Epigenome-Wide Association Study of Fasting Blood Lipids in the Genetics of Lipid-Lowering Drugs and Diet Network Study

Marguerite R. Irvin, PhD*; Degui Zhi, PhD*; Roby Joehanes, PhD*; Michael Mendelson, MD*; Stella Aslibekyan, PhD; Steven A. Claas, MS; Krista S. Thibeault, MS; Nikita Patel, MS; Kenneth Day, PhD; Lindsay Waite Jones, MS; Liming Liang, PhD; Brian H. Chen, PhD; Chen Yao, PhD; Hemant K. Tiwari, PhD; Jose M. Ordovas, PhD; Daniel Levy, MD†; Devin Absher, PhD‡; Donna K. Arnett, PhD†

Background—Genetic research regarding blood lipids has largely focused on DNA sequence variation; few studies have explored epigenetic effects. Genome-wide surveys of DNA methylation may uncover epigenetic factors influencing lipid metabolism.

Methods and Results—To identify whether differential methylation of cytosine-(phosphate)-guanine dinucleotides (CpGs) correlated with lipid phenotypes, we isolated DNA from CD4+ T cells and quantified the proportion of sample methylation at >450 000 CpGs by using the Illumina Infinium HumanMethylation450 Beadchip in 991 participants of the Genetics of Lipid Lowering Drugs and Diet Network. We modeled the percentage of methylation at individual CpGs as a function of fasting very-low-density lipoprotein cholesterol and triglycerides (TGs) by using mixed linear regression adjusted for age, sex, study site, cell purity, and family structure. Four CpGs (cg00574958, cg17058475, cg01082498, and cg09737197) in intron 1 of carnitine palmitoyltransferase 1A (CPT1A) were strongly associated with very-low-density lipoprotein cholesterol (P=1.8×10–21 to 1.6×10–9) and TG (P=1.6×10–26 to 1.5×10–9). Array findings were validated by bisulfite sequencing. We performed quantitative polymerase chain reaction experiments demonstrating that methylation of the top CpG (cg00574958) was correlated with CPT1A expression. The association of cg00574958 with TG and CPT1A expression were replicated in the Framingham Heart Study (P=4.1×10–14 and 3.1×10–13, respectively). DNA methylation at CPT1A cg00574958 explained 11.6% and 5.5% of the variation in TG in the discovery and replication cohorts, respectively.

Conclusions—This genome-wide epigenomic study identified CPT1A methylation as strongly and robustly associated with fasting very-low-density lipoprotein cholesterol and TG. Identifying novel epigenetic contributions to lipid traits may inform future efforts to identify new treatment targets and biomarkers of disease risk. (Circulation. 2014;130:565-572.)

Key Words: cholesterol ■ fatty acids ■ genetics, medical ■ lipids ■ lipoproteins

Genomic studies of lipids and other cardiovascular disease-related traits have traditionally focused on heritable allelic variation, namely, genetic polymorphisms in the nucleotide sequence of DNA in populations. Much progress has been made through research focused on this classical genetic paradigm. For example, multiple large-scale meta-analyses of genome-wide association studies (GWAS) have identified numerous loci associated with fasting blood lipids, many of them constituting novel findings.1 Despite promising discoveries, these loci explain a modest fraction of the observed variance, ≈2% to 12%.1

Clinical Perspective on p 572

Epigenetic changes are functional biochemical alterations in DNA that do not alter the underlying DNA sequence.2 DNA
methylation is an epigenetic process implicated in human disease that involves the methylation of cytosine, usually at cytosine-(phosphate)-guanine ( CpG ) dinucleotides in the promoter region or within genes. This molecular phenotype plays a pivotal role in gene expression by affecting chromatin structure and altering the availability of coding regions to transcription mechanisms. In contrast to DNA sequence variation, epigenetic variation is sensitive to both inherited and environmental inputs.

Genome-wide epigenetic investigation of blood lipids has been largely unexplored. Therefore, we conducted an epigenome-wide association study (EWAS) for fasting very-low-density lipoprotein cholesterol (VLDL-C) and triglycerides (TGs) in 991 participants belonging to the Genetics of Lipid Lowering Drugs and Diet Network (GOLDN) Study. DNA was isolated from CD4+ T cells harvested from stored buffy coats, and methylation was quantified by using the Illumina (San Diego, CA) Infinium Human Methylation450 Beadchip. Our top finding was validated through bisulfite-resequencing and gene expression assays. Subsequently, we sought replication of our top finding for TG in an external study population (the Framingham Heart Study [FHS]) with comparable epigenetic, gene expression, and lipid data.

Methods

Discovery Study Population

The GOLDN study was designed to identify genetic determinants of lipid response to 2 interventions (a high-fat meal challenge and fenofibrate treatment for 3 weeks). In brief, the study ascertained and recruited families from the Family Heart Study at 2 centers, Minneapolis, MN, and Salt Lake City, UT, who self-reported to be white. Only families with at least 2 siblings were recruited for a total of 1327 individuals. Volunteers were required to withhold lipid-lowering agents (pharmaceuticals or nutraceuticals) for at least 4 weeks before the initial visit to be eligible. A total of 1053 individuals met all eligibility requirements. The study protocol was approved by institutional review boards at the University of Minnesota, University of Utah, and Tufts University/New England Medical Center, and the subjects gave informed consent. For the current study, we evaluated fasting TG and VLDL-C among 991 participants for whom epigenetic data were available.

