and imputation R-squared > 0.5. For the purpose of the present study, we only analyzed cis-meQTL variants with p-value < 5×10^{-8}.

2.1.7. Acknowledgements. The FHS is funded by National Institutes of Health contract N01-HC-25195. The laboratory work for this investigation was funded by the Division of Intramural Research, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD. The analytical component of this project was funded by the Division of Intramural Research, National Heart, Lung, and Blood Institute, and the Center for Information Technology, National Institutes of Health, Bethesda, MD.

2.2. Atherosclerosis Risk in Communities (ARIC) Study

2.2.1. Cohort and Study Sample. The ARIC Study is a community-based cohort of 15,792 predominantly black and white men and women aged 45-64 years that was designed to investigate cardiovascular disease and its risk factors (http://www.cscc.unc.edu/aric/). The participants were recruited and enrolled between 1987 and 1989 from 4 U.S. communities: Forsyth County, NC; Jackson, MS; suburbs of Minneapolis, MN; and Washington County, MD. An ethics committee at each site approved the study protocol and study participants provided informed consent. Participants were excluded if they had a history of myocardial infarction (n=592) or had a history of cancer (n=814). After excluding participants without whole blood derived DNA methylation profiling and quality control for FFQ (see below), a total of 3,431 participants (966 Caucasian participants and 2,465 African-American participants) were analyzed.

2.2.2. Genome-wide DNA methylation profiling. Genomic DNA was extracted from peripheral blood leukocyte samples using the Gentra Puregene Blood Kit (Qiagen; Valencia, CA, USA) according to the manufacturer's instructions (www.qiagen.com). Bisulfite conversion of 1 µg genomic DNA was

performed using the EZ-96 DNA Methylation Kit (Deep Well Format) (Zymo Research; Irvine, CA, USA) according to the manufacturer's instructions (www.zymoresearch.com). Bisulfite conversion efficiency was determined by PCR amplification of the converted DNA before proceeding with methylation analyses on the Illumina platform using Zymo Research's Universal Methylated Human DNA Standard and Control Primers.

DNA was isolated from white blood cells as per the standard DNA extraction procedure (Autopure LS, Qiagen). 500 ng of extracted DNA was bisulfite-modified using the EZ DNA Methylation kit (Zymo Research, D5004) following the manufacturer's instructions. The Illumina Infinium HumanMethylation450 Beadchip array (HM450) was used to measure DNA methylation (Illumina Inc.; San Diego, CA, USA). Raw methylation data were extracted using Illumina GenomeStudio software (version 2011.1, Methylation module 1.9.0). The methylation score for each CpG was represented as a beta value according to the fluorescent intensity ratio of methylated to overall signals. Background subtraction was conducted with the GenomeStudio software using built-in negative control bead types on the array.

Subset quantile within array normalization (SWAN) was used to normalize DNA methylation beta-values.[30] Further, in this study, we conducted all analyses at the single probe level, and therefore, any differences in probe type should not strongly influence the results.

2.2.3. Diet assessment. Diet was assessed using a 66-item semi-quantitative food frequency questionnaire (FFQ), modified from the Willett questionnaire, and was administered in person by a trained interviewer at visit 1 (1987-1989).[2,7,31] Participants were excluded who reported extreme total energy intakes (females: <500 or >3,500 kcal/day, males: <700 or >4,500 kcal/day) or were missing responses from more than 15 FFQ items.

2.2.4. Covariate assessment. Covariates were assessed at baseline (visit 1). Body mass index (BMI) was calculated as weight divided by height squared (kg/m^2). Smoking status was categorized as current, former, never, or missing. Physical activity index score ranged from 1 (lowest) to 5 (highest) and was calculated based on intensity and time dedicated to sport and non-sport exercise during leisure time. Heavy alcohol status was defined as at least 2 drinks/day for females and at least 3 drinks/day for males. White blood cell count was determined by automated particle Coulter Counters within 24 hours after venipuncture in local hospital hematology laboratories.

2.2.5. Epigenome-wide association analysis. Three models were implemented as described in the section for the FHS. The association analysis used the diet scores as the independent variables and DNA methylation proportion as outcome, adjusting for age, gender, energy intake, total white blood cell count, and estimated white blood cell proportions. In sensitivity analyses, we additionally adjusted for lifestyle factors (smoking status and physical activity) and BMI.

2.2.6. Acknowledgements. The ARIC Study has been funded in whole or in part with federal funds from the NHLBI, NIH, Department of Health and Human Services (HHSN268201700001I, HHSN2682017000021, HHSN268201700003I, HHSN268201700004I, and HHSN268201700005I). The authors thank the staff and participants of the ARIC Study for their important contributions. Funding was also provided by 5RC2HL102419, R01NS087541, and R01HL131136. Dr. Rebholz is supported by a Mentored Research Scientist Development Award from the National Institute of Diabetes and Digestive and Kidney Diseases (K01 DK107782). Ms. Hu is supported by a training grant from the National Institutes of Health (NIH)/National Heart, Lung, and Blood Institute (NHLBI; T32 HL007024).

2.3. Genetics of Lipid Lowering Drugs and Diet Network (GOLDN)

2.3.1. Cohort and Study Sample. The GOLDN population contains European American families who had at least 2 siblings.[32] These families were recruited from two sites of the National Heart, Lung and Blood

Institute Family Heart Study: Salt Lake City and Minneapolis. Appropriate Institutional Review Boards from all participating institutions approved the study protocol. All participants provided informed consent. Before each study visit, participants abstained from eating for 8 hours, from consuming alcohol for 24 hours, and from using lipid lowering medications or supplements for 4 weeks. Approximately ~1,350 participants were screened. After exclusions based on eligibility criteria [33] as well as availability of methylation data, 994 participants remained in this analysis.

2.3.2. Genome-wide DNA methylation profiling. DNA methylation of GOLDN was measured using the Illumina Infinium HumanMethylation450 Beadchip (Illumina, San Diego, CA) as discussed in.[34,35] Cell type was restricted to CD4+ T-cells that were isolated from frozen buffy coat samples. DNA was isolated using DNeasy kits (Qiagen, Venlo, Netherlands). Genome Studio software (Illumina, San Diego, CA) was used to estimate β scores (proportion of total signal from the methylation-specific probe or color channel) and detection P-values (probability that the probe's total intensity falls within the background signal intensity). Quality control exclusion criteria were: β scores with an associated detection P-value exceeding 0.01, samples with more than 1.5% missing data points across ~470,000 autosomal CpGs, or probes where 10% of samples or more failed to yield adequate intensity.[35] After exclusions, β scores were normalized (separately for Infinium I and II chemistries) using the ComBat package to address batch effects.[30] Finally, methylation loci were removed where the probe sequence mapped to a location that did not match the annotation file or to more than one locus. The final analysis included 463,995 CpG loci.

2.3.3. Diet assessment. Participants' habitual diet was ascertained using a version of the Diet History Questionnaire (DHQ) developed at the National Cancer Institute.[32] The questionnaire consisted of 124 food items, which included questions on portion size and dietary supplements. The DHQ food list and nutrient database was based on national dietary data [US Department of Agriculture (USDA) 1994-6 Continuing Survey of Food Intake by Individuals, available from the USDA Food Surveys Research Group]. Further details are available in prior publications.[32] Energy intake was calculated based on

USDA's 1989-91 Continuing Survey of Food Intakes of Individuals (CSFII). Participants with total energy intake outside of range 800-5500 kcal in men or 600-4500 kcal in women were excluded (n=40).

2.3.4. Covariate assessment. Covariates were ascertained during the clinical examination. Body mass index (BMI) was calculated as weight (kg)/height $(m)^2$. Smokers were defined as participants who had quit smoking less than a year to the day of clinical visit, in addition to current smokers. Physical activity was defined as a continuous score based on intensity and duration of activities.

2.3.5. Epigenome-wide association analysis. We used the lmekin function in the kinship R software package for all analyses (Atkinson B, Therneau T. kinship: mixed-effects Cox models, sparse matrices, and modeling data from large pedigrees. 2007; R package version 1.1.0–17). Three models were used to quantify epigenome-wide associations with the continuous dietary indices. All models included family as a random effect. Fixed effect covariates in the first model included age, sex, and total energy intake. As sensitivity analysis, the second model additionally adjusted for current smoking status and physical activity, while third model additionally adjusted for BMI.

2.4. Multi-Ethnic Study of Atherosclerosis (MESA)

2.4.1 Cohort and Study Sample. The Multi-Ethnic Study of Atherosclerosis (MESA) is a longitudinal study with a multi-ethnic, mixed-gender cohort of 6,814 participants, designed to investigate prevalence, correlates, and progression of subclinical CVD in older adults1 at 6 U.S. sites. Since its inception in 2000, five clinic visits collected extensive clinical, socio-demographic, lifestyle, behavior, laboratory, nutrition, and medication data. 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 in greater detail. The study protocol was approved

by the Institutional Review Board at each site. Exclusion criteria included a history of myocardial infarction, stroke, or cancer. All participants signed informed consent.

2.4.2. Genome-wide DNA methylation profiling. DNA methylation sample, measurement, normalization, and quality control. As previously described,[37] 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 as previously described [37] to calculate the enrichment scores using the gene signature of each

blood cell type from previously defined lists.[38] The final methylation value for each methylation probe was computed as the beta-value, essentially the proportion of the methylated to the total intensity.

2.4.3. Diet assessment. Dietary pattern phenotypes were derived from participant responses to an assessment of diet, administered using a food frequency questionnaire (FFQ) at MESA exam 5. The assessment of diet was a modified Block-style 120-item FFQ, detailing their usual consumption frequency and serving size of specific foods and beverages over the past year. The MESA FFQ was developed for a multi-ethnic population, and its validity has been demonstrated previously.[39] For each food item, the consumption frequency (times/day, week or month) and serving size (small, medium or large) were recorded. Frequency options ranged from 'rare or never' to '2 times/day' for each food item. Related line items were combined to form 47 different food groups. Participants were excluded if implausible data were reported on dietary information (consuming >25104 or <2510 kJ/day).[40]

2.4.4. Covariates. Body mass index (BMI) was calculated as weight (kg)/height (m)2. Smoking status was recorded as never, former, and current smokers. We assessed physical activity using the MESA Typical Week Physical Activity Survey.[41]

2.4.5. Epigenome-wide association analysis. We analyzed the data for the MDS and AHEI scores in parallel, using three linear regression models, for each race separately, with dietary scores as the independent variable. The dependent variable was the DNA methylation beta value after removing the variation due to the batch effect across multiple chips as well as the sample position on the chip. Covariates for model 1 included age, sex, total energy intake, study site, and residual cell estimates of neutrophils, B cells, T cells, monocytes, and natural killer cells to adjust for residual sample contamination with non-monocyte cell types. Model 2 included model 1 covariates and adjusted for smoking status at exam 5, and moderate/vigorous physical activity (MET-MIN/WK). Model 3 included model 2 covariates and adjusted for BMI.

2.4.6. Acknowledgements. MESA and the MESA SHARe project 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 N01-HC-95159, N01-HC-95160, N01-HC-95161, N01-HC-95162, N01-HC-95163, N01-HC-95164, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168, N01-HC-95169, UL1-TR-001079, UL1-TR-000040, and DK063491. The MESA Epigenomics & Transcriptomics Study was funded by NHLBI grant R01HL101250 to Wake Forest University Health Sciences. Analysis of MESA data reported in this publication was also supported by the National Institute On Aging of the National Institutes of Health under Award Number R03AG056959. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

2.5. Rotterdam Study (RS)

2.5.1. Cohort and Study Sample. The RS is a large prospective, population-based cohort study aimed at assessing the occurrence of and risk factors for chronic diseases (cardiovascular, endocrine, hepatic, neurological, ophthalmic, psychiatric, dermatological, oncological, and respiratory) in the middle-aged and elderly. The study comprises 14,926 subjects in total, living in the well-defined Ommoord district in the city of Rotterdam in the Netherlands. In 1989, the first cohort, Rotterdam Study-I (RS-I), was established and comprised of 7,983 subjects with age 55 years or above. In 2000, the second cohort, Rotterdam Study-II (RS-II) was included with 3,011 subjects who had reached an age of 55 years since 1989. In 2006, the third cohort, Rotterdam Study-III (RS-III) was further included with 3,932 subjects with age 45 years and above. DNA methylation was measured in a random sample of 1,454 participants from the first and second visit of the third cohort (RS-III-1, RS-III-2) and third visit of the second cohort (RS-II-3).

2.5.2. Genome-wide DNA methylation profiling. DNA was extracted from whole peripheral blood (stored in EDTA tubes) by standardized salting out methods. Genome-wide DNA methylation levels were measured using the Illumina Human Methylation 450K array (24). 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, samples were hybridized to the arrays according to the manufacturers' protocol. The methylation percentage of a CpG site was reported as a beta-value ranging between 0 (no methylation) and 1 (full methylation). The data preprocessing was additionally performed in both datasets using an R programming pipeline which is based on the pipeline developed by Tost & Toulemat,[42] which includes additional parameters and options to preprocess and normalize methylation data directly from idat files. We excluded probes which had a detection p-value >0.01 in >95% of samples. 11,648 probes at X and Y chromosomes were excluded to avoid gender bias. The raw beta values were then background corrected and normalized using the DASEN option of the WateRmelon R-package.[25] Per individual probe, participants with methylation levels higher than three times the inter-quartiles range (IQR) were excluded.

2.5.3. Diet assessment. Subjects completed a self-administered FFQ containing of 389 items, based on a validated FFQ for Dutch adults that also took into account preparation methods.[6,43] For each item frequency of consumption (in times per month or week), the numbers of servings per day (expressed in standardized household measures, e.g., spoon or cup, and natural units, e.g., piece of fruit or slice of bread), and the preparations methods were included. To derive total energy and nutrient intakes per day the received dietary data were linked to the Dutch Food Composition Table of 2011. We excluded participants with implausible energy intake, for which cut-offs were set at <500 or >5000 kcal/day.

2.5.4. Covariates assessment. Height and weight were measured during the center visit and BMI was calculated (kg/m^2). During home visit interviews, data on tobacco smoking were collected. Information on smoking history was acquired from questionnaires, and categorized as never, former or current smoking. Energy intake was calculated using the FFQs. Data on physical activity were obtained through

the LASA Physical Activity Questionnaire.[44,45] Because of differently extended data generation for physical activity between cohorts, z-scores were calculated to obtain comparable data. For the RS-III-1 we estimated leukocyte proportions (B-cells, CD4+ T-cells, CD8+ T-cells, granulocytes, monocytes and NK-cells) by a formula developed by Houseman.[29,46] For RS-II-3 and RS-III-2 we used white blood cell counts (WBC), i.e. lymphocytes, monocytes, and granulocytes, which were assessed with a Coulter AcT diff2 Hematology Analyzer.

2.5.5. Epigenome-wide association analysis. We modeled cross-sectional associations between Dasen normalized beta-values of the CpG sites as outcome and MDS score or AHEI as exposure using linear mixed effect models. Three models were used. In the first model we adjusted for age, sex, energy intake, white blood cell proportions, array number and position on array. Technical covariates (array number and position on array) were modeled as random effects. The second model was additionally adjusted for smoking and physical activity. In the third model we additionally adjusted for BMI. Finally, we checked all identified CpG sites for cross-reaction or polymorphism.[47] A CpG site was considered polymorphic when a SNP with a minor allele frequency of >0.01 resided at the position of the cytosine or guanine nucleotide, or within 10 bp from the CpG site within the probe binding site.[48]

2.5.6. Acknowledgements. 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. The authors are grateful to the study participants, the staff from the Rotterdam Study and the participating general practitioners and pharmacists.

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

2.6.1. Cohort and Study Sample. The WHI is a multi-center prospective study of risk factors for cardiovascular disease, cancer, osteoporotic fractures, and other causes of morbidity and mortality among postmenopausal women.[49,50] Between 1993 and 1998, 68,132 women aged 50-79 years from forty WHI clinical centers throughout the United States (US) were enrolled in the WHI Clinical Trials (CT). WHI CT participants completed a screening visit (SV) and an annual visit (AV) at one, three, six, and nine years after randomization (AV1, AV3, AV6, AV9). The Epigenetic Mechanisms of PM-Mediated CVD Risk (WHI-EMPC), an ancillary study of epigenetic mechanisms underlying associations between ambient particulate matter air pollution and cardiovascular disease, measured DNA methylation in 2,200 randomly selected WHI CT participants (stage 1: SV, AV3, or AV6), remeasured it in 200 participants at a second visit (stage 2: AV3 or AV6), and remeasured again in 43 participants at a third visit among those who participated in the WHI Long Life Study (stage 3: LLS), yielding 2,443 total observations. The analyses of mortality described herein involve 1,095 WHI-EMPC participants of European ancestry.

2.6.2. Genome-wide DNA methylation profiling. In 91% of whole blood samples from WHI-EMPC participants, DNA was extracted from peripheral blood leukocytes using a Qiagen/5-Prime method. The remaining samples were extracted using a phenol choloroform (8%), Qiagen/Bioserve (1%), or salt precipitation (<1%) method. Extracted DNA was subjected to high throughput bisulfite conversion using the EZ-96 DNA Methylation™ Kit (Zymo Research; Irvine, CA, USA). DNA methylation was measured on a methylome-wide scale at 485,577 CpG sites using the Illumina 450K Infinium Methylation BeadChip (Illumina Inc.; San Diego, CA, USA). Methylation was quantitatively represented by beta, the proportion of methylated cytosines over the sum of methylated and unmethylated cytosines across the same loci. The data were quality controlled, Beta Mixture Quantile (BMIQ)-normalized to adjust for probe bias,[51] and ComBat-adjusted for stage and plate using empirical Bayes methods.[52]

2.6.3. Epigenome-wide association analyses. Associations between each of the 485,5 77 measured CpG sites and all-cause and cause-specific mortality were modeled using Cox-proportional hazards models adjusted for baseline age, principal components of ancestry, smoking status, alcohol use, education, physical activity, BMI, race/ethnicity, and prevalent disease for hypertension, heart failure, coronary heart disease, stroke, and cancer.

2.6.4 Acknowledgements. The Women's Health Initiative (WHI) is funded by the National Heart, Lung and Blood Institute, U.S. Department of Health and Human Services, through contracts HHSN268201100046C, HHSN268201100001C, HHSN268201100002C, HHSN268201100003C, HHSN268201100004C, and HHSN271201100004C. The Epigenetic Mechanisms of PM-Mediated CVD Risk (WHI-EMPC) was supported by National Institute of Environmental Health Science grant R01-ES020836. All contributors to WHI science are listed at https://www.whi.org/researchers/Documents%20%20Write%20a%20Paper/WHI%20Investigator%20Long%20List.pdf.

3. References

1. Feskanich D, Rimm EB, Giovannucci EL, Colditz GA, Stampfer MJ, Litin LB, et al. Reproducibility and validity of food intake measurements from a semiquantitative food frequency questionnaire. *J Am Diet Assoc*. 1993;93(7):790-796.
2. Willett WC, Sampson L, Stampfer MJ, Rosner B, Bain C, Witschi J, et al. Reproducibility and validity of a semiquantitative food frequency questionnaire. *Am J Epidemiol*. 1985;122(1):51-65.
3. Thompson FE, Subar AF, Brown CC, Smith AF, Sharbaugh CO, Jobe JB, et al. Cognitive research enhances accuracy of food frequency questionnaire reports: results of an experimental validation study. *J Am Diet Assoc*. 2002;102(2):212-225.
4. Abiemo EE, Alonso A, Nettleton JA, Steffen LM, Bertoni AG, Jain A, et al. Relationships of the Mediterranean dietary pattern with insulin resistance and diabetes incidence in the Multi-Ethnic Study of Atherosclerosis (MESA). *Br J Nutr*. 2013;109(8):1490-1497.
5. Klipstein-Grobusch K, den Breeijen JH, Goldbohm RA, Geleijnse JM, Hofman A, Grobbee DE, et al. Dietary assessment in the elderly: validation of a semiquantitative food frequency questionnaire. *Eur J Clin Nutr*. 1998;52(8):588-596.
6. Goldbohm RA, van den Brandt PA, Brants HA, van't Veer P, Al M, Sturmans F, et al. Validation of a dietary questionnaire used in a large-scale prospective cohort study on diet and cancer. *Eur J Clin Nutr*. 1994;48(4):253-65.
7. Shimakawa T, Sorlie P, Carpenter MA, Dennis B, Tell GS, Watson R, et al. Dietary intake patterns and sociodemographic factors in the atherosclerosis risk in communities study. *Prev Med*. 1994;23(6):769-780.
8. Rebholz CM, Crews DC, Grams ME, Steffen LM, Levey AS, Miller ER 3rd, et al. DASH (Dietary Approaches to Stop Hypertension) Diet and Risk of Subsequent Kidney Disease. *Am J Kidney Dis*. 2016;68(6):853-861.
9. Trichopoulou A, Costacou T, Bamia C, Trichopoulos D. Adherence to a Mediterranean diet and survival in a Greek population. *N Engl J Med.* 2003;348(26):2599-2608.

