



A Peripheral Blood DNA Methylation Signature of Hepatic Fat Reveals a Potential Causal Pathway for Nonalcoholic Fatty Liver Disease

Jiantao Ma,^{1,2} Jana Nano,^{3,4,5} Jingzhong Ding,⁶ Yinan Zheng,⁷ Rachel Hennein,¹ Chunyu Liu,^{1,8} Elizabeth K. Speliotes,⁹ Tianxiao Huan,¹ Ci Song,^{1,10} Michael M. Mendelson,^{1,11} Roby Joehanes,¹ Michelle T. Long,¹² Liming Liang,^{13,14} Jennifer A. Smith,¹⁵ Lindsay M. Reynolds,⁶ Mohsen Ghanbari,^{3,16} Taulant Muka,³ Joyce B.J. van Meurs,^{3,17} Louise J.M. Alferink,¹⁸ Oscar H. Franco,³ Abbas Dehghan,^{3,19} Scott Ratliff,¹⁵ Wei Zhao,¹⁵ Lawrence Bielak,¹⁵ Sharon L.R. Kardina,¹⁵ Patricia A. Peyser,¹⁵ Hongyan Ning,²⁰ Lisa B. VanWagner,^{20,21} Donald M. Lloyd-Jones,²⁰ John Jeffrey Carr,²² Philip Greenland,²⁰ Alice H. Lichtenstein,²³ Frank B. Hu,²⁴ Yongmei Liu,⁶ Lifang Hou,⁷ Sarwa Darwish Murad,¹⁸ and Daniel Levy¹

<https://doi.org/10.2337/DB18-1193>

Nonalcoholic fatty liver disease (NAFLD) is a risk factor for type 2 diabetes (T2D). We aimed to identify the peripheral blood DNA methylation signature of hepatic fat. We conducted epigenome-wide association studies of hepatic fat in 3,400 European ancestry (EA) participants and in 401 Hispanic ancestry and 724 African

ancestry participants from four population-based cohort studies. Hepatic fat was measured using computed tomography or ultrasound imaging and DNA methylation was assessed at >400,000 cytosine-guanine dinucleotides (CpGs) in whole blood or CD14+ monocytes using a commercial array. We identified 22 CpGs associated

¹Population Sciences Branch, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD, and Framingham Heart Study, Framingham, MA

²Nutrition Data Science, Friedman School of Nutrition Science and Policy, Tufts University, Boston, MA

³Department of Epidemiology, Erasmus University Medical Center, Rotterdam, the Netherlands

⁴Helmholtz Zentrum München, German Research Center for Environmental Health, Institute of Epidemiology, Neuherberg, Germany

⁵German Center for Diabetes Research, München-Neuherberg, Germany

⁶Department of Epidemiology and Prevention, Wake Forest School of Medicine, Winston-Salem, NC

⁷Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL

⁸Department of Biostatistics, Boston University, Boston, MA

⁹University of Michigan Medical School, Ann Arbor, MI

¹⁰Molecular Epidemiology, Department of Medical Sciences, Uppsala University, Uppsala, Sweden

¹¹Department of Cardiology, Boston Children's Hospital, Harvard Medical School, Boston, MA

¹²Section of Gastroenterology, Department of Medicine, Boston University School of Medicine, Boston, MA

¹³Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA

¹⁴Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA

¹⁵Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, MI

¹⁶Department of Genetics, School of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran

¹⁷Department of Internal Medicine, Erasmus University Medical Center, Rotterdam, the Netherlands

¹⁸Department of Gastroenterology and Hepatology, Erasmus University Medical Center, Rotterdam, the Netherlands

¹⁹Department of Biostatistics and Epidemiology, MRC-PHE Centre for Environment and Health, School of Public Health, Imperial College London, London, U.K.

²⁰Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL

²¹Division of Gastroenterology and Hepatology, Department of Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL

²²Department of Radiology and Radiological Sciences, Vanderbilt University Medical Center, Nashville, TN

²³Cardiovascular Nutrition Laboratory, USDA Human Nutrition Research Center on Aging at Tufts University, Boston, MA

²⁴Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA

Corresponding author: Daniel Levy, levy@nhlbi.nih.gov

Received 6 November 2018 and accepted 14 February 2019

This article contains Supplementary Data online at <http://diabetes.diabetesjournals.org/lookup/suppl/doi:10.2337/DB18-1193/-/DC1>.

J.M., J.N., J.D., and Y.Z. contributed equally.

© 2019 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered. More information is available at <http://www.diabetesjournals.org/content/license>.

with hepatic fat in EA participants at a false discovery rate <0.05 (corresponding $P = 6.9 \times 10^{-6}$) with replication at Bonferroni-corrected $P < 8.6 \times 10^{-4}$. Mendelian randomization analyses supported the association of hypomethylation of cg08309687 (*LINC00649*) with NAFLD ($P = 2.5 \times 10^{-4}$). Hypomethylation of the same CpG was also associated with risk for new-onset T2D ($P = 0.005$). Our study demonstrates that a peripheral blood-derived DNA methylation signature is robustly associated with hepatic fat accumulation. The hepatic fat-associated CpGs may represent attractive biomarkers for T2D. Future studies are warranted to explore mechanisms and to examine DNA methylation signatures of NAFLD across racial/ethnic groups.

The estimated global prevalence of nonalcoholic fatty liver disease (NAFLD) in adults is 24% and has increased substantially along with the increasing rates of obesity (1). NAFLD correlates with increased risk of type 2 diabetes (T2D) (2), and it is the second leading contributor to hepatic failure necessitating transplantation (3). A prior study in three family-based cohorts estimated the heritability of hepatic steatosis to be 27%; however, common genetic variants from genome-wide association studies (GWAS) account for $<5\%$ of interindividual variance in hepatic fat (4). Epigenetics may explain part of the interindividual variance of steatosis. Several studies have demonstrated altered DNA methylation profiles in liver biopsy samples collected from individuals with NAFLD (5,6). One study showed that peripheral blood-derived DNA hypermethylation at one cytosine-guanine

dinucleotide (CpG) (cg06690548) located in an intron of *SLC7A11* may be associated with a lower risk of steatosis (7). In general, however, prior studies were limited by small sample sizes to discover DNA methylation sites associated with hepatic fat accumulation.

To bridge this knowledge gap, we examined the epigenome-wide association between DNA methylation at $>400,000$ CpGs and hepatic fat in European ancestry (EA), African ancestry (AA), and Hispanic ancestry (HA) participants from five population-based cohort studies with hepatic fat measurements derived from noninvasive imaging. For hepatic fat-associated CpGs, we further examined their relations to genetic variants, gene expression, and regulatory functions and explored potential relations of DNA methylation to NAFLD and T2D.

