

Early Pregnancy Peripheral Blood DNA Methylation and Risk of Gestational Diabetes Mellitus: A Nested Case-Control Study

Xiaolei Wang

Central South University

Jin Huang

Central South University

Sisi Long

Central South University

Huijun Lin

Central South University

Na Zhang

Central South University

Mengyuan Tian

Central South University

Xinrui Wu

Central South University

Rongjing An

Central South University

Shujuan Ma

Central South University

Hongzhuan Tan (✉ tanhz99@qq.com)

Central South University

Research Article

Keywords: DNA methylation, gestational diabetes mellitus, maternal peripheral blood, risk factors, case-control studies, receiver operator characteristic curve analysis.

Posted Date: May 13th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-511407/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Introduction:

Genome-wide DNA methylation profiling has been used to identify CpG sites relevant to gestational diabetes mellitus (GDM). However, these sites have not been verified in larger samples. Here, our aim was to evaluate the changes in target CpG sites in the peripheral blood of pregnant women with GDM in their first trimester.

Research Design and Methods:

This nested case-control study examined a large cohort of women with GDM in early pregnancy (10–15 weeks; $n = 80$). Target CpG sites were extracted from related published literature and bioinformatics analysis. The DNA methylation levels at 337 CpG sites located in 27 target genes were determined using MethylTarget™ sequencing. The best cut-off levels for methylation of CpG sites were determined using the generated ROC curve. The independent effect of CpG site methylation status on GDM was analyzed using conditional logistic regression.

Results

Methylation levels at 6 CpG sites were significantly higher in the GDM group than in controls, whereas those at 7 CpG sites were significantly lower ($P < 0.05$). The area under the ROC curve at each methylation level of the significant CpG sites ranged between 0.593 and 0.650 for GDM prediction. After adjusting for possible confounders, the hypermethylation status of candidate sites cg68167324 (OR = 3.168, 1.038–9.666) and cg24837915 (OR = 5.232, 1.659–16.506) was identified as more strongly associated with GDM; conversely, the hypermethylation of sites cg157130156 (OR = 0.361, 0.135–0.966) and cg89438648 (OR = 0.206, 0.065–0.655) might indicate lower risk of GDM.

Conclusions

The methylation status of target CpG sites in the peripheral blood of pregnant women during the first trimester is associated with GDM pathogenesis, and has potential as a predictor of GDM.

Introduction

Gestational diabetes mellitus (GDM) is defined as any degree of glucose intolerance with onset or first recognition during pregnancy.¹ Alongside the increasing prevalence of obesity and type 2 diabetes, the incidence of GDM has risen annually.² Specifically, in China, this incidence has been estimated at 14.8% (12.8–16.7%).³ GDM in pregnant women can result in severe adverse pregnancy outcomes

complications, including macrosomia, premature birth, and fetal malformations.⁴ GDM also has serious consequences for the offspring, such as childhood obesity, insulin resistance,⁵ and impaired neurodevelopment.⁶ In addition, GDM can recur, may lead to increased risk of postpartum type 2 diabetes.⁷

The mechanisms underlying GDM include genetic background, inflammatory factors, and oxidative stress. Wu et al. have shown that GDM has a genetic component, and the differences in GDM among races may be due to the interaction between genes and the environment.⁸ Epigenetics bridges the gap between genes and the environment at the molecular level.^{9 10} In particular, DNA methylation is the most common and the earliest form of epigenetic modification. This process is catalyzed by DNA methyltransferases and uses S-adenosylmethionine as the methyl donor to convert the CpG site dinucleotide 5'-cytosine to the 5'-methylcytosine. The CpG site density and methylation degree of the upstream promoter region of the gene directly affect gene activity and expression.¹⁰ DNA methylation therefore plays a key role in regulating genome transcription. An increasing number of studies have shown that DNA methylation levels might affect the occurrence and development of GDM; most such studies investigated the relationship between GDM and DNA methylation levels using cord blood and placental tissues of pregnant women. However, these samples are not representative of the DNA methylation levels during early pregnancy because they can only be collected in the third trimester.

Chim et al. showed that changes in DNA methylation occur during pregnancy using peripheral blood samples from pregnant women.¹¹ Kang et al. conducted genome-wide DNA methylation chip analysis of the peripheral blood samples of eight patients with GDM and eight healthy women in late pregnancy and revealed varying methylation levels in 200 CpG sites across 151 genes.^{12 13} Moreover, Wu et al. evaluated DNA methylation levels in the peripheral blood of pregnant women in the first trimester using genome-wide DNA methylation analysis. They observed varying DNA methylation levels at 100 CpG sites corresponding to 66 genes in the GDM group (n = 11), further suggesting that the DNA methylation status of 5 CpG sites, in the *COPS8*, *PIK3R5*, *HAAO*, *CCDC124*, and *C5orf34* genes could be used as clinical biomarkers of GDM.¹⁴ Similarly, also using genome-wide DNA methylation analysis, Enquobahrie et al. found 17 hypomethylated and 10 hypermethylated CpG sites in the GDM group.¹⁵ However, because these studies employed whole-genome DNA methylation analysis, their sample sizes were necessarily small. The significance of CpG site methylation has not yet been verified in a large study cohort.

