Global DNA Methylation Is Associated With Insulin Resistance

A Monozygotic Twin Study

Jinying Zhao,¹ Jack Goldberg,²,³ James D. Bremner,⁴ and Viola Vaccarino⁵

Insulin resistance (IR), the hallmark of type 2 diabetes, may be under epigenetic control. This study examines the association between global DNA methylation and IR using 84 monozygotic twin pairs. IR was estimated using homeostasis model assessment (HOMA). Global DNA methylation of Alu repeats in peripheral blood leukocytes was quantified by bisulfite pyrosequencing. The association between global DNA methylation and IR was examined using generalized estimating equation (GEE) and within-twin pair analyses, adjusting for potential confounders. Results show that methylation levels at all four CpG sites were individually associated with IR by GEE (all false discovery rate–adjusted P values ≤ 0.026). A 10% increase in mean Alu methylation was associated with an increase of 4.54 units (95% CI 2.38–6.73) in HOMA. Intrapair difference in IR was significantly associated with intrapair difference in global methylation level. A 10% increase in the difference in mean Alu methylation was associated with an increase of 4.54 units (0.34–8.71; P = 0.036) in the difference in HOMA. Confirmation of the results by intrapair analyses suggests that genetic factors do not confound the association between global DNA methylation and IR. Exclusion of twins taking diabetes medication (n = 17) did not change our results. Diabetes 61:542–546, 2012

Insulin resistance is the hallmark of type 2 diabetes and is also a characteristic feature of various cardiometabolic conditions, such as obesity (1), hypertension (2), and cardiovascular disease (CVD) (3). IR results from the complex interplay between genetic and environmental factors. Despite significant effort, the understanding of the molecular processes underlying IR remains a major challenge. Epigenetic mechanisms act at the interface between internal genetic makeup and external environmental factors. Epigenetics provides an appealing explanation for many clinical features of IR (4), such as discordance within monozygotic (MZ) twins, relatively late age of onset, interindividual variation in disease severity, and the close relationship with lifestyle factors. Therefore, epigenetic modifications, especially DNA methylation, may play a critical role in the pathogenesis of IR and its complications (5).

The human genome contains a large portion of repetitive elements, more than one-third of which are methylated (6). Alu repeats, with >1 million copies per haploid genome, are the most plentiful families of repetitive sequences, representing >10% of the human genome sequence (7). Because of their high copy numbers throughout the genome, Alu repeats have been used as surrogate markers for estimating global DNA methylation levels (8). Alterations in global DNA methylation patterns have been associated with various human diseases, such as cancer (9), hypertension (10), CVD (11), and autoimmune disease (12). Since epigenetic mechanisms are in part under genetic control (13), disentangling the role of epigenetic factors in disease etiology could be confounded by genetic factors in unrelated individuals. MZ twins are matched on their DNA as well as many other known and unknown factors, thus providing an ideal sample for epigenetic research. The goal of this study is to examine the association between global DNA methylation and IR within MZ twin pairs.

RESEARCH DESIGN AND METHODS

Study population. Twins included in this study were randomly drawn among those who were born between 1946 and 1956 from the Vietnam Era Twin Registry, one of the largest twin registries in the U.S. (14). A total of 307 twin pairs (who were raised in the same household) were recruited by the Emory Twin Studies (15), which was designed to investigate the role of psychological, behavioral, and biological risk factors for subclinical CVD in twins. The Emory Twin Studies includes samples recruited by two companion studies: the Twins Heart Study (THS) and the Stress and Vascular Evaluation in Twins (SAVEIT). THS enrolled 180 male-female twin pairs (102 MZ pairs, 78 dizygotic pairs) between 2002 and 2006, with the inclusion of a random sample of twin pairs discordant for major depression. SAVEIT included 127 twin pairs (85 MZ pairs, 42 dizygotic pairs) enrolled between 2005 and 2010, with the inclusion of a random sample of twin pairs discordant for posttraumatic stress disorder (PTSD). Both studies followed identical procedures, measurements, and protocols. Both studies were approved by the Emory Institutional Review Board, and all twins signed an informed consent.

As part of the Emory Twin Studies, the current analysis included 84 MZ twin pairs (44 pairs from THS, 40 pairs from SAVEIT) in which both members had DNA samples and phenotype data. Among these, 42 pairs were concordant for depression and 42 pairs were discordant for depression. Seventeen twins (4 pairs and 9 singletons) reported to be receiving diabetes medication (e.g., oral hypoglycemic drug or insulin) at enrollment. All twins were examined in pairs at the Emory University General Clinical Research Center between 2002 and 2010, where their medical history was updated. All twins were white. Zygosity information was determined by DNA analysis.