10. Fung TT, Rexrode KM, Mantzoros CS, Manson JE, Willett WC, Hu FB. Mediterranean diet and incidence of and mortality from coronary heart disease and stroke in women. *Circulation.* 2009;119(8):1093-1100.

11. Chiuve SE, Fung TT, Rimm EB, et al. Alternative dietary indices both strongly predict risk of chronic disease. *J Nutr.* 2012;142(6):1009-1018.

12. Hemani G, Zheng J, Elsworth B, et al. The MR-Base platform supports systematic causal inference across the human phenome. *Elife*. 2018;7.

13. Huan T, Joehanes R, Song C, Peng F, Guo Y, Mendelson M, et al. Genome-wide identification of DNA methylation QTLs in whole blood highlights pathways for cardiovascular disease. *Nat Commun*. 2019;10(1):4267

14. Yengo L, Sidorenko J, Kemper KE, et al. Meta-analysis of genome-wide association studies for height and body mass index in approximately 700000 individuals of European ancestry. *Hum Mol Genet.* 2018;27(20):3641-3649.

15. Shungin D, Winkler TW, Croteau-Chonka DC, et al. New genetic loci link adipose and insulin biology to body fat distribution. *Nature.* 2015;518(7538):187-196.

16. Willer CJ, Schmidt EM, Sengupta S, et al. Discovery and refinement of loci associated with lipid levels. *Nat Genet.* 2013;45(11):1274-1283.

17. Mahajan A, Taliun D, Thurner M, et al. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *Nat Genet.* 2018;50(11):1505-1513.

18. Genomes Project C, Auton A, Brooks LD, et al. A global reference for human genetic variation. *Nature.* 2015;526(7571):68-74.

19. Slieker RC, Bos SD, Goeman JJ, et al. Identification and systematic annotation of tissue-specific differentially methylated regions using the Illumina 450k array. *Epigenetics Chromatin*. 2013;6(1):26.

20. Galperin MY, Fernandez-Suarez XM, Rigden DJ. The 24th annual Nucleic Acids Research database issue: a look back and upcoming changes. *Nucleic Acids Res*. 2017;45(D1):D1-D11.

21. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun*. 2017;8(1):1826.

22. Consortium GT. The Genotype-Tissue Expression (GTEx) project. *Nat Genet*. 2013;45(6):580-585.

23. Kannel WB, Feinleib M, McNamara PM, Garrison RJ, Castelli WP. An investigation of coronary heart disease in families. The Framingham offspring study. *Am J Epidemiol.* 1979;110(3):281-290.

24. Splansky GL, Corey D, Yang Q, 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.* 2007;165(11):1328-1335.

25. Pidsley R, CC YW, Volta M, Lunnon K, Mill J, Schalkwyk LC. A data-driven approach to preprocessing Illumina 450K methylation array data. *BMC Genomics.* 2013;14:293.

26. Liu C, Marioni RE, Hedman AK, et al. A DNA methylation biomarker of alcohol consumption. *Mol Psychiatry.* 2018;23(2):422-433.

27. Rimm EB, Giovannucci EL, Stampfer MJ, Colditz GA, Litin LB, Willett WC. Reproducibility and validity of an expanded self-administered semiquantitative food frequency questionnaire among male health professionals. *Am J Epidemiol.* 1992;135(10):1114-1126; discussion 1127-1136.

28. Kannel WB, Belanger A, D'Agostino R, Israel I. Physical activity and physical demand on the job and risk of cardiovascular disease and death: the Framingham Study. *Am Heart J.* 1986;112(4):820-825.

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

30. Maksimovic J, Gordon L, Oshlack A. SWAN: Subset-quantile within array normalization for illumina infinium HumanMethylation450 BeadChips. *Genome Biol.* 2012;13(6):R44.

31. Stevens J, Metcalf PA, Dennis BH, Tell GS, Shimakawa T, Folsom AR. Reliability of a food frequency questionnaire by ethnicity, gender, age and education. *Nutrition Research.* 1996;16:735-745.

32. Corella D, Arnett DK, Tsai MY, et al. The -256T>C polymorphism in the apolipoprotein A-II gene promoter is associated with body mass index and food intake in the genetics of lipid lowering drugs and diet network study. *Clin Chem.* 2007;53(6):1144-1152.

33. Irvin MR, Kabagambe EK, Tiwari HK, et al. Apolipoprotein E polymorphisms and postprandial triglyceridemia before and after fenofibrate treatment in the Genetics of Lipid Lowering and Diet Network (GOLDN) Study. *Circ Cardiovasc Genet.* 2010;3(5):462-467.

34. Absher DM, Li X, Waite LL, et al. Genome-wide DNA methylation analysis of systemic lupus erythematosus reveals persistent hypomethylation of interferon genes and compositional changes to CD4+ T-cell populations. *PLoS Genet.* 2013;9(8):e1003678.

35. Aslibekyan S, Demerath EW, Mendelson M, et al. Epigenome-wide study identifies novel methylation loci associated with body mass index and waist circumference. *Obesity (Silver Spring).* 2015;23(7):1493-1501.

36. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics.* 2007;8(1):118-127.

37. Liu Y, Aryee MJ, Padyukov L, et al. Epigenome-wide association data implicate DNA methylation as an intermediary of genetic risk in rheumatoid arthritis. *Nat Biotechnol.* 2013;31(2):142-147.

38. Abbas AR, Baldwin D, Ma Y, et al. Immune response in silico (IRIS): immune-specific genes identified from a compendium of microarray expression data. *Genes Immun.* 2005;6(4):319-331.

39. Mayer-Davis EJ, Vitolins MZ, Carmichael SL, et al. Validity and reproducibility of a food frequency interview in a Multi-Cultural Epidemiology Study. *Ann Epidemiol.* 1999;9(5):314-324.

40. Hu T, Jacobs DR, Bazzano LA, Bertoni AG, Steffen LM. Low-carbohydrate diets and prevalence, incidence and progression of coronary artery calcium in the Multi-Ethnic Study of Atherosclerosis (MESA). *Br J Nutr.* 2019:1-8.

41. Bertoni AG, Whitt-Glover MC, Chung H, et al. The association between physical activity and subclinical atherosclerosis: the Multi-Ethnic Study of Atherosclerosis. *Am J Epidemiol.* 2009;169(4):444-454.

42. Touleimat N, Tost J. Complete pipeline for Infinium((R)) Human Methylation 450K BeadChip data processing using subset quantile normalization for accurate DNA methylation estimation. *Epigenomics.* 2012;4(3):325-341.

43. Feunekes GI, Van Staveren WA, De Vries JH, Burema J, Hautvast JG. Relative and biomarker-based validity of a food-frequency questionnaire estimating intake of fats and cholesterol. *Am J Clin Nutr.* 1993;58(4):489-496.

44. Caspersen CJ, Bloemberg BP, Saris WH, Merritt RK, Kromhout D. The prevalence of selected physical activities and their relation with coronary heart disease risk factors in elderly men: the Zutphen Study, 1985. *Am J Epidemiol.* 1991;133(11):1078-1092.

45. Stel VS, Smit JH, Pluijm SM, Visser M, Deeg DJ, Lips P. Comparison of the LASA Physical Activity Questionnaire with a 7-day diary and pedometer. *J Clin Epidemiol.* 2004;57(3):252-258.

46. Aryee MJ, Jaffe AE, Corrada-Bravo H, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics.* 2014;30(10):1363-1369.

47. Chen YA, Lemire M, Choufani S, et al. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics.* 2013;8(2):203-209.

48. Barfield RT, Almli LM, Kilaru V, et al. Accounting for population stratification in DNA methylation studies. *Genet Epidemiol.* 2014;38(3):231-241.

49. Anderson GL, Manson J, Wallace R, et al. Implementation of the Women's Health Initiative study design. Ann Epidemiol. 2003;13(9 Suppl):S5-17.

50. Design of the Women's Health Initiative clinical trial and observational study. The Women's Health Initiative Study Group. Control Clin Trials. 1998;19(1):61-109.

51. Teschendorff AE, Marabita F, Lechner M, et al. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. Bioinformatics. 2013;29(2):189-196.

52. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics. 2007;8(1):118-127.

Supplemental Table 1. Discovery and Replication analysis using MDS in EA participants

				Discovery			Replication			
CpG	CHR	BP	Gene	Beta	SE	P	Beta	SE	P	Replicated
cg01940273	2	233284934		0.0055	0.0010	1.2E-07	0.0044	0.0009	2.12E-06	YES
cg06126421	6	30720080		0.0064	0.0012	2.7E-07	0.0075	0.0025	0.002	YES
cg05575921	5	373378	AHRR	0.0115	0.0016	1.7E-13	0.0107	0.0035	0.002	YES
cg26703534	5	377358	AHRR	0.0031	0.0006	1.5E-06	0.0016	0.0006	0.008	NO
cg21161138	5	399360	AHRR	0.0036	0.0007	4.2E-07	0.0033	0.0013	0.009	NO
cg24859433	6	30720203		0.0028	0.0005	7.9E-08	0.0026	0.0010	0.01	NO
cg03636183	19	17000585	F2RL3	0.0068	0.0011	7.4E-10	0.0051	0.0021	0.01	NO
cg03084350	3	38065265	PLCD1	-0.0023	0.0004	8.2E-08	-0.0010	0.0004	0.03	NO
cg22185977	5	1518133	LPCAT1	-0.0046	0.0009	2.3E-07	-0.0042	0.0020	0.03	NO
cg19859270	3	98251294	GPR15	0.0023	0.0004	1.5E-08	0.0009	0.0006	0.12	NO
cg20761853	17	76850198	TIMP2	-0.0033	0.0007	8.5E-07	-0.0006	0.0007	0.34	NO
cg16755922	17	80536214	FOXK2	-0.0053	0.0009	7.8E-09	-0.0002	0.0004	0.54	NO
cg26897297	13	49796434	MLNR	-0.0039	0.0008	6.2E-07	0.0002	0.0004	0.70	NO

MDS: Mediterranean-style diet score; EA: European ancestry; Genome build 37.

Supplemental Table 2. Significant CpGs (FDR < 0.05) in meta-analysis of association between MDS and DNA methylation in European Ancestry (EA) participants

				Meta-analysis				Discovery			Replication			Internal validation
CpG	CHR	BP	Gene	Beta	SE	P	I_squared	Beta	SE	P	Beta	SE	P	
cg01940273	2	233284934		0.0048	0.0007	1.6E-12	0	0.0055	0.0010	1.2E-07	0.0044	0.0009	2.1E-06	Yes
cg18181703	17	76354621	SOCS3	0.0035	0.0006	3.5E-10	0	0.0035	0.0008	1.0E-05	0.0035	0.0008	8.0E-06	Yes
cg08732950	16	89023389	CBFA2T3	-0.0027	0.0005	2.8E-08	0	-0.0022	0.0007	2.1E-03	-0.0032	0.0007	2.4E-06	Yes
cg05951221	2	233284402		0.0048	0.0009	3.2E-08	0.07	0.0049	0.0011	1.8E-05	0.0050	0.0014	5.0E-04	Yes
cg02097604	17	17750910	TOM1L2	0.0018	0.0003	3.6E-08	0	0.0022	0.0005	4.8E-06	0.0015	0.0005	1.1E-03	Yes
cg12075928	8	141801307	PTK2	0.0032	0.0006	1.1E-07	0	0.0035	0.0008	3.7E-05	0.0029	0.0009	7.0E-04	Yes
cg25189904	1	68299493	GNG12	0.0051	0.0010	2.5E-07	0	0.0044	0.0013	1.1E-03	0.0059	0.0014	4.7E-05	Yes
cg19693031	1	145441552	TXNIP	0.0034	0.0007	3.1E-07	0.14	0.0030	0.0009	5.3E-04	0.0037	0.0010	1.8E-04	Yes
cg02079413	11	2986505	SNORA54;NAP1L4	-0.0020	0.0004	3.1E-07	0.14	-0.0026	0.0007	7.1E-05	-0.0018	0.0005	1.4E-04	Yes
cg04885881	1	11123118		0.0036	0.0007	3.2E-07	0.12	0.0028	0.0009	2.1E-03	0.0043	0.0009	4.5E-06	Yes
cg02716826	9	33447032	SUGT1P1;AQP3	0.0023	0.0005	5.6E-07	0	0.0020	0.0007	4.4E-03	0.0026	0.0006	3.0E-05	Yes
cg03646329	13	48987165	LPAR6;RB1	0.0035	0.0007	1.1E-06	0	0.0032	0.0011	2.8E-03	0.0038	0.0010	1.1E-04	Yes
cg16969872	13	79968324	RBM26	0.0034	0.0007	1.2E-06	0.25	0.0037	0.0008	1.8E-06	0.0032	0.0010	2.1E-03	Yes

Beta/SE: DNA methylation level change according to per-standard deviation change in the Mediterranean-style diet score (MDS) using sex- and age-adjusted model. Position is based on genome build 37. FDR: false discovery rate.

Supplemental Table 3. Discovery and Replication analysis using AHEI in EA participants

				Discovery			Replication			
CpG	CHR	BP	Gene	Beta	SE	P	Beta	SE	P	Replicated
cg18181703	17	76354621	SOCS3	0.0039	0.0008	4.5E-07	0.0037	0.0008	9.5E-07	YES
cg16969872	13	79968324	RBM26	0.0037	0.0007	6.2E-07	0.0026	0.0008	7.2E-04	YES
cg02097604	17	17750910	TOM1L2	0.0024	0.0005	3.2E-07	0.0013	0.0004	0.001	NO
cg03450842	10	80834947	ZMIZ1	0.0025	0.0005	4.2E-06	0.0019	0.0006	0.002	NO
cg00277397	7	71800412	CALN1	0.0033	0.0007	5.3E-07	0.0016	0.0005	0.003	NO
cg19693031	1	145441552	TXNIP	0.0037	0.0008	6.0E-06	0.0023	0.0008	0.003	NO
cg15866367	13	49796387	MLNR	-0.0027	0.0006	1.5E-06	0.0009	0.0003	0.004	NO
cg22185977	5	1518133	LPCAT1	-0.0044	0.0009	2.3E-07	-0.0042	0.0016	0.007	NO
cg24859433	6	30720203		0.0030	0.0005	1.7E-09	0.0023	0.0009	0.01	NO
cg05575921	5	373378	AHRR	0.0113	0.0015	4.7E-14	0.0093	0.0037	0.01	NO
cg01940273	2	233284934		0.0053	0.0010	6.4E-08	0.0041	0.0017	0.01	NO
cg17501210	6	166970252	RPS6KA2	0.0038	0.0008	2.7E-06	0.0040	0.0017	0.02	NO
cg08884571	19	45901453	PPP1R13L	-0.0047	0.0010	5.0E-06	-0.0033	0.0014	0.02	NO
cg06126421	6	30720080		0.0065	0.0012	3.5E-08	0.0063	0.0028	0.02	NO
cg03636183	19	17000585	F2RL3	0.0072	0.0011	1.4E-11	0.0044	0.0023	0.05	NO
cg19157853	11	126286452		-0.0019	0.0004	4.5E-06	-0.0010	0.0006	0.08	NO
cg26703534	5	377358	AHRR	0.0034	0.0006	2.8E-08	0.0016	0.0009	0.10	NO
cg21161138	5	399360	AHRR	0.0033	0.0007	1.5E-06	0.0030	0.0019	0.10	NO
cg15082870	7	36022841		-0.0033	0.0006	1.4E-07	0.0006	0.0004	0.14	NO
cg10563109	6	5087749		-0.0070	0.0014	4.7E-07	-0.0021	0.0014	0.14	NO
cg15344028	2	204801510	ICOS	0.0051	0.0010	3.3E-07	0.0018	0.0014	0.21	NO
cg15614653	2	129104464		0.0030	0.0006	3.3E-06	0.0005	0.0005	0.26	NO
cg25221975	3	13663444	FBLN2	-0.0037	0.0008	4.3E-06	-0.0016	0.0015	0.28	NO
cg01487683	6	5031084		-0.0030	0.0006	3.8E-07	-0.0008	0.0008	0.31	NO
cg05304729	1	158800024	MNDA	-0.0039	0.0009	4.5E-06	-0.0013	0.0014	0.35	NO
cg18315935	1	45120295	TMEM53	-0.0025	0.0005	1.5E-06	0.0003	0.0003	0.41	NO
cg07069844	19	1617027	TCF3	-0.0020	0.0004	5.0E-06	-0.0008	0.0009	0.41	NO
cg00826902	1	3563954	WDR8	-0.0032	0.0007	1.6E-06	-0.0003	0.0004	0.42	NO
cg03084350	3	38065265	PLCD1	-0.0019	0.0004	4.7E-06	-0.0005	0.0007	0.44	NO
cg10639435	8	146104221	ZNF250	-0.0036	0.0007	7.0E-07	-0.0008	0.0012	0.49	NO
cg11092486	6	5087604		-0.0069	0.0014	1.6E-06	-0.0010	0.0016	0.54	NO
cg16755922	17	80536214	FOXK2	-0.0054	0.0009	3.6E-10	0.0002	0.0004	0.68	NO
cg00994936	19	1423902	DAZAP1	-0.0030	0.0006	1.0E-06	0.0001	0.0002	0.71	NO
cg11183227	15	91455407	MAN2A2	-0.0031	0.0006	1.6E-06	-0.0002	0.0006	0.73	NO
cg20761853	17	76850198	TIMP2	-0.0037	0.0007	1.6E-08	-0.0001	0.0004	0.75	NO
cg03440556	10	102107757	SCD	0.0053	0.0010	4.1E-07	0.0004	0.0013	0.76	NO
cg09273716	20	35494835		-0.0029	0.0006	1.6E-06	-0.0002	0.0008	0.83	NO
cg16805291	7	36022575		-0.0035	0.0008	6.0E-06	-0.0002	0.0011	0.86	NO
cg18518074	11	64642316	EHD1	-0.0027	0.0006	2.3E-06	-0.0001	0.0008	0.87	NO
cg03294424	8	127866047		-0.0037	0.0007	4.9E-07	0.0001	0.0013	0.97	NO
cg26002437	10	45473231	C10orf10;RASSF4	0.0021	0.0004	2.3E-06	0.0000	0.0003	0.97	NO

AHEI: Alternative Healthy Eating Index ; EA: European ancestry; position is based on genome build 37.

