Obesity (2015) 23, 1493–1501. doi:10.1002/oby.21111

Introduction

Body mass index (BMI) and waist circumference (WC) are quantitative traits characterized by multifactorial etiology, relevance to disease risk, and high heritability (1). However, known genetic polymorphisms explain only a modest fraction of inherited variation in both obesity-related traits (2). Emerging evidence suggests that epigenetic changes play a key role in both heritable and environmental

1 Department of Epidemiology, School of Public Health, University of Alabama at Birmingham, Birmingham, Alabama, USA. Correspondence: Stella Aslibekyan (saslibek@uab.edu)2 Division of Epidemiology and Community Health, School of Public Health, University of Minnesota, Minneapolis, Minnesota, USA3 Population Sciences Branch, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Maryland, USA4 Framingham Heart Study, Framingham, Massachusetts, USA5 Department of Cardiology, Boston Children’s Hospital, Boston, Massachusetts, USA6 Department of Biostatistics, School of Public Health, University of Alabama at Birmingham, Birmingham, Alabama, USA7 Division of Biostatistics, School of Public Health, University of Minnesota, Minneapolis, Minnesota, USA8 Department of Epidemiology, School of Public Health, University of Alabama at Birmingham, Birmingham, Alabama, USA9 Department of Biostatistics, School of Public Health, University of Alabama at Birmingham, Birmingham, Alabama, USA10 Department of Biostatistics, Boston University, Boston, Massachusetts, USA11 Brown Foundation Institute of Molecular Medicine, University of Texas Health Science Center, Houston, Texas, USA12 Human Genetics Center, School of Public Health, University of Texas Health Science Center at Houston, Texas, USA13 HudsonAlpha Institute for Biotechnology, Huntsville, Alabama, USA14 Division of Laboratory Medicine and Pathology, University of Minnesota, Minneapolis, Minnesota, USA15 School of Medicine, University of Utah, Salt Lake City, Utah, USA16 Division of Statistical Genomics, Department of Genetics, Washington University in St Louis, St Louis, Missouri, USA17 Department of Epidemiology, Atherothrombosis and Imaging, Centro Nacional De Investigaciones Cardiovasculares, Madrid, Spain18 Instituto Madrileño De Estudios Avanzados Alimentacion, Madrid, Spain19 Jean Mayer USDA Human Nutrition Research Center on Aging, Tufts University, Boston, Massachusetts, USA.

Disclosure: The authors declared no conflict of interest.

Author contributions: Substantial contributions to conception and design: DMA, DKA, SA, EB, JB, EW, MFI, DL, JMO, HKT; Acquisition of data: DMA, DKA, EB, IBB, JB, EW, MF, MLG, PNH, DL, JMO, JSP, KST, KHT, MYT; Analysis: DMA, DKA, SA, EW, WG, LL, LL, CL, MM, JS, KST, DZ; Interpretation of data: DMA, DKA, SA, EW, WG, BH, LL, LL, CM, JSP, JS, KST, DZ; Drafting of article: SA. All authors were involved in revising the manuscript critically for important intellectual content, and all granted final approval of the submitted and published versions.

* Devin M. Absher and Donna K. Arnett are regarded as joint senior authors.

Additional Supporting Information may be found in the online version of this article.

Received: 23 July 2014; Accepted: 6 March 2015; Published online 25 June 2015. doi:10.1002/oby.21111
influences on obesity (3). Consistent with that paradigm, recent studies identified DNA methylation patterns in several biologically relevant regions as important correlates of BMI (4-6).

Epigenetic processes such as DNA methylation underlie the gene–environment interactions in complex traits by changing gene expression (7). This relationship between DNA methylation and gene expression is further complicated by DNA sequence variation, which contributes to both processes through expression- (eQTL) and methylation-quantitative trait loci (meQTL). In the pathogenesis of obesity, changes in DNA methylation may represent both the cause and the consequence, acting through inflammation, oxidative stress, hypoxia, or other biological pathways (4,8,9). Beyond changes in mean DNA methylation levels, obesity is also associated with increased variance in DNA methylation, possibly reflecting adaptation to changing environmental conditions (5).

Because most population studies to date have pursued the candidate gene approach, many potentially novel associations between DNA methylation patterns and obesity remain unexplored. Large-scale studies with well-characterized methylome data are poised to fill this gap. To that end, we quantified methylation levels at ~470,000 cytosine-phosphate-guanine (CpG) sites in CD4+ T-cells from 991 participants of the Genetics of Lipid Lowering Drugs and Diet Network (GOLDN) study (https://dsgweb.wustl.edu/goldn/). Our choice of cell type was informed by evidence of association between obesity-related immune dysfunction and DNA methylation (6) as well as relative abundance of CD4+ T-cells. Subsequently, we performed an epigenome-wide association study of BMI and WC and replicated several of our top findings in blood leukocyte samples from the Framingham Heart Study (FHS) and the Atherosclerosis Risk in Communities (ARIC) study.