Data Collection

TG and VLDL-C were measured before the diet and drug intervention. Participants were asked to fast for ≥12 hours and abstain from alcohol intake for ≥24 hours. TGs were measured by a glycercol-blanked enzymatic method (Trig/GB, Roche Diagnostics Corporation, Indianapolis, IN). Nuclear magnetic resonance spectroscopy measured VLDL-C (Liposcience, Raleigh, NC). Data on medical history, physical activity, and other lifestyle factors such as alcohol intake, smoking status, and diet were collected by using an interviewer-administered questionnaire.

Methylation Assays

DNA was extracted from CD4+ T cells harvested from stored buffy coats with the use of antibody-linked Invitrogen Dynabeads. DNA was isolated from CD4+ T cells harvested from stored buffy coats, and methylation was quantified by using the Illumina (San Diego, CA) Infinium Human Methylation450 Beadchip. Our top finding was validated through bisulfite-resequencing and gene expression assays. Subsequently, we sought replication of our top finding for TG in an external study population (the Framingham Heart Study [FHS]) with comparable epigenetic, gene expression, and lipid data.

Expression Data

Quantitative polymerase chain reaction was conducted with Life Technologies (Grand Island, NY) TaqMan probes for CPT1A (at 2 different sites [ exon 11–12 and exon 2–3 boundaries ]) and for 5 control genes inuffy coats from 87 GOLDN individuals. These assays yielded nearly identical estimates of CPT1A expression and did not distinguish between the 2 known splice variants of CPT1A. RNA was extracted with Trizol, followed by QIAGEN (Venlo, Netherlands) RNAeasy purification. Two micrograms of input RNA was used in a 40–μL reaction with the use of an iScript CDNA synthesis kit (Bio-Rad, Hercules, CA). Template CDNAs were diluted 1:4, and relative mRNA expression levels were quantified in triplicate 10–μL reactions for fatty acid–metabolizing enzymes including CPT1, ACA2, and SLC25A20. Therefore, we hypothesized that this cell type should reflect underlying epigenetic variation influencing blood lipids while minimizing potential confounding. Third, blood collection is the most viable tissue collection method among healthy individuals.

Stored buffy coats were collected at the same time lipid concentrations were measured. We lysed cells captured on the beads and extracted DNA by using DNase kits (Qiagen, Venlo, Netherlands). We used the Illumina Infinium Human Methylation450 Beadchip (Illumina Inc, San Diego, CA) to interrogate ≈470000 autosomal CpG sites across the genome. A description of the array, and CpG site nomenclature conventions, as well, can be found at http://www.illumina.com/products/methylation_450_beadchip_kits.ilmn. For each assay, 500 ng of DNA was treated with sodium bisulfite (EZ DNA, Zymo Research, Irvine, CA) before standard Illumina amplification, hybridization, and imaging steps. The resulting intensity files were analyzed with Illumina’s GenomeStudio, which generated β-scores (ie, the proportion of total signal from the methylation-specific probe or color channel) and detection P values (defined as the 1 – P value computed from the background model characterizing the chance that the target sequence signal was distinguishable from the negative controls). β-Scores with an associated detection P value of >0.01 were removed, and samples with >1.5% missing data points were eliminated from further analysis. Furthermore, any CpG probes where >10% of samples failed to yield adequate intensity were removed. A total of 58 samples were removed. The quality-control filters applied are comparable to other published reports. The filtered β-scores were then subjected to batch normalization with the ComBat package (R software for nonparametric mode (http://www.bu.edu/jlab/wp-assets/ComBat/Abstract.html)). We performed the normalization in parallel on random subsets of 20000 CpGs per run where each array of 12 samples was used as a batch. Our methods have been extensively described in Absher et al, and the utility of ComBat to correct for batch effects in comparison with other programs is reported. To correct for differing probe chemistry on the Illumina Infinium Human Methylation450 Beadchip, we separately normalized probes from the Infinium I and II chemistries and subsequently adjusted the β-scores for Infinium II probes by using the equation derived from fitting a second-order polynomial to the observed methylation values across all pairs of probes located <50 bp apart (within-chromosome correlations >0.99), where 1 probe was Infinium I and 1 probe was Infinium II. Finally, we eliminated any CpGs where the probe sequence mapped either to a location that did not match the annotation file or to >1 locus. We identified such markers by realigning all probes (with unconverted Cs) to the human reference genome. After these quality-control procedures, there were methylation data from 461281 CpGs. Principal components (PCs) based on the β-scores of all autosomal CpGs passing quality control were generated by using the precomp function in R (V 2.12.1) and used to adjust for cell purity in association analysis similarly to Hidalgo et al. Convolution-estimated CD4+ T-cell percentages were generated by using cell type–specific methylation data from external reference samples by adapting the method of Abbas et al. Predicted CD4+ T-cell percentage purity was highly correlated with PC1 ( r = 0.85, P = 4 × 10–291) but not other PCs, thus supporting the usefulness of methylation PCs in adjusting for cell purity in our analysis. PCs have the added benefit that they can control for unknown confounding, and we chose 4 PCs based on the scree plot of the eigenvalues (see Figure I in the online-only Data Supplement).
by TaqMan gene expression assays with a 7900HT Real Time PCR system according to standard polymerase chain reaction cycling conditions (Life Technologies). All gene expression assays were purchased from Life Technologies, including 2 assays for CPTIA (Hs00912671_m1 and Hs00912801_m1), plus 5 internal control genes for baseline normalization (Hs00168719_m1 [FPBP], Hs00154728_m1 [DECRT], Hs02786624_g1 [GAPDH], Hs00981420_m1 [B2M], Hs00951083_m1 [TFRC]). Because methylation was assayed in CD4+ T cells, but CPTIA expression was measured in buffy coats, the control genes were selected for (1) T-cell expression, (2) neutrophil expression, or (3) red blood cell expression (HBZ), the last 2 of which are to be our greatest source of confounding, so these controls served as both loading controls and as estimators of buffy coat composition.