RESEARCH DESIGN AND METHODS

Study Population

The current study included multiethnic participants from five population-based cohorts including the Coronary Artery Risk Development Study in Young Adults (CARDIA), the Framingham Heart Study (FHS), the Genetic Epidemiology Network of Arteriopathy (GENOA), the Multi-Ethnic Study of Atherosclerosis (MESA), and the Rotterdam Study (RS) (Supplementary Table 1). All cohorts excluded participants who consumed a high amount of alcohol, equivalent to ≥ 21 drinks/week in men or ≥ 14 drinks/week in women (8). Cohort-specific exclusion is detailed in Supplementary Data. Due to potential differences in DNA methylation patterns between different ethnicities (9), we analyzed the association between

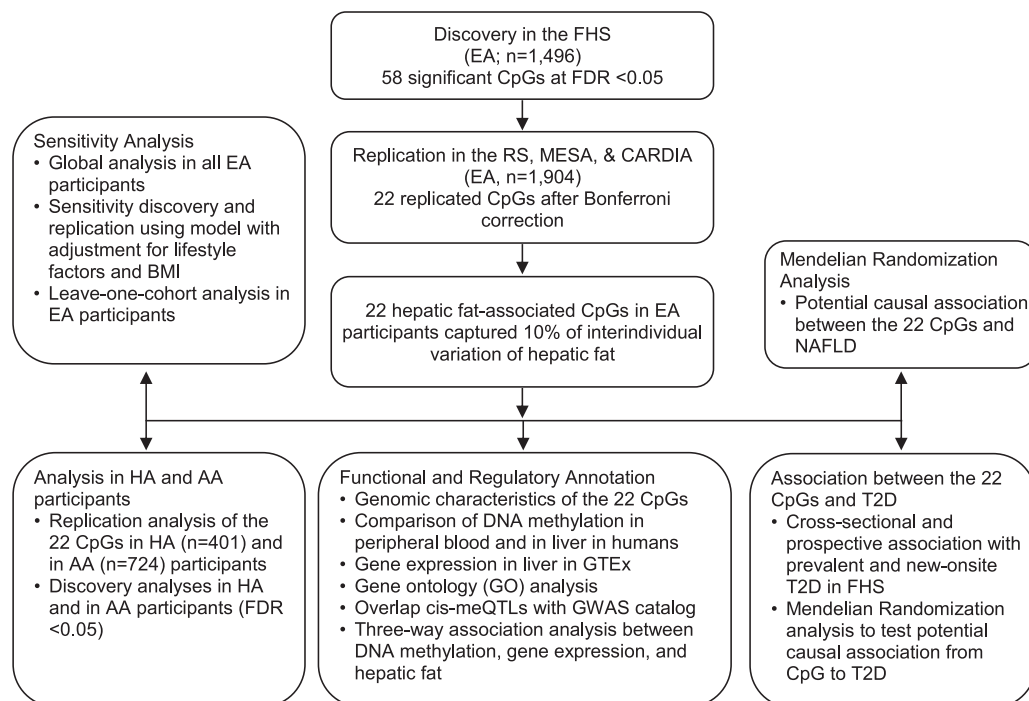


Figure 1—Study design flowchart.

DNA methylation and hepatic fat separately in EA ($n = 3,400$), HA ($n = 401$), and AA ($n = 724$) participants. The protocol for each study was approved by the institutional review board in each cohort. All participants provided written informed consent.

Study Design

The study design flowchart is presented in Fig. 1. We first conducted the epigenome-wide association studies of hepatic fat among EA participants, including both discovery and replication. We then examined differential DNA methylation in relation to hepatic fat in the HA and AA participants. We further examined the functional and regulatory annotations for the replicated CpGs and tested the potential associations of the identified CpGs with NAFLD and T2D.

Hepatic Fat Assessment

A detailed description for hepatic fat assessment in each cohort is presented in Supplementary Data. The RS used ultrasound to estimate hepatic fat and diagnosed steatosis on a dichotomized scale. The other cohorts used computed tomography (CT) to quantify hepatic fat on a continuous scale by using either mean Hounsfield units of the liver image or the ratio of the Hounsfield units of the liver image to that of a control.

DNA Methylation Profiling

Methylation profiles were measured by Illumina BeadChip using DNA derived from all leukocytes or CD14+ monocytes in peripheral blood (Supplementary Table 1). Details for DNA preparation, bisulfite conversion, methylation profiling, quality control procedures, and raw data normalization in each cohort are provided in Supplementary Data. We analyzed either methylation signal M values (in CARDIA [calculated as logit transformation of β values]) or β values (in all other cohorts [calculated as methylated signals divided by the sum of methylated and unmethylated signals]). Nonautosomal probes were excluded from the current study.

Statistical Analyses

Epigenome-Wide Association Study of Hepatic Fat

We conducted the discovery epigenome-wide association study (EWAS) in FHS and interrogated the differentially methylated CpGs at false discovery rate (FDR) < 0.05 in the replication cohorts (EA samples in CARDIA, MESA, and RS). Linear regression models or mixed models with consideration of family structures were conducted to examine directionality and calculate P values in each cohort. Because hepatic fat was measured using different scales, we meta-analyzed the P values in the replication cohorts using logit method based on the general fixed effect model. The statistical significance in the replication analysis was determined using the Bonferroni-corrected P value threshold, defined as 0.05 divided by the number of significant CpGs in the discovery phase. We used sex- and

age-adjusted models (model 1) in the discovery and replication analyses. Estimated leukocyte composition (10) and technical variables were also adjusted for in a cohort-specific manner (Supplementary Data). Included in the sensitivity analyses, we conducted global meta-analyses in all EA participants to examine the impact of potential confounders. We performed the same sex- and age-adjusted model (model 1). We additionally adjusted for lifestyle factors including smoking status, physical activity levels, and alcohol intake in model 2. We further adjusted for BMI in model 3. We also performed a discovery and replication analysis using model 3 in EA participants. Covariates are assessed using cohort-specific methods (Supplementary Data). We conducted similar EWAS with adjustment for the same covariates to identify hepatic fat-related CpGs in MESA HA participants and AA participants in CARDIA, GENOA, and MESA. We calculated adjusted R^2 using multiple linear regression models with and without adjustment for the replicated CpGs and used the difference of the two adjusted R^2 values to represent the interindividual variation of hepatic fat from the replicated CpGs.

Gene Expression Association Analysis

In FHS, we examined the associations between DNA methylation, hepatic fat, and gene expression. To prioritize genes in these analyses, we selected Illumina-annotated genes. For CpGs without annotated genes, we identified a set of genes by overlapping *cis*-meQTLs (methylation quantitative trait loci ± 500 kilobases [kb] from CpG) with *cis*-eQTLs (expression quantitative trait loci residing within 500 kb of a nearby gene) at P value $< 5 \times 10^{-7}$. The association between DNA methylation and gene expression was analyzed in 4,561 participants in the FHS as previously described (11). For genes significantly associated with CpGs (P value $< 5 \times 10^{-7}$), we further examined their association with hepatic fat using similar statistical procedures in 2,317 FHS participants. For genes that associated with both CpGs and hepatic fat, we conducted mediation tests to estimate the proportion of mediation by gene expression on the association of CpGs and hepatic fat. We used the R package mediation with sex- and age-adjusted linear mixed models as described above and used the quasi-Bayesian Monte Carlo method with 1,000 simulations to calculate CIs (12).

Association With T2D

Owing to the well-documented association between NAFLD and T2D (13), we conducted cross-sectional and prospective analyses between replicated CpGs and T2D. We defined T2D as fasting glucose ≥ 7.0 mmol/L, HbA_{1c} $\geq 6.5\%$ (48 mmol/mol), or current use of insulin or an oral hypoglycemic drug. In the cross-sectional analysis, we analyzed 4,068 FHS participants who attended the eighth Offspring cohort examination (2005–2008) or the second examination of the Third Generation cohort (2008–2011). In the prospective analysis, we analyzed 1,764 FHS participants who attended both the eighth

and ninth (2011–2014) Offspring cohort examinations. In the prospective analysis, prevalent T2D cases at baseline were excluded. The first model in the cross-sectional analysis adjusted for sex, age, alcohol intake, smoking status, physical activity level, DNA methylation measurement laboratory and technical variables, top three principal components for DNA methylation measurement, and estimated leukocyte composition. The first model in the prospectively analysis adjusted for the same covariates and baseline fasting glucose and HbA_{1c}. In the second model for both analyses, we additionally adjusted for BMI. We used generalized estimating equations to estimate the odds ratio (OR) after accounting for the family structure in this study sample.