Supplemental Table 4. Significant CpGs (FDR < 0.05) in meta-analysis of association between AHEI and DNA methylation in European Ancestry (EA) participants

				Meta-analysis				Discovery			Replication			Internal
CpG	CHR	BP	Gene	Beta	SE	P	I_square	Beta	SE	P	Beta	SE	P	validation
cg18181703	17	76354621	SOCS3	0.0038	0.0005	2.0E-12	0	0.0039	0.0008	4.5E-07	0.0037	0.0008	9.5E-07	YES
cg16969872	13	79968324	RBM26	0.0032	0.0005	3.0E-09	0	0.0037	0.0008	6.2E-07	0.0026	0.0008	7.2E-04	YES
cg27118035	16	31891978	ZNF267	-0.0028	0.0005	4.9E-09	0	-0.0025	0.0006	7.8E-05	-0.0031	0.0007	1.3E-05	YES
cg02097604	17	17750910	TOM1L2	0.0018	0.0003	6.6E-09	0	0.0024	0.0005	3.2E-07	0.0013	0.0004	1.3E-03	YES
cg25953130	10	63753550	ARID5B	0.0044	0.0008	1.2E-08	0	0.0045	0.0011	3.6E-05	0.0042	0.0011	1.9E-04	YES
cg16936953	17	57915665	VMP1	0.0042	0.0007	1.5E-08	0	0.0043	0.0011	1.2E-04	0.0040	0.0010	3.3E-05	YES
cg05232694	20	48809539		0.0037	0.0007	3.1E-08	0	0.0039	0.0009	4.4E-05	0.0037	0.0011	6.5E-04	YES
cg20842915	7	39665132	RALA	0.0029	0.0005	8.1E-08	0	0.0037	0.0010	9.2E-05	0.0025	0.0007	1.4E-04	YES
cg03190891	10	97201172	SORBS1	-0.0025	0.0005	9.0E-08	0	-0.0031	0.0007	3.6E-05	-0.0021	0.0006	4.2E-04	YES
cg08884571	19	45901453	PPP1R13L	-0.0042	0.0008	4.6E-07	0	-0.0047	0.0010	5.0E-06	-0.0033	0.0014	2.1E-02	YES
cg25909064	11	120082805	OAF	0.0018	0.0004	8.0E-07	0	0.0023	0.0006	3.2E-04	0.0016	0.0005	4.7E-04	YES
cg24694018	1	145457621	POLR3GL	0.0017	0.0003	8.3E-07	0	0.0017	0.0005	1.2E-03	0.0017	0.0005	2.0E-04	YES
cg19202384	17	79894511	PYCR1	0.0019	0.0004	9.9E-07	0.04	0.0023	0.0007	6.7E-04	0.0018	0.0006	7.2E-04	YES
cg27039118	8	116575902	TRPS1	0.0036	0.0007	1.2E-06	0	0.0046	0.0011	5.6E-05	0.0029	0.0010	3.2E-03	YES
cg13074055	14	106329206		0.0047	0.0010	1.3E-06	0	0.0054	0.0015	2.1E-04	0.0042	0.0013	1.5E-03	YES
cg01294327	19	2291373	LINGO3	0.0049	0.0010	1.4E-06	0	0.0051	0.0015	6.0E-04	0.0047	0.0014	6.9E-04	YES
cg11468085	11	67435577	ALDH3B2	-0.0022	0.0005	1.4E-06	0.06	-0.0024	0.0007	3.9E-04	-0.0022	0.0007	1.8E-03	YES
cg11250194	11	61601937	FADS2	0.0025	0.0005	1.5E-06	0	0.0026	0.0007	1.9E-04	0.0024	0.0008	2.4E-03	YES
cg03646329	13	48987165	LPAR6;RB1	0.0034	0.0007	1.5E-06	0	0.0040	0.0010	1.2E-04	0.0029	0.0010	2.8E-03	YES
cg24735226	1	65096537	CACHD1	-0.0037	0.0008	1.6E-06	0	-0.0031	0.0011	5.2E-03	-0.0041	0.0010	7.9E-05	YES
cg02959282	3	136667373	NCK1	0.0024	0.0005	1.8E-06	0	0.0019	0.0010	6.0E-02	0.0026	0.0006	9.9E-06	NO
cg19157853	11	126286452		-0.0016	0.0003	2.3E-06	0	-0.0019	0.0004	4.5E-06	-0.0010	0.0006	8.2E-02	NO
cg26470501	19	45252955	BCL3	0.0019	0.0004	2.4E-06	0.03	0.0020	0.0006	5.0E-04	0.0020	0.0007	3.6E-03	YES
cg02508743	8	56903623	LYN	-0.0024	0.0005	2.5E-06	0	-0.0022	0.0008	4.4E-03	-0.0026	0.0008	9.8E-04	YES
cg09940677	14	103415458	CDC42BPB	-0.0014	0.0003	2.9E-06	0	-0.0014	0.0004	6.6E-04	-0.0014	0.0004	1.4E-03	YES
cg07805029	1	92953256	GFI1	0.0027	0.0006	3.1E-06	0	0.0018	0.0009	4.8E-02	0.0032	0.0007	1.2E-05	YES

Beta/SE: DNA methylation level change according to per-standard deviation change in the Alternative Healthy Eating Index (AHEI) using sex- and age-adjusted model. Position is based on genome build 37. FDR: false discovery rate.

Supplemental Table 5. Pairwised Pearson correlation coefficients of 30 diet associated CpGs in EA participants

CHR	BP	CpG	cg04885881	cg24735226	cg07805029	cg19693031	cg24694018	cg01940273	cg20842915	cg02508743	cg27039118	cg02716826	cg25953130	cg03190891	cg02079413	cg11250194	cg11468085	cg25909064	cg03646329	cg16969872	cg09940677	cg13074055	cg27118035	cg08732950	cg02097604	cg16936953	cg18181703	cg19202384	cg01294327	cg26470501	cg08884571	cg05232694
1	11123118	cg04885881	1.00																												
1	65096537	cg24735226	0.14	1.00																											
1	92953256	cg07805029	0.02	-0.23	1.00																										
1	145441552	cg19693031	0.17	0.16	0.00	1.00																									
1	145457621	cg24694018	0.23	-0.06	0.30	0.23	1.00																								
2	233284934	cg01940273	0.34	-0.09	0.13	0.13	0.34	1.00																							
7	39665132	cg20842915	-0.15	-0.24	0.34	-0.08	0.19	0.19	1.00																						
8	56903623	cg02508743	0.16	0.27	0.05	0.23	0.02	-0.16	-0.24	1.00																					
8	116575902	cg27039118	0.33	0.13	0.14	0.18	0.38	0.20	0.12	0.18	1.00																				
9	33447032	cg02716826	0.56	0.32	-0.15	0.17	0.17	0.18	-0.18	0.22	0.30	1.00																			
10	63753550	cg25953130	0.30	0.18	-0.06	0.08	0.19	0.25	0.13	-0.01	0.35	0.43	1.00																		
10	97201172	cg03190891	0.25	0.50	-0.34	0.16	-0.19	-0.22	-0.40	0.45	0.09	0.41	0.15	1.00																	
11	2986505	cg02079413	-0.02	-0.11	0.33	0.09	0.22	0.10	0.11	0.12	0.07	-0.19	-0.09	-0.18	1.00																
11	61601937	cg11250194	0.23	-0.05	0.35	0.19	0.41	0.19	0.06	0.17	0.19	0.12	-0.05	-0.09	0.22	1.00															
11	67435577	cg11468085	0.16	0.36	-0.47	0.00	-0.31	-0.17	-0.34	0.25	0.00	0.34	0.29	0.56	-0.26	-0.35	1.00														
11	120082805	cg25909064	0.20	-0.11	0.30	0.13	0.53	0.40	0.29	-0.13	0.26	0.07	0.14	-0.28	0.25	0.35	-0.33	1.00													
13	48987165	cg03646329	0.01	0.21	-0.22	-0.03	-0.17	0.07	0.07	-0.06	0.05	0.28	0.42	0.22	-0.22	-0.27	0.35	-0.10	1.00												
13	79968324	cg16969872	0.14	0.26	-0.29	0.13	0.03	0.12	-0.12	0.16	0.17	0.34	0.30	0.28	-0.18	-0.13	0.36	-0.07	0.26	1.00											
14	103415458	cg09940677	0.02	0.25	-0.14	0.12	-0.06	-0.18	-0.19	0.46	0.07	0.14	0.06	0.41	0.03	-0.04	0.46	-0.18	0.08	0.17	1.00										
14	106329206	cg13074055	0.07	0.04	0.13	0.07	0.25	0.17	0.13	0.11	0.27	0.10	0.22	0.00	0.09	0.07	-0.02	0.14	0.05	0.15	0.07	1.00									
16	31891978	cg27118035	0.11	0.29	-0.21	0.05	-0.29	-0.23	-0.26	0.26	-0.05	0.17	0.03	0.49	-0.11	-0.12	0.36	-0.27	0.15	0.13	0.27	-0.09	1.00								
16	89023389	cg08732950	0.17	0.19	0.02	0.24	0.07	0.02	-0.09	0.23	0.10	0.19	0.06	0.21	0.07	0.17	0.09	0.13	-0.04	0.08	0.21	0.02	0.14	1.00							
17	17750910	cg02097604	0.38	0.14	0.17	0.18	0.42	0.24	0.10	0.10	0.35	0.40	0.27	0.14	0.11	0.36	-0.08	0.48	0.03	0.14	0.03	0.19	-0.01	0.24	1.00						
17	57915665	cg16936953	0.39	0.31	-0.28	0.10	-0.03	0.05	-0.22	0.19	0.18	0.66	0.37	0.46	-0.22	-0.04	0.43	-0.07	0.33	0.38	0.14	0.07	0.20	0.08	0.25	1.00					
17	76354621	cg18181703	0.51	0.19	-0.03	0.16	0.37	0.37	-0.03	0.06	0.35	0.66	0.40	0.18	-0.06	0.17	0.10	0.31	0.17	0.29	-0.03	0.16	-0.02	0.16	0.51	0.52	1.00				
17	79894511	cg19202384	0.11	-0.19	0.22	0.04	0.32	0.42	0.34	-0.26	0.12	0.02	0.21	-0.35	0.16	0.14	-0.26	0.54	0.05	-0.05	-0.19	0.14	-0.28	0.06	0.29	-0.09	0.26	1.00			
19	2291373	cg01294327	0.20	0.04	0.06	0.02	0.27	0.38	0.12	-0.11	0.13	0.06	0.17	-0.08	0.08	0.06	-0.09	0.31	0.05	0.12	-0.08	0.20	-0.05	0.13	0.24	0.00	0.32	0.34	1.00		
19	45252955	cg26470501	0.44	0.08	0.13	0.12	0.50	0.34	0.11	0.05	0.39	0.44	0.31	0.05	0.02	0.28	-0.06	0.40	-0.02	0.19	-0.06	0.23	-0.10	0.07	0.45	0.31	0.59	0.24	0.26	1.00	
19	45901453	cg08884571	0.04	0.04	0.09	0.08	0.16	0.11	-0.03	0.10	0.02	-0.02	-0.10	0.00	0.07	0.15	-0.11	0.03	-0.18	-0.03	-0.02	-0.02	0.00	0.12	0.03	-0.06	0.06	-0.03	0.11	0.10	1.00
20	48809539	cg05232694	0.25	-0.04	0.34	0.11	0.38	0.31	0.23	0.06	0.22	0.16	0.18	-0.13	0.16	0.28	-0.22	0.44	-0.03	0.01	-0.10	0.21	-0.12	0.05	0.35	0.05	0.32	0.37	0.23	0.37	0.05	1.00

Grey colored region represents CpGs located in the same chromosome; EA: European ancestry; position is based on genome build 37.

Supplemental Table 6. Functional description for protein coding genes annotated to diet-associated CpGs

CpG	CHR	Position	Gene	Full name	Description
cg24735226	1	65096537	CACHD1	Cache domain containing 1	This gene encodes protein cache domain containing 1, which may regulate voltage-dependent calcium channels. Alias name is Von Willebrand Factor Type A And Cache Domain Containing 1.
cg07805029	1	92953256	GFI1	Growth factor independent 1 transcriptional repressor	This gene encodes a nuclear zinc finger protein that functions as a transcriptional repressor. This protein plays a role in diverse developmental contexts, including hematopoiesis and oncogenesis. It functions as part of a complex along with other cofactors to control histone modifications that lead to silencing of the target gene promoters. Mutations in this gene cause autosomal dominant severe congenital neutropenia, and also dominant nonimmune chronic idiopathic neutropenia of adults, which are heterogeneous hematopoietic disorders that cause predispositions to leukemias and infections.
cg19693031	1	145441552	TXNIP	Thioredoxin interacting protein	This gene encodes a thioredoxin-binding protein that is a member of the alpha arrestin protein family. Thioredoxin is a thiol-oxidoreductase that is a major regulator of cellular redox signaling which protects cells from oxidative stress. This protein inhibits the antioxidative function of thioredoxin resulting in the accumulation of reactive oxygen species and cellular stress. This protein also functions as a regulator of cellular metabolism and of endoplasmic reticulum (ER) stress. This protein may also function as a tumor suppressor.
cg24694018	1	145457621	POLR3GL	RNA polymerase III subunit G like	Among its related pathways are RNA Polymerase III Transcription Initiation and RIG-I/MDA5 mediated induction of IFN-alpha/beta pathways. Gene Ontology (GO) annotations related to this gene include *RNA polymerase III activity*.
cg20842915	7	39665132	RALA	RAS like proto-oncogene A	The product of this gene belongs to the small GTPase superfamily, Ras family of proteins. GTP-binding proteins mediate the transmembrane signaling initiated by the occupancy of certain cell surface receptors. This gene encodes a low molecular mass ras-like GTP-binding protein that shares about 50% similarity with other ras proteins
cg02508743	8	56903623	LYN	LYN proto-oncogene, Src family tyrosine kinase	This gene encodes a tyrosine protein kinase, which maybe involved in the regulation of mast cell degranulation, and erythroid differentiation.
cg27039118	8	116575902	TRPS1	Transcriptional repressor GATA binding 1	This gene encodes a transcription factor that represses GATA-regulated genes and binds to a dynein light chain protein. Binding of the encoded protein to the dynein light chain protein affects binding to GATA consensus sequences and suppresses its transcriptional activity. Defects in this gene are a cause of tricho-rhino-phalangeal syndrome (TRPS) types I-III
cg02716826	9	33447032	AQP3	Aquaporin 3 (Gill blood group)	This gene encodes the water channel protein aquaporin 3. Aquaporins are a family of small integral membrane proteins related to the major intrinsic protein, also known as aquaporin 0. Aquaporin 3 is localized at the basal lateral membranes of collecting duct cells in the kidney. In addition to its water channel function, aquaporin 3 has been found to facilitate the transport of nonionic small solutes such as urea and glycerol, but to a smaller degree. It has been suggested that water channels can be functionally heterogeneous and possess water and solute permeation mechanisms.
cg25953130	10	63753550	ARID5B	AT-rich interaction domain 5B	This gene encodes a member of the AT-rich interaction domain (ARID) family of DNA binding proteins. The encoded protein forms a histone H3K9Me2 demethylase complex with PHD finger protein 2 and regulates the transcription of target genes involved in adipogenesis and liver development. This gene also plays a role in cell growth and differentiation of B-lymphocyte progenitors, and single nucleotide polymorphisms in this gene are associated with acute lymphoblastic leukemia.
cg03190891	10	97201172	SORBS1	Sorbin and SH3 domain containing 1	This gene encodes a CBL-associated protein which functions in the signaling and stimulation of insulin. Mutations in this gene may be associated with human disorders of insulin resistance.
cg02079413	11	2986505	NAP1L4	Nucleosome assembly protein 1 like 4	This gene encodes a member of the nucleosome assembly protein (NAP) family which can interact with both core and linker histones. It can shuttle between the cytoplasm and nucleus, suggesting a role as a histone chaperone. This gene is one of several located near the imprinted gene domain of 11p15.5, an important tumor-suppressor gene region. Alterations in this region have been associated with the Beckwith-Wiedemann syndrome, Wilms tumor, rhabdomyosarcoma, adrenocortical carcinoma, and lung, ovarian, and breast cancer.
cg11250194	11	61601937	FADS2	Fatty acid desaturase 2	The protein encoded by this gene is a member of the fatty acid desaturase (FADS) gene family. Desaturase enzymes regulate unsaturation of fatty acids through the introduction of double bonds between defined carbons of the fatty acyl chain. FADS family members are considered fusion products composed of an N-terminal cytochrome b5-like domain and a C-terminal multiple membrane-spanning desaturase portion, both of which are characterized by conserved histidine motifs. This gene is clustered with family members at 11q12-q13.1; this cluster is thought to have arisen evolutionarily from gene duplication based on its similar exon/intron organization.

Supplemental Table 6. Functional description for protein coding genes annotated to diet-associated CpGs (continued)

CpG	CHR	Position	Gene	Full name	Description
cg11468085	11	67435577	ALDH3B2	Aldehyde dehydrogenase 3 family member B2	This gene encodes a member of the aldehyde dehydrogenase family, a group of isozymes that may play a major role in the detoxification of aldehydes generated by alcohol metabolism and lipid peroxidation. The gene of this particular family member is over 10 kb in length. Altered methylation patterns at this locus have been observed in spermatozoa derived from patients exhibiting reduced fecundity.
cg25909064	11	120082805	OAF	Out at first homolog	Diseases associated with OAF include Chronic Maxillary Sinusitis.
cg03646329	13	48987165	LPAR6	Lysophosphatidic acid receptor 6	The protein encoded by this gene belongs to the family of G-protein coupled receptors, that are preferentially activated by adenosine and uridine nucleotides. This gene aligns with an internal intron of the retinoblastoma susceptibility gene in the reverse orientation.
			RB1	RB transcriptional corepressor 1	The protein encoded by this gene is a negative regulator of the cell cycle and was the first tumor suppressor gene found. The encoded protein also stabilizes constitutive heterochromatin to maintain the overall chromatin structure. The active, hypophosphorylated form of the protein binds transcription factor E2F1. Defects in this gene are a cause of childhood cancer retinoblastoma (RB), bladder cancer, and osteogenic sarcoma.
cg16969872	13	79968324	RBM26	RNA binding motif protein 26	Gene Ontology (GO) annotations related to this gene include *nucleic acid binding* and *nucleotide binding*.
cg09940677	14	103415458	CDC42BPB	CDC42 binding protein kinase beta	This gene encodes a member of the serine/threonine protein kinase family. The encoded protein contains a Cdc42/Rac-binding p21 binding domain resembling that of PAK kinase. The kinase domain of this protein is most closely related to that of myotonic dystrophy kinase-related ROK. Studies of the similar gene in rat suggested that this kinase may act as a downstream effector of Cdc42 in cytoskeletal reorganization.
cg27118035	16	31891978	ZNF267	Zinc finger protein 267	Among its related pathways are Gene Expression. Gene Ontology (GO) annotations related to this gene include nucleic acid binding and DNA-binding transcription factor activity.
cg08732950	16	89023389	CBFA2T3	CBFA2/RUNX1 translocation partner 3	This gene encodes a member of the myeloid translocation gene family which interact with DNA-bound transcription factors and recruit a range of corepressors to facilitate transcriptional repression. The t(16;21)(q24;q22) translocation is one of the less common karyotypic abnormalities in acute myeloid leukemia. The translocation produces a chimeric gene made up of the 5'-region of the runt-related transcription factor 1 gene fused to the 3'-region of this gene. This gene is also a putative breast tumor suppressor.
cg02097604	17	17750910	TOM1L2	Target of myb1 like 2 membrane trafficking protein	This gene belongs to a small gene family whose members have an N-terminal VHS domain followed by a GAT domain; domains which typically participate in vesicular trafficking. The canonical protein encoded by this gene also has a C-terminal clathrin binding motif. This protein has been shown to interact with Tollip, clathrin and ubiquitin and is thought to play a role in endosomal sorting. This gene resides in the 3.7 Mb deletion of chromosome region 17p11.2 that is associated with Smith-Magenis syndrome.
cg16936953	17	57915665	VMP1	Vacuole membrane protein 1	This gene encodes a transmembrane protein that plays a key regulatory role in the process of autophagy. The ectopic overexpression of the encoded protein in cultured cells triggers autophagy even under nutrient-rich conditions. This gene is overexpressed in pancreatitis affected acinar cells where the encoded protein mediates sequestration and degradation of potentially deleterious activated zymogen granules in a process termed, zymophagy.
cg18181703	17	76354621	SOCS3	Suppressor of cytokine signaling 3	This gene encodes a member of the STAT-induced STAT inhibitor (SSI), also known as suppressor of cytokine signaling (SOCS), family. SSI family members are cytokine-inducible negative regulators of cytokine signaling. The expression of this gene is induced by various cytokines, including IL6, IL10, and interferon (IFN)-gamma. The protein encoded by this gene can bind to JAK2 kinase, and inhibit the activity of JAK2 kinase. Studies of the mouse counterpart of this gene suggested the roles of this gene in the negative regulation of fetal liver hematopoiesis, and placental development.
cg19202384	17	79894511	PYCR1	Pyrroline-5-carboxylate reductase 1	This gene encodes an enzyme that catalyzes the NAD(P)H-dependent conversion of pyrroline-5-carboxylate to proline. This enzyme may also play a physiologic role in the generation of NADP(+) in some cell types. The protein forms a homopolymer and localizes to the mitochondrion.
cg01294327	19	2291373	LINGO3	Leucine rich repeat and Ig domain containing 3	

Supplemental Table 6. Functional description for protein coding genes annotated to diet-associated CpGs (continued)

CpG	CHR	Position	Gene	Full name	Description
cg26470501	19	45252955	BCL3	BCL3 transcription coactivator	This gene is a proto-oncogene candidate. It is identified by its translocation into the immunoglobulin alpha-locus in some cases of B-cell leukemia. The protein encoded by this gene contains seven ankyrin repeats, which are most closely related to those found in I kappa B proteins. This protein functions as a transcriptional co-activator that activates through its association with NF-kappa B homodimers. The expression of this gene can be induced by NF-kappa B, which forms a part of the autoregulatory loop that controls the nuclear residence of p50 NF-kappa B.
cg08884571	19	45901453	PPP1R13L	Protein phosphatase 1 regulatory subunit 13 like	PPP1R13L is one of the most evolutionarily conserved inhibitors of p53 (TP53; MIM 191170), whereas ASPP1 (MIM 606455) and ASPP2 (MIM 602143) are activators of p53.