Reaction cycle threshold (Ct) values for the 2 CPTIA probes yielded nearly identical results and were averaged. Individual Ct values for each endogenous control were median centered. Endogenous control Ct values were averaged per sample and subtracted from median-centered CPTIA target Ct values to generate ΔΔCt input for relative quantitative method (2−ΔΔCt). These normalized CPTIA expression values were compared with the methylation level at cg00574958 corrected for methylation PCs but not other covariates (because the samples were from the same individuals) by using linear regression.

**Bisulfite Resequencing**

To validate the array results we used CATCH-Seq (Ubiquity Genomics, Huntsville, AL) target enrichment to perform bisulfite sequencing of ∼200 kb around CPT1A in 154 participants chosen at the extremes of the observed cg00574958 β-value. We used the same T-cell DNA samples assayed on the Methylation450 array. The 154 DNA samples were prepared for Illumina sequencing by the use of custom methylated adapters to uniquely barcode each sample during library creation. The library creation followed standard Illumina protocols of shearing, end-repair, and adapter ligation. Before bisulfite treatment (QIAEPIGEN Epitect) and polymerase chain reaction amplification, pools of 12 samples were captured with biotinylated probes that were generated from human BAC clones mapping to 200 kb of each endogenous control were median centered. Endogenous control individual Ct values were included in the comparison with the array data. Methylation levels at each CpG across the target region were estimated by using Bowtie2 for alignment and Bismark for the calculation of the methylation proportions.

**Statistical Analysis**

In the discovery stage, we modeled β-score at each CpG site as a function of the log of fasting TG or VLDL-C by using mixed linear regression models adjusted for age, sex, study site, and 4 methylation PCs as fixed effects, and family structure as a random effect. After removing an additional 4 observations because of missing phenotype or covariate data (ie, 1053-58-4), a total of 991 participants were considered in the association analysis. A Bonferroni correction for multiple testing was applied to the discovery models adjusted for 4 lipid traits where \( \sigma = 0.05/(461, 281 \times 4) = 4.5 \times 10^{-8} \) (note results for high-density lipoprotein cholesterol and low-density lipoprotein cholesterol are not presented, because no CpG was significant after correction for multiple testing). Because environmental factors can directly confound EWAS by affecting both the epigenotype and phenotype, sensitivity analysis adjusting for potential major sources of confounding including current alcohol use and current smoking use was conducted. A second sensitivity analysis considered additional sources of confounding including physical activity (hours of moderate to heavy physical activity per week), diet (% energy from carbohydrate and % energy from total fat), and education (as a proxy for socioeconomic status). We have previously described this population to be very genetically homogeneous and have not adjusted for substructure. However, principle components based on single-nucleotide polymorphism (SNP) data from GWAS (SNP PCs) calculated in

Eigensrat (as described by Aslibekyan et al\(^2\)) were available for 72% of our study population, and the first 10 SNP PCs were included in the second sensitivity analysis described. All models were implemented in the R kinship package (lmekin function).\(^2\) The lmekin function fits a linear mixed-effects model that uses the kinship coefficient to define the correlation of random effects, whereas the fixed effects are used to test for associations and adjust for potential confounders. We selected the maximum number of unrelated individuals from our data (n=278) and report the difference in the R² estimate from a linear model with and without the top CpG term (adjusted for the same covariates described for the primary model) as the variance explained by that term. A similar nested-models approach was used to determine the variance explained by the combination of the top 4 CpGs reported.