Risk factor measurements. All measurements were performed in the morning after an overnight fast, and both members of a pair were tested at the same time. A medical history and a physical exam were obtained from all twins. Weight and height were used to calculate BMI as weight in kilograms divided by height in meters squared. Systolic blood pressure (SBP) and diastolic blood pressure were measured by mercury sphygmomanometer on the right arm with
the subject in sitting position after 10 min of rest. The average of two measurements 5 min apart was used in the statistical analyses. Cigarette smoking was classified into current smoker (any number of cigarettes) versus never or past smoker. A pack year was calculated by multiplying the number of packs of cigarettes a person has smoked per day by the number of years the person has smoked. Venous blood samples were drawn for the measurements of glucose and lipid profiles after an overnight fast. Plasma samples were separated and stored at −80°C until analysis. The Emory Lipid Research Laboratory, a participant in the Centers for Disease Control and Prevention/National Heart, Lung, and Blood Institute Lipid Standardization Program, performed all analyses from freshly isolated EDTA plasma. Total triglycerides and cholesterol were determined by enzymatic methods (Beckman Coulter Diagnostics, Fullerton, CA). Direct HDL and direct LDL cholesterol were obtained using homogeneous assays (Equal Diagnostics, Exton, PA). Fasting glucose levels were measured on the Beckman CX7 chemistry autoanalyzer (Beckman Coulter Diagnostics). All biochemical assays for each twin pair were processed in the same analytical run. An index of IR, the homeostasis model assessment (HOMA), was calculated as (fasting glucose [mmol/L]) × (fasting insulin [μU/mL])/22.5 (16). Diabetes mellitus was defined as having a fasting glucose level of >126 mg/dL or being treated with antidiabetes medications.

Global DNA methylation by quantitative bisulfite pyrosequencing. Global DNA methylation of Alu repetitive elements was quantify by a commercial service (EpigenDx, Inc., Worcester, MA). In brief, genomic DNA isolated from peripheral blood leukocytes was bisulfite treated using EZ DNA Methylation-Gold Kit (Zymo Research Corp., Irvine, CA) according to the manufacturer’s recommendations. DNA methylation of Alu repetitive sequences was quantified by PCR pyrosequencing using primers designed toward a consensus Alu sequence (human Alu National Center for Biotechnology Information: X55933.1) based on methods described by Yang et al. (8).

The methylation status of each CpG site was analyzed individually as an artificial TC single nucleotide polymorphism using Q CpG software (Pyrosequencing, Qiagen, Valencia, CA). The methylation level at each CpG site for each sample was calculated as the percentage of the methylated alleles over the sum of the methylated and unmethylated alleles. The mean methylation level was calculated using methylation levels of all measured CpG sites within the PCR region of a gene. The methylation level at each CpG site ranges from 0 (unmethylated) to 1 (fully methylated). EpigenDx performed the assay validation, bisulfite conversion, and pyrosequencing. We included high, medium, and low methylated DNA controls as in each run.

The Alu assay was validated by PCR bias testing using pyrosequencing. Low methylated control DNA and in vitro methylated DNA were mixed at different ratios (0, 20, 40, up to 100%) followed by bisulfite modification, PCR, and pyrosequencing analysis. The percent methylation obtained from the mixing study was highly correlated with an r² value of 0.99. The assay was also validated by sensitivity and reproducibility testing. Different amounts of genomic DNA were used for the analysis in triplicates. Reproducible results were obtained by using as low as 5 ng of genomic DNA in the bisulfite modification reaction.

There is a high degree of variability within the 1.4 million Alu sites, and mutations can often occur as a result of spontaneous deamination of 5-methylcytosine to thymidine. Of the number of potential CpG sites in the Alu assay, only ~30% are actual CpG sites that can be methylated. In this study, the pyrosequencing assay covers four potential CpG sites within Alu sequences. Because PCR primers were designed to recognize the conserved sequence within Alu repeat elements and the PCR products generated represent the combination of all Alu repeats in the genome, the four CpG sites examined in this study should reflect the global methylation level of all Alu repeats in the human genome.

Statistical analyses. Statistical analyses were performed to examine whether global DNA methylation was associated with IR, accounting for known risk factors. Prior to analysis, IR and methylation data were logarthimically transformed. Multiple testing was performed by controlled-adjusted discovery rate (FDR).