Here, we evaluated the DNA methylation levels in the peripheral blood of women in early pregnancy using specific target gene DNA methylation detection in order to verify the relationship between the methylation status of the targeted CpG sites and GDM.

Material And Methods

Study design and population

This was a nested case-control study based on an early pregnancy follow-up cohort. The cohort was established in Hunan Province Maternal and Child Health Hospital (ChiCTR1900020652) between March 2017 and December 2018, and a total of 890 pregnant women were enrolled. All the eligible participants agreed to participate in this study and provided written informed consent. The study protocol was approved by the Medical Ethical Committee of the Hunan Provincial Maternal and Child Health Hospital in South China (approval number: EC201624 on January 11, 2017) and all the methods were carried out in compliance with relevant guidelines. Pregnant women were recruited in their first trimester (10–14 weeks) and followed up for 42 days post-partum. The inclusion criteria were: (i) singleton pregnancy and natural conception; (ii) diabetes-free at recruitment; (iii) had not received any antibiotic treatment throughout the current pregnancy; (iv) no acute infection in the 2 weeks before sample collection; (v) planned to attend for all obstetric examinations and delivery at the above hospital. We collected questionnaire data and venous blood samples; additional patient information and data concerning their clinical examinations were collected through the hospital's electronic recodes system. The venous blood samples (5 ml) were collected during early pregnancy by certified nurses in the morning following a 10-h overnight fast. Serum and blood cells were separated by centrifugation at 3500 rpm for 15 min and stored at -80°C until further use.

Diagnostic criteria for GDM and selection of controls

All subjects underwent a 2-h standard 75g oral glucose tolerance (OGTT)¹⁶ in the hospital outpatient department at 24–28 weeks of gestation. The oxidase method was used to estimate blood glucose levels, with measurement completed using an automatic biochemical analyzer (Hitachi 7600) at the hospital. GDM was defined according to the International Association of Diabetes and Pregnancy Study Groups (IADPSG) standard. That is, GDM was considered to be present when at least one of the following blood glucose concentrations was obtained: ≥ 5.1 mmol/L (fasting), ≥ 10.0 mmol/L (after 1 h), and ≥ 8.5 mmol/L (after 2 h).¹⁶ The controls were selected from among women in the same cohort who had normal blood glucose levels throughout the pregnancy. A 1:1 pair match for each GDM patient was identified, based on age (± 3 years) and gestational age (± 1 week), resulting in a final study population of 80 eligible GDM patients and 80 healthy controls.

Selection of candidate CpG sites

Suitable target CpG sites were identified from the literature and from bioinformatics analysis aimed at finding CpG sites related to GDM. Published studies have used whole-genome methylation sequencing methods to identify the associations of CpG sites. In the literature, we selected the 21 target CpG sites that were most likely to be related to the occurrence of GDM.^{12–15 17} We also found two target CpG sites related to GDM through bioinformatics methods. Moreover, the literature^{14 18 19} provided four additional genes closely related to the occurrence of GDM, namely *RDH12*, *GCK*, *PPARG*, and *IL6*. Since these reports did not indicate any selectable CpG sites, we tested all CpG islands in the promoter regions of these four genes, generating a total of 6 CpG islands, with one or two per promoter region. Overall, this procedure therefore identified a total of 29 CpG islands, containing 337 CpG sites. We obtained the sequences for

each of the 29 CpG island target fragments. The primers used and their sources are shown in Supplementary Tables S1, S2, and S3.

DNA extraction

Genomic DNA was extracted from frozen samples using Genomic Tip-500 columns (Qiagen, Valencia, CA, USA) and from bisulfite-converted samples using the EZ DNA Methylation™-GOLD Kit (Zymo Research, CA, USA) in accordance with the manufacturer's instructions. Genomic DNA integrity was measured using agarose gel electrophoresis and quality control was ensured using a NanoDrop 2000 (NanoDrop technologies, Wilmington, DE, USA), which requires that the DNA concentration ≥ 20 ng/ μ L, and that the total amount of DNA ≥ 1 μ g.

DNA methylation analysis

The DNA methylation level of the target CpG site is defined as the number of methylated reads at that site (i.e., the number of reads with base C detected) divided by the total number of reads at that site, and was obtained by MethylTarget™ sequencing (Genesky Biotechnologies Inc. Shanghai, China), a method based on next-generation sequencing-based multiple targeted CpG methylation analysis. Primer design and validation were performed using bisulfate-converted DNA samples on the Methylation Primer software. The primer sets were designed to flank each target CpG site by 100–300 nucleotides and are summarized in Supplementary Table S1. After PCR amplification (HotStarTaq polymerase kit, TAKARA, Tokyo, Japan) and library construction, paired-end sequencing was performed (Illumina HiSeq Benchtop Sequencer, CA, USA) in accordance with the manufacturer's protocol. To ensure the accuracy of sequencing results, we added a quality control step to the DNA methylation sequencing method; for details, see Supplementary Text S1.

Covariates

In this study, we collected information on maternal demographics, lifestyle, and pregnancy history through structured questionnaires during each follow-up. This included factors with the potential to confound the exposure-outcome relationship, including pre-pregnancy body mass index (BMI) (continuous), history of drinking (yes/no), history of smoking (yes/no), parity (continuous), pregnancy times (continuous), polycystic ovary syndrome (PCOS) (yes/no), and waist circumference at enrolment (continuous). Many study participants were primiparas, so history of GDM was not considered in the analysis.