Supplemental Table 7. Differential gene expression of annotated genes for the 30 diet-associated CpGs

Category	Tissue	N_genes	N_overlap	P	adjusted P	Genes
Twoside	Spleen	4334	18	3.9E-06	0.0001	GFI1,POLR3GL,LYN,TRPS1,AQP3,ARID5B,SORBS1,FADS2,OAF,LPAR6,RBM26,ZNF267,CBFA2T3,TOM1L2,VMP1,LINGO3,BCL3,PPP1R13L
Twoside	Blood	5380	18	9.8E-05	0.003	GFI1,POLR3GL,RALA,LYN,TRPS1,AQP3,ARID5B,OAF,LPAR6,CDC42BPB,ZNF267,CBFA2T3,TOM1L2,VMP1,SOCS3,LINGO3,BCL3,PPP1R13L
Twoside	Lung	2898	13	1.1E-04	0.003	GFI1,RALA,LYN,AQP3,SORBS1,OAF,LPAR6,ZNF267,CBFA2T3,TOM1L2,VMP1,SOCS3,BCL3
Up	Spleen	2993	12	7.1E-04	0.02	GFI1,POLR3GL,LYN,AQP3,OAF,LPAR6,RBM26,ZNF267,CBFA2T3,VMP1,LINGO3,BCL3
Up	Lung	2157	10	7.9E-04	0.02	GFI1,RALA,LYN,AQP3,LPAR6,ZNF267,CBFA2T3,VMP1,SOCS3,BCL3
Twoside	Muscle	5676	17	8.5E-04	0.03	CACHD1,POLR3GL,RALA,AQP3,ARID5B,SORBS1,FADS2,OAF,LPAR6,RBM26,CDC42BPB,ZNF267,VMP1,SOCS3,PYCR1,BCL3,PPP1R13L
Up	Salivary_Gland	1113	7	1.1E-03	0.03	LYN,TRPS1,AQP3,ALDH3B2,LPAR6,PYCR1,PPP1R13L
Twoside	Small_Intestine	3161	12	1.2E-03	0.04	CACHD1,GFI1,LYN,TRPS1,AQP3,FADS2,OAF,CBFA2T3,TOM1L2,LINGO3,BCL3,PPP1R13L
Down	Testis	2284	10	1.2E-03	0.04	CACHD1,LYN,AQP3,ARID5B,SORBS1,OAF,VMP1,SOCS3,BCL3,PPP1R13L

Analysis was performed using GTEx v6 of 30 general tissue types. P was calculated using hypergeometric test through FUMA GWAS platform. Bonferroni correction was used to calculate adjusted P. Genes are annotated genes overlapping with differentially expressed genes in a given category

Supplemental Table 8. Traits associated with *cis*-meQTLs in NHGRI-EBI GWAS catalog

CpG	SNP	cismeQTL.P	GWAS.Reported.Gene	GWAS.Reported.Risk.Allele	GWAS.P	GWAS.Trait	GWAS.Trait.Description	PUBMD.ID
cg11250194	rs102274	2.0E-263	TMEM258	rs102274-C	8.0E-21	Serum metabolite concentrations in chronic kidney disease	(Phosphatidylcholine with diacyl residue sum C36:4)	29545352
cg11250194	rs102275	2.1E-255	TMEM258	rs102275-C	8.0E-17	Serum metabolite ratios in chronic kidney disease	(Phosphatidylcholine with acyl-alkyl residue sum C38:5/Phosphatidylcholine with acyl-alkyl residue sum C44:4)	29545352
cg11250194	rs102275	2.1E-255	TMEM258	rs102275-C	3.0E-32	Serum metabolite concentrations in chronic kidney disease	(Lysophosphatidylcholine with acyl residue C20:4)	29545352
cg11250194	rs102275	2.1E-255	C11orf0, C11orf9, FADS1, FADS2, FEN1	rs102275-T	7.0E-13	Palmitoleic acid (16:1n-7) levels		23362303
cg11250194	rs102275	2.1E-255	FADS1, FADS2, FADS3	rs102275-?	1.0E-203	Phospholipid levels (plasma)	(levels)	22359512
cg11250194	rs102275	2.1E-255	FTH1, INCENP, FADS1, FADS2, SCGB2A1, SCGB1D1, SCGB2A2, RAB3IL1, AHNAK, C11orf9, DAGLA, FEN1, FADS3, SYT7, BEST1	rs102275-C	7.0E-147	Plasma omega-6 polyunsaturated fatty acid levels (arachidonic acid)	(Conditioned on rs174547)	24823311
cg11250194	rs102275	2.1E-255	FADS1, FADS2, FADS3	rs102275-?	4.0E-264	Metabolite levels		22916037
cg07805029	rs11810217	2.7E-09	EVI5	rs11810217-A	6.0E-15	Multiple sclerosis		21833088
cg11468085	rs12288023	8.3E-13	ACY3	rs12288023-C	9.0E-16	Serum metabolite levels	(N-acetylphenylalanine)	24625756
cg16936953	rs1292053	8.6E-74	RPS6KB1, TUBD1	rs1292053-G	9.0E-13	Inflammatory bowel disease		23128233
cg27039118	rs13271228	2.6E-94	TRPS1	rs13271228-G	2.0E-24	Monocyte count		27863252
cg02508743	rs13273123	1.3E-76	PLAG1	rs13273123-G	1.0E-09	Height		19396169
cg11250194	rs1535	8.1E-267	FADS2	rs1535-G	3.0E-14	Serum metabolite ratios in chronic kidney disease	(Phosphatidylcholine with acyl-alkyl residue sum C36:3/Hydroxysphingomyelin with acyl residue sum C22:2)	29545352
cg11250194	rs1535	8.1E-267	FADS1, FADS2, FADS3	rs1535-A	5.0E-45	LDL cholesterol levels	(Trans-ethnic initial)	28334899
cg11250194	rs174479	4.8E-81	FADS1, FADS2, FADS3	rs174479-?	2.0E-14	Sphingolipid levels	(levels)	22359512
cg11250194	rs174528	2.5E-225	MYRF	rs174528-C	9.0E-14	Serum metabolite ratios in chronic kidney disease	(Phosphatidylcholine with diacyl residue sum C34:1/Phosphatidylcholine with diacyl residue sum C40:5)	29545352
cg11250194	rs174530	2.0E-227	MYRF	rs174530-G	2.0E-12	Serum metabolite ratios in chronic kidney disease	(Phosphatidylcholine with diacyl residue sum C32:3/Phosphatidylcholine with diacyl residue sum C34:4)	29545352
cg11250194	rs174533	1.6E-260	MYRF	rs174533-A	5.0E-17	Serum metabolite ratios in chronic kidney disease	(Phosphatidylcholine with acyl-alkyl residue sum C34:2/Phosphatidylcholine with acyl-alkyl residue sum C38:6)	29545352
cg11250194	rs174533	1.6E-260	FADS1, FADS2, FADS3	rs174533-A	1.0E-09	LDL cholesterol levels	(Asian initial, BMI unadjusted)	28334899
cg11250194	rs174533	1.6E-260	MYRF, TMEM258	rs174533-A	4.0E-21	Red blood cell count		27863252
cg11250194	rs174535	5.4E-261	FADS1	rs174535-T	3.0E-36	Blood metabolite levels	(1-linoleoylglycerophosphoethanolamine)	24816252
cg11250194	rs174536	1.8E-261	FEN1, FADS2, TMEM258, MYRF	rs174536-C	2.0E-30	Resting heart rate		27798624
cg11250194	rs174537	7.6E-263	MYRF	rs174537-T	2.0E-19	Serum metabolite ratios in chronic kidney disease	(Phosphatidylcholine with acyl-alkyl residue sum C36:5/Sphingomyelin with acyl residue sum C16:0)	29545352
cg11250194	rs174537	7.6E-263	NR	rs174537-G	9.0E-21	Colorectal cancer		26965516
cg11250194	rs174537	7.6E-263	MYRF	rs174537-T	5.0E-10	Serum metabolite concentrations in chronic kidney disease	(Phosphatidylcholine with acyl-alkyl residue sum C38:4)	29545352
cg11250194	rs174538	1.5E-273	FADS1	rs174538-A	9.0E-14	Blood metabolite levels	(docosapentaenoate (n3 DPA; 22:5n3))	24816252
cg11250194	rs174541	9.8E-254	FEN1, FADS1	rs174541-C	2.0E-18	Serum metabolite ratios in chronic kidney disease	(Phosphatidylcholine with acyl-alkyl residue sum C36:3/Phosphatidylcholine with acyl-alkyl residue sum C40:5)	29545352
cg11250194	rs174545	6.0E-265	FADS1	rs174545-G	7.0E-21	Serum metabolite ratios in chronic kidney disease	(Phosphatidylcholine with acyl-alkyl residue sum C36:3/Phosphatidylcholine with acyl-alkyl residue sum C40:1)	29545352
cg11250194	rs174546	6.2E-265	FADS1, FADS2, FADS3	rs174546-T	8.0E-28	HDL cholesterol		24097068
cg11250194	rs174546	6.2E-265	FADS1, FADS2, FADS3	rs174546-T	3.0E-37	Cholesterol, total		24097068
cg11250194	rs174546	6.2E-265	FADS1, FADS2, FADS3	rs174546-C	2.0E-22	Cholesterol, total		20686565
cg11250194	rs174546	6.2E-265	FADS1, FADS2, FADS3	rs174546-T	7.0E-38	Triglycerides		24097068
cg11250194	rs174546	6.2E-265	FADS1, FADS2, FADS3	rs174546-T	2.0E-22	HDL cholesterol		20686565
cg11250194	rs174546	6.2E-265	FADS1, FADS2, FADS3	rs174546-T	2.0E-39	LDL cholesterol		24097068
cg11250194	rs174546	6.2E-265	FADS1, FADS2, FADS3	rs174546-T	5.0E-24	Triglycerides		20686565
cg11250194	rs174546	6.2E-265	FADS1, FADS2, FADS3	rs174546-T	1.0E-21	LDL cholesterol		20686565
cg11250194	rs174547	6.7E-265	FADS1	rs174547-C	9.0E-116	Metabolic traits	(SM-3 + 152 other traits)	21886157
cg11250194	rs174548	6.1E-311	FADS1	rs174548-G	5.0E-14	Triglycerides		20864672
cg11250194	rs174548	6.1E-311	FADS1	rs174548-G	5.0E-12	Serum metabolite ratios in chronic kidney disease	(Phosphatidylcholine with acyl-alkyl residue sum C36:3/Phosphatidylcholine with acyl-alkyl residue sum C42:2)	29545352
cg11250194	rs174548	6.1E-311	FADS1	rs174548-G	1.0E-12	HDL cholesterol		20864672
cg11250194	rs174548	6.1E-311	FADS1	rs174548-C	1.0E-84	Blood metabolite levels	(arachidonate (20:4n6))	24816252
cg11250194	rs174549	3.9E-312	FADS1	rs174549-A	2.0E-18	Serum metabolite ratios in chronic kidney disease	(Phosphatidylcholine with diacyl residue sum C36:3/Phosphatidylcholine with diacyl residue sum C42:4)	29545352

Supplemental Table 8. Traits associated with *cis*-meQTLs in NHGRI-EBI GWAS catalog (continued)

CpG	SNP	cismeQTL.P	GWAS.Reported.Gene	GWAS.Reported.Risk.Allele	GWAS.P	GWAS.Trait	GWAS.Trait.Description	PUBMD.ID
cg11250194	rs174549	3.9E-312	FADS1, FADS2, FADS3, FEN1	rs174549-?	1.0E-20	Laryngeal squamous cell carcinoma		25194280
cg11250194	rs174549	3.9E-312	FADS1	rs174549-A	2.0E-12	Eosinophil counts		27863252
cg11250194	rs174550	4.9E-265	FADS1, FADS2, FADS3	rs174550-T	6.0E-30	HDL cholesterol levels	(Trans-ethnic initial)	28334899
cg11250194	rs174550	4.9E-265	FADS1	rs174550-T	7.0E-24	Blood metabolite levels	(adrenate (22:4n6))	24816252
cg11250194	rs174550	4.9E-265	FTH1, INCENP, FADS1, FADS2, SCGB1D1, SCGB2A1, C11orf9, DAGLA, FEN1, RAB3IL, C11orf10, FADS3, SYT7, BEST1, SCGB1D2	rs174550-T	4.0E-274	Plasma omega-6 polyunsaturated fatty acid levels (linoleic acid)		24823311
cg11250194	rs174554	1.3E-265	FADS1	rs174554-G	3.0E-17	Serum metabolite ratios in chronic kidney disease	(Phosphatidylcholine with diacyl residue sum C38:3/Phosphatidylcholine with acyl-alkyl residue sum C40:1)	29545352
cg11250194	rs174556	3.7E-312	FADS1	rs174556-T	2.0E-22	Blood metabolite levels	(eicosapentaenoate (EPA; 20:5n3))	24816252
cg11250194	rs174560	3.0E-311	FADS1	rs174560-C	4.0E-15	Serum metabolite ratios in chronic kidney disease	(Phosphatidylcholine with diacyl residue sum C40:5/Lysophosphatidylcholine with acyl residue C20:3)	29545352
cg11250194	rs174560	3.0E-311	FADS1	rs174560-C	3.0E-10	Serum metabolite concentrations in chronic kidney disease	(Phosphatidylcholine with diacyl residue sum C36:3)	29545352
cg11250194	rs174561	3.9E-312	MIR1908	rs174561-C	3.0E-17	Serum metabolite ratios in chronic kidney disease	(Phosphatidylcholine with diacyl residue sum C36:3/Phosphatidylcholine with diacyl residue sum C42:1)	29545352
cg11250194	rs174562	1.1E-265	FADS1	rs174562-G	3.0E-22	Serum metabolite ratios in chronic kidney disease	(Phosphatidylcholine with diacyl residue sum C36:3/Phosphatidylcholine with acyl-alkyl residue sum C40:5)	29545352
cg11250194	rs174564	9.6E-263	FADS1, FADS2	rs174564-G	3.0E-16	Serum metabolite ratios in chronic kidney disease	(Phosphatidylcholine with diacyl residue sum C30:0/Phosphatidylcholine with diacyl residue sum C32:2)	29545352
cg11250194	rs174566	4.2E-266	FADS1, FADS2	rs174566-G	2.0E-22	Serum metabolite ratios in chronic kidney disease	(Phosphatidylcholine with diacyl residue sum C36:5/Phosphatidylcholine with acyl-alkyl residue sum C36:2)	29545352
cg11250194	rs174566	4.2E-266	FADS1, FADS2	rs174566-G	5.0E-14	Serum metabolite concentrations in chronic kidney disease	(Phosphatidylcholine with diacyl residue sum C36:5)	29545352
cg11250194	rs174566	4.2E-266	FADS1, FADS2, MYRF, FEN1	rs174566-G	1.0E-12	Delta-5 desaturase activity response to n3-polyunsaturated fat supplement		29246731
cg11250194	rs174567	6.9E-265	FADS1, FADS2	rs174567-G	2.0E-13	Serum metabolite ratios in chronic kidney disease	(Phosphatidylcholine with acyl-alkyl residue sum C42:5/Phosphatidylcholine with acyl-alkyl residue sum C44:3)	29545352
cg11250194	rs174568	1.6E-263	FADS1, FADS2	rs174568-T	3.0E-08	Serum metabolite concentrations in chronic kidney disease	(Phosphatidylcholine with diacyl residue sum C34:4)	29545352
cg11250194	rs174568	1.6E-263	FADS1, FADS2	rs174568-T	3.0E-14	Serum metabolite ratios in chronic kidney disease	(Phosphatidylcholine with diacyl residue sum C38:0/Phosphatidylcholine with acyl-alkyl residue sum C38:6)	29545352
cg11250194	rs174570	9.3E-25	FADS2, FADS3	rs174570-G	4.0E-13	LDL cholesterol		19060911
cg11250194	rs174570	9.3E-25	FADS2, FADS3	rs174570-G	2.0E-10	Cholesterol, total		19060911
cg11250194	rs174574	5.1E-262	FADS2	rs174574-C	1.0E-13	Serum metabolite ratios in chronic kidney disease	(Phosphatidylcholine with acyl-alkyl residue sum C36:2/Sphingomyelin with acyl residue sum C16:1)	29545352
cg11250194	rs174576	1.3E-265	FADS2	rs174576-A	1.0E-30	Serum metabolite ratios in chronic kidney disease	(Phosphatidylcholine with diacyl residue sum C36:1/Phosphatidylcholine with diacyl residue sum C36:4)	29545352
cg11250194	rs174577	5.0E-265	FADS2	rs174577-A	2.0E-17	Iron status biomarkers (transferrin levels)		25352340
cg11250194	rs174577	5.0E-265	C11orf9, FADS2	rs174577-A	3.0E-08	P wave duration		24850809
cg11250194	rs174578	1.6E-263	FADS2	rs174578-A	9.0E-18	Serum metabolite ratios in chronic kidney disease	(Phosphatidylcholine with diacyl residue sum C36:1/Phosphatidylcholine with acyl-alkyl residue sum C38:5)	29545352
cg11250194	rs174578	1.6E-263	FADS1	rs174578-A	3.0E-24	Blood metabolite levels	(1-arachidonoylglycerophosphoinositol)	24816252
cg11250194	rs174580	1.1E-262	FADS2	rs174580-G	2.0E-28	Serum metabolite ratios in chronic kidney disease	(Phosphatidylcholine with diacyl residue sum C34:2/Phosphatidylcholine with acyl-alkyl residue sum C38:5)	29545352
cg11250194	rs174581	1.3E-262	FADS2	rs174581-A	3.0E-12	Serum metabolite ratios in chronic kidney disease	(Phosphatidylcholine with diacyl residue sum C34:2/Phosphatidylcholine with diacyl residue sum C40:4)	29545352
cg11250194	rs174581	1.3E-262	FADS1, FADS2, FEN1, MYRF, TMEM258	rs174581-A	1.0E-08	Male-pattern baldness		28196072
cg11250194	rs174583	2.3E-262	FADS2	rs174583-T	7.0E-18	Serum metabolite ratios in chronic kidney disease	(Phosphatidylcholine with acyl-alkyl residue sum C38:5/Sphingomyelin with acyl residue sum C16:0)	29545352
cg11250194	rs174592	1.3E-242	FADS2	rs174592-G	3.0E-13	Serum metabolite ratios in chronic kidney disease	(Phosphatidylcholine with acyl-alkyl residue sum C38:4/Sphingomyelin with acyl residue sum C16:0)	29545352
cg11250194	rs174594	4.1E-233	FADS2	rs174594-?	4.0E-09	Hemoglobin A1c levels		29403010
cg11250194	rs174594	4.1E-233	FADS2	rs174594-A	5.0E-13	Serum metabolite ratios in chronic kidney disease	(Phosphatidylcholine with acyl-alkyl residue sum C36:2/Phosphatidylcholine with acyl-alkyl residue sum C42:5)	29545352
cg11250194	rs174601	8.1E-237	FADS1	rs174601-T	8.0E-16	Blood metabolite levels	(stearidonate (18:4n3))	24816252
cg27039118	rs1905376	5.0E-87	TRPS1	rs1905376-G	1.0E-13	Red blood cell count		27863252
cg27039118	rs2049865	1.1E-91	TRPS1	rs2049865-A	6.0E-09	Diverticular disease		30177863