**GOLDN GWAS, Methyl Quantitative Trait Loci, and Other Annotations**

As described above, the majority of the 991 participants in this study had been previously genotyped at 906,600 loci by using the Affymetrix Genome-Wide Human SNP Array 6.0 (Santa Clara, CA).\(^3\) MACH software (Version 1.0.16) was used to impute nongenotyped SNPs with the use of HapMap Phase II (release 22, Human Genome build 36, hg18) as a reference, resulting in genotypes for 2,529,001 SNPs. To investigate an underlying genetic component to observed epigenetic associations, we performed cis-methyl-quantitative trait locus (QTL) analyses of SNPs within 1 Mb of top CpG findings similarly to that described by Zhi et al.\(^5\) In brief, we used linear mixed models, fit using the lmekin function, to regress the methylation level (β-score) of a CpG site on the genetic variant, adjusting for covariates (age, sex, and study site) as fixed effects and family structure as a random effect. Additionally, we used publicly available regional ENCODE annotations accessed December 13, 2013 to evaluate transcription factor binding sites, chromatin modifications, histone acetylation, and consistency of tissue-specific methylation patterns in the region of our association results. We downloaded the most recent Global Lipids Consortium GWAS data (N=188,577) and used LocusZoom (http://csg.sph.umich.edu/locuszoom) to plot regional SNP results from meta-analysis on Chromosome 11 in the CPTIA region.\(^1\) Finally, we examined expression QTL (eQTL) SNPs in the region of CPTIA available at the blood eQTL browser (http://genenetwork.uw.edu/bloodeqtlbrowser/).

**Replication Population**

The FHS Offspring cohort, initiated in 1971, includes 5124 children and spouses of children of the Original cohort as previously described.\(^2\) Genome-wide DNA methylation and gene expression were assessed from peripheral blood samples (n=2846) collected during examination cycle 8 (2005–2008). Genomic DNA was extracted from buffy coat preparations by using the Gentra Puregene DNA extraction kit (Qiagen). Bisulfite conversion was conducted on genomic DNA by using the EZ DNA Methylation kit, and bisulfite-converted DNA was then hybridized to the Illumina Infinium Human Methylation450 Beadchip in 2 laboratory batches (n=576 and 2270). RNA was extracted from whole blood by using the PAXgene Blood RNA System Kit (Qiagen) with mRNA expression profiling from the Affymetrix Human Exon 1.0 ST GeneChip platform. Rigorous quality-control measures were conducted. For the DNA methylation arrays, 71 samples were excluded owing to poor SNP matching of control positions; 45 were removed because of a missing rate >1%, and 73 were removed as outliers by using multidimensional scaling, for a total of 135 samples excluded (ie, considering overlap among quality-control measures), leaving 2711 samples. We excluded 833 CpG probes because of a missing rate >20%. Methylation data were normalized within laboratory batches by using a DASEN method from the watermelon package.\(^7\) Robust multiplex average method was used to normalize the gene expression values with quality-control measures as previously reported.\(^8\) Data from 2280 Third Generation FHS participants with both measured cell counts and gene expression obtained during the second examination cycle (2008–2011) was used to predict cell count proportions in the current sample by using partial least squares regression.\(^9\) Internal validation using training and testing data sets achieved an r² >0.8 in
the majority of cell lines (except basophils). Then, the prediction was extended to the Offspring cohort by using the resulting coefficients and Offspring cohort gene expression data. We excluded Offspring cohort participants on lipid-lowering therapy and restricted our analyses to participants with both DNA methylation and gene expression, leaving 1261 participants for analyses. We tested associations between levels of log-transformed triglycerides and (1) CpG-specific DNA methylation at cg00574958 and (2) gene-level CPT1A mRNA expression (transcript cluster ID#8337964), using linear mixed regression models, adjusted for age, sex, and estimated cell count as fixed effects, technical covariates (ie, chip) as random effects, and kinship among family members as a random correlation structure, using the pedigreemm package of R. DNA methylation and gene expression were specified as dependent variables in the regression models. Pooled DNA methylation data from the 2 laboratory batches, as opposed to meta-analyzed data, were used to avoid exaggeration of the standard error in the meta-analyzed data attributable to a large population imbalance (10:1) between the 2 groups. Effect sizes of the pooled results match those of the majority of cell lines (except basophils). Then, the prediction was extended to the Offspring cohort by using the resulting coefficients and Offspring cohort gene expression data. We excluded Offspring cohort participants on lipid-lowering therapy and restricted our analyses to participants with both DNA methylation and gene expression, leaving 1261 participants for analyses. We tested associations between levels of log-transformed triglycerides and (1) CpG-specific DNA methylation at cg00574958 and (2) gene-level CPT1A mRNA expression (transcript cluster ID#8337964), using linear mixed regression models, adjusted for age, sex, and estimated cell count as fixed effects, technical covariates (ie, chip) as random effects, and kinship among family members as a random correlation structure, using the pedigreemm package of R. DNA methylation and gene expression were specified as dependent variables in the regression models. Pooled DNA methylation data from the 2 laboratory batches, as opposed to meta-analyzed data, were used to avoid exaggeration of the standard error in the meta-analyzed data attributable to a large population imbalance (10:1) between the 2 groups. Effect sizes of the pooled results match those of the meta-analyzed results. The association between the DNA methylation and gene expression was performed on the residuals after the removal of the fixed and random covariates, along with the kinship correlation structure by using simple linear models, primarily to avoid potential confounding by blood count. Mediation analyses were conducted by using nonparametric estimation methods to model the indirect association between CPT1A DNA methylation at cg00574958 and log-transformed TG levels through changes in CPT1A gene expression. In the mediation analyses, a pathway is specified a priori in which a hypothesized causal factor (cg00574958 DNA methylation) influences a mediator (CPT1A gene expression), which in turn affects the outcome of interest (plasma triglyceride levels). The model assumes no unmeasured confounding or effect modification between the included elements. The proportion mediated describes the average magnitude of indirect association between CPT1A DNA methylation at cg00574958 and log-transformed triglycerides attributed through changes in CPT1A gene expression relative to the average total association. Asymptotic 95% confidence intervals were obtained from nonparametric bootstrapping with 20000 iterations implemented in the mediation package in R. Finally, the results for GOLDN and FHS were subsequently meta-analyzed by using the Fisher method, obtaining the Chernoff bound of the \( \chi^2 \) cumulative distribution function for the \( P \) value.