Mendelian Randomization Analysis

We conducted Mendelian randomization (MR) analyses (Supplementary Fig. 1) to test for association of the replicated CpGs with NAFLD. We also examined the association between these CpGs and T2D. We performed MR analysis using a two-sample MR approach (14). We used independent *cis*-meQTLs ($n = 4,170$), with pairwise linkage disequilibrium $r^2 < 0.1$ and F statistics > 10 , as instrumental variables (IVs). Using the TwoSampleMR R package, we performed the primary analysis using the inverse variance weighted (IVW) method and sensitivity analysis using the MR-Egger method when association was significant using IVW. The effect sizes and SEs for IVs-CpG were obtained from the FHS and the effect size and SEs were obtained for IVs-NAFLD and for IVs-T2D from published GWAS (4,15).

Functional and Regulatory Annotation

We conducted hypergeometric tests with Bonferroni correction to compare the genomic characteristics of the replicated CpGs with the whole set of CpGs that we analyzed and calculated the probability for these characteristics, e.g., whether CpGs were more likely to reside in a CpG island or gene region, using the Infinium Human-Methylation450 BeadChip annotation files. We queried *cis*-meQTLs in the platform of Functional Mapping and Annotation of Genome-Wide Association Studies (FUMA GWAS) (16). Using this platform, we examined the overlap between *cis*-meQTLs with signals in the NHGRI-EBI Catalog of published GWAS (17). We also studied genes using differentially expressed genes measured in human whole blood and liver samples in the Genotype-Tissue Expression (GTEx v6) database (18) and visualized the genomic region for *cis*-meQTLs. To assess the relevance of the identified peripheral blood-derived CpGs in liver, we compared DNA methylation levels in whole blood with that in liver using data deposited in the Gene Expression Omnibus (GEO Series accession number GSE48472) (19). Gene Ontology (GO) biological process enrichment analysis was performed using the GO Consortium website (<http://www.geneontology.org/>).

Data and Resource Availability

The data sets analyzed in the current study are available at https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000285.v3.p2 (CARDIA), https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000007.v29.p10 (FHS), https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs001238.v1.p1 (GENOA), and https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000209.v13.p3 (MESA). Data can be obtained upon request. Requests should be directed toward the management team of the RS (secretariat.epi@erasmusmc.nl), which has a protocol for approving data requests. Because of restrictions based on privacy regulations and informed consent of the participants, data cannot be made freely available in a public repository (RS).

RESULTS

EWAS of Hepatic Fat in EA Participants

Of the 400,129 CpGs analyzed in age- and sex-adjusted models, 58 CpGs were significantly associated with hepatic fat in the FHS discovery cohort ($n = 1,496$) at FDR < 0.05

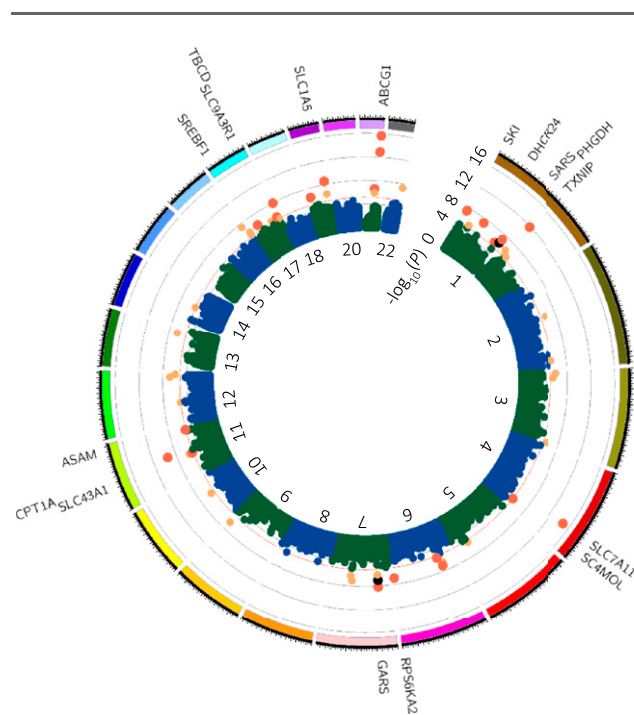


Figure 2—Circular Manhattan plot of the epigenome-wide association of peripheral blood-derived DNA methylation with hepatic fat. Plot was generated using results from the sex- and age-adjusted model in the discovery cohort (FHS). Chromosome numbers are presented in the inner circle. Red, orange, and black dots are 58 CpGs with FDR < 0.05 (corresponding P value 6.9×10^{-6} [red track]) in the discovery stage analysis. Red dots are the 22 independent CpGs with Bonferroni-corrected P value < 0.05 in the replication cohorts (CARDIA, MESA, and RS). Black dots are two nonsentinel CpGs with Bonferroni-corrected P value < 0.05 in the replication cohorts. Genes annotated to the replicated CpGs were presented in the outer ring. Plot was drawn in Circos.

(corresponding P value = 6.9×10^{-6} [Manhattan plot in Fig. 2 and Supplementary Fig. 2, quantile-quantile plot in Supplementary Fig. 3 and Supplementary Table 2). The t statistics for the 58 CpGs were correlated between the FHS and each of the replication cohorts (Supplementary Fig. 4). The associations between the 58 CpGs and hepatic fat remained materially unchanged following additional adjustment for the time gap between assessments of DNA methylation and hepatic fat in FHS (Supplementary Fig. 5). In CARDIA, hepatic fat was measured at the year 25 examination. The t statistics calculated between hepatic fat and DNA methylation measured at the year 25 examination were highly correlated with those calculated using DNA methylation measured at the year 15 examination (Supplementary Fig. 5). Twenty-four (41%) CpGs replicated (Bonferroni-corrected P value $< 0.05/58 = 8.6 \times 10^{-4}$) in the meta-analysis of the replication cohorts ($n = 1,904$) (Table 1). We removed two CpGs from the replication analysis because they were highly correlated ($|r| \geq 0.7$) and close to other CpGs with lower P values in the same chromosome (Supplementary Table 3). Forest plots for all 22 CpGs using regression coefficients and SEs (standardized by the cohort-specific SD of the regression coefficients) are shown in Supplementary Fig. 6.

Sensitivity Analysis

In the global meta-analysis of all EA participants, additional adjustment for lifestyle factors did not materially change the association between DNA methylation levels and hepatic fat (Fig. 3). Further adjustment for BMI, which is correlated with hepatic fat (Spearman $r = 0.45$ in FHS), reduced the strength of the associations (Fig. 3); however, all 22 CpGs remained nominally associated with hepatic fat (P value < 0.05) (Table 1). Additionally, two CpGs, cg06690548 (*SLC7A11*) and cg19693031 (*TXNIP*), remained significant in FHS at FDR < 0.05 (corresponding P value = 1.2×10^{-7}) (Supplementary Table 4) and in the replication samples (Bonferroni-corrected P value < 0.05). Both CpGs were among the replicated CpGs in the sex- and age-adjusted analysis. Leave-one-cohort-out analysis in EA participants showed that P values in the global meta-analysis with exclusion of one cohort were highly correlated with those in all samples, with r ranging from 0.83 to 0.87 (Supplementary Fig. 7). We further examined whether lipid traits, serum fasting triglycerides (TGs), and total cholesterol concentrations (TC) affect the associations of the 22 CpGs with hepatic fat. As shown in Supplementary Fig. 8, additional adjustment for TG but not TC attenuated the associations for most of CpGs in FHS. Additionally, after Bonferroni correction (corresponding P value < 0.002), TGs significantly mediated the association with hepatic fat for 12 CpGs, with a mediation percent ranging from 13 to 39% (Supplementary Table 5). However, the association between all 22 CpGs and hepatic fat remained significant after additional adjustment for TG (all P values < 0.002) (Supplementary Fig. 8). Additional adjustment for glycemic traits (T2D, fasting

glucose, and HbA_{1c}) attenuated the associations between the 22 CpGs and hepatic fat; however, all associations remained significant (Supplementary Fig. 8). We found that adjustment for caloric intake and total dietary fat intake did not materially alter the association between the 22 CpGs and hepatic fat (Supplementary Fig. 9).