Supplemental Table 8. Traits associated with *cis*-meQTLs in NHGRI-EBI GWAS catalog (continued)

CpG	SNP	cismeQTL.P	GWAS.Reported.Gene	GWAS.Reported.Risk.Allele	GWAS.P	GWAS.Trait	GWAS.Trait.Description	PUBMD.ID
cg11250194	rs2071213	2.6E-11	FADS1, FADS2	rs2071213-A	3.0E-09	Moyamoya disease		29273593
cg27039118	rs2293889	1.5E-92	TRPS1	rs2293889-T	5.0E-19	HDL cholesterol levels	(Trans-ethnic initial)	28334899
cg27039118	rs2293889	1.5E-92	TRPS1	rs2293889-T	4.0E-17	HDL cholesterol		24097068
cg27039118	rs2293889	1.5E-92	TRPS1	rs2293889-T	6.0E-11	HDL cholesterol		20686565
cg11250194	rs2727270	5.4E-24	FTH1, INCENP, FADS1, FADS2, SCGB1D1, SCGB2A1, C11orf9, DAGLA, FEN1, RAB3IL, C11orf10, FADS3, SYT7, BEST1, SCGB1D2	rs2727270-T	3.0E-21	Plasma omega-6 polyunsaturated fatty acid levels (linoleic acid)	(Conditioned on rs174547)	24823311
cg11250194	rs2727271	2.6E-24	FADS1	rs2727271-A	3.0E-11	Blood metabolite levels	(X-12990--docosapentaenoic acid (n6-DPA))	24816252
cg11250194	rs28456	8.3E-315	FADS1, FADS2	rs28456-G	3.0E-15	Serum metabolite ratios in chronic kidney disease	(Phosphatidylcholine with diacyl residue sum C42:1/Phosphatidylcholine with acyl-alkyl residue sum C36:2)	29545352
cg11250194	rs28456	8.3E-315	C11orf9, DAGLA, C11orf10, FEN1, FADS1, FADS2	rs28456-G	6.0E-09	Bipolar disorder	(Japanese)	28115744
cg07805029	rs41286801	5.4E-09	EVI5	rs41286801-A	1.0E-26	Multiple sclerosis		24076602
cg25953130	rs4245595	2.6E-20	ARID5B	rs4245595-C	2.0E-09	Acute lymphoblastic leukemia (childhood)		25310577
cg11250194	rs4246215	1.2E-253	C11orf9, FADS1, FADS2	rs4246215-T	2.0E-15	Inflammatory bowel disease		23128233
cg26470501	rs519113	9.7E-10	APOE, PVRL2, TOMM40	rs519113-C	8.0E-11	HDL cholesterol		21909109
cg05232694	rs6512645	3.6E-08	intergenic	rs6512645-A	2.0E-13	Mean corpuscular volume		27863252
cg05232694	rs6512645	3.6E-08	intergenic	rs6512645-A	1.0E-14	Mean corpuscular hemoglobin		27863252
cg07805029	rs6689470	5.2E-09	EVI5	rs6689470-A	4.0E-10	Multiple sclerosis		27386562
cg27039118	rs6982226	3.7E-09	NR	rs6982226-T	1.0E-10	Male-pattern baldness		28196072
cg25953130	rs7089424	1.3E-20	ARID5B	rs7089424-T	2.0E-62	Acute lymphoblastic leukemia (childhood)		29348612
cg25953130	rs7089424	1.3E-20	ARID5B	rs7089424-C	7.0E-19	Acute lymphoblastic leukemia (childhood)		19684604
cg16936953	rs7212590	3.2E-40	VMP1	rs7212590-T	2.0E-30	Monocyte count		27863252
cg02508743	rs7833986	8.7E-76	PLAG1	rs7833986-A	8.0E-10	Height		20189936
cg11250194	rs7943728	1.9E-274	MYRF	rs7943728-A	4.0E-19	Serum metabolite ratios in chronic kidney disease	(Phosphatidylcholine with acyl-alkyl residue sum C34:2/Phosphatidylcholine with acyl-alkyl residue sum C36:3)	29545352
cg16969872	rs7988232	1.3E-10	RBM26	rs7988232-A	3.0E-09	Systolic blood pressure		30224653
cg27039118	rs800890	4.7E-08	TRPS1	rs800890-C	2.0E-08	Male-pattern baldness		28196072
cg16936953	rs8070345	1.4E-75	VMP1	rs8070345-A	2.0E-23	Multiple sclerosis		24076602
cg05232694	rs913678	5.5E-24	intergenic	rs913678-C	2.0E-36	Monocyte count		27863252
cg11250194	rs968567	6.2E-281	FADS1, FADS2, FADS3	rs968567-C	2.0E-08	Rheumatoid arthritis	(EA)	24390342
cg11250194	rs968567	6.2E-281	FADS1	rs968567-T	3.0E-19	Blood metabolite levels	(1-eicosatrienoylglycerophosphocholine)	24816252
cg11250194	rs968567	6.2E-281	FADS2	rs968567-T	5.0E-14	Serum metabolite ratios in chronic kidney disease	(Phosphatidylcholine with diacyl residue sum C42:0/Phosphatidylcholine with acyl-alkyl residue sum C44:6)	29545352
cg11250194	rs97384	8.2E-223	FADS2	rs97384-C	4.0E-13	Serum metabolite ratios in chronic kidney disease	(Phosphatidylcholine with diacyl residue sum C34:2/Phosphatidylcholine with acyl-alkyl residue sum C42:5)	29545352
cg11250194	rs99780	6.5E-265	FADS2	rs99780-T	2.0E-16	Serum metabolite ratios in chronic kidney disease	(Phosphatidylcholine with diacyl residue sum C34:4/Phosphatidylcholine with acyl-alkyl residue sum C36:2)	29545352

Supplemental Table 9. Association between diet-associated CpGs with cardiovascular risk factors reported in EWAS catalog

CpG	Direction. Diet	Trait	Direction. Trait	Beta	SE	P	N	Age	Tissue	PMID
cg01294327	+	Smoking	-	-0.0404	0.0055	2.0E-13	9389	60.3	Whole blood, CD4+ T cells or monocytes	27651444
cg01294327	+	Smoking	-	-0.61493	0.0913	1.6E-11	1597	67	Peripheral blood	28686328
cg01294327	+	Smoking	-	-0.236	0.046823	4.7E-07	596	53.9	Whole blood	26756918
cg01940273	+	Smoking	-	-1.80094	0.0886	6.8E-92	1597	67	Peripheral blood	28686328
cg01940273	+	Smoking				1.8E-78	721	52.6	Peripheral blood	26864933
cg01940273	+	Smoking	-	-0.61	0.0336	7.4E-74	582	54	Whole blood	25556184
cg01940273	+	Tobacco smoking				1.4E-64	454	52	Whole blood	28225026
cg01940273	+	Smoking	-	-0.08675	0.0058	1.0E-50	535	70	Whole blood	25249537
cg01940273	+	Smoking				5.5E-43	468	53	Whole blood	23691101
cg01940273	+	Smoking				5.6E-43	645	63.2	Peripheral blood	26829059
cg01940273	+	Smoking	-	-0.0815	0.0067	2.0E-34	9389	60.3	Whole blood, CD4+ T cells or monocytes	27651444
cg01940273	+	Smoking	-	-0.539	0.044994	4.6E-33	596	53.9	Whole blood	26756918
cg01940273	+	Smoking				3.7E-32	464	55.39	Whole Blood	25424692
cg01940273	+	Smoking				2.5E-10	111	48.4	Peripheral Blood	24559495
cg01940273	+	Alcohol consumption per day	-	-9.79E-05	1.56E-05	3.8E-10	2423	56	Whole blood	27843151
cg01940273	+	Smoking				5.9E-09	190	55	Peripheral blood	23175441
cg01940273	+	Alcohol consumption der day	-	-0.00046	7.99E-05	1.1E-08	9643	64.2	Whole blood	27843151
cg01940273	+	Smoking				1.9E-08	100		Peripheral blood	26441693
cg02079413	-	Body mass index	+	7.4	1.7	2.1E-05	4851	54.5	Whole blood	28002404
cg02508743	-	Smoking	+	0.011	0.0016	8.1E-12	9389	60.3	Whole blood, CD4+ T cells or monocytes	27651444
cg02716826	+	C-reactive protein	-	-0.0037	0.000463	1.3E-15	8863	65	Whole blood	27955697
cg02716826	+	Age				5.4E-15	1264	60	Monocytes	25404168
cg02716826	+	C-reactive protein	-	-0.031	0.005	2.7E-08	1741	60.9	Peripheral blood	27824951
cg02716826	+	Body mass index	-	-14.6	2.75	1.1E-07	2707	57.9	Whole blood	28002404
cg02716826	+	Smoking	-	-0.0096	0.002	2.5E-06	9389	60.3	Whole blood, CD4+ T cells or monocytes	27651444
cg02716826	+	Body mass index	-	-0.0236	0.00957	1.5E-02	442	67	Whole blood	28095459
cg03190891	-	Smoking	+	0.0114	0.0016	1.8E-12	9389	60.3	Whole blood, CD4+ T cells or monocytes	27651444
cg03190891	-	Smoking	+	0.581587	0.0928	3.6E-10	1597	67	Peripheral blood	28686328
cg03190891	-	Alcohol consumption per day	+	5.05E-05	9.72E-06	2.0E-07	2423	56	Whole blood	27843151
cg03190891	-	Smoking				1.5E-06	721	52.6	Peripheral blood	26864933
cg03190891	-	Alcohol consumption der day	+	0.000151	3.99E-05	1.6E-04	9643	64.2	Whole blood	27843151
cg03646329	+	Smoking	-	-0.0244	0.0039	5.6E-10	9389	60.3	Whole blood, CD4+ T cells or monocytes	27651444
cg03646329	+	Smoking	-	-0.196	0.039837	8.7E-07	596	53.9	Whole blood	26756918
cg03646329	+	Smoking				2.0E-06	721	52.6	Peripheral blood	26864933
cg03646329	+	Smoking				3.1E-03	1531	61.5	Whole blood	23691101
cg04885881	+	Smoking	-	-0.29	0.0277	1.4E-25	582	54	Whole blood	25556184
cg04885881	+	Age	-	-0.0209	0.002119	3.0E-23	1366	73	Whole blood	28811542
cg04885881	+	Tobacco smoking				8.5E-20	454	52	Whole blood	28225026
cg04885881	+	Smoking	-	0	0	3.6E-19	13474	62.9	Whole blood, CD4+ T cells or monocytes	27651444
cg04885881	+	Smoking				6.6E-13	468	53	Whole blood	23691101

Supplemental Table 9. Association between diet-associated CpGs with cardiovascular risk factors reported in EWAS catalog (continued)

CpG	Direction. Diet	Trait	Direction. Trait	Beta	SE	P	N	Age	Tissue	PMID
cg04885881	+	Smoking				4.1E-11	717	53.4	Peripheral blood	26864933
cg04885881	+	Smoking	-	-0.58137	0.0903	1.2E-10	1597	67	Peripheral blood	28686328
cg04885881	+	Smoking	-	-0.022	0.0036	1.5E-09	421	43.8	Whole Blood	24334605
cg04885881	+	Alcohol consumption per day	-	-9.12E-05	1.54E-05	3.1E-09	2423	56	Whole blood	27843151
cg04885881	+	Smoking				6.9E-09	449	63.1	Peripheral blood	26829059
cg04885881	+	Alcohol consumption der day	-	-0.00023	5.95E-05	1.6E-04	9643	64.2	Whole blood	27843151
cg05232694	+	C-reactive protein	-	-0.0053	0.000894	3.0E-09	8863	65	Whole blood	27955697
cg05232694	+	Smoking	-	0	0	7.3E-07	13474	62.9	Whole blood, CD4+ T cells or monocytes	27651444
cg05232694	+	Alcohol consumption der day	-	-0.0005	0.000136	2.2E-04	9643	64.2	Whole blood	27843151
cg08732950	-	Smoking	+	0.0088	0.0016	2.3E-08	9389	60.3	Whole blood, CD4+ T cells or monocytes	27651444
cg09940677	-	Smoking	+	0.0081	9.00E-04	1.0E-18	9389	60.3	Whole blood, CD4+ T cells or monocytes	27651444
cg09940677	-	Alcohol consumption per day	+	5.70E-05	9.00E-06	2.4E-10	2423	56	Whole blood	27843151
cg09940677	-	C-reactive protein	+	0.0023	0.000856	7.2E-03	4111	59.7	Whole blood	27955697
cg11468085	-	Alcohol consumption per day	+	4.59E-05	8.87E-06	2.3E-07	2423	56	Whole blood	27843151
cg13074055	+	Smoking	-	-0.322	0.049434	7.3E-11	596	53.9	Whole blood	26756918
cg13074055	+	Smoking	-	-0.59383	0.0928	1.5E-10	1597	67	Peripheral blood	28686328
cg13074055	+	Smoking	-	-0.0403	0.0067	1.6E-09	9389	60.3	Whole blood, CD4+ T cells or monocytes	27651444
cg13074055	+	Smoking				2.6E-09	721	52.6	Peripheral blood	26864933
cg13074055	+	Tobacco smoking				3.0E-05	454	52	Whole blood	28225026
cg16936953	+	C-reactive protein	-	-0.0077	0.000816	3.7E-21	8863	65	Whole blood	27955697
cg16936953	+	Age				1.5E-09	227		T cells	25404168
cg16936953	+	Smoking	-	-0.0118	0.0022	7.0E-08	9389	60.3	Whole blood, CD4+ T cells or monocytes	27651444
cg16969872	+	Smoking				2.7E-08	721	52.6	Peripheral blood	26864933
cg16969872	+	Smoking	-	-0.017	0.0031	5.2E-08	9389	60.3	Whole blood, CD4+ T cells or monocytes	27651444
cg16969872	+	Tobacco smoking				3.5E-06	454	52	Whole blood	28225026
cg16969872	+	Alcohol consumption der day	-	-0.00026	5.87E-05	9.3E-06	9643	64.2	Whole blood	27843151
cg16969872	+	Smoking				2.6E-02	1531	61.5	Whole blood	23691101
cg18181703	+	Body mass index				4.0E-35	10238	54.5	Whole blood	28002404
cg18181703	+	C-reactive protein	-	-0.0091	0.001268	7.1E-13	4111	59.7	Whole blood	27955697
cg18181703	+	Alcohol consumption per day	-	-8.36E-05	1.29E-05	9.2E-11	2423	56	Whole blood	27843151
cg18181703	+	Body mass index	-	-0.0022	0.00033	1.1E-10	510	29	Whole blood	27826092
cg18181703	+	Type II diabetes	-	-0.05129	0.0099	2.1E-07	2664	50	Peripheral blood	26095709
cg18181703	+	Body mass index	-	-0.15	0.031	1.8E-06	1052	39	Peripheral blood mononuclear cells	27564309
cg18181703	+	Smoking	-	-0.0036	0.0012	2.4E-03	13474	62.9	Whole blood, CD4+ T cells or monocytes	27651444
cg18181703	+	Alcohol consumption der day	-	-0.00013	4.53E-05	4.8E-03	9643	64.2	Whole blood	27843151
cg19202384	+	Alcohol consumption der day	-	-0.00026	4.78E-05	4.6E-08	9643	64.2	Whole blood	27843151
cg19202384	+	Smoking	-	-0.0068	0.0016	1.3E-05	9389	60.3	Whole blood, CD4+ T cells or monocytes	27651444
cg19202384	+	Tobacco smoking				5.0E-05	454	52	Whole blood	28225026
cg19202384	+	Smoking				5.1E-05	721	52.6	Peripheral blood	26864933
cg19693031	+	Serum triglycerides	-	-0.039	0.004	0.0E+00	3187	65.7	Whole blood	28173150

Supplemental Table 9. Association between diet-associated CpGs with cardiovascular risk factors reported in EWAS catalog (continued)

CpG	Direction. Diet	Trait	Direction. Trait	Beta	SE	P	N	Age	Tissue	PMID
cg19693031	+	HbA1c	-	-0.129	0.016	7.3E-16	268	74	Peripheral blood	26643952
cg19693031	+	Alcohol consumption der day	-	-0.00044	6.46E-05	1.0E-11	9643	64.2	Whole blood	27843151
cg19693031	+	Triglycerides	-	-0.0133	0.0025	6.7E-08	1485	64	Whole blood	28194238
cg19693031	+	Hypertriglyceridemic waist	+	0.302844	0.05726	1.2E-07	850	46.8	Peripheral blood	26798409
cg19693031	+	Serum triglycerides	-	-0.01093	0.002263	1.4E-06	1955	61	Whole blood	28213390
cg19693031	+	Type II diabetes	-	-0.3	0.0621	1.4E-06	167	74.8	Peripheral blood	26643952
cg19693031	+	Alcohol consumption per day	-	-0.0005	0.000115	1.7E-05	1251	60	CD14+ monocytes	27843151
cg19693031	+	Type II diabetes	-	-0.031	0.0072	1.9E-05	527	63	Whole blood	26433941
cg19693031	+	Type II diabetes	-	-0.0408	0.0097	2.5E-05	1141	59	Peripheral blood	26095709
cg19693031	+	Triglycerides				5.7E-03	971	62	Whole blood	25583993
cg20842915	+	Smoking	-	-0.52763	0.0952	3.0E-08	1597	67	Peripheral blood	28686328
cg20842915	+	Smoking	-	-0.0088	0.0028	1.8E-03	9389	60.3	Whole blood, CD4+ T cells or monocytes	27651444
cg24694018	+	Smoking	-	-0.0052	0.0014	2.0E-04	9389	60.3	Whole blood, CD4+ T cells or monocytes	27651444
cg24735226	-	Smoking	+	0.0052	0.0015	6.0E-04	13474	62.9	Whole blood, CD4+ T cells or monocytes	27651444
cg25909064	+	Age	-	-0.0055	0.001464	8.6E-05	1366	73	Whole blood	28811542
cg25953130	+	Smoking	-	-0.21	0.0362	6.5E-09	582	54	Whole blood	25556184
cg25953130	+	Smoking				2.1E-07	468	53	Whole blood	23691101
cg25953130	+	Smoking				4.7E-07	721	52.6	Peripheral blood	26864933
cg25953130	+	Smoking	-	-0.0084	0.0019	8.1E-06	13474	62.9	Whole blood, CD4+ T cells or monocytes	27651444
cg26470501	+	C-reactive protein	-	-0.0045	0.000488	2.9E-20	8863	65	Whole blood	27955697
cg26470501	+	Body mass index				8.3E-17	10238	54.5	Whole blood	28002404
cg26470501	+	Body mass index				2.6E-10	3743	69.2	Whole blood	28095459
cg26470501	+	Smoking	-	-0.09	0.0168	7.8E-08	582	54	Whole blood	25556184
cg26470501	+	Smoking	-	-0.0123	0.0024	2.3E-07	9389	60.3	Whole blood, CD4+ T cells or monocytes	27651444
cg26470501	+	Alcohol consumption per day	-	-4.35E-05	9.91E-06	1.2E-05	2423	56	Whole blood	27843151
cg26470501	+	Tobacco smoking				5.0E-05	454	52	Whole blood	28225026
cg27039118	+	Smoking	-	-0.0158	0.003	1.4E-07	9389	60.3	Whole blood, CD4+ T cells or monocytes	27651444
cg27118035	-	Smoking	+	0.0136	0.0014	2.4E-22	9389	60.3	Whole blood, CD4+ T cells or monocytes	27651444
cg27118035	-	Alcohol consumption per day	+	5.91E-05	9.02E-06	5.9E-11	2423	56	Whole blood	27843151
cg27118035	-	Smoking	+	0.538722	0.0935	8.4E-09	1597	67	Peripheral blood	28686328
cg27118035	-	Alcohol consumption der day	+	0.000103	2.52E-05	4.6E-05	9643	64.2	Whole blood	27843151