DNA isolation and bisulfite conversion
In the GOLDN study, DNA was isolated from CD4+ T-cells harvested from frozen buffy coat samples from peripheral blood using positive selection by antigen-specific magnetic beads (Invitrogen, Carlsbad, CA). We chose CD4+ T-cells for reasons of both biological plausibility, as many metabolism-related genes are expressed in immune cells (12), as well as logistics, as they are the most abundant lymphocyte in humans. We lysed cells captured on the beads and extracted DNA using DNeasy kits (Qiagen, Venlo, The Netherlands). Bisulfite conversion was performed on 500 ng of genomic DNA using the EZ DNA Methylation Kit (Zymo Research, Irvine, CA) prior to quantifying epigenome-wide methylation (13).

Epigenome-wide association study
All three studies used the Infinium HumanMethylation450 array (Illumina, San Diego, CA) to quantify genome-wide DNA methylation. The GOLDN analysis is described in extensive detail in a manuscript from our group (13). After the standard steps of whole genome amplification, hybridization, and imaging, GOLDN, ARIC, and FHS investigators used Illumina GenomeStudio software to estimate β scores, defined as the proportion of total signal from the methylation-specific probe or color channel, and detection P-values, defined as the probability that the total intensity for a given probe falls within the background signal intensity.

Normalization and quality control
During the quality control stage in GOLDN, we removed any β scores with an associated detection P-value greater than 0.01 and samples with more than 1.5% missing data points, as well as CpGs where the probe sequence mapped either to a location that did not match the annotation file, or to more than one locus (13). Additionally, we excluded any CpG probes where more than 10% of samples failed to yield adequate intensity. A total of 991 samples and 461,281 CpGs were used in the analysis. The resulting β scores were normalized using the ComBat package (http://www.bio.uci.edu/jlab/wp-assets/ComBat/Download.html) for R software (14). A detailed published description of our normalization methods and the performance of ComBat is available (13). Briefly, we performed the normalization on random subsets of 20,000 CpGs per run, with each array of 12 samples used as a “batch,” adjusting for both plate and position on the plate. We separately normalized probes from the Infinium I and II chemistries and subsequently adjusted the β scores for Infinium II probes as described earlier (13). To adjust for cell purity, we generated principal components (PCs) based on the β scores of all autosomal CpGs that passed quality control. Additionally, we estimated predicted CD4+ T-cell purity using a linear model based deconvolution method derived from the methods described by Abbas et al. (15) and Houseman et al. (16). To estimate the percentage of CD4+ T-cells versus other cells in each sample, we used a reference data set of CD4+ T-cells and granulocytes extracted from fresh blood sample. Predicted CD4+ T-cell purity (%) was robustly associated with the first PC ($r^2 = 0.85$) but not other PCs (17).

Measurement of obesity traits and covariates
On the day of anthropometric data collection, GOLDN participants wore light clothes and no shoes. Weight was measured by a beam balance and height was ascertained by a stadiometer. BMI was calculated as weight in kilograms divided by height in meters squared. WC was measured over the unclothed abdomen at the umbilicus at the end of a normal expiration (11). Smoking (current vs. not) and demographic covariates were self-reported via questionnaire.

Methods

Study subjects
The design of the family-based GOLDN study is described in detail in prior publications (10). Briefly, we recruited families with at least two siblings from the participants of the National Heart, Lung, and Blood Institute Family Heart Study at the genetically homogeneous sites of Minneapolis and Salt Lake City. All participants (n = 1,327) self-identified as European Americans. The goal of the trial was to identify genetic and epigenetic factors that mediated response to acute lipid-raising (i.e., postprandial lipemia challenge) or lipid-lowering (3 week fenofibrate therapy) interventions. Before each acute lipid-raising (i.e., postprandial lipemia challenge) or lipid-lowering (i.e., postprandial lipemia challenge) or lipid-lowering (3 week fenofibrate therapy) interventions. Before each study visit, participants abstained from eating for at least 8 h, drinking alcohol for 24 h, and using lipid-lowering medications for 4 weeks. We collected DNA samples and BMI/WC measurements for this epigenome-wide study at the baseline visit. The descriptions of the FHS and ARIC populations, as well as other methods pertaining to replication analyses, are summarized in Supporting Information. GOLDN, FHS, and ARIC study protocols were approved by Institutional Review Boards at each participating university and all participants provided written informed consent.

Measurement of obesity traits and covariates
On the day of anthropometric data collection, GOLDN participants wore light clothes and no shoes. Weight was measured by a beam balance and height was ascertained by a stadiometer. BMI was calculated as weight in kilograms divided by height in meters squared. WC was measured over the unclothed abdomen at the umbilicus at the end of a normal expiration (11). Smoking (current vs. not) and demographic covariates were self-reported via questionnaire.