Generalized estimating equation analyses. We first performed statistical analyses using generalized estimating equation (GEE) modeling to investigate the association between DNA methylation and IR. GEE was used to account for the within-twin pair correlations. In the GEE model, IR was the dependent variable and methylation level at each CpG site or the mean methylation level was the independent variable, adjusting for age, current smoking (yes/no), BMI, HDL, LDL, and sBP. To determine whether combining data from the two studies would affect the study results, we also included study affiliation (THS or SAVEIT) as a covariate in all statistical analyses. As described above, our sample included twins recruited by two projects, with oversampling of twins with either lifetime history of major depression or PTSD. To examine whether the oversampling scheme would influence our results, we performed separate analyses by further adjusting for major depression or PTSD.

Pairwise analyses. Because twins in a pair are not independent, we conducted separate analyses by considering twins as members of a twin pair. First, we calculated intrapair difference in mean methylation level within a twin pair, defined as the absolute difference in mean methylation level between two members of each twin pair. The intrapair difference in IR was similarly calculated. Then we performed linear regression by regressing the intrapair difference (independent variable) on the intrapair differences in mean Alu methylation level (independent variable), adjusting for intrapair differences in pack years, BMI, LDL, HDL, sBP, and depressive symptoms (assessed by Beck Depression Inventory–II scores).

Sensitivity analyses. To examine whether treatment of type 2 diabetes (n = 17, including 4 pairs and 9 singletons) had an influence on our results, we conducted sensitivity analysis by removing twins who were taking diabetes medication (e.g., oral hypoglycemic drug or insulin). All 17 twins receiving diabetes treatment were removed from GEE analysis. The cotwins of the 9 singletons were also removed from pairwise analysis. The final sample size included in sensitivity analyses was 151 for GEE and 142 (71 pairs) for intrapair analysis.

RESULTS

Table 1 presents the demographic characteristics of the twins included in this analysis. The age of the twins ranged from 48 to 61 years with a mean age of 55 years. Methylation levels at the four studied CpG sites were highly correlated (r² ranged from 0.62 to 0.69; all P values <0.0001). The mean methylation level of the four CpG sites was 24.7%, with the highest and lowest methylation levels being 32.1 and 15.0%, respectively.

Multivariate GEE analyses demonstrated that DNA methylation levels at all four CpG sites showed significant individual association with IR after adjustments for age, smoking, BMI, LDL, HDL, sBP, study affiliation (THS or SAVEIT), and depression. The mean methylation level of Alu elements was also significantly associated with IR (P < 0.0001) after controlling for the potential confounders listed above. On average, a 10% increase in the mean Alu methylation level was associated with a 4.55-unit increase (95% CI 2.38–6.73) in HOMA. All associations between IR and global DNA methylation remained statistically significant after correction for multiple testing by FDR (all FDR-adjusted P values ≤0.026). With the exception of BMI, none of the covariates (age, smoking, LDL, HDL, sBP, depression, study affiliation, and receiving diabetes medication) was significant for any of the studied CpG sites in GEE analysis. Table 2 shows the results for the association between global DNA methylation and IR.

DNA methylation levels of the two members within a pair were highly correlated at each of the four CpG sites (all P values <0.001). The mean global methylation levels

TABLE 1
Demographic, clinical, and laboratory characteristics of the twins

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Mean ± SD or %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>55.1 ± 2.8</td>
</tr>
<tr>
<td>Type 2 diabetes (%)</td>
<td>11.4</td>
</tr>
<tr>
<td>Hypertension (%)</td>
<td>36.5</td>
</tr>
<tr>
<td>Current smoking (%)</td>
<td>59.6</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>29.4 ± 4.8</td>
</tr>
<tr>
<td>sBP (mmHg)</td>
<td>129.3 ± 17.5</td>
</tr>
<tr>
<td>dBP (mmHg)</td>
<td>81.4 ± 11.6</td>
</tr>
<tr>
<td>HDL (mg/dL)</td>
<td>37.6 ± 10.7</td>
</tr>
<tr>
<td>LDL (mg/dL)</td>
<td>123.3 ± 36.8</td>
</tr>
<tr>
<td>Total triglyceride (mg/dL)</td>
<td>177.5 ± 102.5</td>
</tr>
<tr>
<td>Fasting glucose (mg/dL)</td>
<td>102.4 ± 19.5</td>
</tr>
<tr>
<td>Fasting insulin (μU/mL)</td>
<td>8.3 ± 5.8</td>
</tr>
<tr>
<td>HOMA</td>
<td>2.2 ± 1.87</td>
</tr>
<tr>
<td>Mean Alu methylation level (%)</td>
<td>24.7 ± 0.88</td>
</tr>
</tbody>
</table>