Variation in Hepatic Fat Explained by Differentially Methylated CpGs

In CARDIA, the 22 replicated CpGs captured 10% of interindividual variation (i.e., the adjusted R^2) in hepatic fat after we accounted for sex and age (P value = 1.4×10^{-7}) and 4.3% of interindividual variation after we additionally accounted for lifestyle factors and BMI (P value = 0.001). The proportion of variation explained by the 22 replicated CpGs was similar to that observed in the discovery cohort (FHS): 14.9% and 5.8%, respectively. With additional adjustment for serum alanine transaminase, serum aspartate transaminase, and a genetic risk score for NAFLD, the 22 replicated CpGs explained 4.6% of interindividual variation in hepatic fat in FHS (P value = 1.4×10^{-12}).

DNA Methylation Profiles in HA and AA Participants

For the 22 CpGs that replicated in the EA participants, cg19693031 (*TXNIP*) remained significant in HA participants after Bonferroni correction (P value $< 2.3 \times 10^{-3}$) (Supplementary Table 6). Additionally, four CpGs were nominally significant in HA participants (P value < 0.05) (Supplementary Table 6) and three CpGs were nominally significant in the meta-analysis of AA participants ($P < 0.05$) (Supplementary Table 7). In epigenome-wide analysis, no CpG was detected at FDR < 0.05 in HA participants. We discovered 26 CpGs at FDR < 0.05 (corresponding P value = 2.7×10^{-6}) (Supplementary Table 8) in the meta-analysis of epigenome-wide studies in AA participants, of which two CpGs were nominally significant (P value < 0.05) in the global meta-analysis of EA participants.

Functional and Regulatory Annotation of Hepatic Fat-Associated CpGs

The mean whole blood DNA methylation levels of the 22 replicated CpGs in EA participants were moderately correlated with those measured in liver tissue (19) ($r = 0.59$) (Supplementary Fig. 10). These CpGs were more likely to reside in the south shore (0–2 kb downstream of CpG island; P value = 5.5×10^{-4}) or south shelf (2–4 kb downstream of CpG islands; P value = 4.1×10^{-4}), in DNase I hypersensitivity sites (P value = 1.7×10^{-3}), in reprogramming-specific differentially methylated regions (P value = 3.8×10^{-4}), and in gene body regions (P value = 2.4×10^{-4}).

Of the 22 CpGs, 19 CpGs have been annotated to 18 unique genes (Table 1 and Supplementary Table 9). Based on liver-specific differentially expressed genes in

Table 1—Significant CpGs in the EWAS for hepatic fat in EA participants

CpG	CHR	Position	Gene	Has cis-meQTLs	Primary discovery-replication analysis			Global meta-analysis		
					Discovery P value	Replication P value	Direction	P value*	P value*	P value***
cg09469355	1	2161886	SKI	Yes	2.93E-08	4.07E-04	-, -, -, -	6.66E-09	7.60E-08	5.14E-05
cg17901584	1	55353706	DHCR24	Yes	4.76E-08	2.10E-04	-, -, -, -	5.28E-09	1.48E-08	2.94E-03
cg03725309	1	109757585	SARS	Yes	1.37E-06	1.29E-06	-, -, -, -	6.21E-10	4.60E-10	1.02E-04
cg114476101	1	120255992	PHGDH	Yes	7.54E-08	1.10E-05	-, +, -, -	6.67E-10	2.10E-09	9.84E-05
cg19693031	1	145441552	TXNIP	No	1.33E-12	3.29E-04	-, +, -, -	1.96E-11	4.32E-11	1.66E-07
cg06690548	4	139162808	SLC7A11	No	1.78E-15	5.27E-06	-, -, -, -	6.77E-14	2.30E-14	9.25E-09
cg05119988	4	166251189	SC4MOL	Yes	6.91E-06	7.75E-05	-, -, -, -	5.50E-08	2.06E-07	8.51E-04
cg03957124	6	37016869		No	9.08E-08	9.80E-05	-, -, -, -	4.25E-09	1.31E-08	9.10E-03
cg18120259	6	43894639		Yes	3.35E-06	3.06E-04	-, +, -, -	1.12E-07	1.22E-06	7.93E-03
cg17501210	6	166970252	RPS6KA2	Yes	1.30E-07	3.38E-07	-, -, -, -	5.53E-11	4.23E-10	1.76E-03
cg21429551	7	30635762	GARS	Yes	7.29E-09	1.85E-04	-, +, -, -	1.53E-09	2.94E-09	1.42E-04
cg11376147	11	57261198	SLC43A1	Yes	4.88E-06	8.76E-05	-, +, -, -	4.87E-08	7.11E-08	2.70E-03
cg00574958	11	68607622	CPT1A	No	3.27E-10	9.93E-06	-, -, -, -	3.05E-11	6.28E-11	4.46E-03
cg26894079	11	122954435	ASAM	Yes	5.58E-06	8.57E-04	-, -, -, -	3.94E-07	2.81E-06	3.83E-03
cg11024682	17	17730094	SREBF1	Yes	6.58E-07	2.68E-07	+, +, +, +	1.10E-10	1.67E-10	1.08E-03
cg14020176	17	72764985	SLC9A3R1	Yes	2.36E-06	7.97E-04	+, +, +, +	2.05E-07	2.73E-07	3.36E-05
cg19016694	17	80821826	TBCD	Yes	9.49E-09	2.42E-05	-, -, -, -	3.74E-10	4.92E-10	7.63E-04
cg15860624	19	3811194	ZFR2	Yes	3.46E-07	5.12E-05	+, +, +, +	5.73E-09	4.74E-09	1.76E-06
cg02711608	19	47287964	SLC1A5	Yes	2.76E-09	1.21E-04	-, +, -, -	6.23E-10	1.06E-09	2.44E-03
cg08309687	21	35320596		Yes	2.48E-07	3.57E-06	-, -, -, -	5.37E-10	5.29E-10	4.65E-05
cg27243685	21	43642366	ABCG1	Yes	1.15E-13	1.14E-05	+, +, +, +	6.77E-13	2.08E-12	1.20E-06
cg06500161	21	43656587	ABCG1	Yes	1.95E-16	3.35E-09	+, +, +, +	3.45E-16	7.33E-16	2.45E-09

The discovery-replication analysis used sex- and age-adjusted models. The global meta-analysis was conducted in all EA participants. The direction column shows sign of regression coefficients in order of FHS, RS, CARDIA, and MESA: the + sign represents hypermethylation, associated with increased hepatic fat, and - sign represents hypomethylation, associated with increased hepatic fat. CHR, chromosome. *Model 1 adjusted for sex and age. **Model 2 adjusted for sex, age, and lifestyle covariates including smoking, physical activity, and alcohol. ***Model 3 adjusted for sex, age, lifestyle covariates, and BMI.