DNA isolation and bisulfite conversion
In the GOLDN study, DNA was isolated from CD4+ T-cells harvested from frozen buffy coat samples from peripheral blood using positive selection by antigen-specific magnetic beads (Invitrogen, Carlsbad, CA). We chose CD4+ T-cells for reasons of both biological plausibility, as many metabolism-related genes are expressed in immune cells (12), as well as logistics, as they are the most abundant lymphocyte in humans. We lysed cells captured on the beads and extracted DNA using DNeasy kits (Qiagen, Venlo, The Netherlands). Bisulfite conversion was performed on 500 ng of genomic DNA using the EZ DNA Methylation Kit (Zymo Research, Irvine, CA) prior to quantifying epigenome-wide methylation (13).

Epigenome-wide association study
All three studies used the Infinium HumanMethylation450 array (Illumina, San Diego, CA) to quantify genome-wide DNA methylation. The GOLDN analysis is described in extensive detail in a manuscript from our group (13). After the standard steps of whole genome amplification, hybridization, and imaging, GOLDN, ARIC, and FHS investigators used Illumina GenomeStudio software to estimate β scores, defined as the proportion of total signal from the methylation-specific probe or color channel, and detection P-values, defined as the probability that the total intensity for a given probe falls within the background signal intensity.

Normalization and quality control
During the quality control stage in GOLDN, we removed any β scores with an associated detection P-value greater than 0.01 and samples with more than 1.5% missing data points, as well as CpGs where the probe sequence mapped either to a location that did not match the annotation file, or to more than one locus (13). Additionally, we excluded any CpG probes where more than 10% of samples failed to yield adequate intensity. A total of 991 samples and 461,281 CpGs were used in the analysis. The resulting β scores were normalized using the ComBat package (http://www.bio.uci.edu/jlab/wp-assets/ComBat/Download.html) for R software (14). A detailed published description of our normalization methods and the performance of ComBat is available (13). Briefly, we performed the normalization on random subsets of 20,000 CpGs per run, with each array of 12 samples used as a “batch,” adjusting for both plate and position on the plate. We separately normalized probes from the Infinium I and II chemistries and subsequently adjusted the β scores for Infinium II probes as described earlier (13). To adjust for cell purity, we generated principal components (PCs) based on the β scores of all autosomal CpGs that passed quality control. Additionally, we estimated predicted CD4+ T-cell purity using a linear model based deconvolution method derived from the methods described by Abbas et al. (15) and Houseman et al. (16). To estimate the percentage of CD4+ T-cells versus other cells in each sample, we used a reference data set of CD4+ T-cells and granulocytes extracted from fresh blood sample. Predicted CD4+ T-cell purity (%) was robustly associated with the first PC ($r^2 = 0.85$) but not other PCs (17).
TABLE 1 Demographic and anthropometric characteristics of the study populations

|                      | GOLDN (n = 991) | ARIC (n = 2,097) | FHS case-control (n = 1,935) | FHS random sample (n = 442) |
|----------------------|-----------------|-----------------|-----------------------------|----------------------------|
| Age, years\(^a\)     | 49 ± 16         | 56 ± 6          | 65 ± 9                      | 71 ± 8                     |
| Sex, % female        | 52              | 64              | 61                          | 30                         |
| Race, %              |                 |                 |                             |                            |
| European American, % | 100             | -               | 100                         | 100                        |
| African American, %  | -               | 100             | -                           | -                          |
| Current smokers, %   | 7               | 24              | 9                           | 6                          |
| Body mass index, kg/m\(^2\) | 28 ± 6 | 30 ± 6          | 28 ± 6                      | 29 ± 5                     |
| Waist circumference, cm | 97 ± 16 | 101 ± 15        | -                           | -                          |

\(^a\)Values are shown as mean ± SD or %.

Statistical analysis

In the GOLDN study, we modeled associations between methylation scores at each CpG site and obesity traits using linear mixed models, adjusted for age, sex, study site, current smoking, the first 4 CD4+ T-cell purity PCs, and pedigree as a random effect using the lmekin function of the kinship package (http://cran.r-project.org/src/contrib/Archive/kinship/) in R (18). Previously, we tested genetic ancestry as a potential confounder and did not find associations with the outcome, likely due to the homogeneity of our study population. We implemented a stringent Bonferroni correction to address the multiple testing problem, setting the statistical significance level at 0.05/470,000 = 1.1 × 10^−7. We constructed a Manhattan plot to visualize the results.