Statistical analysis

Continuous data and categorical data were represented by the mean \pm standard deviation (SD) and frequency (percentage), respectively. Paired-samples T test were used to compare normally distributed continuous data, whereas Wilcoxon signed rank test were used to analyze non-normally distributed continuous data. Dichotomous variables were analyzed using McNemar χ^2 test. $P < 0.05$ was considered statistically significant, and all statistical tests were two-sided. ROC curve analysis was used to assess the possible predictive value of the methylation level of individual CpG site for the occurrence of GDM.

CpG sites with positive correlation between methylation and GDM were considered risk sites, and those with inverse correlation were considered protective sites. We used the control group to provide the state variable value in the ROC curve analysis for the protective CpG sites. Through the ROC curve, the methylation status (high or low) of the target CpG site was classified based on the best cut-off value, defined as the DNA methylation level with the highest Youden index. Conditional logistic regression analysis was used to determine the independent influence of target CpG site methylation status on GDM. Covariates with significant differences were included in the model for correction. The model variable selection criterion was $\alpha_{in} = 0.05$; the variable elimination criterion was $\alpha_{out} = 0.10$; the Wald forward method was used to establish a conditional logistic regression model to screen CpG sites with independent effects. All the statistical analyses were performed using SPSS software v25.0 (IBM Corporation, Armonk, NY, USA).

Results

Patient characteristics

The participant characteristics are summarized in Table 1. The age of the GDM patients ranged between 23 and 43 years (mean: 31.6 years), whereas that of the healthy controls ranged between 24 and 45 years (mean: 32.0 years). No significant difference ($P > 0.05$) was observed between the two groups in terms of gravidity, parity, PCOS, smoking history, alcohol intake history, age, or gestational age. The GDM group had higher fasting glucose, 1-h post-OGTT glucose, 2-h post-OGTT glucose, pre-pregnant BMI, and waist circumference than the control group.

Table 1
Study participant's characteristics

Variable	Control (N = 80)		GDM (N = 80)		P-value
Basic characteristics					
Gravidity, M (Q1-Q3)	2 (1–3)		2 (1–3)		0.750
Parity, M (Q1-Q3)	1 (0–1)		1 (0–1)		0.950
PCOS, n (%)	2 (2.5)		7 (8.8)		0.086
Smoking history, n (%)	2 (2.5)		1 (1.3)		0.560
Drink history, n (%)	3 (3.8)		3 (3.8)		1
Demographic characteristics, Mean ± SD, Range					
Maternal age	31.6 ± 4.3	23.0 ~ 43.0	32.0 ± 4.5	24.0 ~ 45.0	0.617
Gestational age (weeks)	12.7 ± 0.7	10.0 ~ 14.4	12.6 ± 0.8	9.9 ~ 14.3	0.634
Pre-pregnant BMI (kg/m ²)	20.4 ± 2.6	16.2 ~ 30.9	22.4 ± 3.2	16.7 ~ 31.9	< 0.001
Waist circumference (cm)	77.3 ± 7.5	63.0 ~ 108.0	82.0 ± 9.2	58.0 ~ 107.5	< 0.001
75g OGTT in second trimester, Mean ± SD, Range					
Fasting glucose (mmol/L)	4.4 ± 0.3	3.8 ~ 5.0	4.9 ± 0.5	3.9 ~ 5.9	< 0.001
1-h post-OGTT (mmol/L)	6.4 ± 0.9	3.5 ~ 7.7	9.3 ± 1.6	5.0 ~ 13.0	< 0.001
2-h post-OGTT (mmol/L)	5.7 ± 0.6	3.8 ~ 6.6	8.2 ± 1.4	4.0 ~ 11.3	< 0.001
Abbreviation: M, median; (Q1-Q3), interquartile ranges; PCOS: polycystic ovary syndrome.					

Varying DNA methylation levels at target CpG sites

We assessed the DNA methylation levels at 337 CpG sites within 27 genes, and the results of the paired-samples T test or Wilcoxon signed rank test analysis for all the CpG sites are summarized in Supplementary Table S4. There were 13 CpG sites with results differing significantly between the group (Table 2; scatter diagram in Supplementary Figure S1). Briefly explanations the functions of the genes where these sites are located can be found in Supplementary Table S5. Missing data were replaced by the average (Supplementary Table S4). The methylation levels at 6 CpG sites within the *ARHGAP40*, *STAT1*, *C5orf34*, *RDH12*, and *YAP1* genes were higher in the GDM group than in the control group, whereas those at 7 CpG sites within the *HAPLN3*, *IFNGR2*, *NAGA*, *YAP1*, *NFATC4*, and *DNAJB6* genes were lower in the GDM group than in the control group.