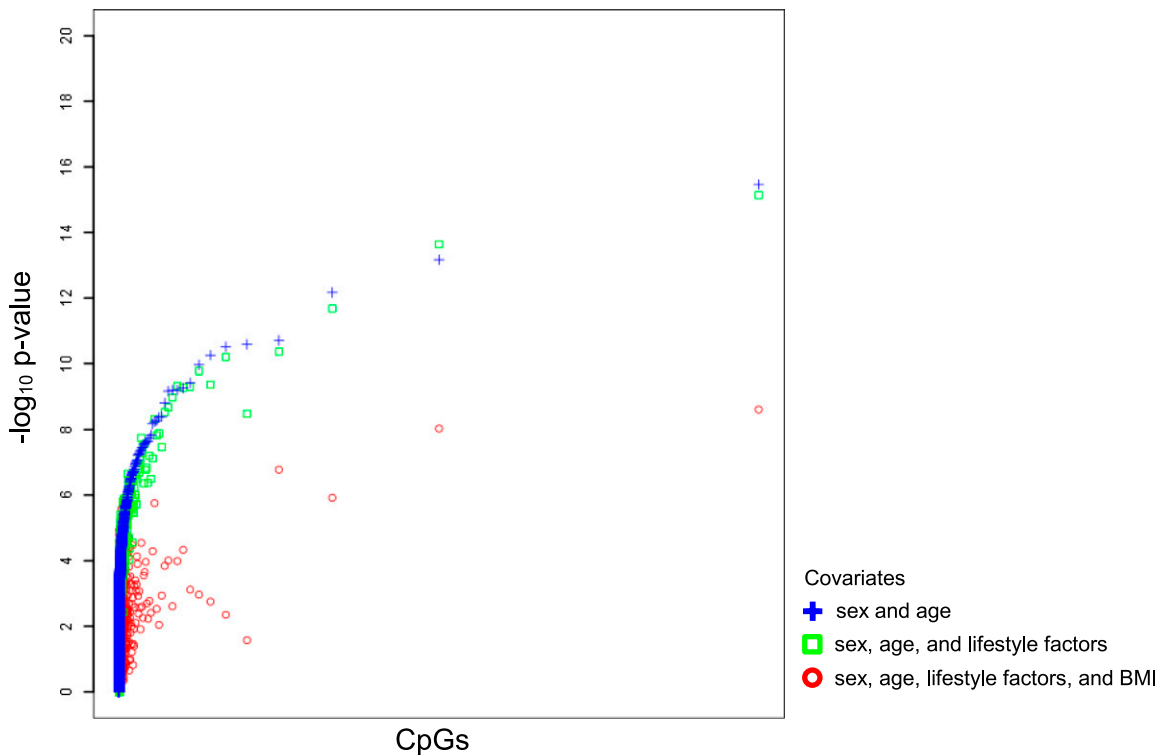


Figure 3—Comparisons of sequential adjustment models in EA participants. y-axis values are observed $-\log_{10}(P)$ values in the global meta-analysis of all EA participants. x-axis is ordered by CpGs (symbols at same vertical position are same CpG). Blue cross, sex- and age-adjusted model; green square, model with additional adjustment for lifestyle factors (smoking, physical activity, and alcohol intake); red circle, the fully adjusted model including sex, age, smoking, physical activity, alcohol intake, and BMI.

GTE_x (18), the 18 annotated genes were enriched with genes that are upregulated in the liver, including *DHCR24*, *SLC43A1*, *CPT1A*, *SREBF1*, *SC4MOL*, and *SLC9A3R1* (Bonferroni-corrected P value = 0.005) (Supplementary Table 10). GO analysis showed enrichment for 18 biological processes (Supplementary Table 11). The most significant enriched pathway was positive regulation of the cholesterol biosynthesis (GO:0045542; *ABCG1* and *SREBF1*; >100-fold enrichment; FDR adjusted P value = 0.02).

GWAS Analysis

By overlapping *cis*-meQTL variants with GWAS results in the NHGRI-EBI Catalog of GWAS (17), we found that *cis*-meQTLs or their proxies (linkage disequilibrium $R^2 > 0.8$) for nine CpGs were associated with 26 unique traits in GWAS (Supplementary Table 12). For example, rs637868 for cg14476101 (*PHDHG*) was associated in GWAS with alanine aminotransferase levels (20) and rs2834288 for cg08309687 (LINC00649) was associated in GWAS with abundance of gut microbiota (21).

Three-Way Association and Mediation Analysis of CpGs, Gene Expression, and Liver Fat

In the FHS, 7 of the 22 replicated CpGs were associated with peripheral blood-derived expression of six annotated genes (e.g., cg17901584 with *DHCR24*) at a P value threshold of $< 5 \times 10^{-7}$ (Supplementary Table 13).

Additionally, cg17501210 was associated with expression of one nonannotated *cis*-gene, *RNASET2* (transcription start site residing 301 kb downstream from cg17501210; P value = 9.4×10^{-11}). Among these seven genes, expression levels of *ABCG1* and *CPT1A* were significantly associated with liver fat (P value = 1.2×10^{-30} and 2.0×10^{-17} , respectively). Expression of *ABCG1* and *CPT1A* mediated the association between corresponding CpGs and hepatic fat by ~20% and 10%, respectively (Fig. 4).

We identified eight *cis*-genes for cg08309687 (LINC00649) by overlapping their *cis*-meQTLs (P value $< 5 \times 10^{-7}$) with *cis*-eQTLs (P value $< 5 \times 10^{-7}$) (regional plot in Supplementary Fig. 11). None of the CpG was associated with gene expression at the predefined threshold (P value $< 5 \times 10^{-7}$). Among the five nominally significant genes (P value < 0.05) (Supplementary Table 13), expression of *TMEM50B* mediated 8% (95% CI 2, 16; P value $< 2.2 \times 10^{-16}$) of the association of cg08309687 with hepatic fat (Fig. 4).

MR Analysis for a Potential Role of DNA Methylation in NAFLD

The IVW analysis (Supplementary Table 14) showed that hypomethylation at cg08309687 (LINC00649) was significantly associated with NAFLD (P value = 2.5×10^{-4}) (Fig. 5). Hypomethylation at cg14476101 (*PHDHG*) was nominally associated with NAFLD (P value = 0.01)