CpG sites that reached a P-value < 1.0 × 10^−7 (8 for BMI and 5 for WC) were tested in the replication stage. Consequently, the Bonferroni-corrected P-value in the replication stage was 0.006 for BMI and 0.01 for WC. BMI replication analyses were performed in both FHS and ARIC, while WC data were only available in ARIC. Further details on statistical methods in the replication stage are available in Supporting Information. We meta-analyzed P-values from GOLDN, 2 batches of FHS, and ARIC. Because of the different magnitude of effects between FHS and the other two cohorts at the CPT1A locus, the meta-analysis P-value calculation for cg00574958 used the Chernoff bound of the chi-squared cumulative distribution function. At all other loci, we calculated the summary P-values exactly.

Results

Table 1 summarizes the general characteristics of the GOLDN, FHS, and ARIC populations. Approximately half of GOLDN participants were recruited at each of the Minnesota and Salt Lake City study sites, while 93% of the ARIC cohort came from the Minnesota site, with only 7% recruited at the North Carolina Site; all FHS participants were recruited in Framingham, Massachusetts. The discovery and replication cohorts differ most strikingly on the selection of samples for methylation analysis, with the former restricted to CD4+ T-cells and the latter both using whole blood. FHS participants were considerably older than either GOLDN or ARIC participants. With the exception of the sex distribution, the two FHS laboratory batches were demographically and anthropometrically similar. The main distinction between GOLDN/FHS and ARIC is the racial composition of the cohorts; while GOLDN and FHS are entirely comprised of European Americans, the methylation data in ARIC was only available on African Americans. Compared to GOLDN, ARIC participants were on average older, more likely to be female, and more likely to report current smoking. The mean values for BMI and WC in all three cohorts slightly exceeded the clinical guidelines for healthy weight (19, 20).

Using the Bonferroni-corrected statistical significance level of 1.1 × 10^−7, we identified eight loci where methylation status was associated with BMI (Table 2, Figure 1) and five loci for WC (Table 3, Figure 2). In the BMI case–control batch from FHS, we successfully replicated significant associations with four out of the eight BMI loci, located in CPT1A, PHGDH, CD38, and a site on chromosome 10 located in the promoter of a long non-coding intergenic RNA (lincRNA). Of those, associations with CpG sites in CPT1A and PHGDH also replicated in the larger FHS random sample. In ARIC, we replicated associations for both BMI and WC with CpG sites in CPT1A, CD38, and the lincRNA on chromosome 10, as well as the association of WC with the methylation of the locus in PHGDH (borderline significant) and the association of BMI with the methylation of the locus in AHRR. The directions of the associations were also confirmed: at a CpG site in the first intron of CPT1A, the β score was inversely associated with BMI and WC, while the methylation status of a CpG site adjacent to the transcriptional start site of CD38 was positively associated with both phenotypes in all three studies. The direction of the associations between the methylation status of the lincRNA and PHGDH CpG sites and obesity traits was also consistent across studies.

Using information provided by the Illumina array manufacturer, we functionally annotated the top hits and found that four of them (cg26164488, cg07504977, cg26140475, and cg25349939) represent enhancer elements (bioinformatically determined), one (cg14476101) is located within a reprogramming-specific differentially methylated region, and one (cg26140475) is located within a general differentially methylated region. The cg26140475 hit is also located at a DNAse hypersensitive site. Finally, Illumina has bioinformatically determined that the hits in CD38, AHRR, and the intergenic region on chromosome 16 are implicated in promoter activity.

www.obesityjournal.org
Table 2: Top CpG methylation sites associated with body mass index in GOLDN (n = 951), HFS (n = 1,935) at the level of genome-wide significance (P < 1.1 x 10^-7) in the discovery phase

| Marker | Chr | Gene | Beta ^±| SE | P | Meta P |
|--------|-----|------|--------|-----|----|--------|
| cg0564586 | 11  | CPT1A | -0.00932 ± 0.0001 | 7.6 x 10^-10 | 6.9 x 10^-10 | 2.7 x 10^-10 |
| cg12824395 | 4  | CD38 | 0.0013 ± 0.0002 | 7.5 x 10^-10 | 7.5 x 10^-10 | 2.7 x 10^-10 |
| cg01728715 | 4  | AHRR | 0.0012 ± 0.0002 | 7.5 x 10^-10 | 7.5 x 10^-10 | 2.7 x 10^-10 |
| cg07504977 | 10  | PHGDH | 0.0013 ± 0.0002 | 7.5 x 10^-10 | 7.5 x 10^-10 | 2.7 x 10^-10 |
| cg26140475 | 8  | PHGDH | 0.0013 ± 0.0002 | 7.5 x 10^-10 | 7.5 x 10^-10 | 2.7 x 10^-10 |

**Discussion**

Our study contributes to the growing body of evidence in support of epigenetic correlates of complex disease phenotypes. Specifically, we have discovered associations between DNA methylation measured in immune cells at several CpG sites and two obesity-related traits, BMI and WC, in a cohort of 991 healthy adults of European American descent. The top associations were replicated in whole blood samples taken from two independent study samples including both European Americans and African Americans, indicating the robustness of our findings.