Table 2
Varying DNA methylation Target CpG Site and Related Genes

CpG site	Gene Symbol	Chr	Distance to TSS	Methylation level (Mean \pm SD)		Fold Change ^a	P-value
				GDM (80)	Control (80)		
cg37274257	ARHGAP40	20	(+) 43681	0.601 \pm 0.046	0.584 \pm 0.048	1.03	0.015
cg43487508	C5orf34	5	(+) 27765	0.456 \pm 0.078	0.414 \pm 0.105	1.10	0.012
cg157130156	DNAJB6	7	(+) 446	0.010 \pm 0.006	0.012 \pm 0.005	0.83	0.013
cg157130085	DNAJB6	7	(+) 375	0.006 \pm 0.003	0.007 \pm 0.004	0.86	0.018
cg101980999	YAP1	11	(-) 192	0.009 \pm 0.003	0.007 \pm 0.003	1.29	0.001
cg101981060	YAP1	11	(-) 131	0.011 \pm 0.004	0.012 \pm 0.004	0.92	0.028
cg34775358	IFNGR2	21	(+) 157	0.024 \pm 0.038	0.036 \pm 0.039	0.67	0.012
cg42466321	NAGA	22	(+) 525	0.049 \pm 0.014	0.055 \pm 0.016	0.89	0.008
cg68167324	RDH12	14	(-) 1278	0.273 \pm 0.009	0.270 \pm 0.009	1.01	0.019
cg68167386	RDH12	14	(-) 1216	0.314 \pm 0.009	0.311 \pm 0.013	1.01	0.042
cg89438648	HAPLN3	15	(+) 209	0.011 \pm 0.004	0.013 \pm 0.006	0.85	0.048
cg24837915	NFATC4	14	(-) 288	0.015 \pm 0.003	0.014 \pm 0.003	1.07	0.018
cg191879104	STAT1	2	(-) 128	0.008 \pm 0.003	0.009 \pm 0.005	0.89	0.037

Abbreviation: Chr, chromosome; TSS, transcription start site and strand; Fold Change ^a, the ratio of the mean methylation level in the GDM group to that in the Control group.

Predictive value of the methylation status of the CpG site

For the 13 significantly different DNA methylation CpG sites, we further estimated the possible predictive value of the methylation level of individual CpG site for the occurrence of GDM using the ROC curve. The

ROC curve parameters and the cut-off value are summarized in Table 3. The largest area under the curve (AUC) reached 0.650.

Table 3
The AUC and Cut-off value of CpG site DNA methylation

CpG site	Gene Symbol	Cut-off value	Sensitivity	Specificity	AUC	95%CI	P-value
Risk CpG sites of GDM							
cg37274257	ARHGAP40	0.600	0.550	0.638	0.594	0.506–0.682	0.040*
cg43487508	C5orf34	0.420	0.775	0.475	0.602	0.514–0.691	0.026*
cg24837915	NFATC4	0.014	0.713	0.525	0.628	0.541–0.715	0.005*
cg68167386	RDH12	0.315	0.475	0.750	0.601	0.513–0.689	0.028*
cg68167324	RDH 12	0.278	0.550	0.563	0.589	0.501–0.677	0.051
cg101980999	YAP1	0.009	0.525	0.713	0.650	0.566–0.735	0.001*
Protective CpG sites of GDM							
cg42466321	NAGA	0.054	0.588	0.608	0.615	0.529–0.703	0.012*
cg89438648	HAPLN3	0.013	0.688	0.532	0.597	0.509–0.688	0.034*
cg157130156	DNAJB6	0.013	0.747	0.463	0.608	0.521–0.696	0.019*
cg191879104	STAT1	0.008	0.663	0.525	0.583	0.495–0.672	0.069
cg157130085	DNAJB6	0.004	0.506	0.688	0.593	0.505–0.681	0.043*
cg101981060	YAP1	0.014	0.550	0.575	0.582	0.494–0.670	0.074
cg34775358	IFNGR2	0.028	0.671	0.566	0.614	0.533–0.708	0.014*
*:P < 0.05.							

Comparison of the DNA methylation status of target CpG sites

To clearly show the effect of DNA methylation at the target CpG sites on GDM occurrence, we classified the DNA methylation levels into hypermethylation and hypomethylation statuses based on the best cut-off value (Table 3). Table 4 presents the differences in the DNA methylation statuses of the CpG sites between the GDM and control groups. Significant differences were observed in 8 CpG sites based on the McNemar χ^2 test ($P < 0.05$).

Table 4
The difference of DNA methylation status of CpG sites between GDM group and Control group

CpG site	Gene Symbol	GDM		Control		χ^2	P-value
		hypo-	hyper-	hypo-	hyper-		
cg37274257	ARHGAP40	38	42	51	29	1.21	0.268
cg89438648	HAPLN3	61	19	42	38	5.34	0.027*
Cg24837915	NFATC4	23	57	40	40	7.57	0.043*
cg43487508	C5orf34	16	64	36	44	4.59	0.000*
cg34775358	IFNGR2	55	25	33	47	0.63	0.488
cg42466321	NAGA	54	26	39	41	1.78	0.218
cg68167324	RDH12	55	25	69	11	29.33	0.000*
cg68167386	RDH12	42	38	59	21	7.00	0.011*
cg101980999	YAP1	38	42	57	23	3.69	0.072
cg101981060	YAP1	65	15	56	24	18.89	0.000*
cg157130156	DNAJB6	46	34	62	18	12.25	0.006*
cg191879104	STAT1	42	38	27	53	1.27	0.305
cg157130085	DNAJB6	25	55	16	64	17.09	0.000*

*:P < 0.05; hypo-, hypomethylation; hyper-, hypermethylation.