Supplemental Table 10. Mendelian Randomization analysis for association between diet-associated CpGs and cardiovascular risk factors (CpG to CVD trait)

		BMI						WHRadjBMI						TG						HDL						T2D					
CpG	Direction.Diet	Direction	Beta	SE	P	n.SNP	P for Pleiotropy	Direction	Beta	SE	P	n.SNP	P for Pleiotropy	Direction	Beta	SE	P	n.SNP	P for Pleiotropy	Direction	Beta	SE	P	n.SNP	P for Pleiotropy	Direction	Beta	SE	P	n.SNP	P for Pleiotropy
cg16969872	+	-	-0.11	0.20	0.59	3	0.67	-	-0.22	0.50	0.65	3	0.83	+	0.02	1.49	0.99	2		+	0.05	0.80	0.95	2		-	-0.10	0.70	0.88	4	0.70
cg01940273	+	-	-0.42	0.22	0.06	3	0.32	-	-0.17	0.46	0.71	3	0.78	+	0.09	0.55	0.87	2		-	-0.83	0.56	0.14	2		-	-0.23	0.66	0.73	3	0.45
cg02716826	+	+	0.75	0.71	0.29	1		+	0.75	1.78	0.67	1		+	0.40	1.94	0.84	1		+	0.04	1.98	0.98	1		-	-0.76	2.61	0.77	1	
cg04885881	+	-	-5.60	0.85	4.7E-11	1		+	4.21	2.06	0.04	1		-	-1.48	2.33	0.53	1		-	-1.97	2.37	0.41	1		+	12.09	3.13	0.0001	1	
cg16936953	+	-	-0.35	0.11	0.001	5	0.36	+	0.51	0.25	0.04	6	0.57	+	0.50	0.26	0.06	5	0.77	-	-0.49	0.27	0.07	5	0.81	+	1.34	0.79	0.09	3	0.91
cg25909064	+	+	0.57	0.74	0.44	3	0.73	-	-0.48	0.69	0.49	4	0.33	+	1.01	0.67	0.13	4	0.61	-	-0.41	0.81	0.61	4	0.93	+	0.27	2.38	0.91	3	0.64
cg25953130	+	+	0.78	0.27	0.004	1		-	-1.65	0.67	0.01	1		+	0.33	0.75	0.66	1		+	0.72	0.75	0.34	1							
cg01294327	+	-	-0.01	0.13	0.93	2																				-	-0.16	0.97	0.87	2	
cg05232694	+	+	0.22	0.19	0.24	5	0.12	-	-0.68	0.36	0.06	5	0.38	-	-0.33	0.39	0.40	5	0.86	+	0.51	0.40	0.21	5	0.88	-	-1.89	1.40	0.18	4	0.35
cg07805029	+	-	-2.15	0.40	6.7E-08	1		+	3.58	4.58	0.43	1		+	4.48	3.03	0.14	1		+	4.16	3.03	0.17	1							
cg11250194	+	-	-0.33	0.16	0.03	12	0.69	+	0.26	0.24	0.28	13	0.63	-	-3.01	0.39	1.5E-14	18	0.86	+	2.33	0.41	1.7E-08	18	1.00	+	1.12	0.38	0.003	17	0.71
cg26470501	+	+	0.77	0.19	6.5E-05	4	0.63	+	0.58	0.50	0.25	4	0.95	-	-0.87	0.65	0.18	4	0.71	-	-0.44	1.13	0.69	4	0.87	+	2.72	0.81	0.0009	2	
cg27039118	+	+	0.63	0.55	0.25	4	0.23	-	-0.03	0.50	0.95	4	0.53	+	0.58	0.65	0.37	4	0.12	-	-0.71	0.60	0.24	4	0.17	+	0.92	0.65	0.16	5	0.49
cg08732950	-	-	-0.71	0.26	0.007	5	0.41	-	-0.81	0.67	0.22	6	0.54	-	-2.15	1.11	0.05	5	0.39	+	2.63	1.03	0.01	5	0.21	-	-3.68	1.35	0.006	7	0.75
cg02079413	-	+	0.73	0.15	1.0E-06	5	0.62	+	0.43	0.37	0.25	5	0.77	-	-0.14	0.41	0.73	5	0.81	+	2.14	0.69	0.002	5	0.30	-	-0.77	0.75	0.30	4	0.51
cg27118035	-							-	-2.99	2.08	0.15	1		-	-2.62	2.28	0.25	1		-	-2.38	2.42	0.32	1							
cg02508743	-	+	0.28	0.12	0.02	6	0.95	-	-0.27	0.30	0.37	6	0.60	-	-0.33	0.34	0.33	5	0.72	+	0.06	0.35	0.86	5	0.90	+	1.08	0.55	0.047	9	0.19
cg03190891	-	+	0.62	0.77	0.42	3	0.67	+	0.97	0.82	0.23	3	0.73	-	-0.05	0.77	0.95	3	0.53	-	-0.84	0.79	0.29	3	0.97	+	3.96	1.05	0.0002	3	0.88
cg08884571	-	+	0.01	0.39	0.97	1		+	2.40	1.08	0.03	1														-	-3.68	1.21	0.002	1	
cg09940677	-	-	-0.04	1.27	0.98	1		+	0.83	3.06	0.79	1		+	2.86	3.29	0.39	1		+	1.11	3.33	0.74	1		+	1.31	4.76	0.78	1	
cg11468085	+	+	0.31	0.28	0.27	2		-	-1.25	0.79	0.12	2		-	-0.11	0.78	0.89	2		+	0.21	1.15	0.85	2		+	0.32	1.88	0.86	2	
cg24735226	-	+	0.13	0.12	0.27	5	0.32	-	-0.21	0.26	0.42	6	0.41	+	0.20	0.30	0.50	5	0.63	-	-0.75	0.40	0.06	5	0.09	-	-0.17	0.33	0.61	11	0.51

Supplemental Table 11. Mendelian Randomization analysis for association between diet-associated CpGs and cardiovascular risk factors (CVD trait to CpG)

		BMI					WHR					TG					HDL					T2D				
CpG	Direction.Diet	Direction	Beta	SE	P	P for Pleiotropy	Direction	Beta	SE	P	P for Pleiotropy	Direction	Beta	SE	P	P for Pleiotropy	Direction	Beta	SE	P	P for Pleiotropy	Direction	Beta	SE	P	P for Pleiotropy
cg01294327	+	+	6.E-20	3.E-18	0.99	0.07	-	-8.E-21	8.E-18	1.00	0.74	-	-7.E-18	7.E-18	0.36	0.75	-	-8.E-19	4.E-18	0.84	0.91	-	-2.E-18	2.E-18	0.30	0.22
cg01940273	+	-	-3.E-03	2.E-03	0.11	0.72	+	5.E-04	4.E-03	0.90	0.45	+	3.E-03	4.E-03	0.45	0.68	+	2.E-03	2.E-03	0.33	0.94	-	-7.E-04	8.E-04	0.35	0.83
cg02079413	-	+	2.E-04	7.E-04	0.77	0.21	+	1.E-03	2.E-03	0.46	0.35	+	9.E-04	1.E-03	0.51	0.27	-	-1.E-03	7.E-04	0.15	0.44	+	5.E-04	3.E-04	0.12	0.07
cg02097604	+	+	3.E-05	8.E-04	0.97	0.43	+	5.E-04	2.E-03	0.79	0.09	+	7.E-04	2.E-03	0.70	0.92	+	1.E-03	9.E-04	0.19	0.40	-	-7.E-05	4.E-04	0.85	0.52
cg02508743	-	-	-1.E-04	1.E-03	0.92	0.07	+	1.E-03	3.E-03	0.60	0.85	+	1.E-03	2.E-03	0.60	0.46	+	5.E-04	1.E-03	0.69	0.65	+	2.E-04	5.E-04	0.73	0.64
cg02716826	+	-	-2.E-03	9.E-04	0.06	0.48	+	4.E-03	2.E-03	0.06	0.94	-	-5.E-04	2.E-03	0.81	0.79	+	4.E-04	1.E-03	0.69	0.11	-	-5.E-04	4.E-04	0.19	0.79
cg03190891	-	+	1.E-03	1.E-03	0.18	0.20	-	-1.E-03	2.E-03	0.57	0.11	+	4.E-04	2.E-03	0.86	0.42	+	2.E-03	1.E-03	0.10	0.92	-	-5.E-04	5.E-04	0.28	1.00
cg03646329	+	+	2.E-03	2.E-03	0.24	0.50	+	6.E-03	5.E-03	0.24	0.60	-	-4.E-04	4.E-03	0.93	0.34	-	-3.E-03	2.E-03	0.14	0.04	+	5.E-04	9.E-04	0.57	0.80
cg04885881	+	-	-6.E-04	1.E-03	0.57	0.23	+	6.E-04	3.E-03	0.82	0.32	+	8.E-04	2.E-03	0.74	0.11	+	9.E-05	1.E-03	0.95	0.03	-	-3.E-04	5.E-04	0.59	0.11
cg05232694	+	-	-9.E-04	2.E-03	0.67	0.66	+	6.E-03	5.E-03	0.26	0.42	-	-4.E-03	5.E-03	0.42	0.82	+	9.E-04	2.E-03	0.71	0.08	-	-7.E-04	1.E-03	0.48	0.87
cg07805029	+	-	-3.E-03	2.E-03	0.28	0.84	+	1.E-02	6.E-03	0.07	0.20	-	-5.E-03	5.E-03	0.32	0.73	+	2.E-03	3.E-03	0.56	0.44	+	2.E-03	1.E-03	0.11	0.77
cg08732950	-	+	5.E-04	6.E-04	0.43	0.30	+	1.E-03	1.E-03	0.34	0.26	-	-2.E-03	1.E-03	0.13	0.82	+	2.E-04	7.E-04	0.75	0.54	-	-8.E-05	3.E-04	0.78	0.40
cg08884571	-	+	7.E-04	1.E-03	0.63	0.83	+	2.E-03	3.E-03	0.58	0.31	-	-1.E-03	3.E-03	0.68	0.76	-	-1.E-03	2.E-03	0.49	0.87	+	7.E-04	6.E-04	0.28	0.84
cg09940677	-	+	2.E-04	7.E-04	0.81	0.48	+	8.E-05	2.E-03	0.96	0.48	-	-2.E-03	1.E-03	0.27	0.15	-	-3.E-04	8.E-04	0.69	0.62	+	3.E-05	3.E-04	0.92	0.55
cg11250194	+	-	-3.E-04	8.E-04	0.69	0.50	+	2.E-03	2.E-03	0.38	0.87	-	-2.E-02	1.E-02	0.20	0.74	+	5.E-04	8.E-04	0.52	0.26	+	3.E-04	3.E-04	0.39	0.24
cg11468085	+	+	8.E-04	1.E-03	0.45	0.70	-	-1.E-03	2.E-03	0.61	0.21	+	2.E-03	2.E-03	0.47	0.39	-	-1.E-03	1.E-03	0.37	0.18	-	-2.E-04	4.E-04	0.66	0.48
cg13074055	+	+	4.E-03	2.E-03	0.07	0.66	+	3.E-03	5.E-03	0.49	0.80	+	6.E-03	4.E-03	0.16	0.40	-	-2.E-03	2.E-03	0.39	0.48	-	-4.E-05	9.E-04	0.96	0.45
cg16936953	+	-	-1.E-03	2.E-03	0.56	0.91	+	1.E-04	4.E-03	0.97	0.75	+	3.E-03	3.E-03	0.37	0.66	+	8.E-04	2.E-03	0.65	0.92	+	4.E-04	7.E-04	0.60	0.28
cg16969872	+	-	-5.E-04	1.E-03	0.63	0.83	+	1.E-03	2.E-03	0.66	0.29	-	-4.E-03	2.E-03	0.09	0.18	-	-5.E-05	1.E-03	0.97	0.88	-	-5.E-04	5.E-04	0.32	0.78
cg18181703	+	-	-2.E-03	1.E-03	0.04	0.35	+	3.E-03	3.E-03	0.21	0.96	-	-6.E-04	2.E-03	0.81	0.78	+	7.E-05	1.E-03	0.96	0.85	+	1.E-04	5.E-04	0.81	0.99
cg19202384	+	-	-5.E-04	1.E-03	0.68	0.22	+	3.E-03	3.E-03	0.31	0.13	-	-5.E-03	2.E-03	0.07	0.27	+	1.E-03	1.E-03	0.47	0.19	-	-5.E-04	5.E-04	0.31	0.38
cg19693031	+	+	3.E-04	1.E-03	0.83	0.76	-	-1.E-03	3.E-03	0.66	0.74	+	2.E-03	3.E-03	0.52	0.26	+	6.E-04	2.E-03	0.72	0.19	-	-1.E-03	6.E-04	0.08	0.65
cg20842915	+	+	1.E-04	2.E-03	0.96	0.27	-	-4.E-03	5.E-03	0.45	0.27	-	-1.E-03	7.E-03	0.84	0.56	-	-2.E-03	3.E-03	0.50	0.88	-	-3.E-04	1.E-03	0.76	0.23
cg24694018	+	+	8.E-04	9.E-04	0.33	0.78	-	-6.E-04	2.E-03	0.77	0.98	+	2.E-04	2.E-03	0.93	0.47	+	2.E-03	9.E-04	0.08	0.41	-	-4.E-04	4.E-04	0.33	0.08
cg24735226	-	-	-1.E-06	9.E-04	1.00	0.16	-	-1.E-03	2.E-03	0.56	0.44	-	-3.E-03	2.E-03	0.08	0.12	+	1.E-03	9.E-04	0.24	0.48	+	6.E-04	4.E-04	0.13	0.20
cg25909064	+	-	-8.E-04	8.E-04	0.30	0.73	-	-1.E-03	2.E-03	0.41	0.39	-	-7.E-04	2.E-03	0.65	0.60	-	-6.E-04	8.E-04	0.50	0.53	-	-2.E-05	3.E-04	0.94	0.55
cg25953130	+	+	1.E-03	2.E-03	0.57	0.75	-	-1.E-02	4.E-03	0.02	0.56	+	2.E-03	4.E-03	0.70	0.87	+	3.E-03	2.E-03	0.20	0.96	-	-7.E-04	8.E-04	0.42	0.55
cg26470501	+	-	-1.E-03	1.E-03	0.22	0.27	-	-5.E-04	3.E-03	0.85	0.16	-	-3.E-03	2.E-03	0.16	0.77	-	-2.E-04	1.E-03	0.84	0.68	+	3.E-04	5.E-04	0.57	0.91
cg27039118	+	+	2.E-03	2.E-03	0.20	0.87	-	-3.E-03	4.E-03	0.34	0.88	+	1.E-04	3.E-03	0.97	0.39	-	-3.E-03	5.E-03	0.57	0.25	-	-3.E-04	7.E-04	0.62	0.15
cg27118035	-	-	-4.E-04	1.E-03	0.73	0.54	+	3.E-03	3.E-03	0.24	0.61	-	-3.E-03	2.E-03	0.27	0.19	+	1.E-03	1.E-03	0.23	0.77	+	6.E-04	5.E-04	0.22	0.83

Supplemental Table 12. Association of the 30 diet-associated CpGs in European Ancestry (EA) participants with diet scores in African Ancestry (AA) and Hispanic Ancestry (HA) participants

					Direction			AA [1]			HA [2]		
CpG	CHR	Position	Gene	Diet	EA	AA	HA	Beta	SE	P	Beta	SE	P
cg04885881	1	11123118		MDS	+	+	-	0.0005	0.0010	0.60	-0.0018	0.0034	0.61
cg24735226	1	65096537	CACHD1	AHEI	+	+	-	0.0008	0.0015	0.60	-0.0036	0.0055	0.51
cg07805029	1	92953256	GFI1	AHEI	+	+	+	0.0010	0.0009	0.28	0.0004	0.0016	0.79
cg19693031	1	145441552	TXNIP	MDS	-	-	+	-0.0009	0.0014	0.52	0.0022	0.0035	0.54
cg24694018	1	145457621	POLR3GL	AHEI	+	+	-	0.0031	0.0036	0.40	-0.0045	0.0023	0.05
cg01940273	2	233284934		MDS	+	+	+	0.0025	0.0026	0.34	0.0025	0.0027	0.36
cg20842915	7	39665132	RALA	AHEI	+	+	+	0.0003	0.0007	0.68	0.0015	0.0022	0.49
cg02508743	8	56903623	LYN	AHEI	-	-	-	-0.0036	0.0032	0.26	-0.0042	0.0025	0.10
cg27039118	8	116575902	TRPS1	AHEI	+	+	-	0.0009	0.0018	0.62	-0.0039	0.0052	0.46
cg02716826	9	33447032	SUGT1P1;AQP3	MDS	+	+	+	0.0015	0.0012	0.20	0.0009	0.0028	0.74
cg25953130	10	63753550	ARID5B	AHEI	+	+	-	0.0030	0.0026	0.26	-0.0038	0.0046	0.41
cg03190891	10	97201172	SORBS1	AHEI	+	+	+	0.0001	0.0022	0.95	0.0026	0.0021	0.21
cg02079413	11	2986505	SNORA54;NAP1L4	MDS	+	+	-	0.0001	0.0005	0.86	-0.0015	0.0005	0.002
cg11250194	11	61601937	FADS2	AHEI	+	+	-	0.0003	0.0008	0.70	-0.0006	0.0017	0.71
cg11468085	11	67435577	ALDH3B2	AHEI	-	-	-	-0.0014	0.0007	0.06	-0.0026	0.0019	0.18
cg25909064	11	120082805	OAF	AHEI	+	+	+	0.0010	0.0006	0.11	0.0019	0.0025	0.45
cg03646329	13	48987165	LPAR6;RB1	AHEI	-	-	-	-0.0011	0.0016	0.48	-0.0055	0.0031	0.07
				MDS	-	-	-	-0.0025	0.0016	0.12	-0.0032	0.0032	0.32
cg16969872	13	79968324	RBM26	MDS	+	+	+	0.0030	0.0054	0.58	0.0090	0.0033	0.006
				AHEI	+	+	+	0.0033	0.0056	0.56	0.0084	0.0033	0.01
cg09940677	14	103415458	CDC42BPB	AHEI	-	-	-	-0.0013	0.0018	0.45	-0.0029	0.0014	0.04
cg13074055	14	106329206		AHEI	+	+	-	0.0047	0.0054	0.38	-0.0082	0.0066	0.22
cg27118035	16	31891978	ZNF267	AHEI	-	-	-	-0.0018	0.0018	0.31	0.0000	0.0025	0.99
cg08732950	16	89023389	CBFA2T3	MDS	-	-	-	-0.0003	0.0013	0.81	-0.0029	0.0028	0.30
cg02097604	17	17750910	TOM1L2	AHEI	+	+	+	0.0019	0.0023	0.41	0.0028	0.0017	0.09
				MDS	+	+	+	0.0018	0.0023	0.42	0.0020	0.0017	0.25
cg16936953	17	57915665	VMP1	AHEI	-	-	+	-0.0007	0.0033	0.84	0.0004	0.0041	0.91
cg18181703	17	76354621	SOCS3	MDS	+	+	+	0.0041	0.0030	0.17	0.0019	0.0030	0.54
				AHEI	+	+	+	0.0040	0.0034	0.24	0.0010	0.0030	0.74
cg19202384	17	79894511	PYCR1	AHEI	-	-	+	-0.0004	0.0006	0.54	0.0004	0.0027	0.87
cg01294327	19	2291373	LINGO3	AHEI	+	+	+	0.0012	0.0029	0.67	0.0072	0.0061	0.24
cg26470501	19	45252955	BCL3	AHEI	+	+	+	0.0024	0.0015	0.11	0.0065	0.0023	0.005
cg08884571	19	45901453	PPP1R13L	AHEI	+	+	-	0.0030	0.0028	0.29	-0.0085	0.0052	0.10
cg05232694	20	48809539		AHEI	-	-	+	-0.0001	0.0017	0.96	0.0023	0.0044	0.60