The biological significance of the CPT1A locus is hard to overstate. Using data from GOLDN and FHS, we recently identified a CpG site in CPT1A as an important epigenetic determinant of fasting triglyceride and very low-density lipoprotein cholesterol levels (17). Interestingly, the observed associations with BMI and WC did not persist upon adjustment for triglycerides (data not shown). CPT1A encodes carnitine palmitoyltransferase 1A, the rate-limiting enzyme for mitochondrial fatty acid oxidation. Public databases show that CPT1A is expressed in CD4+ T-cells (21) and is implicated in several metabolic processes. Prior studies have linked genetic variation in CPT1A and indices of obesity. For example, a large-scale candidate gene study found that the L479 allele is associated with reduced adiposity in Yup’ik individuals (22). Another population-based study conducted in French Canadians found that CPT1A genotype modified the association between dietary fat intake and adiposity variables (BMI, weight, and WC) (23). In light of these studies and the well-established connection between sequence variation and epigenetic changes (24), we have previously conducted a genome-wide study to search for CPT1A methylation quantitative trait loci; however, no evidence of the latter was found (17).

The InfiniumHuman Methylation450 array measurement of CPT1A methylation in GOLDN was validated using bisulfite sequencing and found to be inversely correlated with CPT1A expression in buffy coats (17). Based on previously published reports and our findings, the following paradigm emerges: higher methylation status of CPT1A results in decreased expression of the gene, which in turn is negatively correlated with BMI and WC. Because this is a cross-sectional study, we cannot infer temporality or causality; future studies should consider repeated measurement analysis to establish temporality of the association. Additionally, as dietary factors such as intake of long-chain monounsaturated fatty acids have also been shown to regulate CPT1A expression (25) as well as DNA methylation patterns (26), future investigations would benefit from incorporating dietary information to provide further insight into these findings.

The second CpG methylation locus that achieved epigenome-wide significance and was replicated in ARIC (as well as in the case-control study from FHS) is located 289 bp upstream of the transcription start site of CD38, an immunologically relevant gene expressed in CD4+ T-cells. It falls in the upstream shore of a CpG island, has a strong polymerase II (Pol II) ChIP-seq signal and a minor peak of the histone mark H3K27Ac in immune cells, and is differentially methylated by cell type (Figure 3). These bioinformatics data likely indicate that methylation at this locus is driving Pol II activity and may be correlated with CD38 expression in CD4+ T-cells. Several lines of evidence link the gene product of CD38 with metabolic traits. First, CD38-knockout mice are resistant to developing...
diet-induced obesity, liver steatosis, and glucose intolerance via a mechanism that involves an increase in intracellular NAD(+) levels and decreased protein acetylation through sirtuin activation (27,28). In human macrophages and adipocytes, quercetin, which is a CD38 inhibitor, attenuates inflammation and insulin resistance (29). In our study, we observed a positive relationship between CD38 methylation and adiposity traits. Because we did not measure CD38 RNA directly, we cannot fully describe the mechanisms linking changes in DNA methylation and adiposity traits, especially in light of recent evidence suggesting that DNA methylation may be correlated with both positive and negative changes in gene expression (24). Because of the immunologic relevance of CD38, it is likely that the difference between the magnitude of estimates in the discovery and replication cohorts (as well as non-replication in the random sample from FHS) is due to using CD4+ T-cells vs. whole blood samples, in which differential immune cell composition could dilute the association. Our findings emphasize the relevance of CD38 to human metabolic traits and may lay the groundwork for understanding the epigenetic mechanisms of emerging therapies for obesity and inflammation.

Another promising finding emerged from an intergenic region on chromosome 10. The cg07504977 locus is located on the north shelf of a CpG island and maps to the promoter of a lincRNA, LINC00263 (Figure 4). The methylation status of this locus is positively associated with BMI in all three studies, as well as with WC in both GOLDN and ARIC. Although the specific function of LINC00263 is unknown, evidence points to the functional importance of long non-coding RNAs in adipogenesis (30) and obesity-related syndromes (31). In particular, a recent study demonstrated that 10% of all transcriptionally active regions in mature adipocytes map to chromosome 10, with many of those regions clustering proximally to LINC00263 (32). As shown in Figure 4, that region also contains H3K27Ac marks, providing further evidence for its regulatory relevance.

We identified associations between the methylation status of a CpG locus in PHGDH and BMI that were robust in samples from GOLDN, FHS, and borderline significant in ARIC. PHGDH encodes the phosphoglycerate dehydrogenase enzyme, which catalyzes the first step in the phosphorylation pathway of serine biosynthesis.