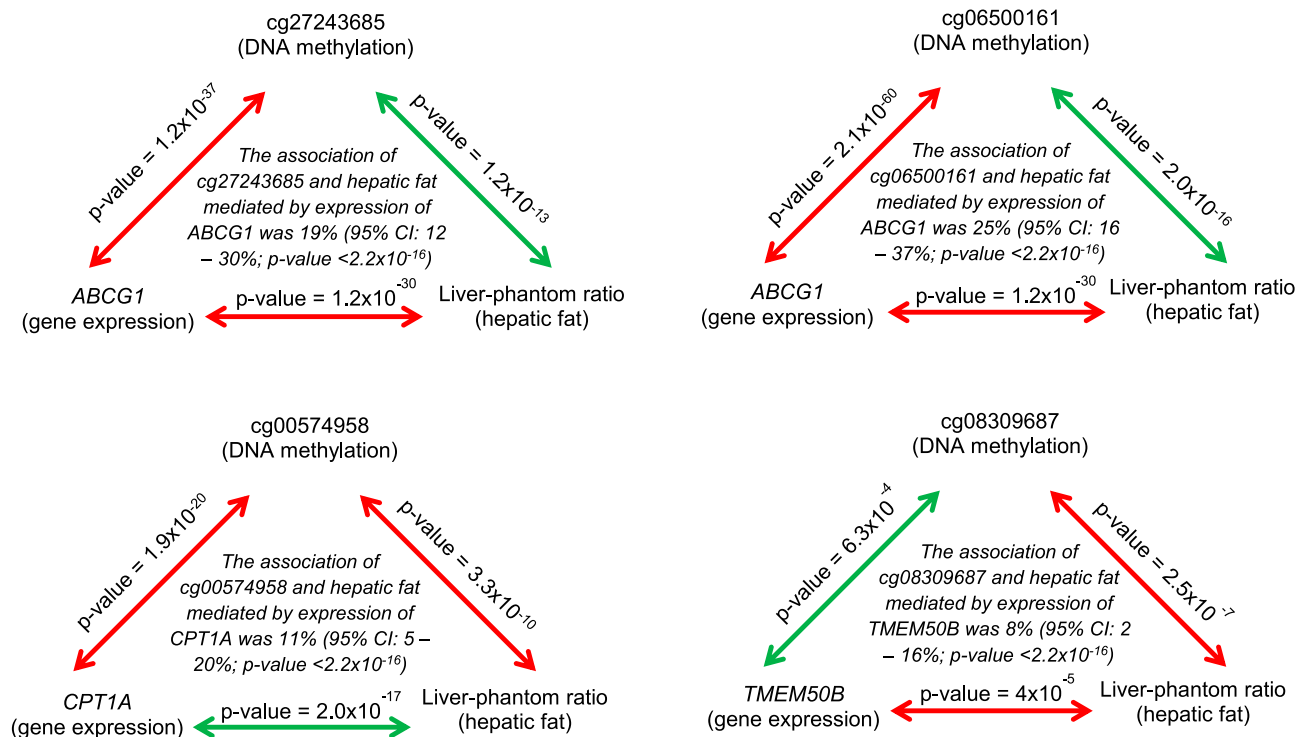


Figure 4—Three-way association of whole blood–derived DNA methylation level at cg06500161, cg27243685, cg00574958, and cg08309687; whole blood–derived gene expression for ABCG1, CPT1A, and TMEM50B; and CT-derived liver-to-phantom ratio (hepatic fat) in the FHS. Green, positive association; red, inverse association.

(Supplementary Fig. 12). Neither CpG was significant in the sensitivity analysis using the MR-Egger method (P value = 0.17 and 0.97, respectively). No horizontal pleiotropy effect was detected (P value = 0.66 and 0.15).

Association With T2D

In the FHS, 18 of the 22 replicated CpGs in EA participants were cross-sectionally associated with T2D after Bonferroni

correction (P value < 0.002) (Supplementary Table 15). In addition, 6 of the 18 CpGs were prospectively associated with incident T2D (P value < 0.05) (Table 2). In analysis with additional adjustment for BMI, associations remained significant (P value < 0.05) for 13 CpGs in the cross-sectional analysis (Supplementary Table 15) and for 4 CpGs in the prospective analysis (Table 2). For all significant CpG–T2D pairs, the direction of the association was concordant with the expected direction; e.g., hypermethylation at cg08309687 (LINC00649) was expected to be associated with decreased risk of T2D because hypermethylation at this locus was associated with lower hepatic fat. In concordance with the prospective analysis using a BMI-adjusted model, MR analysis showed consistent results for cg21429551 (GARS); i.e., hypermethylation at this locus was nominally associated with lower risk of T2D (P value = 0.03) (Supplementary Table 16).

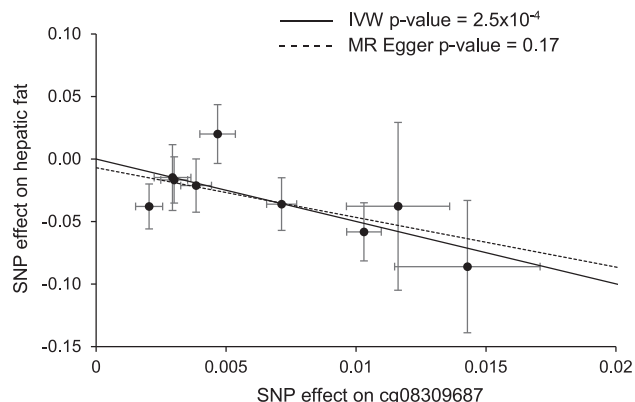


Figure 5—Association of cg08309687 with NAFLD using MR analysis. Plot depicts IVW (solid line) and MR-Egger estimate (dashed line). No horizontal pleiotropy effect was detected (P value = 0.66). SNP, single nucleotide polymorphism.

DISCUSSION

We found that differential methylation of peripheral blood–derived DNA at 22 CpG sites was associated with hepatic fat in EA participants. These CpGs reside at several loci regulating key biological processes relevant to the development of steatosis. The 22 CpGs explained $\sim 10\%$ of interindividual variation in the sex- and age-adjusted model, which was larger than that explained by single nucleotide polymorphisms identified from GWAS (4). In addition, the current study associated several hepatic

Table 2—Prospective association between hepatic fat–associated CpGs and incident T2D in 1,764 FHS participants

CpG	CHR	MAPINFO	Gene	Model 1		Model 2	
				OR (95% CI)	P	OR (95% CI)	P
cg09469355	1	2161886	<i>SKI</i>	0.88 (0.61, 1.27)	0.50	0.89 (0.61, 1.30)	0.56
cg17901584	1	55353706	<i>DHCR24</i>	1.19 (0.84, 1.68)	0.32	1.29 (0.91, 1.81)	0.15
cg03725309	1	109757585	<i>SARS</i>	0.54 (0.36, 0.83)	0.005	0.55 (0.36, 0.83)	0.005
cg14476101	1	120255992	<i>PHGDH</i>	0.90 (0.69, 1.19)	0.47	0.93 (0.71, 1.22)	0.60
cg19693031	1	145441552	<i>TXNIP</i>	0.90 (0.70, 1.16)	0.41	0.88 (0.68, 1.15)	0.36
cg06690548	4	139162808	<i>SLC7A11</i>	1.05 (0.86, 1.28)	0.63	1.08 (0.88, 1.32)	0.47
cg05119988	4	166251189	<i>SC4MOL</i>	1.28 (0.94, 1.75)	0.11	1.30 (0.95, 1.78)	0.10
cg03957124	6	37016869		1.05 (0.62, 1.77)	0.86	1.11 (0.66, 1.87)	0.70
cg18120259	6	43894639		1.18 (0.79, 1.77)	0.42	1.23 (0.81, 1.86)	0.33
cg17501210	6	166970252	<i>RPS6KA2</i>	0.73 (0.52, 1.03)	0.07	0.78 (0.55, 1.10)	0.15
cg21429551	7	30635762	<i>GARS</i>	0.71 (0.53, 0.93)	0.01	0.72 (0.55, 0.95)	0.02
cg11376147	11	57261198	<i>SLC43A1</i>	0.80 (0.50, 1.28)	0.35	0.82 (0.51, 1.31)	0.41
cg00574958	11	68607622	<i>CPT1A</i>	0.83 (0.56, 1.23)	0.35	0.90 (0.60, 1.34)	0.59
cg26894079	11	122954435	<i>ASAM</i>	0.69 (0.48, 0.98)	0.04	0.71 (0.49, 1.01)	0.06
cg11024682	17	17730094	<i>SREBF1</i>	1.44 (1.14, 1.81)	0.002	1.39 (1.10, 1.76)	0.006
cg14020176	17	72764985	<i>SLC9A3R1</i>	1.35 (0.95, 1.93)	0.10	1.31 (0.91, 1.88)	0.15
cg19016694	17	80821826	<i>TBCD</i>	0.90 (0.57, 1.43)	0.67	0.94 (0.59, 1.49)	0.79
cg15860624	19	3811194	<i>ZFR2</i>	0.86 (0.62, 1.19)	0.36	0.82 (0.59, 1.16)	0.26
cg02711608	19	47287964	<i>SLC1A5</i>	0.80 (0.56, 1.15)	0.23	0.82 (0.57, 1.18)	0.28
cg08309687	21	35320596		0.68 (0.52, 0.91)	0.008	0.70 (0.53, 0.93)	0.02
cg27243685	21	43642366	<i>ABCG1</i>	1.35 (1.00, 1.83)	0.049	1.34 (0.99, 1.81)	0.06
cg06500161	21	43656587	<i>ABCG1</i>	1.32 (1.02, 1.71)	0.03	1.28 (0.98, 1.66)	0.07

Model 1 adjusted for sex; age; alcohol intake; smoking status; physical activity level; baseline glucose; baseline HbA_{1c} level; laboratory for DNA methylation assessment; DNA methylation chip ID, row, and column; estimated leukocyte composition; and top three principal components. Model 2 adjusted for model 1 covariates and BMI. CHR, chromosome.

fat–associated CpGs with increased risk of new-onset T2D; e.g., hypomethylation at cg08309687 was associated with increased hepatic fat and higher risk of T2D.