Results

The GOLDN study population (N=991) was on average 48.8±16 years of age. The sample was 52% female, and 49% were recruited from the Minnesota field center as opposed to the Utah site. Mean fasting VLDL-C (±standard deviation) was 105.3±93 mg/dL and mean fasting TG (±standard deviation) was 137.0±95 mg/dL. Mean TG levels were within the average fasting TG level of 112.3±66 mg/dL. In FHS, the mediation analysis found that 13.5% (95% confidence interval, 7.4–22.2%) of the association of CPT1A expression with ENCODE annotations are displayed in Figure 3.

After quantitative polymerase chain reaction in 87 GOLDN participants, the percentage of methylation (corrected for methylation PCs) at cg00574958 was significantly negatively correlated \( (r=-0.378, P=0.00031) \) with relative CPT1A expression (Figure II in the online-only Data Supplement). CPT1A expression in GOLDN was positively correlated with TG \( (r=0.19, P=0.04, n=87) \). Bisulfite resequencing among 121 participants of the region surrounding CPT1A (coverage \( \geq 100\times \)) demonstrated strong correlation with the percentage of methylation at cg00574958 observed by the methyl450 array with \( r=0.83 \) and \( P=5.4\times10^{-31} \) (see Figure III in the online-only Data Supplement), validating variable methylation at that site. In FHS, increased methylation at cg00574958 was also negatively associated with relative CPT1A expression \( (\beta\text{(se)}=-3.812 (0.52), P=3.1\times10^{-11}) \), whereas CPT1A expression was positively associated with TG after adjustment for age, sex, and family structure \( (\beta\text{(se)}=0.113 (0.02), P=5.6\times10^{-12}) \). In FHS, the mediation analysis found that 13.5% (95% confidence interval, 7.4–22.2%) of the association of CPT1A DNA methylation at cg00574958 with log TG can be attributed to changes in gene expression of CPT1A.

Using GOLDN data, we also evaluated cis-methyl-QTLs in the region of the 4 top CpGs in CPT1A. Results are displayed in Figure IV in the online-only Data Supplement. We show that

\[
-c\log_{10}(P)
\]

Figure 1. Epigenome-wide association Manhattan plot for TG in the discovery data set (n=991). TG indicates triglyceride.
there is little genetic association between SNPs and the CpGs of interest within a 1-Mb window surrounding the CpGs. Additionally, Figure V in the online-only Data Supplement shows a regional plot of GWAS results from the Global Lipids Consortium in the region of CPT1A locus. Data show that the most significant SNP \( P \) value for the TG outcome was nearer to a neighboring gene and not statistically significant \((MRPL21, P=4.0\times10^{-4})\). Finally, there are several strong eQTLs from peripheral blood samples near the 3’ region of CPT1A (Figure VI in the online-only Data Supplement).

### Discussion

The current study reports the top results from an epigenome-wide analysis of fasting TGs and VLDL-C by using the Illumina Infinium Human Methylation450 Beadchip. DNA was isolated from CD4+ T cells harvested from stored lymphocytes. Over 450,000 CpGs were evaluated in 991 participants from the GOLDN Study. Four CpGs in CPT1A were significantly associated with both lipid traits. Adjustment for potential confounders did not materially change the findings. The top CpG marker in CPT1A was validated by bisulfite resequencing and was associated with CPT1A gene expression in a subset of GOLDN participants. The top CpG association and expression results were robustly replicated in the well-characterized FHS population. This study demonstrates the potential importance of epigenetic variation in CPT1A in interindividual differences in fasting TG and VLDL-C levels.

Carnitine plays an essential role in the transfer of long-chain fatty acids across the mitochondrial membrane. CPT1 is a key enzyme in the carnitine-dependent transport of long-chain fatty acids into the mitochondria, and its deficiency results in a decreased rate of fatty acid \( \beta \)-oxidation.33 Three tissue-specific CPT1 isoforms exist, including the liver (CPTIA), muscle (CPTIB), and brain (CPTIC) forms.34 Publicly available gene expression data show that CPTIA is also expressed in CD4+ T cells.35 CPTIA deficiency is a very rare autosomal recessive disorder of fatty acid \( \beta \)-oxidation caused by functional mutations in the gene that have been both directly and indirectly linked to alterations in the active enzyme site.36 Carriers of functional mutations may be at risk for lipid and other metabolic disorders.37-41 However, to the best of our knowledge, no large SNP GWAS or epigenetic studies have highlighted the gene.