**TABLE 3** Top CpG methylation sites associated with waist circumference in GOLDN (n = 991) and ARIC (n = 2,097) at the level of genome-wide significance (P < 1.1 × 10⁻⁷) in the discovery phase

| Marker   | Chr | Gene   | β      | SE      | P            | β      | SE      | P            | Meta P  |
|----------|-----|--------|--------|---------|--------------|--------|---------|--------------|---------|
| cg00574958 | 11  | CPT1A  | -0.0003| 0.00005 | 2.6 × 10⁻¹⁰ | -0.003 | 0.0004 | 5.8 × 10⁻¹⁷ | 9.9 × 10⁻²³ |
| cg04332373 | 4   | CD38   | 0.0005 | 0.00008 | 6.2 × 10⁻¹⁰ | 0.003  | 0.0009 | 8.0 × 10⁻⁵  | 1.6 × 10⁻¹² |
| cg07504977 | 10  | NA     | 0.0004 | 0.00007 | 7.6 × 10⁻⁹  | 0.009  | 0.002  | 3.3 × 10⁻⁷  | 8.9 × 10⁻¹⁴ |
| cg25349939  | 2   | GTDC1  | 0.0005 | 0.00008 | 1.6 × 10⁻⁴  | 0.002  | 0.001  | 0.05        | 1.8 × 10⁻⁸  |
| cg14476101  | 1   | PHGDH  | -0.0006| 0.0001  | 1.7 × 10⁻⁸  | -0.005 | 0.002  | 0.01        | 4.0 × 10⁻⁹  |
Previously, the methylation of the same locus has been linked to blood concentrations of 4-androsten-3beta,17beta-diol, a steroid hormone upregulated in obesity (33,34). The metabolic implications of PHGDH are further underscored by evidence of its overexpression in approximately 70% of estrogen receptor-negative breast cancers, as it drives metabolic changes leading to rapid cellular growth (35). While the effects of phosphoglycerate dehydrogenase on oncogenesis are well-characterized, our findings warrant further examination of this enzyme in other metabolic and obesity-related phenotypes.

Interestingly, we did not observe statistically significant associations between BMI and the methylation status of three sites in HIF3A, which was recently identified and validated as a novel epigenetic obesity determinant in a genome-wide analysis (4). Specifically, in

![Figure 2](https://example.com/figure2.png)

**Figure 2** Manhattan plot of epigenome-wide results of testing for association between epigenome-wide methylation and waist circumference. The X-axes display the chromosome on which the CpG site is located; the Y-axes display –log10(P-value). The red horizontal line indicates the threshold for epigenome-wide statistical significance after a Bonferroni correction. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

![Figure 3](https://example.com/figure3.png)

**Figure 3** ENCODE annotation of the genomic region containing the transcription start site of CD38. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]
the analysis of the GOLDN data the P-values for β scores at cg22891070, cg27146050, and cg16672562 were 0.13, 0.35, and 0.16, respectively. Differences could be due to cohort makeup (healthy individuals in GOLDN vs. both healthy blood donors and MI patients in the Cardiogenics Consortium or venous thrombosis patients in one of the two replication cohorts), choice of tissue (CD4+ T-cells vs. whole blood samples), or chance. Although findings in our study as well as in Cardiogenics were independently replicated, these discrepancies highlight the need for future functional studies as follow-up to large-scale methylation assays.
Our study documents novel associations between CD4+ T-cell DNA methylation in several biologically plausible genomic regions and human obesity traits, i.e. BMI and WC. The associations with the CPT1A locus were robust in all replication analyses despite the differences in sample collection types (isolated CD4+ T-cells in GOLDN vs. whole blood in FHS and ARIC) and racial composition. The sample size, high resolution of epigenetic phenotyping, and successful replication represent distinguishing strengths of our study. However, the results presented here need to be interpreted in light of several potential limitations. First, because DNA methylation can both precede phenotypic changes as well as result from environmental influences such as obesity (36), follow-up analyses of repeated measurements are required to further causal understanding of the observed associations. Second, the effect of our findings on gene expression has been previously established for CPT1A but not for other genes that contain the top differentially methylated CpG sites in this study. Third, because both DNA methylation and adiposity traits are associated with a range of environmental inputs, residual confounding remains a possible explanation for our findings despite covariate adjustment. For example, the third top hit in our BMI analysis is a CpG site in AHR, the methylation status of which was compellingly linked to smoking (37,38). When we restricted the GOLDN analyses to never smokers (n = 703, data not shown), the P-values for AHR increased from 7.5 × 10⁻⁹ to 2.2 × 10⁻⁷, beyond the threshold for epigenome-wide significance. The same attenuation did not take place for other top hits, namely in CPT1A and CD38, suggesting that those associations are less likely to be a mere artifact of smoking. Fourth, the CpG sites in CPT1A, CD38, and AHR contain known sequence variants (rs78442314, rs77882313, and rs6899226 respectively), which could influence the binding potential of the probe and affect methylation measurements. However, prior genome-wide analyses did not yield evidence of association between any sequence polymorphism and the methylation score at the top CPT1A locus (17). Fifth, the between-cohort differences in statistical methods, especially in normalization approaches, may contribute to non-replication; while that is unlikely to impact our top findings, it may present problems for loci with more modest signals (39). Finally, the methylation patterns in CD4+ T-cells used in the study may have been altered during the freezing and thawing process and may not generalize to other cell types. The differences between CD4+ T-cells and other cell types present in whole blood samples are particularly relevant as such heterogeneity can confound epigenomic association results, mirroring statistical issues created by population stratification. Despite that concern, we are reassured by successful replication of our top findings in whole blood samples from ARIC and FHS.