1. Meta-analysis of AA participants in ARIC and MESA. 2. In MESA

Supplemental Table 13. CpGs with FDR < 0.05 in transethnic meta-analysis

CpGs	CHR	Position	Gene	Diet	Beta	SE	P	I_squared
Unique CpGs identified in trans-ethnicity analysis								
cg17901584	1	55353706	DHCR24	AHEI	0.002	0.001	9.90E-07	0
cg25189904	1	68299493	GNG12	MDS	0.004	0.001	2.64E-07	0.01
cg02959282	3	136667373	NCK1	AHEI	0.002	0.0004	8.09E-07	0.08
cg18860310	4	87752504	SLC10A6	AHEI	0.002	0.0003	2.80E-07	0
cg22185977	5	1518133	LPCAT1	AHEI	-0.004	0.001	3.72E-09	0
				MDS	-0.004	0.001	4.71E-07	0.06
cg16567172	7	75931606	HSPB1	MDS	-0.001	0.0003	1.16E-07	0
cg00285394	8	126011954	SQLE	AHEI	0.003	0.001	1.05E-06	0
cg07986378	12	11898284	ETV6	MDS	0.004	0.001	1.09E-07	0.36
cg02790691	18	32289550	DTNA	MDS	0.002	0.0005	6.98E-07	0
cg19751789	19	11199944	LDLR	AHEI	0.001	0.0002	2.49E-08	0
CpGs also identified in European Ancestry participants								
cg01940273	2	233284934		MDS	0.005	0.001	4.50E-12	0
cg20842915	7	39665132	RALA	AHEI	0.002	0.001	8.74E-07	0
cg02716826	9	33447032	SUGT1P1;AQP3	MDS	0.002	0.0004	5.58E-07	0
cg25953130	10	63753550	ARID5B	AHEI	0.004	0.001	1.24E-07	0.08
cg02079413	11	2986505	SNORA54;NAP1L4	MDS	-0.002	0.0003	1.49E-08	0
cg25909064	11	120082805	OAF	AHEI	0.002	0.0003	3.29E-07	0
cg16969872	13	79968324	RBM26	AHEI	0.003	0.001	7.41E-08	0.16
cg27118035	16	31891978	ZNF267	AHEI	-0.002	0.0004	1.31E-08	0
cg08732950	16	89023389	CBFA2T3	MDS	-0.002	0.0005	2.25E-07	0
cg16936953	17	57915665	TMEM49	AHEI	0.003	0.001	1.46E-07	0
cg18181703	17	76354621	SOCS3	AHEI	0.003	0.001	2.04E-11	0.04
				MDS	0.003	0.001	4.56E-08	0.30

Supplemental Table 14. Pearson correlation of the 9 unique CpGs in multiethnic meta-analysis with the 32 CpGs significant in European Ancestry participants

		CHR	1	1	3	4	5	7	8	12	18	19
		Position	55353706	68299493	136667373	87752504	1518133	75931606	126011954	11898284	32289550	11199944
CHR	Position	CpG	cg17901584	cg25189904	cg02959282	cg18860310	cg22185977	cg16567172	cg00285394	cg07986378	cg02790691	cg19751789
1	11123118	cg04885881	0.17	0.30	-0.03	0.14	0.23	0.23	0.47	0.29	-0.02	0.18
1	65096537	cg24735226	-0.08	0.06	-0.12	-0.13	-0.05	0.21	0.12	0.01	-0.06	0.01
1	92953256	cg07805029	0.38	0.03	0.23	0.35	0.38	-0.18	0.00	0.40	0.41	0.10
1	145441552	cg19693031	0.23	0.19	-0.05	0.04	0.15	0.32	0.08	0.22	0.11	0.08
1	145457621	cg24694018	0.36	0.17	0.16	0.30	0.39	0.13	0.30	0.49	0.20	0.18
2	233284934	cg01940273	0.15	0.32	0.20	0.19	0.16	0.27	0.19	0.33	0.10	0.23
7	39665132	cg20842915	0.04	-0.04	0.47	0.10	-0.06	-0.10	-0.12	0.10	0.24	0.05
8	56903623	cg02508743	0.20	0.12	-0.22	0.05	0.14	-0.03	0.21	0.25	0.14	0.00
8	116575902	cg27039118	0.23	0.20	0.10	0.18	0.12	0.13	0.24	0.28	0.20	-0.01
9	33447032	cg02716826	0.08	0.29	-0.11	0.10	0.03	0.29	0.45	0.16	-0.12	0.27
10	63753550	cg25953130	-0.08	0.18	0.18	0.05	-0.05	0.18	0.26	0.02	-0.03	0.12
10	97201172	cg03190891	-0.11	0.10	-0.24	-0.17	0.03	0.09	0.20	0.01	-0.18	0.05
11	2986505	cg02079413	0.28	0.02	0.05	0.23	0.39	0.01	-0.01	0.26	0.23	0.08
11	61601937	cg11250194	0.49	0.17	0.05	0.32	0.33	0.08	0.19	0.48	0.28	0.31
11	67435577	cg11468085	-0.40	0.00	-0.18	-0.31	-0.13	0.07	0.18	-0.28	-0.39	-0.06
11	120082805	cg25909064	0.26	0.05	0.30	0.24	0.25	0.33	0.20	0.31	0.14	0.24
13	48987165	cg03646329	-0.33	0.02	0.19	-0.20	-0.29	0.12	0.02	-0.25	-0.14	0.13
13	79968324	cg16969872	-0.10	0.19	-0.11	-0.03	-0.15	0.08	0.25	0.00	-0.12	-0.09
14	103415458	cg09940677	-0.12	0.00	-0.08	-0.19	0.09	0.00	0.14	0.03	-0.11	-0.07
14	106329206	cg13074055	0.18	0.17	0.01	0.16	0.11	-0.07	0.22	0.24	0.11	0.05
16	31891978	cg27118035	-0.10	-0.04	-0.16	-0.18	-0.08	-0.01	0.04	-0.06	-0.03	-0.05
16	89023389	cg08732950	0.12	0.09	-0.06	0.03	0.27	0.30	0.06	0.18	0.07	0.12
17	17750910	cg02097604	0.29	0.18	0.11	0.30	0.31	0.21	0.36	0.33	0.14	0.12
17	57915665	cg16936953	0.00	0.19	-0.15	0.14	-0.11	0.21	0.31	-0.04	-0.10	0.15
17	76354621	cg18181703	0.20	0.28	0.02	0.24	0.07	0.32	0.48	0.26	0.03	0.29
17	79894511	cg19202384	0.07	0.03	0.31	0.20	0.12	0.18	0.14	0.14	0.12	0.14
19	2291373	cg01294327	0.12	0.13	0.12	0.07	0.03	0.18	0.19	0.18	0.08	0.07
19	45252955	cg26470501	0.23	0.23	0.11	0.22	0.15	0.22	0.43	0.33	0.06	0.13
19	45901453	cg08884571	0.21	0.09	-0.03	0.12	0.09	0.11	0.04	0.21	0.13	0.03
20	48809539	cg05232694	0.32	0.09	0.21	0.28	0.10	0.13	0.24	0.34	0.28	0.24

Supplemental Table 15. Functional description for protein coding genes annotated to diet-associated CpGs

CpG	CHR	Position	Gene	Full name	Description
cg17901584	1	55353706	DHCR24	24-dehydrocholesterol reductase	This gene encodes a flavin adenine dinucleotide (FAD)-dependent oxidoreductase which catalyzes the reduction of the delta-24 double bond of sterol intermediates during cholesterol biosynthesis. The protein contains a leader sequence that directs it to the endoplasmic reticulum membrane. Missense mutations in this gene have been associated with desmosterolosis. Also, reduced expression of the gene occurs in the temporal cortex of Alzheimer disease patients and overexpression has been observed in adrenal gland cancer cells.
cg25189904	1	68299493	GNG12	G protein subunit gamma 12	Among its related pathways are Beta-Adrenergic Signaling and Development Dopamine D2 receptor transactivation of EGFR. Gene Ontology (GO) annotations related to this gene include obsolete signal transducer activity and phosphate ion binding
cg02959282	3	136667373	NCK1	NCK adaptor protein 1	The protein encoded by this gene is one of the signaling and transforming proteins containing Src homology 2 and 3 (SH2 and SH3) domains. It is located in the cytoplasm and is an adaptor protein involved in transducing signals from receptor tyrosine kinases to downstream signal recipients such as RAS.
cg18860310	4	87752504	SLC10A6	Solute carrier family 10 member 6	Among its related pathways are Transport of glucose and other sugars, bile salts and organic acids, metal ions and amine compounds. Gene Ontology (GO) annotations related to this gene include bile acid:sodium symporter activity and sodium-dependent organic anion transmembrane transporter activity.
cg22185977	5	1518133	LPCAT1	Lysophosphatidylcholine acyltransferase 1	This gene encodes a member of the 1-acyl-sn-glycerol-3-phosphate acyltransferase family of proteins. The encoded enzyme plays a role in phospholipid metabolism, specifically in the conversion of lysophosphatidylcholine to phosphatidylcholine in the presence of acyl-CoA. This process is important in the synthesis of lung surfactant and platelet-activating factor (PAF). Elevated expression of this gene may contribute to the progression of oral squamous cell, prostate, breast, and other human cancers.
cg16567172	7	75931606	HSPB1	Heat shock protein family B (small) member 1	This gene encodes a member of the small heat shock protein (HSP20) family of proteins. In response to environmental stress, the encoded protein translocates from the cytoplasm to the nucleus and functions as a molecular chaperone that promotes the correct folding of other proteins. This protein plays an important role in the differentiation of a wide variety of cell types. Expression of this gene is correlated with poor clinical outcome in multiple human cancers, and the encoded protein may promote cancer cell proliferation and metastasis, while protecting cancer cells from apoptosis. Mutations in this gene have been identified in human patients with Charcot-Marie-Tooth disease and distal hereditary motor neuropathy.
cg00285394	8	126011954	SQLE	Squalene epoxidase	Squalene epoxidase catalyzes the first oxygenation step in sterol biosynthesis and is thought to be one of the rate-limiting enzymes in this pathway.
cg07986378	12	11898284	ETV6	ETS variant 6	This gene encodes an ETS family transcription factor. The product of this gene contains two functional domains: a N-terminal pointed (PNT) domain that is involved in protein-protein interactions with itself and other proteins, and a C-terminal DNA-binding domain. Gene knockout studies in mice suggest that it is required for hematopoiesis and maintenance of the developing vascular network. This gene is known to be involved in a large number of chromosomal rearrangements associated with leukemia and congenital fibrosarcoma.
cg02790691	18	32289550	DTNA	Dystrobrevin alpha	The protein encoded by this gene belongs to the dystrobrevin subfamily of the dystrophin family. This protein is a component of the dystrophin-associated protein complex (DPC), which consists of dystrophin and several integral and peripheral membrane proteins, including dystroglycans, sarcoglycans, syntrophins and alpha- and beta-dystrobrevin. The DPC localizes to the sarcolemma and its disruption is associated with various forms of muscular dystrophy. Mutations in this gene are associated with left ventricular noncompaction with congenital heart defects.
cg19751789	19	11199944	LDLR	Low density lipoprotein receptor	The low density lipoprotein receptor (LDLR) gene family consists of cell surface proteins involved in receptor-mediated endocytosis of specific ligands. Low density lipoprotein (LDL) is normally bound at the cell membrane and taken into the cell ending up in lysosomes where the protein is degraded and the cholesterol is made available for repression of microsomal enzyme 3-hydroxy-3-methylglutaryl coenzyme A (HMG CoA) reductase, the rate-limiting step in cholesterol synthesis. At the same time, a reciprocal stimulation of cholesterol ester synthesis takes place. Mutations in this gene cause the autosomal dominant disorder, familial hypercholesterolemia.

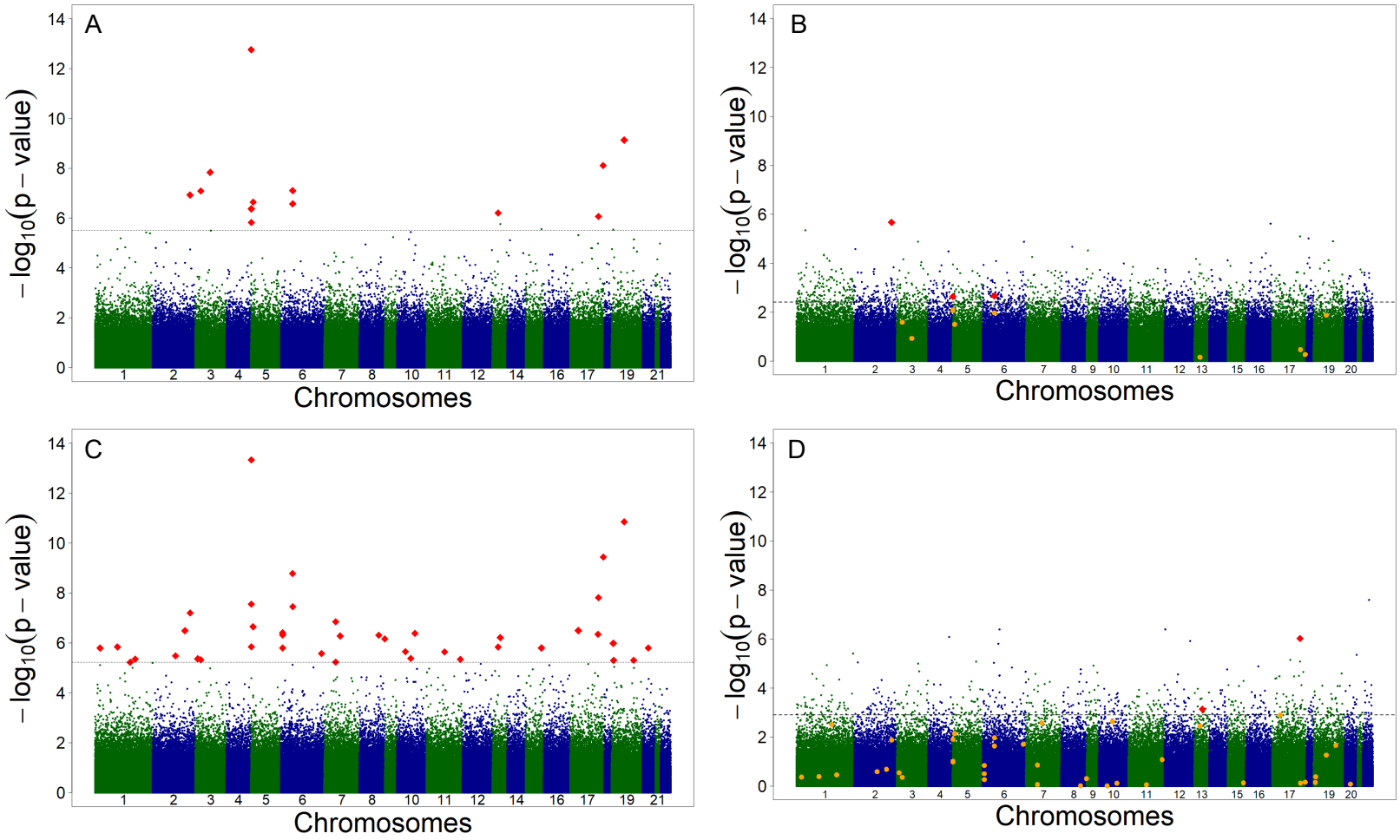
Supplemental Table 16. Enrichment analysis for the 10 unique CpGs identified in transethnic analysis

Gene set	Pathway	N_genes	N_overlap	P	adjusted P	Genes
KEGG	KEGG_STEROID_BIOSYNTHESIS	17	2	4.8E-08	8.9E-06	DHCR24, SQLE
Reactome	REACTOME_CHOLESTEROL_BIOSYNTHESIS	22	2	1.1E-07	7.3E-05	DHCR24, SQLE
	REACTOME_METABOLISM_OF_LIPIDS_AND_LIPOPROTEINS	468	4	9.6E-07	6.5E-04	DHCR24, LDLR, LPCAT1, SQLE

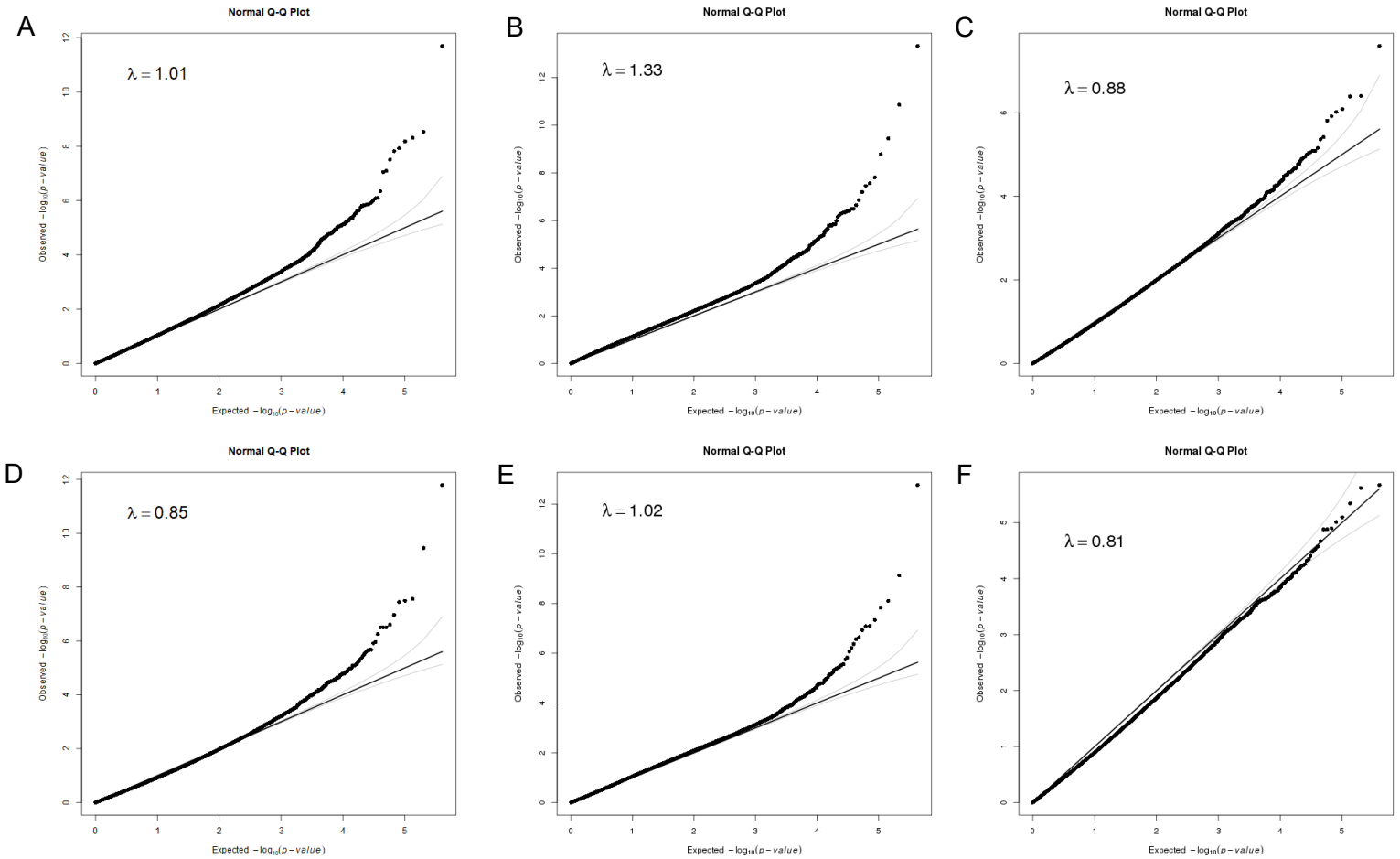
P was calculated using hypergeometric test through FUMA GWAS platform. Bonferroni correction was used to calculate adjusted P

Supplemental Table 17. Unique CpGs identified in transethnic analysis also reported in EWAS catalog