Conditional logistic regression analysis for DNA methylation status and GDM

Conditional logistic regression analysis was used to analyze the independent effect of the methylation status of the individual site on GDM occurrence. The independent variables included the methylation status of the eight significantly different CpG sites listed in Table 4 (0= "hypomethylation"; 1= "hypermethylation"). The confounding variable included waist circumference and pre-pregnancy BMI.

We found that the methylation status of four CPG sites influenced GDM occurrence. Specifically, the hypermethylation of cg68167324 (OR = 3.168; 95%CI: 1.038–9.666), and cg24837915 (OR = 5.232;

95%CI: 1.659-116.506) may indicate increased risk of GDM occurrence. In contrast, the hypermethylation of cg157130156 (OR = 0.361; 95%CI: 0.135–0.966) and cg89438648 (OR = 0.206; 95%CI: 0.065–0.655) may indicate decreased risk of GDM occurrence (Table 5).

Table 5
Conditional logistic analysis of methylation status of CPG sites on GDM

CpG site	Gene Symbol	β	Wals 2	P-value	OR	95%CI
cg89438648	HAPLN3	-1.581	7.164	0.007	0.206	0.065 ~ 0.655
cg24837915	NFATC4	1.655	7.970	0.005	5.232	1.659 ~ 16.506
c68167324	RDH12	1.153	4.104	0.043	3.168	1.038 ~ 9.666
cg157130156	DNAJB6	-1.020	4.116	0.042	0.361	0.135 ~ 0.966
OR: odds ratio; β : regression coefficients; 95%CI: 95% confidence interval of OR.						

Discussion

An increasing number of studies have explored the pathogenesis of GDM from the perspective of epigenetics. However, most of these were small (< 30 GDM cases), and they mainly observed the associations between GDM occurrence and the DNA methylation level of cord blood or placental tissue.^{20 21} In this study, we evaluated the DNA methylation status of GDM-related CpG sites in the peripheral blood of women in early pregnancy using MethylTarget™ sequencing. In addition, we verified the associations between target CpG sites and GDM using a larger sample size ($n = 80$). Overall, we identified 13 CpG sites with significant differences in DNA methylation levels between the GDM and control groups based on quantitative analysis. The AUCs of the ROC curve for each methylation level of the significant CpG sites ranged from 0.593 to 0.650 predictive utility in relation to GDM. The methylation status of eight individual CpG sites were identified as differing significantly between GDM and control groups by qualitative analysis, and these were located in the promoter regions of *RDH 12*, *HAPLN3*, *NFATC4*, *YAP1*, and *DNAJB6*, and the intron region of *C5orf34*. Importantly, we found that the methylation statuses of four CpG sites were significantly associated with GDM occurrence, namely cg89438648 (*HAPLN3*), cg68167324 (*RDH12*), cg157130156 (*DNAJB6*), and cg24837915 (*NFATC4*), using conditional logistic regression analysis.

In this study, hypermethylation of the CpG site cg89438648, located in the promoter region of *HAPLN3*, was found to suggest a lower risk of GDM (OR = 0.206; 95%CI: 0.065 ~ 0.655). *HAPLN3* codes for hyaluronan and proteoglycan link protein 3 (HAPLN3), and the connexin 3 belong to the hyaluronic acid and proteoglycan connexin (HAPLN) family, which plays roles in the aggregation of proteoglycans and hyaluronic acid, and in cell adhesion.²² *HAPLN3* is involved in the organization and stability of the hyaluronic acid (HA)-dependent extracellular matrix (ECM) in many tissues. HA is one component of the ECM within the islet tissue of humans and mice.²³ It can cause islet amyloid deposition, which is associated with decreased β -cell area and an increase in β cell apoptosis.²⁴ Hull et al. suggested that islet

amyloid deposition could reduce the number of β -cells.^{24 25} Hypermethylation of the CpG site cg89438648 located in the HAPLN3 promoter region, could reduce the level of HAPLN3, in turn reducing the stability of the HA-ECM, and consequently reducing the impact amyloid deposition on β cells.

We found that the hypermethylation status of CpG site cg68167324 located in *RDH 12*, can increase the risk of GDM (OR = 3.168; 95%CI: 1.038–9.666). *RDH 12* encodes retinol dehydrogenase 12 (RDH12), a member of the short-chain dehydrogenases/reductases (SDRs) family,²⁶ which participates in steroid and retinol metabolism.²⁷ RDH12, a NADPH-dependent all-trans retinol dehydrogenase, is the key enzyme in the metabolism of retinoids.²⁸ Two oxidation products of retinoids, 9-cis-retinoic acid and all-trans retinoic acid, function to stimulate insulin secretion.²⁹ In adipocytes, retinoic acid induces the expression of the insulin signaling gene *PDK-1* and that of the glucose transporter GLUT4. Activating retinoic acid induces the expression of genes involved in lipid and glucose metabolism, thereby improving insulin action.³⁰ Thus, hypermethylation of the CpG cg68167386 located upstream of the promoter region of *RDH12* may inhibit its transcriptional activity and reduce RDH12 levels in peripheral blood. Subsequently, the retinoic acid metabolic pathway would be inhibited, affecting insulin secretion, and reducing its effectiveness.