On balance, our study highlights the importance of DNA methylation in complex human traits and identifies several candidate regions as promising epigenetic correlates of obesity. Further investigations of these regions will further characterize the heritable components of BMI and WC beyond sequence variants, potentially leading to novel therapeutic and preventative approaches to the challenge posed by the global obesity epidemic.

Acknowledgments

The authors thank the staff and participants of ARIC, FHS, and GOLDN for their important contributions. The work on the GOLDN study has been funded by the National Institutes of Health (NIH) National Heart, Lung, and Blood Institute (NHLBI), grant U01HL072524-04. The Framingham Heart Study is supported by NHLBI contract N01HC25195. The ARIC study is carried out as a collaborative study supported by NHLBI contracts (HHSN268201100005C, HHSN268201100006C, HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN268201100010C, HHSN268201100011C, and HHSN268201100012C), R01HL087641, R01HL59367, and R01HL86694; National Human Genome Research Institute contract U01HG004042; and NIH contract HHSN268200625226C. Infrastructure for ARIC was partly supported by grant UL1RR025005, a component of the NIH and specifically the NIH Roadmap for Medical Research. Support for the ARIC exome chip genotyping was provided by Building on GWAS for NHLBI-diseases: the U.S. CHARGE consortium through the NIH American Recovery and Reinvestment Act of 2009 (ARRA) (5RC2HL102419).

© 2015 The Obesity Society

References

1. Hemani G, Yang J, Vinkhuyzen A, et al. Inference of the genetic architecture underlying BMI and weight with the use of 20,240 sibling pairs. Am J Hum Genet 2013;93:865-875.
2. Vattikuti S, Guo J, Chow CC. Heritability and genetic correlations explained by common SNPs for metabolic syndrome traits. PLoS Genet 2012;8:e1002637.
3. Schwenk RW, Vogel H, Schurmann A. Genetic and epigenetic control of metabolic health. Mol Metab 2013;2:337-347.
4. Dick KJ, Nelson CP, Tsaprouni L, et al. DNA methylation and body-mass index: a genome-wide analysis. Lancet 2014;383:1990-1998.
5. Xu X, Su S, Barnes VA, et al. A genome-wide methylation study on obesity: differential variability and differential methylation. Epigenetics 2013;8:522-533.
6. Wang X, Zhu H, Snieder H, et al. Obesity related methylation changes in DNA of peripheral blood leukocytes. BMC Med 2010;8:87.
7. Franks PW, Ling C. Epigenetics and obesity: the devil is in the details. BMC Med 2010;8:88.
8. Milagro FI, Campion J, Cordero P, et al. A dual epigenomic approach for the search of obesity biomarkers: DNA methylation in relation to diet-induced weight loss. FASEB J 2011;25:1378-1389.
9. Gregor MF, Hotamisligil GS. Inflammatory mechanisms in obesity. Annu Rev Immunol 2011;29:415-445.
10. Corella D, Arnett DK, Tsai MY, et al. The 26ST−C polymorphism in the apolipoprotein a-II gene promoter is associated with body mass index and food intake in the genetics of lipid lowering drugs and diet network study. Clin Chem 2007;53:1144-1152.
11. Warodomwichit D, Shen J, Arnett DK, et al. ADIPQQ polymorphisms, monounsaturated fatty acids, and obesity risk: the GOLDN study. Obesity (Silver Spring) 2009;17:510-517.
12. Bouwens M, Afman LA, Muller M. Activation of peroxisome proliferator-activated receptor alpha in human peripheral blood mononuclear cells reveals an individual gene expression profile response. BMC Genomics 2008;9:262.
13. Absher DM, Li X, Waite LL, et al. Genome-wide DNA methylation analysis of systemic lupus erythematosus reveals persistent hypomethylation of interferon genes and compositional changes to CD4+ T-cell populations. PLoS Genet 2013;9:e1003678.
14. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics 2007;8:118-127.
15. Abbas AR, Woldelegel K, Seshasayee D, Modrusan Z, Clark HF. Deconvolution of blood microarray data identifies cellular activation patterns in systemic lupus erythematosus. PLoS One 2009;4:e6908.
16. Housman EA, Accamando WP, Koestler DC, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinformatics 2012;13:86.
17. Irvin MR, Zhu D, Joehanes R, et al. Epigenome-wide association study of fasting blood lipids in the genetics of lipid lowering drugs and diet network study. Circulation 2014;130:567-572.
18. Sinnwell JP1, Therneau TM, Schaid DJ. The kinship2 R package for pedigree data. Hum Hered 2014;78:91–93. doi: 10.1159/000363105.
19. Clinical Guidelines on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults–the Evidence Report. National Institutes of Health. Obes Res 1998;6 (Suppl 2):51s-209s.
20. Lean ME, Han TS, Morrison CE. Waist circumference as a measure for indicating need for weight management. BMJ 1995;311:158-161.
21. Su AI, Wiltshire T, Batalov S, et al. A gene atlas of the mouse and human protein-encoding transcripts. Proc Natl Acad Sci USA 2004;101:6062-6067.
22. Lemas DJ, Wiener HW, O’Brien DM, et al. Genetic polymorphisms in carnitine palmitoyltransferase 1A gene are associated with variation in body composition and fasting lipid traits in yup’ik eskimos. *J Lipid Res* 2012;53:175-184.