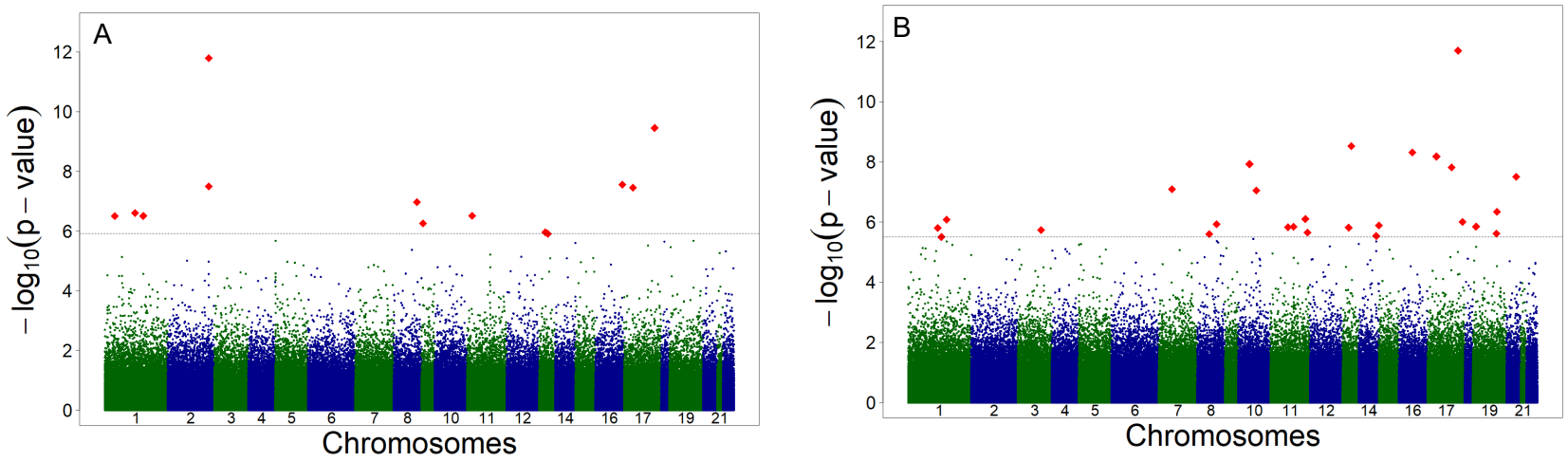
CpG	Direction. Diet	Trait	Direction. Trait	Beta	SE	P	N	Age	Tissue	PMID
cg17901584	+	Serum high-density lipoprotein cholesterol	+	7.17E-04	5.21E-05	5.47E-43	2306	68	Whole blood	28213390
cg25189904	+	Smoking				1.71E-39	1011	60.8	Whole blood	23691101
cg25189904	+	Smoking	-	-5.60E-01	0.0471	1.25E-32	582	54	Whole blood	25556184
cg25189904	+	Smoking				2.43E-23	721	52.6	Peripheral blood	26864933
cg25189904	+	Smoking	-	-7.68E-02	0.008	5.2E-22	9389	60.3	Whole blood, CD4+ T cells or monocytes	27651444
cg17901584	+	Serum triglycerides	-	-1.86E-02	0.001963	3.07E-21	2306	68	Whole blood	28213390
cg07986378	+	Smoking				8.98E-20	721	52.6	Peripheral blood	26864933
cg07986378	+	Smoking	-	-3.10E-01	0.0355	2.57E-18	582	54	Whole blood	25556184
cg25189904	+	Tobacco smoking				1.23E-16	454	52	Whole blood	28225026
cg25189904	+	Smoking	-	-4.00E-02	0.0049	3E-16	400	53	Buccal cells	26181258
cg25189904	+	Smoking	-	-7.64E-01	0.0942	4.72E-16	1597	67	Peripheral blood	28686328
cg07986378	+	Smoking	-	-7.39E-01	0.0922	1.08E-15	1597	67	Peripheral blood	28686328
cg25189904	+	Smoking	-	-7.70E-02	0.0098	3.77E-15	535	70	Whole blood	25249537
cg25189904	+	Smoking				8.26E-14	449	63.1	Peripheral blood	26829059
cg25189904	+	Smoking	-	-0.296	0.044253	2.25E-11	596	53.9	Whole blood	26756918
cg17901584	+	Body mass index	-	-13.7	2	1.4E-10	5387	54.5	Whole blood	28002404
cg25189904	+	Serum cotinine	-	-0.0009	0.00014	2.23E-10	500	62	Whole blood	26826776
cg18860310	+	C-reactive protein	-	-0.0031	0.000492	2.96E-10	8863	65	Whole blood	27955697
cg00285394	+	Serum total cholesterol	+	0.000213	3.44E-05	5.98E-10	2306	68	Whole blood	28213390
cg07986378	+	Smoking	-	-0.029	0.0047	1E-09	9389	60.3	Whole blood, CD4+ T cells or monocytes	27651444
cg00285394	+	Serum low-density lipoprotein cholesterol	+	0.00023	0.000039	4.12E-09	2304	68	Whole blood	28213390
cg25189904	+	Smoking				4.89E-09	464	55.39	Whole Blood	25424692
cg17901584	+	Triglycerides	-	-0.0162	0.0029	1.56E-08	725	60	Whole blood	28194238
cg25189904	+	Smoking	-	-0.027	0.0049	3.13E-08	421	43.8	Whole Blood	24334605
cg07986378	+	Tobacco smoking				4.44E-08	454	52	Whole blood	28225026
cg17901584	+	Serum total cholesterol	+	0.000149	2.73E-05	4.73E-08	2306	68	Whole blood	28213390
cg17901584	+	High-density lipoprotein cholesterol	+	0.0196	0.0036	6.44E-08	725	60	Whole blood	28194238
cg17901584	+	Waist circumference	-	-0.008	0.0015	8.34E-08	2097	56.2	Leukocytes	25935004
cg02790691	+	Alcohol consumption der day	-	-0.00026	5.25E-05	9.26E-07	9643	64.2	Whole blood	27843151
cg07986378	+	Alcohol consumption per day	-	-0.00053	0.00011	1.52E-06	1251	60	CD14+ monocytes	27843151
cg25189904	+	Alcohol consumption per day	-	-0.0001	2.13E-05	1.97E-06	2423	56	Whole blood	27843151
cg25189904	+	Smoking				2.84E-06	184	50.5	Peripheral blood	23175441
cg22185977	-	Alcohol consumption per day	+	4.46E-05	9.74E-06	4.76E-06	2423	56	Whole blood	27843151
cg17901584	+	Body mass index	-	-0.0433	0.0095	1.05E-05	442	67	Whole blood	28095459
cg07986378	+	Smoking	-	-0.185	0.042041	1.08E-05	596	53.9	Whole blood	26756918
cg18860310	+	Alcohol consumption der day	-	-0.00015	3.46E-05	2.61E-05	9643	64.2	Whole blood	27843151
cg16567172	-	Smoking	+	0.0055	0.0014	0.000072	9389	60.3	Whole blood, CD4+ T cells or monocytes	27651444
cg25189904	+	Alcohol consumption der day	-	-0.00023	5.94E-05	0.00012	9643	64.2	Whole blood	27843151
cg19751789	+	Smoking	-	-0.0033	0.0009	0.0004	9389	60.3	Whole blood, CD4+ T cells or monocytes	27651444
cg16567172	-	Alcohol consumption der day	-	-8.9E-05	2.61E-05	0.000686	9643	64.2	Whole blood	27843151



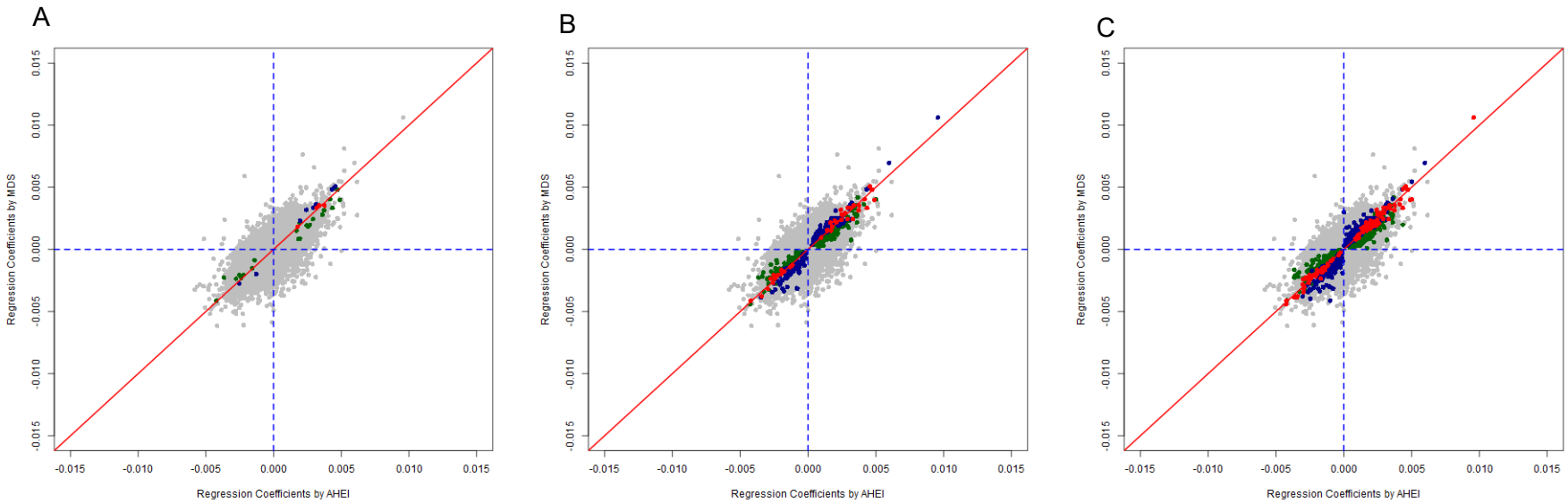
Supplemental Figure 1. Manhattan plots in discovery (A for MDS and C for AHEI) and in replication (B for MDS and D for AHEI) analyses in European ancestry participants. Models were adjusted for sex, age, and energy intake. Dashed lines represent FDR < 0.05 (corresponding p-value of 1.5×10^{-6} for MDS and 6×10^{-6} for AHEI). In B and D, orange colored dots are CpGs with FDR < 0.05 in discovery analysis and red colored dots are replicated CpGs. Dashed lines represent Bonferroni corrected p-value threshold (0.05/13 for MDS and 0.05/24 for AHEI). AHEI: Alternative Healthy Eating Index. MDS: Mediterranean-style diet score. FDR: false discovery rate.



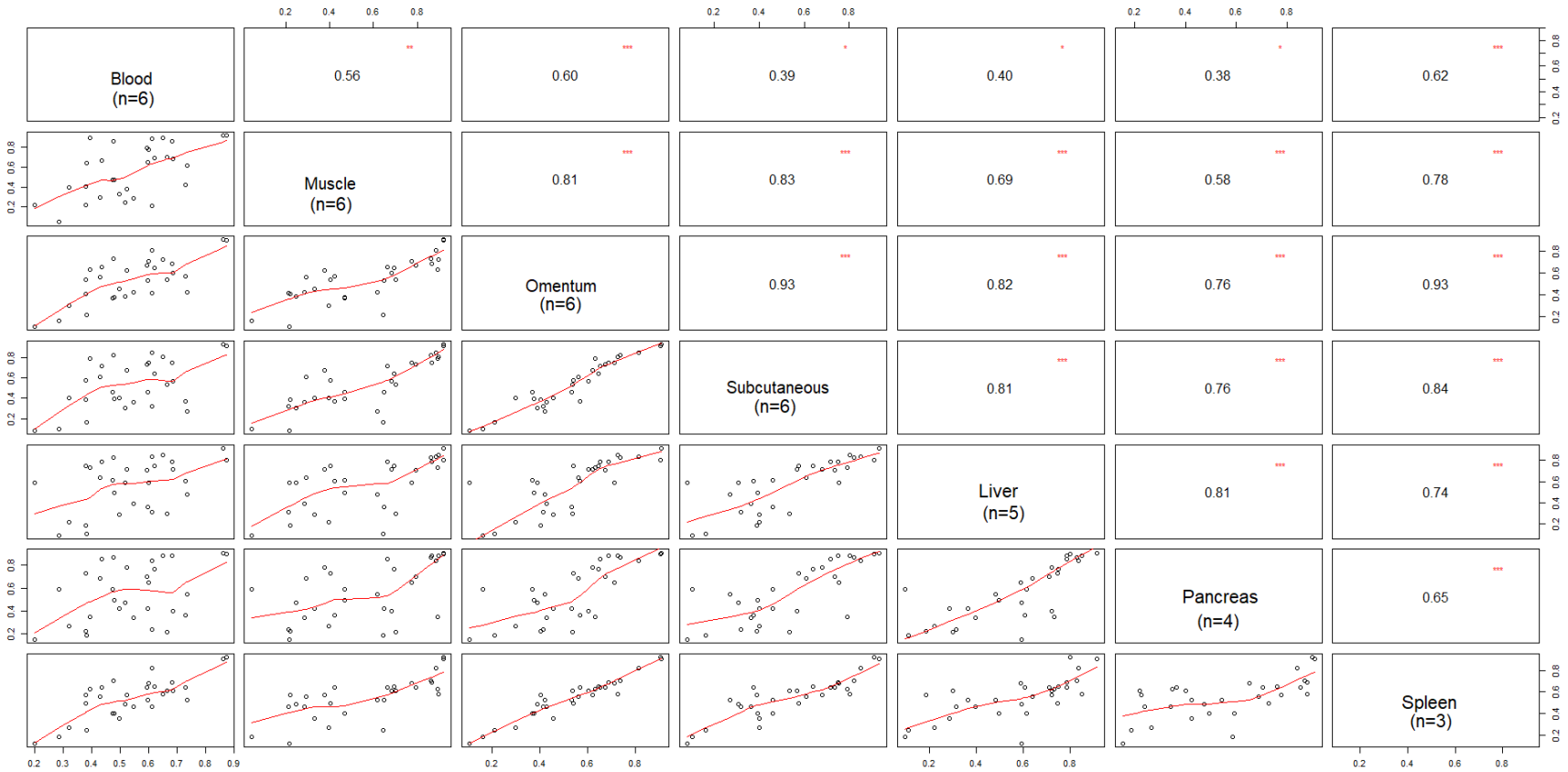
Supplemental Figure 2. Quantile-Quantile (QQ) plots with lambda values generated using the sex- and age-adjusted model in European Ancestry (EA) participants. Plots A to C are for analyses using AHEI in all EA participants, in discovery cohort (FHS), and in replication cohorts (ARIC, GOLDN, MESA, and RS), respectively. Plots D to F are for analyses using MDS in all EA participants, in discovery cohort, and in replication cohorts. AHEI: Alternative Healthy Eating Index. MDS: Mediterranean-style diet score.



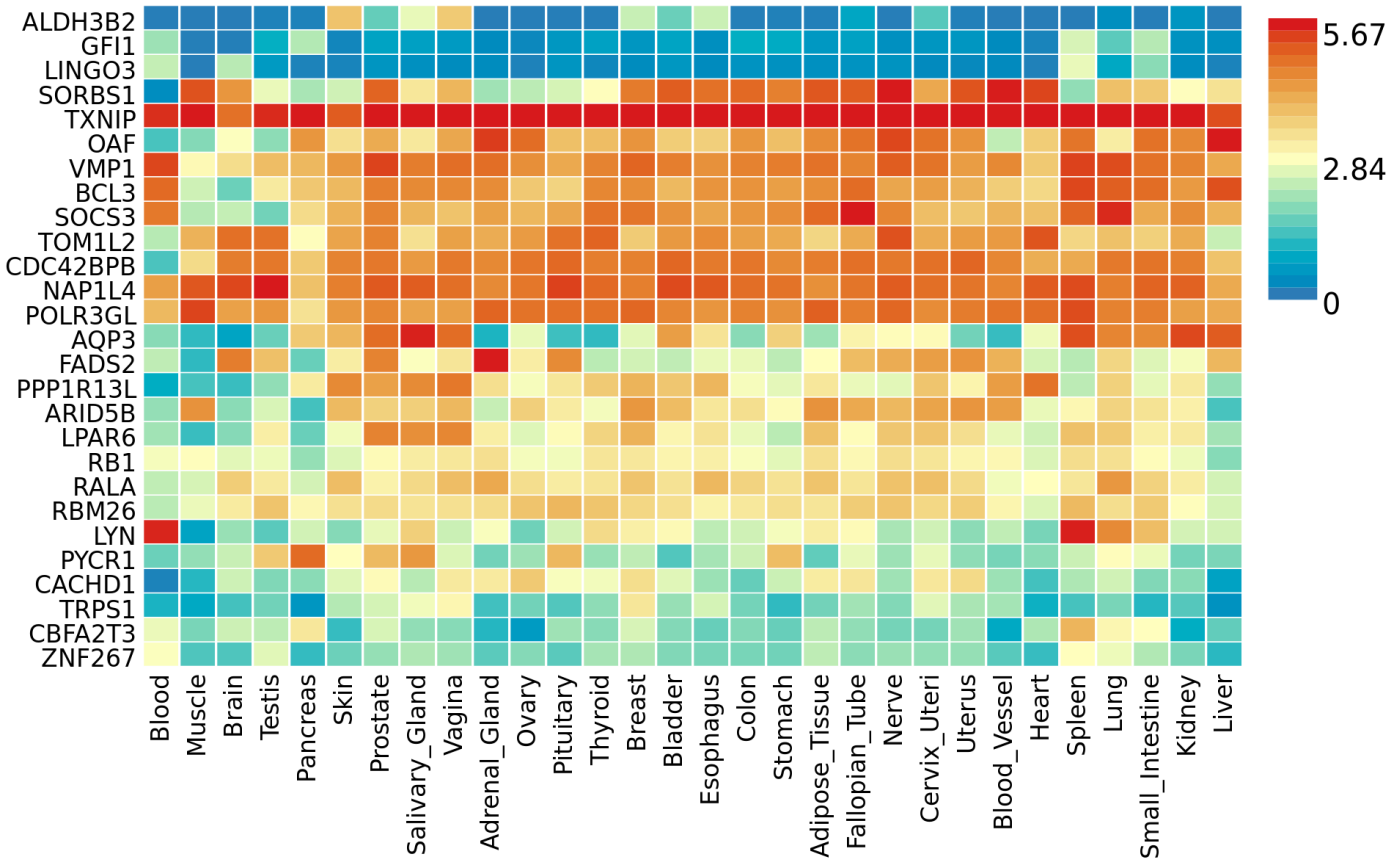
Supplemental Figure 3. Manhattan plots generated from meta-analyses using the sex, age, and energy intake adjusted models in European ancestry participants. A is for MDS and B is for AHEI. Dotted lines represent FDR < 0.05, corresponding p-value of 3.1×10^{-6} for AHEI and 1.2×10^{-6} for MDS. CpGs close to (within ± 500kb) and correlated with (Pearson |r| > 0.7) a CpG with a lower p-value and CpGs with p-value > 0.05 and discordant direction in the discovery and replication cohorts were excluded in subsequent analysis. AHEI: Alternative Healthy Eating Index. MDS: Mediterranean-style diet score. FDR: false discovery rate.



Supplemental Figure 4. Correlation of regression coefficients generated from analysis using AHEI with that from analysis using MDS. Grey dots are for all CpGs and colored (green, blue, and red) dots are for CpGs with FDR < 0.05 in A, top 200 CpGs in B, and top 500 CpGs in C. Green dots represent analysis using AHEI, blue dots represent analysis using MDS, and red dots represent overlapped CpGs, e.g., in plot A, green dots are CpGs with FDR < 0.05 in AHEI analysis, blue dots are CpGs with FDR < 0.05 in MDS analysis, and red dots are CpGs with FDR < 0.05 in both analyses using AHEI and MDS. AHEI: Alternative Healthy Eating Index. MDS: Mediterranean-style diet score. FDR: false discovery rate.



Supplemental Figure 5. Correlation of mean DNA methylation levels in peripheral whole blood with that measured in six internal tissues including muscle, omentum, subcutaneous adipose tissue, liver, pancreas, and spleen. *** < 0.0001; ** < 0.01, * < 0.05.



Supplemental Figure 6. Heatmap for average expression of the annotated 29 protein coding genes in 30 specific tissue types. Colors represents average expression value (log2 transformed Reads Per Kilobase per Million per tissue per gene, winsorization at 50). Higher expression is represented using darker red and lower expression is represented by darker blue. Figure downloaded from GTEx through FUMA GWAS (www.fuma.ctglab.nl/). GTEx: Genotype-Tissue Expression. FUMA GWAS: Functional Mapping and Annotation of Genome-Wide Association Studies.