Our study reports an inverse relationship between methylation at multiple loci in intron 1 and TG and VLDL-C. We replicated the top marker’s association with TG, including the direction, in 1261 participants from FHS. Gene expression studies in a subset of our GOLDN data show increased methylation at our top CpG site (cg00574958) is associated with decreased relative gene expression (Figure II in the online-only Data Supplement). These gene expression findings in relation to methylation at cg00574958 were validated in FHS. In both GOLDN and FHS, increased CPTIA expression was associated with increased TG. A previous study in an animal model suggests decreased CPTIA activity correlates with increased lipid levels in myocytes.42 Differences in the animal model and our results could be attributable to our cross-sectional design or the cell types examined. Importantly, all associations including the direction of effect replicated across our studies. Mediation analysis in FHS suggests that \( \approx 14\% \) of the association between the top CpG and TG can be attributed to changes in gene expression. Further evaluation of these trends in \( >1 \) tissue type in humans and animals are warranted to help make causal inferences about the relationships between methylation, gene expression, and lipid levels.

Following the replication of our results, we evaluated potential functional genomic mechanisms underlying the epigenetic associations. GOLDN methyl-QTL data demonstrate that nearby SNPs are not strongly associated with the CpGs of interest.

---

**Table. Significant EWAS Signals Observed for VLDL-C and TG in the GOLDN Discovery Data Set (n=991)**

<table>
<thead>
<tr>
<th>Marker</th>
<th>Chr</th>
<th>Location</th>
<th>( \beta ) (SE)†</th>
<th>( P ) Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>VLDL-C</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cg00574958</td>
<td>11</td>
<td>68607622</td>
<td>−0.022 (2.28\times10^{-3})</td>
<td>1.79\times10^{-21}</td>
</tr>
<tr>
<td>cg17058475</td>
<td>11</td>
<td>68607737</td>
<td>−0.025 (3.15\times10^{-3})</td>
<td>3.32\times10^{-15}</td>
</tr>
<tr>
<td>cg10182498</td>
<td>11</td>
<td>68607675</td>
<td>−0.008 (1.24\times10^{-3})</td>
<td>1.06\times10^{-10}</td>
</tr>
<tr>
<td>cg09737197</td>
<td>11</td>
<td>68608225</td>
<td>−0.019 (3.31\times10^{-3})</td>
<td>1.57\times10^{-4}</td>
</tr>
<tr>
<td>TG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cg00574958</td>
<td>11</td>
<td>68607622</td>
<td>−0.032 (2.99\times10^{-3})</td>
<td>1.56\times10^{-20}</td>
</tr>
<tr>
<td>cg17058475</td>
<td>11</td>
<td>68607737</td>
<td>−0.035 (4.14\times10^{-3})</td>
<td>8.90\times10^{-17}</td>
</tr>
<tr>
<td>cg10182498</td>
<td>11</td>
<td>68608225</td>
<td>−0.011 (1.63\times10^{-3})</td>
<td>8.79\times10^{-11}</td>
</tr>
<tr>
<td>cg09737197</td>
<td>11</td>
<td>68607675</td>
<td>−0.027 (4.36\times10^{-3})</td>
<td>1.53\times10^{-9}</td>
</tr>
</tbody>
</table>

EWAS indicates epigenome-wide association study; SE, standard error; TG, triglyceride; and VLDL-C, very-low-density lipoprotein cholesterol.

†The regression parameter estimate (\( \beta \)) represents the change in methylation \( \beta \)-score for each unit change in log (TG) or log (VLDL-C).
Additionally, a regional examination of CPT1A and nearby loci from the Global Lipids Consortium GW AS data found no SNPs (or lipid QTLs) that reach genome-wide significance for the TG trait. Taken together these results suggest that common SNP variation represented by GWAS does not underlie the observed associations. We also investigated peripheral blood eQTLs in the CPT1A region with the use of published data. Several strong eQTLs for CPT1A were identified near the 3′ region of the gene. However, the eQTL SNPs are distant (∼100 kb), and most likely exert an effect on gene expression independent from the CpGs highlighted by our study. Results from ENCODE suggest that the region of interest is an active regulatory site. There are several transcription factor binding sites near our highlighted CpGs (red dotted rectangle in Figure 3). More than one of the annotated transcription factors is involved in lipid metabolism including sterol regulatory element-binding proteins, peroxisome proliferator–activated receptor γ, and upstream transcription factor 1.43–45 The 4 CpGs are just upstream of an active promoter region according to Chromatin State Segmentation by HMM in the HepG2 cell line (bright red region on Figure 3) and in between 2 CpG islands. Additionally, evidence of increased acetylation of lysine 27 of the H3 histone protein (H3K27Ac) also indicates enhanced transcription in the region. Given this functional evidence and observed associations between methylation and gene expression in GOLDN and FHS, we speculate that the CpG methylation state in intron 1 may facilitate an epigenetic program including open chromatin and histone-related enhanced binding of transcription factors, although future studies are needed to evaluate this function.