23. Robitaille J, Houde A, Lemieux S, Perusse L, Gaudet D, Vohl MC. Variants within the muscle and liver isoforms of the carnitine palmitoyltransferase I (cpt1) gene interact with fat intake to modulate indices of obesity in French-canadians. *J Mol Med (Berlin, Germany)* 2007;85:129-137.

24. Bell JT, Pai AA, Pickrell JK, et al. DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. *Genome Biol* 2011;12:R10.

25. Yang ZH, Miyahara H, Iwasaki Y, Takeo J, Katayama M. Dietary supplementation with long-chain monounsaturated fatty acids attenuates obesity-related metabolic dysfunction and increases expression of PPAR gamma in adipose tissue in type 2 diabetic KK-ay mice. *Nutr Metab (Lond)* 2013;10:16.

26. Zhang FF, Morabia A, Carroll J, et al. Dietary patterns are associated with levels of global genomic DNA methylation in a cancer-free population. *J Natr* 2011;141:1165-1171.

27. Barbosa MT, Soares SM, Novak CM, et al. The enzyme cd38 (a NAD glycohydrolase, EC 3.2.2.5) is necessary for the development of diet-induced obesity. *FASEB J* 2007;21:3629-3639.

28. Escande C, Nin V, Price NL, et al. Flavonoid apigenin is an inhibitor of the NAD cde38; implications for cellular NAD+ metabolism, protein acetylation, and treatment of metabolic syndrome. *Diabetes* 2013;62:1084-1093.

29. Overman A, Chuang CC, McIntosh M. Quercetin attenuates inflammation in human macrophages and adipocytes exposed to macrophage-conditioned media. *Int J Obes (Lond)* 2011;35:1165-1172.

30. Sun L, Goff LA, Trapnell C, et al. Long noncoding RNAs regulate adipogenesis. *Proc Natl Acad Sci USA* 2013;110:3387-3392.

31. Powell WT, Coulson RL, Crary FK, et al. A Prader-willi locus IncRNA cloud modulates diurnal genes and energy expenditure. *Hum Mol Genet* 2013;22:4318-4328.

32. Yi F, Yang F, Liu X, et al. RNA-seq identified a super-long intergenic transcript functioning in adipogenesis. *RNA Biol* 2013;10:991-1001.

33. Petersen AK, Zeilinger S, Kastenmuller G, et al. Epigenetics meets metabolomics: an epigenome-wide association study with blood serum metabolic traits. *Hum Mol Genet* 2014;23:534-545.

34. Azziz R, Potter HD, Bradley EL Jr, Boots LR. delta 5-Androstene-3 beta,17 beta-diol in healthy eumenorrheic women: relationship to body mass and hormonal profile. *Fertil Steril* 1994;62:321-326.

35. Possemato R, Marks KM, Shaul YD, et al. Functional genomics reveal that the serine synthesis pathway is essential in breast cancer. *Nature* 2011;476:346-350.

36. Ge ZZ, Lao SM, Lin F, et al. DNA methylation in oocytes and liver of female mice and their offspring: effects of high-fat-diet-induced obesity. *Environ Health Perspect* 2014;122:159-164.

37. Shenker NS, Polidoro S, van Veldhoven K, et al. Epigenome-wide association study in the european prospective investigation into cancer and nutrition (EPIC-turin) identifies novel genetic loci associated with smoking. *Hum Mol Genet* 2013;22:843-851.

38. Zeilinger S, Kuhnel B, Klopp N, et al. Tobacco smoking leads to extensive genome-wide changes in DNA methylation. *PLoS One* 2013;8:e63812.

39. Wu MC, Joubert BR, Kuan PF, et al. A systematic assessment of normalization approaches for the infinium 450K methylation platform. *Epigenetics* 2014;9:318-329.