Previous studies using array-based methods have examined epigenome-wide DNA methylation patterns in liver samples of individuals with biopsy-proven NAFLD (5,6). While these studies showed a strong contrast between DNA methylation profiles of individuals with non-alcoholic steatohepatitis compared with control subjects (i.e., individuals without steatosis or steatohepatitis), differences in DNA methylation patterns of individuals with steatosis (fatty liver alone) versus control subjects were less obvious (5,6). In contrast to prior studies using liver biopsies, we used noninvasive imaging to assess hepatic fat. We therefore had a much larger sample size and statistical power to detect epigenetic signals associated with elevated hepatic fat.

Two recent large EWAS identified CpGs associated with BMI (11,22). The majority (17 of the 22) of the replicated hepatic fat–associated CpGs in the current study were also observed in at least one of the two large-scale BMI-DNA methylation studies. Several CpGs associated with both liver fat and BMI are annotated to key genes involved in

lipid metabolism pathways, namely, *SREBF1*, *CPT1A*, *ABCG1*, and *DHCR24* (23,24). This is concordant with our observation that TGs mediated the associations for half of the replicated CpGs. These genes are in a top gene network associated with general adiposity identified in a previous analysis in MESA (25). Therefore, the overlap between hepatic fat–associated CpGs and those linked to BMI supports a key role of adiposity and adiposity-related pathways in the pathogenesis of steatosis.

Our MR analyses implicated cg08309687 as a putatively causal factor for NAFLD. This CpG is located in a long intergenic noncoding region (LINC00649) that may play a role in transcription regulation relevant to steatosis. We could not examine whether LINC00649 mediated the observed association between cg08309687 and hepatic fat because we lacked expression data for LINC00649 in our study cohorts. However, we observed that *cis*-meQTL variants for cg08309687 coincide with *cis*-eQTLs of several nearby genes. We also established three-way associations of cg08309687 methylation, hepatic fat, and gene expression levels for a nearby gene, *TMEM50B* (Fig. 3). Together with data from GTEx, our analysis indicates that DNA methylation at cg08309687 may affect LINC00649 and

alter expression of nearby genes (e.g., *TMEM50B*) and impact hepatic fat accumulation. In addition, our data suggest that cg08309687 may be involved in the regulation of the gut microbiome, which has been postulated to play a critical role in the development of NAFLD (26). Future studies are needed to explore pathways underlying the observed association between DNA hypomethylation at cg08309687 and NAFLD.

Using models with additional adjustment for BMI, we showed that two CpGs, cg06690548 (*SLC7A11*) and cg19693031 (*TXNIP*), are independently associated with hepatic fat in the EA population. Although cg19693031 (*TXNIP*) has not been previously linked to obesity-related traits, several studies have shown that hypomethylation at cg19693031 (*TXNIP*) was associated with increased risk of T2D (27). As NAFLD is strongly associated with T2D (13), these observations are in accordance with our finding that DNA hypomethylation at this locus is associated with increased hepatic fat.

The present analyses included participants from multiple ancestries, and differences in DNA methylation patterns associated with hepatic fat were observed between ethnic groups. Such variability is consistent with other observations that DNA methylation levels differ by race and/or ethnicity (9). However, the sample size in the current study may not be sufficient for robust comparisons of EA with AA and HA participants. The MR analyses in the current study may be limited for a few reasons. First, the effect sizes and SEs for the IVs were estimated using relatively small-scale GWAS. Second, effect sizes and SEs for instrument-exposure and instrument-outcome associations were estimated in partially overlapping study samples, which may lead to instrument bias (28). Moreover, one participating cohort (RS) used ultrasound, which could not be used to quantify liver fat on a continuous scale as other participating cohorts did. This difference may partly explain the inconsistency in the direction of regression coefficients for some of the replicated CpGs in RS compared with the other cohorts. Further, because not all cohorts measured liver fat using the same scales, we meta-analyzed *P* values rather than regression coefficients and SEs and could not estimate the overall effect size for the association of DNA methylation and hepatic fat accumulation. However, we observed that DNA methylation signals were correlated among the cohorts of EA participants. As demonstrated by accounting for the time gap between assessments of DNA methylation and hepatic fat in FHS and CARDIA, DNA methylation signals were relatively stable over time (Supplementary Fig. 5). In addition, CpGs measured in monocytes were correlated well with those measured using whole blood samples (i.e., all leukocytes). This finding is consistent with observations from one prior BMI-DNA methylation study in which the observed associations between DNA methylation and BMI were shared across different cell subsets (22).

In conclusion, the current study demonstrates a unique DNA methylation pattern related to hepatic fat in EA participants. In addition, our study showed that several

hepatic fat-associated CpGs may be useful biomarkers for prediction of new-onset T2D. Future studies with larger and more ethnically diverse sample sizes are needed to validate our findings and to explore the biological relevance of peripheral DNA methylation in the development and progression of NAFLD and T2D.

Acknowledgments. The authors thank Michael Verbiest, Mila Jhamai, Sarah Higgins, Marijn Verkerk, and Lisette Stolk (Erasmus Medical Center and Erasmus University, Rotterdam, Netherlands) for their help in creating the methylation database. The authors are grateful to the study participants, the staff from the RS, and the participating general practitioners and pharmacists. The authors also thank the families that participated in GENOA.

Funding. Infrastructure for the CHARGE (Cohorts for Heart and Aging Research in Genomic Epidemiology) Consortium is supported in part by the National Heart, Lung, and Blood Institute (NHLBI), National Institutes of Health (NIH), grant R01-HL-105756. This work was supported in part by the Intramural Research Program of the NIH: NHLBI; National Institute on Aging (NIA); and National Institute of Environmental Health Sciences. The FHS is funded by NIH contract N01-HC-25195. The laboratory work for this investigation was funded by the Division of Intramural Research, NHLBI, NIH, Bethesda, MD. The analytical component of this project was funded by the Division of Intramural Research, NHLBI, and the Center for Information Technology, NIH. The generation and management of the Illumina HumanMethylation450K BeadChip array data (EWAS data) for the RS was executed by the Human Genotyping Facility of the Genetic Laboratory of the Department of Internal Medicine, Erasmus University Medical Center. The EWAS data were funded by the Genetic Laboratory of the Department of Internal Medicine, Erasmus University Medical Center, and by the Netherlands Organization for Scientific Research (NWO) (project number 184021007) and made available as a Rainbow Project (RP3 [Biobank-based Integrative Omics Study (BIOS)]) of the Biobanking and Biomolecular Research Infrastructure Netherlands (BBMRI-NL). The RS is funded by Erasmus University Medical Center and Erasmus University, Rotterdam; Netherlands Organization for Health Research and Development (ZonMw); the Research Institute for Diseases in the Elderly; the Ministry of Education, Culture and Science; the Ministry for Health, Welfare and Sports; the European Commission (Directorate-General XII); and the Municipality of Rotterdam. MESA and the MESA SHARe (SNP Health Association Resource) project are conducted and supported by the NHLBI in collaboration with MESA investigators. Support for MESA is provided by contracts NIH 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 and Transcriptomics Study was funded by NIA grant 1R01HL101250-01 to Wake Forest University Health Sciences. CARDIA is supported by contracts HHSN268201300025C, HHSN268201300026C, HHSN268201300027C, HHSN268201300028C, HHSN268201300029C, and HHSN268200900041C and grant R01-HL098445 for the year 25 CT exam (J.J.C.) from the NHLBI. The laboratory work and analytical component were funded by the American Heart Association (17SFRN33700278 and 14SFRN20790000 [L.H.]). LVW is supported by NHLBI grant K23-HL-136891. Support for GENOA was provided by the NHLBI of the NIH (HL-054464, HL-054457, HL-054481, HL-100185, HL-119443, HL-085571, and HL-133221).