Data are mean ± SD or %. dBP, diastolic blood pressure.
of the two twins within a pair were also significantly correlated ($r = 0.46$, $P < 0.0001$). Regression analysis showed that intrapair difference in IR was significantly associated with intrapair difference in global DNA methylation level at two of the four examined CpG residues, after adjusting for differences in pack years, BMI, lipids, sBP, and depressive symptoms. On average, a 10% increase in the difference in mean Alu methylation level was associated with an increase of 4.54 units in the difference in HOMA (95% CI 0.34–8.71; $P = 0.036$). For both GEE and pairwise analyses, exclusion of twins receiving diabetes treatment did not change our results (Table 2).

### DISCUSSION

Using a well-matched MZ twin sample, we demonstrated that global DNA methylation in peripheral blood leukocytes was significantly associated with IR, independent of established risk factors. On average, a 10% increase in global DNA methylation was associated with a 4.55-unit increase in HOMA after controlling for multiple covariates.

Epigenetic factors have been suggested to play a critical role in the etiologies of metabolic disorders (5) and diabetes complications (17). IR, the primary pathophysiologic mechanism underlying diabetes and CVD, may also be under epigenetic control. Several genes that are critical to glucose and lipid metabolism, such as peroxisome proliferator-activated receptor $\gamma$ coactivator 1-a ($PGC-1\alpha$) (18–20) and insulin gene (21), have been shown to be epigenetically misregulated. Mitochondrial dysfunction plays a central role in the pathogenesis of IR (22). Promoter methylation of the mitochondrial transcription factor A ($TFAM$) gene, an important gene essential for mitochondrial DNA maintenance, was also associated with IR in adolescents (23). In addition, epigenetic aberrations have been implicated in obesity (24), hypertension (25), and inflammation (26), all of which are known to be related to IR (27,28). The association between global DNA methylation and IR detected in our study is in line with previous findings, lending further support for a potential role of aberrant DNA methylation in IR.

Although the molecular mechanism linking global DNA methylation and IR is unknown, a number of explanations are possible. Alu elements are known to create genomic instability and affect gene expression (7) and have been implicated in numerous human diseases (29). Recent studies have demonstrated interindividual variability in DNA methylation profiles at specific Alu elements (30). Epigenetic alterations arising during the lifetime of MZ twins in Alu repeats were also reported (31). It is possible that alterations in DNA methylation of Alu repetitive sequences would promote genomic instability (32–34) and, consequently, cause IR and type 2 diabetes. This hypothesis is supported by both animal and human studies (35,36). Alternatively, it is possible that epigenetic changes in Alu elements could lead to inflammation (26), a mechanism known to be involved in IR and its complications (28). Another possible biological mechanism underlying the association between global DNA methylation and IR could be attributed to telomere shortening, which has been shown to be epigenetically regulated (37) and associated with both genomic instability (38) and IR (39).

Our study has some limitations. First, because of practical difficulties in obtaining tissues from living individuals, methylation levels were tested in peripheral blood leukocytes but not directly from the primary affected organs in IR and diabetes (e.g., pancreata, skeletal muscles, or adipocytes). Therefore, our results may not provide a direct index of DNA methylation in the system of IR or insulin secretion. However, there is increasing evidence that leukocytes may be a useful cell model to evaluate epigenetic changes because epimutations may not be limited to the affected tissue but can also be detected in peripheral blood leukocytes (40). It is important that blood samples are much easier to obtain and could be used for large-scale epidemiologic studies, tracking changes associated with IR and diabetes as the disease develops. Second, our epigenetic data were collected from DNA derived from whole blood leukocytes, which is a mixture of many cell types; as such, we were unable to assess blood cell–specific differences in methylation changes, which would have to be accounted for when replicating the present findings in future studies. Third, global DNA methylation measurement provides an estimate of average methylation level across the entire genome and, as such, does not have a locus-specific resolution necessary to pinpoint the individual gene responsible for diabetes risk. Genome-wide array- or sequence-based interrogation techniques would be needed to provide such resolution. Fourth, our sample included twins with oversampling of either major depression or PTSD, both of which may influence diabetes risk (41,42). However, the observed association between aberrant DNA methylation and IR is unlikely to be confounded by these two psychiatric conditions because further adjustments did not change the results. Moreover, depression or PTSD is associated with neither DNA methylation nor IR in our study sample. Therefore, the observed impact of global DNA methylation on IR should be through biochemical pathways that are independent of these two psychiatric disorders. Fifth, our twin sample was derived from a middle-aged sample of male military veterans; therefore, the generalizability to females and other younger or older