The DNAJB6 (DnaJ homolog, subfamily B, member 6) protein is a member of the heat shock protein 40 (HSP40) family³¹ and acts as a molecular chaperone for various cellular processes. While observing insulin resistant and diabetic patients, Kurucz et al. found that *HSP* expression was significantly changed without diabetes, and that the mRNA level of HSP72-inducible subtypes was significantly reduced in patients with type 2 diabetes.³² Additionally, the expression of *HSP70* in the skeletal muscle of patients with type 2 diabetes is reduced and has been shown to correlate with the degree of insulin resistance.³³ These HSP molecular chaperones are related to diabetes.³⁴ However, the exact association between DNAJB6 and type 2 diabetes needs further study. In this study, hypermethylation of CpG sites cg157130156, located in the promoter region of DNAJB6, was observed in the GDM group. This might result in increased DNAJB6 levels via the up-regulation of *DNABJ6* transcription, thereby reducing the risk of GDM (OR = 0.361; 95%CI: 0.135–0.966).

NFATC4 codes the nuclear factor of activated T cells 4 (NFATC4), which is a member of the transcription factor family under the control of calcineurin (a Ca^{2+} -dependent phosphatase).³⁵ In adipose tissue, NFATC4 has been shown to promote the secretion of inflammatory factors,³⁶ and to act as a transcriptional repressor in regulating adiponectin gene expression, suggesting that adiponectin expression is down-regulated in obesity and type 2 diabetes.³⁷ In this study, hypermethylation of the CpG site cg24837915 located in the promoter region of *NFATC4*, was associated with the presence of GDM (OR = 5.232; 95% CI, 1.659–16.506).

During pregnancy, early anabolism increases and mild insulin resistance occurs.³⁸ When insulin secretion fails to balance insulin resistance, impaired glucose tolerance develops, which might subsequently lead to GDM.³⁹ Therefore, impaired secretion by β cells is also a key factor in GDM pathogenesis. Here, we

explored the pathogenesis of GDM from an epigenetic perspective and identified 13 CpG sites that had methylation levels showing associations with GDM pathogenesis. Furthermore, conditional logistic regression analysis showed that the methylation status of four CpG sites located in the promoter regions of four genes was associated with GDM pathogenesis. These CpG sites are located in genes that could contribute to the development of GDM. Of these four CpG sites, hypermethylation of cg24837915 and cg68167324 was shown to be associated with GDM, whereas that of cg89438648 and cg157130156 could indicate reduced risk of GDM. Thus, the methylation status of these genes may function as predictors of GDM. No publications reporting on the relationship between methylation of these four CpG sites and GDM have been found, so our suggestion of such a relationship is based on the known modes of action of the genes concerned.

However, the study also had some limitations. First, our limited sample size, further verification of our results using a larger sample size is needed. Second, the selection of our target CpG sites is based on published literature, and we did not screen for differential CpG sites in the same population in this study, so there may be other CpG sites related to the pathogenesis of GDM that have not been verified.

Conclusions

In summary, by determining the DNA methylation of target CpG sites in the peripheral blood of women in early pregnancy, we found that the methylation levels of 13 CpG sites were related to the occurrence of GDM by quantitative analysis. After adjusting for possible confounding factors by conditional logistic regression, four CpG sites showed independent effects on the occurrence of GDM. These findings indicate that the methylation status of these CpG sites in the peripheral blood of pregnant women during the first trimester was related to the pathogenesis of GDM, and may be a potential predictor of GDM.

Declarations

Data availability statements: Raw data of patient characteristics and the data of target fragment DNA methylation sequencing generated during the current study is available from the corresponding author upon reasonable request.

Finding: This work was supported by Natural Science Foundation of Hunan Province (grant #2020JJ5789), National Natural Science Foundation of China (grant #81773535), Key Research and Development Program of Hunan Province (grant #2018SK2061), Hunan Provincial Grant for Innovative Province Construction (2019SK4012).

Duality of interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Contribution statement: XW wrote this article, HT and SM designed the study, and HT supervised the study and guaranteed the work. XW, MS, JH, SL and HL analyzed the data. XW, SM, HL, JH, NZ, MT and XW participated in the collection of basic data and samples. XW, SM, SL, HL, JH, NZ, MT and JA

participated in the discussion and reviewed/edited the manuscript. The final manuscript read and approved by all authors.

Acknowledgments: The authors would like to thank all research participants for their contributions and time.