The views expressed in this manuscript are those of the authors and do not necessarily represent the views of the NHLBI, the NIH, or the U.S. Department of Health and Human Services. All authors have completed the Unified Competing Interest form (available on request from the corresponding author) and declare no support from any organization for the submitted work, no financial relationships with any organizations that might have an interest in the submitted work in the previous 3 years, and no other relationships or activities that could appear to have influenced the submitted work.

Duality of Interest. No potential conflicts of interest relevant to this article were reported.

Author Contributions. J.M. and D.L. designed the analysis. J.M., J.N., J.D., Y.Z., T.H., R.J., and L.B. analyzed data. C.L. provided statistical support. J.M., J.N., J.D., Y.Z., and J.A.S. wrote the manuscript. R.H., E.K.S., C.S., M.M.M., M.T.L., L.L., L.M.R., M.G., T.M., J.B.J.v.M., L.J.M.A., O.H.F., A.D., S.R., W.Z., S.L.R.K., P.A.P., H.N., L.B.V., D.M.L.-J., J.J.C., P.G., A.H.L., F.B.H., Y.L., L.H., S.D.M., and D.L. provided critical editorial comments. D.L. had primary responsibility for the final content. D.L. is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. All authors read and approved the final manuscript.

References

1. Younossi Z, Anstee QM, Marietti M, et al. Global burden of NAFLD and NASH: trends, predictions, risk factors and prevention. *Nat Rev Gastroenterol Hepatol* 2018;15:11–20
2. Targher G, Byrne CD, Lonardo A, Zoppini G, Barbui C. Non-alcoholic fatty liver disease and risk of incident cardiovascular disease: a meta-analysis. *J Hepatol* 2016;65:589–600
3. Wong RJ, Aguilar M, Cheung R, et al. Nonalcoholic steatohepatitis is the second leading etiology of liver disease among adults awaiting liver transplantation in the United States. *Gastroenterology* 2015;148:547–555
4. Speliotes EK, Yerges-Armstrong LM, Wu J, et al.; NASH CRN; GIANT Consortium; MAGIC Investigators; GOLD Consortium. Genome-wide association analysis identifies variants associated with nonalcoholic fatty liver disease that have distinct effects on metabolic traits. *PLoS Genet* 2011;7:e1001324
5. de Mello VD, Matte A, Perflyev A, et al. Human liver epigenetic alterations in non-alcoholic steatohepatitis are related to insulin action. *Epigenetics* 2017;12:287–295
6. Murphy SK, Yang H, Moylan CA, et al. Relationship between methylome and transcriptome in patients with nonalcoholic fatty liver disease. *Gastroenterology* 2013;145:1076–1087
7. Nano J, Ghanbari M, Wang W, et al. Epigenome-wide association study identifies methylation sites associated with liver enzymes and hepatic steatosis. *Gastroenterology* 2017;153:1096–1106.e2
8. Chalasani N, Younossi Z, Lavine JE, et al. The diagnosis and management of nonalcoholic fatty liver disease: practice guidance from the American Association for the Study of Liver Diseases. *Hepatology* 2018;67:328–357
9. Galanter JM, Gignoux CR, Oh SS, et al. Differential methylation between ethnic sub-groups reflects the effect of genetic ancestry and environmental exposures. *eLife* 2017;6:e20532
10. Houseman EA, Accomando WP, Koestler DC, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics* 2012;13:86
11. Mendelson MM, Marioni RE, Joehanes R, et al. Association of body mass index with DNA methylation and gene expression in blood cells and relations to cardiometabolic disease: a Mendelian randomization approach. *PLoS Med* 2017;14:e1002215
12. Dustin T, Tepper Y, Kentaro H, Luke K, Kosuke I. Mediation: R Package for Causal Mediation Analysis. R package version 4.4.2. Available from <http://CRAN.R-project.org/package=mediation>. Accessed 11 March 2019
13. Anstee QM, Targher G, Day CP. Progression of NAFLD to diabetes mellitus, cardiovascular disease or cirrhosis. *Nat Rev Gastroenterol Hepatol* 2013;10:330–344
14. Hartwig FP, Davies NM, Hemani G, Davey Smith G. Two-sample Mendelian randomization: avoiding the downsides of a powerful, widely applicable but potentially fallible technique. *Int J Epidemiol* 2016;45:1717–1726
15. 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;50:1505–1513
16. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun* 2017;8:1826
17. Galperin MY, Fernández-Suárez XM, Rigden DJ. The 24th annual Nucleic Acids Research database issue: a look back and upcoming changes. *Nucleic Acids Res* 2017;45:D1–D11
18. GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* 2013;45:580–585
19. Sliker 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;6:26
20. Liu Y, Fernandez CA, Smith C, et al. Genome-wide study links PNPLA3 variant with elevated hepatic transaminase after acute lymphoblastic leukemia therapy. *Clin Pharmacol Ther* 2017;102:131–140
21. Bonder MJ, Kurilshikov A, Tigchelaar EF, et al. The effect of host genetics on the gut microbiome. *Nat Genet* 2016;48:1407–1412
22. Wahl S, Drong A, Lehne B, et al. Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. *Nature* 2017;541:81–86
23. Strable MS, Ntambi JM. Genetic control of de novo lipogenesis: role in diet-induced obesity. *Crit Rev Biochem Mol Biol* 2010;45:199–214
24. Frisdal E, Le Lay S, Hooton H, et al. Adipocyte ATP-binding cassette G1 promotes triglyceride storage, fat mass growth, and human obesity. *Diabetes* 2015;64:840–855
25. Ding J, Reynolds LM, Zeller T, et al. Alterations of a cellular cholesterol metabolism network are a molecular feature of obesity-related type 2 diabetes and cardiovascular disease. *Diabetes* 2015;64:3464–3474
26. Leung C, Rivera L, Furness JB, Angus PW. The role of the gut microbiota in NAFLD. *Nat Rev Gastroenterol Hepatol* 2016;13:412–425
27. Soriano-Tárraga C, Jiménez-Conde J, Giralte-Steinhauer E, et al.; GENESTROKE Consortium. Epigenome-wide association study identifies TXNIP gene associated with type 2 diabetes mellitus and sustained hyperglycemia. *Hum Mol Genet* 2016;25:609–619
28. Hartwig FP, Davies NM. Why internal weights should be avoided (not only) in MR-Egger regression. *Int J Epidemiol* 2016;45:1676–1678