<table>
<thead>
<tr>
<th>CpG site</th>
<th>Methylation level (%)</th>
<th>Increase in HOMA (95% CI)*</th>
<th>$P^{2}$</th>
<th>Increase in HOMA (95% CI)**</th>
<th>$P^{**}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>23.44 ± 0.91</td>
<td>3.46 (1.39–5.52)</td>
<td>0.0003</td>
<td>3.05 (1.12–4.92)</td>
<td>0.004</td>
</tr>
<tr>
<td>2</td>
<td>15.62 ± 0.82</td>
<td>3.85 (1.93–5.76)</td>
<td>&lt;0.0001</td>
<td>3.29 (1.32–5.26)</td>
<td>0.005</td>
</tr>
<tr>
<td>3</td>
<td>27.58 ± 1.27</td>
<td>3.18 (1.63–4.73)</td>
<td>&lt;0.0001</td>
<td>1.91 (0.46–3.36)</td>
<td>0.017</td>
</tr>
<tr>
<td>4</td>
<td>32.15 ± 1.10</td>
<td>1.68 (0.20–3.15)</td>
<td>0.026</td>
<td>1.88 (0.49–3.26)</td>
<td>0.011</td>
</tr>
<tr>
<td>Mean</td>
<td>27.70 ± 0.89</td>
<td>4.55 (2.38–6.73)</td>
<td>&lt;0.0001</td>
<td>3.25 (1.33–5.16)</td>
<td>0.002</td>
</tr>
</tbody>
</table>

Data are mean ± SD unless otherwise indicated. *Adjusted for age, BMI, smoking, LDL, HDL, sBP, study affiliation, and depression. **Exclusion of twins taking diabetes medication (adjusted for same covariates). $\dagger$All $P$ values remained statistically significant after correction for multiple testing by FDR (FDR-adjusted $P \leq 0.026$).
populations is not known. Finally, because our study was cross-sectional, we are unable to determine whether the observed epigenomic signatures are epiphenomena or part of the causal pathways leading to IR. If causally related, approaches to mitigate the development of these epigenomic signatures would be expected.

Nonetheless, this is the first study to demonstrate the association of global DNA methylation in Alu elements with IR in a well-matched MZ twin sample. Previous studies show that interindividual epigenetic processes are under genetic control (43,44). Therefore, the study of the causal impact of epigenotype on disease outcome requires careful matching on or control for genetic background and other potential confounding variables. MZ twin pairs provide a natural experiment to investigate the effects of epigenetics on disease outcome because MZ twins match exactly on genetic background, age, and sex, thus eliminating the effects of these important confounders (4,31). In addition, identical twins in general share raising environment, providing further control for confounding by early life experience, which has a long-lasting impact on the epigenetic plasticity of the human genome (45).

In summary, this study provides the first evidence that alteration in global DNA hypermethylation is associated with increased risk of IR independent of established risk factors. Because epigenetic modifications are possibly reversible, and IR usually precedes the onset of diabetes, our results, if confirmed, would suggest the potential for lifestyle or therapeutic interventions for IR and a wide range of metabolic disorders.

ACKNOWLEDGMENTS
This study was supported by American Heart Association Grant 0730100N and National Institutes of Health grants R21-HL-092363, K01-AG-034259, K24-HL-077506, R01-HL-68630, and R01-AG-026255. The U.S. Department of Veterans Affairs has provided financial support for the development and maintenance of the Vietnam Era Twin (VET) Registry.

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

J.Z. designed the study, analyzed data, and wrote the manuscript. J.G. and J.D.B. contributed to discussion and reviewed and edited the manuscript. V.V. designed the study, contributed to discussion, and reviewed and edited the manuscript. J.Z. is the guarantor of this work, and as such, had full access to all data that in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Numerous organizations have provided invaluable assistance in the conduct of this study, including the Department of Defense, the National Personnel Records Center, the National Archives and Records Administration, the Internal Revenue Service, the National Institutes of Health, the National Opinion Research Center, the National Research Council, the National Academy of Sciences, and the Institute for Survey Research, Temple University. Most important, the authors gratefully acknowledge the continued cooperation and participation of the members of the VET Registry and their families. Without their contribution, this research would not have been possible.

REFERENCES