References

1. Abell SK, De Courten B, Boyle JA, et al. Inflammatory and Other Biomarkers: Role in Pathophysiology and Prediction of Gestational Diabetes Mellitus. *Int J Mol Sci* 2015;16(6):13442-73. doi: 10.3390/ijms160613442 [published Online First: 2015/06/26]
2. Anna V, van der Ploeg HP, Cheung NW, et al. Sociodemographic correlates of the increasing trend in prevalence of gestational diabetes mellitus in a large population of women between 1995 and 2005. *Diabetes Care* 2008;31(12):2288-93. doi: 10.2337/dc08-1038 [published Online First: 2008/09/24]
3. Gao C, Sun X, Lu L, et al. Prevalence of gestational diabetes mellitus in mainland China: A systematic review and meta-analysis. *J Diabetes Investig* 2019;10(1):154-62. doi: 10.1111/jdi.12854 [published Online First: 2018/04/24]
4. Chiefari E, Arcidiacono B, Foti D, et al. Gestational diabetes mellitus: an updated overview. *J Endocrinol Invest* 2017;40(9):899-909. doi: 10.1007/s40618-016-0607-5 [published Online First: 2017/03/12]
5. Dana Dabelea RLH, Robert S. Lindsay, David J. Intrauterine Exposure to Diabetes Conveys Risks for Type 2 Diabetes and Obesity-A Study of Discordant Sibships. *DIABETES* December 2000; 49
6. Fraser A, Nelson SM, Macdonald-Wallis C, et al. Associations of existing diabetes, gestational diabetes, and glycosuria with offspring IQ and educational attainment: the Avon Longitudinal Study of Parents and Children. *Exp Diabetes Res* 2012;2012:963735. doi: 10.1155/2012/963735 [published Online First: 2012/08/29]
7. Coustan DR. Gestational diabetes mellitus. *Clin Chem* 2013;59(9):1310-21. doi: 10.1373/clinchem.2013.203331 [published Online First: 2013/03/29]
8. Wu L, Cui L, Tam WH, et al. Genetic variants associated with gestational diabetes mellitus: a meta-analysis and subgroup analysis. *Sci Rep* 2016;6:30539. doi: 10.1038/srep30539 [published Online First: 2016/07/30]
9. Moore LD, Le T, Fan G. DNA methylation and its basic function. *Neuropsychopharmacology* 2013;38(1):23-38. doi: 10.1038/npp.2012.112 [published Online First: 2012/07/12]
10. Kim JK, Samaranyake M, Pradhan S. Epigenetic mechanisms in mammals. *Cell Mol Life Sci* 2009;66(4):596-612. doi: 10.1007/s00018-008-8432-4 [published Online First: 2008/11/06]
11. Chim SS, Shing TK, Hung EC, et al. Detection and characterization of placental microRNAs in maternal plasma. *Clin Chem* 2008;54(3):482-90. doi: 10.1373/clinchem.2007.097972 [published Online First: 2008/01/26]

12. Kang J, Lee CN, Li HY, et al. Genome-wide DNA methylation variation in maternal and cord blood of gestational diabetes population. *Diabetes Res Clin Pract* 2017;132:127-36. doi: 10.1016/j.diabres.2017.07.034 [published Online First: 2017/08/24]
13. Kang J, Lee CN, Li HY, et al. Association of Interleukin-10 Methylation Levels With Gestational Diabetes in a Taiwanese Population. *Front Genet* 2018;9:222. doi: 10.3389/fgene.2018.00222 [published Online First: 2018/07/11]
14. Wu P, Farrell WE, Haworth KE, et al. Maternal genome-wide DNA methylation profiling in gestational diabetes shows distinctive disease-associated changes relative to matched healthy pregnancies. *Epigenetics* 2018;13(2):122-28. doi: 10.1080/15592294.2016.1166321 [published Online First: 2016/03/29]
15. Enquobahrie DA, Moore A, Muhie S, et al. Early Pregnancy Maternal Blood DNA Methylation in Repeat Pregnancies and Change in Gestational Diabetes Mellitus Status-A Pilot Study. *Reprod Sci* 2015;22(7):904-10. doi: 10.1177/1933719115570903 [published Online First: 2015/02/14]
16. International Association of D, Pregnancy Study Groups Consensus P, Metzger BE, et al. International association of diabetes and pregnancy study groups recommendations on the diagnosis and classification of hyperglycemia in pregnancy. *Diabetes Care* 2010;33(3):676-82. doi: 10.2337/dc09-1848 [published Online First: 2010/03/02]
17. Li E, Luo T, Wang Y. Identification of diagnostic biomarkers in patients with gestational diabetes mellitus based on transcriptome gene expression and methylation correlation analysis. *Reprod Biol Endocrinol* 2019;17(1):112. doi: 10.1186/s12958-019-0556-x [published Online First: 2019/12/29]
18. Naidoo V, Naidoo M, Ghai M. Cell- and tissue-specific epigenetic changes associated with chronic inflammation in insulin resistance and type 2 diabetes mellitus. *Scandinavian journal of immunology* 2018;88(6):e12723. doi: 10.1111/sji.12723 [published Online First: 2018/12/28]
19. de la Garza AL, Etxeberria U, Palacios-Ortega S, et al. Modulation of hyperglycemia and TNF α -mediated inflammation by helichrysum and grapefruit extracts in diabetic db/db mice. *Food Funct* 2014;5(9):2120-8. doi: 10.1039/c4fo00154k [published Online First: 2014/07/09]
20. Yang IV, Zhang W, Davidson EJ, et al. Epigenetic marks of in utero exposure to gestational diabetes and childhood adiposity outcomes: the EPOCH study. *Diabet Med* 2018;35(5):612-20. doi: 10.1111/dme.13604 [published Online First: 2018/02/21]
21. Haertle L, El Hajj N, Dittrich M, et al. Epigenetic signatures of gestational diabetes mellitus on cord blood methylation. *Clin Epigenetics* 2017;9:28. doi: 10.1186/s13148-017-0329-3 [published Online First: 2017/04/01]
22. Spicer AP, Joo A, Bowling RA, Jr. A hyaluronan binding link protein gene family whose members are physically linked adjacent to chondroitin sulfate proteoglycan core protein genes: the missing links. *J Biol Chem* 2003;278(23):21083-91. doi: 10.1074/jbc.M213100200 [published Online First: 2003/03/29]
23. Bogdani M, Korpos E, Simeonovic CJ, et al. Extracellular matrix components in the pathogenesis of type 1 diabetes. *Curr Diab Rep* 2014;14(12):552. doi: 10.1007/s11892-014-0552-7 [published Online