Few EWAS of lipid traits have been reported and, to the best of our knowledge, have not highlighted CPT1A.46,47 However, a recent metabolomic EWAS provides further in silico validation of our results. Specifically, 649 blood metabolomic traits from 1814 participants of the Kooperative Gesundheitsforschung in der Region Augsburg (KORA) study were assessed for association with methylation at 457,004 CpG sites determined on the Infinium HumanMethylation450 BeadChip platform.48 Our top finding in CPT1A (cg00574958) was among the top findings for VLDL-A (P = 9.23×10−14), and adjustment of the epigenetic association result for nearby SNP variation measured in GWAS did not alter the result. The authors concluded that the association did not exhibit an underlying genetic signal. Validation studies, external replication, and a discussion of the biological relevance of the finding were not provided by that report. Also relevant to our findings is a 2006 publication by Shen et al49 which showed that hypermethylation of the promoter region of CPT1A affects gene expression during differentiation of human embryonic stem cells into neural progenitor/stem cells and suggested a link to lipid metabolism.

The study findings implicate a role for CPT1A methylation in interindividual variation in blood lipid levels beyond DNA sequence variants. Because methylation alters gene expression, findings suggest that future drug development for lipid lowering could be centered around therapeutically altering expression or action of CPT1A. However, our results are several steps upstream of therapeutic implications. First, both GOLDN and FHS have a cross-sectional design limiting causal inference. Specifically, we cannot determine if the observed associations are attributable to methylation effects on lipids or vice versa. The functional mechanism of methylation variation on gene expression also needs further evaluation. Finally, because both GOLDN and FHS represent healthy whites, further replication can expand the generalizability of our results to other ethnic and clinical populations. Despite these limitations, our study is supported by several strengths, including genome-wide epigenetic testing with a dense panel of CpG markers on a large sample of healthy adults with replication complimented by expression and validation studies.
impactful findings, yet known predictors explain only a limited portion of the observed variability. Methylation at CpG sites is an important genomic regulatory mechanism that few studies have evaluated on a genome-wide level in the context of cardiovascular disease–related traits. The current study quantified methylation at >450,000 CpG sites in participants from the GOLDN study. EWAS identified 4 CpGs in CPT1A significantly associated with TG and VLDL-C. Results were validated in an independent study population. The association does not appear to be influenced by surrounding genetic variants, and ENCODE suggests that the region is an active epigenetic regulatory site for the gene. In conclusion, this study emphasizes the importance of expanding genomic studies of lipid-related traits beyond sequence variants and of identifying additional loci that could become useful in the prevention of cardiovascular disease.

Sources of Funding

The GOLDN epigenetics study is funded by National Institutes of Health (NIH)/National Heart, Lung, and Blood Institute (NHLBI) grant R01 HL104135-01. The Framingham Heart Study is supported by NIH/NHLBI contract N01-HC-25195. The Framingham Heart Study is supported by National Institutes of Health (NIH), National Heart, Lung, and Blood Institute (NHLBI), grant R01 HL104135-01. The Framingham Heart Study is supported by NIH/NHLBI contract N01-HC-25195. The Framingham Heart Study is supported by NIH/NHLBI contract N01-HC-25195.


**CLINICAL PERSPECTIVE**

The findings implicate a role for CPT1A methylation in interindividual variation in triglycerides and very-low-density lipoprotein cholesterol beyond DNA sequence variants. Specifically, we report strong association between methylation in the promoter region of CPT1A in CD4+ T cells and triglycerides and very-low-density lipoprotein cholesterol in >900 healthy white adults from the Genetics of Lipid Lowering Drugs and Diet Network Study (GOLDN). Common genetic variation did not underlie the epigenetic signal in GOLDN and resequencing studies validated the array findings for the top cytosine-(phosphate)-guanine site. Results were robustly replicated in >1200 adults from the Framingham Heart Study (FHS) by using DNA from buffy coat preparations. The top cytosine-(phosphate)-guanine site was associated with gene expression in both populations. Because we found that methylation at this site alters gene expression, our findings suggest future lipid-lowering treatment could alter expression or action of CPT1A. However, our findings are based on cross-sectional data in healthy white populations, and further research is needed before CPT1A can be targeted for drug development. Importantly, further replication can expand the generalizability of our results to other ethnic and clinical populations. Longitudinal studies are needed to determine whether the observed associations are attributable to methylation effects on lipids or vice versa. Finally, the functional mechanism of methylation variation on gene expression also needs further evaluation. In summary, the current study represents a promising finding in genomics for the potential future treatment of hypertriglyceridemia, but much additional work is needed before drug development can ensue.
Epigenome-Wide Association Study of Fasting Blood Lipids in the Genetics of Lipid-Lowering Drugs and Diet Network Study

_Circulation_. 2014;130:565-572; originally published online June 11, 2014;
doi: 10.1161/CIRCULATIONAHA.114.009158

_Circulation_ is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2014 American Heart Association, Inc. All rights reserved.
Print ISSN: 0009-7322. Online ISSN: 1524-4539

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circ.ahajournals.org/content/130/7/565

Data Supplement (unedited) at:
http://circ.ahajournals.org/content/suppl/2014/06/11/CIRCULATIONAHA.114.009158.DC1

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in _Circulation_ can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to _Circulation_ is online at:
http://circ.ahajournals.org/subscriptions/
**Supplemental Materials**

Epigenome Wide Association Study of Fasting Blood Lipids in the Genetics of Lipid Lowering Drugs and Diet Network Study


---

Supplemental Table 1. Association models with and without adjustment for current smoking and alcohol use in the discovery dataset.

<table>
<thead>
<tr>
<th>Marker</th>
<th>Main model* Beta</th>
<th>se</th>
<th>P-Value</th>
<th>Sensitivity analysis§ Beta</th>
<th>se</th>
<th>P-Value</th>
<th>Trait</th>
<th>N</th>
<th>Sensitivity analysis† Beta</th>
<th>se</th>
<th>P-Value</th>
<th>Trait</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>cg00574958</td>
<td>-0.0328</td>
<td>2.99E-3</td>
<td>1.56E-26</td>
<td>-0.0326</td>
<td>2.98E-3</td>
<td>2.05E-26</td>
<td>TG</td>
<td>991</td>
<td>-0.0333</td>
<td>3.58E-3</td>
<td>1.45E-21</td>
<td>TG</td>
<td>716</td>
</tr>
<tr>
<td>cg17058475</td>
<td>-0.0351</td>
<td>4.14E-3</td>
<td>8.90E-17</td>
<td>-0.0347</td>
<td>4.12E-3</td>
<td>1.31E-16</td>
<td>TG</td>
<td>991</td>
<td>-0.03792</td>
<td>4.93E-3</td>
<td>5.07E-14</td>
<td>TG</td>
<td>716</td>
</tr>
<tr>
<td>cg01082498</td>
<td>-0.0107</td>
<td>1.63E-3</td>
<td>8.79E-11</td>
<td>-0.0106</td>
<td>1.62E-3</td>
<td>1.08E-10</td>
<td>TG</td>
<td>989</td>
<td>-0.01115</td>
<td>1.91E-3</td>
<td>7.83E-9</td>
<td>TG</td>
<td>714</td>
</tr>
<tr>
<td>cg09737197</td>
<td>-0.0266</td>
<td>4.36E-3</td>
<td>1.53E-9</td>
<td>-0.0262</td>
<td>4.35E-3</td>
<td>2.46E-9</td>
<td>TG</td>
<td>991</td>
<td>-0.0298</td>
<td>5.15E-3</td>
<td>1.14E-8</td>
<td>TG</td>
<td>716</td>
</tr>
<tr>
<td>cg00574958</td>
<td>-0.0223</td>
<td>2.29E-3</td>
<td>1.79E-21</td>
<td>-0.0222</td>
<td>2.28E-3</td>
<td>1.90E-21</td>
<td>VLDL-C</td>
<td>985</td>
<td>-0.0238</td>
<td>2.76E-3</td>
<td>3.98E-17</td>
<td>VLDL-C</td>
<td>712</td>
</tr>
<tr>
<td>cg17058475</td>
<td>-0.0252</td>
<td>3.15E-3</td>
<td>3.32E-15</td>
<td>-0.025</td>
<td>3.13E-3</td>
<td>3.81E-15</td>
<td>VLDL-C</td>
<td>985</td>
<td>-0.0265</td>
<td>3.77E-3</td>
<td>4.72E-12</td>
<td>VLDL-C</td>
<td>712</td>
</tr>
<tr>
<td>cg01082498</td>
<td>-0.0081</td>
<td>1.24E-3</td>
<td>1.06E-10</td>
<td>0.0080</td>
<td>1.23E-3</td>
<td>1.11E-10</td>
<td>VLDL-C</td>
<td>983</td>
<td>-0.0084</td>
<td>1.46E-3</td>
<td>1.25E-8</td>
<td>VLDL-C</td>
<td>710</td>
</tr>
<tr>
<td>cg09737197</td>
<td>-0.0189</td>
<td>3.31E-3</td>
<td>1.57E-8</td>
<td>-0.0186</td>
<td>3.30E-3</td>
<td>2.11E-8</td>
<td>VLDL-C</td>
<td>985</td>
<td>-0.0209</td>
<td>3.93E-3</td>
<td>1.30E-7</td>
<td>VLDL-C</td>
<td>712</td>
</tr>
</tbody>
</table>

*adjusted for age, sex, center, methylation PCs, and family relationship

§ additionally adjusted for current smoking and current alcohol use

† additionally adjusted for education, hours of moderate to heavy physical activity a week, % energy from carbohydrate, % energy from total fat, and 10 SNP PCs
Supplemental Figure 1. Scree plot of 20 eigenvalues.

Supplemental Figure 2. Correlation between relative CPT1A expression measured using qPCR and % methylation at cg00574958 measured on the Illumina methyl450 array.
Supplemental Figure 3. Correlation between the beta score for cg00574958 measured on the Illumina methyl450 array (y-axis) and by bisulfite resequencing.

Supplemental Figure 4: cis-meQTLs describing the SNP-CpG association near CPT1A. X-coordinate is the location of the CpG site. The red arrow points to the region of the top four CpG EWAS hits.
Supplemental Figure 5: Regional plot of genome-wide association results from the Global Lipids Consortium in the chromosome 11 region of CPT1A.
Supplemental Figure 6: Expression quantitative trait loci (eQTL) single-nucleotide polymorphisms in the region of CPT1A from meta-analysis in non-transformed peripheral blood samples from 5,311 individuals with replication in 2,775 individuals near CPT1A recently published by Westra et al. 1

References