First: 2014/10/27]

24. Jurgens CA, Toukatly MN, Fligner CL, et al. beta-cell loss and beta-cell apoptosis in human type 2 diabetes are related to islet amyloid deposition. *Am J Pathol* 2011;178(6):2632-40. doi: 10.1016/j.ajpath.2011.02.036 [published Online First: 2011/06/07]
25. Hull RL, Bogdani M, Nagy N, et al. Hyaluronan: A Mediator of Islet Dysfunction and Destruction in Diabetes? *J Histochem Cytochem* 2015;63(8):592-603. doi: 10.1369/0022155415576542 [published Online First: 2015/07/29]
26. Keller B, Adamski J. RDH12, a retinol dehydrogenase causing Leber's congenital amaurosis, is also involved in steroid metabolism. *J Steroid Biochem Mol Biol* 2007;104(3-5):190-4. doi: 10.1016/j.jsbmb.2007.03.015 [published Online First: 2007/05/22]
27. Jörnvall H, Persson B, Krook M, et al. Short-chain dehydrogenases/reductases (SDR). *Biochemistry* 1995;34(18):6003-13. doi: 10.1021/bi00018a001 [published Online First: 1995/05/09]
28. Duester G. Families of retinoid dehydrogenases regulating vitamin A function: production of visual pigment and retinoic acid. *European journal of biochemistry* 2000;267(14):4315-24. doi: 10.1046/j.1432-1327.2000.01497.x [published Online First: 2000/07/06]
29. Blumentrath J, Neye H, Verspohl EJ. Effects of retinoids and thiazolidinediones on proliferation, insulin release, insulin mRNA, GLUT 2 transporter protein and mRNA of INS-1 cells. *Cell biochemistry and function* 2001;19(3):159-69. doi: 10.1002/cbf.907 [published Online First: 2001/08/09]
30. Berry DC, Noy N. All-trans-retinoic acid represses obesity and insulin resistance by activating both peroxisome proliferation-activated receptor beta/delta and retinoic acid receptor. *Mol Cell Biol* 2009;29(12):3286-96. doi: 10.1128/MCB.01742-08 [published Online First: 2009/04/15]
31. Elistratova IV, Ivanchenko OB, Grechko AV, et al. Heat shock protein HSP40 family chaperone DNAJB6/MRJ expression analysis in blood cells obtained from patients with atopic dermatitis in different phases. *Patologicheskaja fiziologija i eksperimental'naja terapija* 2016;60(3):23-30. [published Online First: 2016/07/01]
32. Kurucz I, Morva A, Vaag A, et al. Decreased expression of heat shock protein 72 in skeletal muscle of patients with type 2 diabetes correlates with insulin resistance. *Diabetes* 2002;51(4):1102-9. doi: 10.2337/diabetes.51.4.1102 [published Online First: 2002/03/28]
33. Stokov IA, Manukhina EB, Bakhtina LY, et al. The function of endogenous protective systems in patients with insulin-dependent diabetes mellitus and polyneuropathy: effect of antioxidant therapy. *Bulletin of experimental biology and medicine* 2000;130(10):986-90. [published Online First: 2001/02/15]
34. Zilae M, Shirali S. Heat Shock Proteins and Diabetes. *Can J Diabetes* 2016;40(6):594-602. doi: 10.1016/j.jcjd.2016.05.016 [published Online First: 2016/08/23]
35. Rao A, Luo C, Hogan PG. TRANSCRIPTION FACTORS OF THE NFAT FAMILY: Regulation and Function. *1997;15(1):707-47.* doi: 10.1146/annurev.immunol.15.1.707
36. Kim HB, Kumar A, Wang L, et al. Lipin 1 represses NFATc4 transcriptional activity in adipocytes to inhibit secretion of inflammatory factors. *Mol Cell Biol* 2010;30(12):3126-39. doi:

10.1128/MCB.01671-09 [published Online First: 2010/04/14]

37. Kim HB, Kong M, Kim TM, et al. NFATc4 and ATF3 negatively regulate adiponectin gene expression in 3T3-L1 adipocytes. *Diabetes* 2006;55(5):1342-52. doi: 10.2337/db05-1507 [published Online First: 2006/04/29]
38. Galtier-Dereure F, Boegner C, Bringer J. Obesity and pregnancy: complications and cost. *Am J Clin Nutr* 2000;71(5 Suppl):1242s-8s. doi: 10.1093/ajcn/71.5.1242s [published Online First: 2000/05/09]
39. Di Cianni G, Miccoli R, Volpe L, et al. Intermediate metabolism in normal pregnancy and in gestational diabetes. *Diabetes Metab Res Rev* 2003;19(4):259-70. doi: 10.1002/dmrr.390 [published Online First: 2003/07/25]

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryMaterial.docx](#)