Observational Study

Medicine®

OPEN

Comprehensive analysis of DNA methylation and gene expression profiles in gestational diabetes mellitus

Jing He, MM[a], Kang Liu, MM[a], Xiaohong Hou, MM[b], Jieqiang Lu, PhD[b,*]

Abstract

Gestational diabetes mellitus (GDM) has a high prevalence during pregnancy. This research aims to identify genes and their pathways related to GDM by combining bioinformatics analysis.

The DNA methylation and gene expression profiles data set was obtained from Gene Expression Omnibus. Differentially expressed genes (DEG) and differentially methylated genes (DMG) were screened by R package limma. The methylation-regulated differentially expressed genes (MeDEGs) were obtained by overlapping the DEGs and DMGs. A protein–protein interaction network was constructed using the search tool for searching interacting genes. The results are visualized in Cytoscape. Disease-related miRNAs and pathways were retrieved from Human MicroRNA Disease Database and Comparative Toxic Genome Database. Real-time quantitative PCR further verified the expression changes of these genes in GDM tissues and normal tissues.

After overlapping DEGs and DMGs, 138 MeDEGs were identified. These genes were mainly enriched in the biological processes of the "immune response," "defense response," and "response to wounding." Pathway enrichment shows that these genes are involved in "Antigen processing and presentation," "Graft-versus-host disease," "Type I diabetes mellitus," and "Allograft rejection." Six mRNAs (including superoxide dismutase 2 (*SOD2*), mitogen-activated protein kinase kinase kinase kinase 3 (*MAP4K3*), dual specificity phosphatase 5 (*DUSP5*), p21-activated kinases 2 (*PAK2*), serine protease inhibitor clade E member 1 (*SERPINE1*), and protein phosphatase 1 regulatory subunit 15B (*PPP1R15B*)) were identified as being related to GDM. The results obtained by real-time quantitative PCR are consistent with the results of the microarray analysis.

This study identified new types of MeDEGs and discovered their related pathways and functions in GDM, which may be used as molecular targets and diagnostic biomarkers for the precise diagnosis and treatment of GDM.

Abbreviations: CTD = comparative toxicogenomics database, DEGs = differentially expressed genes, DMGs = differentially methylated genes, GDM = gestational diabetes mellitus, GO = gene ontology, KEGG = Kyoto Encyclopedia of Genes and Genomes, MeDEGs = methylation-regulated differentially expressed genes, PPI = protein–protein interaction.

Keywords: bioinformatics, differentially expressed genes, differentially methylated genes, gestational diabetes mellitus, methylation-regulated differentially expressed genes

1. Introduction

Gestational diabetes mellitus (GDM) refers to the impaired glucose tolerance first discovered during pregnancy and is one of the most common pregnancy complications.[1,2] The prevalence of GDM is between 1% and 20%, and the prevalence of GDM in the high-risk study population exceeds 25%.[3–5] Perovic et al[4] found that the mean fetal liver length in GDM was significantly higher than that of healthy pregnant women. Moreover, GDM can increase the morbidity and mortality of mothers and fetuses, and is associated with macrosomia and various perinatal complications.[6] However, the screening and diagnosis of GDM lack uniform standards, and the missed diagnosis rate is very high. It is reported that GDM is caused by increased insulin resistance and pancreatic β cell dysfunction, involving genes related to insulin signal transduction, insulin secretion, diabetes onset in young adults, and lipid and glucose metabolism.[7,8]

DNA methylation is an epigenetic mechanism, which is essential for regulating gene transcription. So far, studies have reported that DNA methylation plays a vital role in the occurrence and development of many diseases.[9–11] Hajj et al found that the DNA methylation level of the maternally imprinted mesoderm-specific transcript (MEST) gene in the placenta and cord blood tissue of women with gestational diabetes was significantly lower than that of women with nongestational diabetes. In addition, obese adults have a lower degree of MEST methylation compared with normal-weight controls.[12,13] Nazari et al[14] found that GDM may adversely affect the pancreatic β-cells of the offspring through the hypomethylation of the CDKN2A/B promoter. Strakovsky et al[15] studied the high-fat diet during pregnancy, which has nothing to do with the occurrence of maternal obesity and diabetes. They found

Editor: Milan Perovic.

The authors have no conflicts of interest to disclose.

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

[a] Department of Obstetrics and Gynecology, Shanxi Bethune Hospital, Shanxi Medical University, Taiyuan, Shanxi, [b] Department of Obstetrics and Gynecology, The 2nd Affiliated Hospital of Wenzhou Medical University, Zhejiang, P. R. China.

[*] Correspondence: Jieqiang Lu, Department of Obstetrics and Gynecology, The 2nd Affiliated Hospital of Wenzhou Medical University, 306 Hualongqiao Road, Wenzhou, Zhejiang 325000, P. R. China (e-mail: jieqianglu666@163.com).

How to cite this article: He J, Liu K, Hou X, Lu J. Comprehensive analysis of DNA methylation and gene expression profiles in gestational diabetes mellitus. Medicine 2021;100:26(e26497).

Received: 5 January 2021 / Received in final form: 26 May 2021 / Accepted: 30 May 2021

http://dx.doi.org/10.1097/MD.0000000000026497

for the first time that the increase in mRNA expression of several genes is related to the hepatic gluconeogenesis pathway in the liver of the offspring of the fetus, which corresponds to the increase in the level of glucose in the offspring during childbirth.

In this study, we performed a bioinformatics analysis based on the microarray data of GDM. The methylation-regulated differentially expressed genes (MeDEGs) were identified, and enrichment analysis was performed on these MeDEGs. In addition, a protein–protein interaction (PPI) network was constructed to understand the molecular mechanism of GDM and provide candidate biomarkers for diagnosis and treatment.

2. Materials and methods

2.1. Microarray data

Microarray dataset GSE70494 (including gene expression dataset (GSE70493) and methylation dataset [GSE70453]) deposited by Binder et al were downloaded from the Gene Expression Omnibus database[16] (https://www.ncbi.nlm.nih.gov/). The GSE70453 methylation dataset was obtained from the Affymetrix Human Transcriptome Array 2.0 platform and includes 32 GDM placenta tissue samples and 31 healthy control placenta tissue samples. The GSE70493 gene expression profile dataset was obtained from the Illumina HumanMethylation450 BeadChip platform and contains 41 GDM placenta tissue samples and 41 healthy control placenta tissue samples. A total of 55 samples with both methylation and expression levels were selected, including 25 healthy control placenta tissue samples and 30 GDM placenta tissue samples.

2.2. Data preprocessing and analyzing

The limma package Version 3.34.0 in R3.4.1[17] (https://bioconductor.org/packages/release/bioc/html/limma.html) was used to analyze gene methylation profile data and gene expression profile data to identify differentially methylated genes (DMGs) and differentially expressed genes (DEGs). The FDR $<$ 0.05 and $|\log2FC| > 0.263$ was used as the threshold for screening DMGs and DEGs. Then, through the pheatmap package Version1.0.8 in R3.4.1[18] (https://cran.r-project.org/package=pheatmap), two-direction hierarchical clustering based on Euclidean Distance was performed on the expression level of DERs and the methylation level of DMRs. Finally, Venn diagrams were used to identify MeDEGs.

2.3. Functional and pathway enrichment analysis

DAVID 6.8[19,20] (https://david.ncifcrf.gov/) was used to perform Gene ontology (GO) analysis enrichment and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis on selected MeDEGs. The $P < .05$ were defined as significant terms.

2.4. Protein–protein interaction (PPI) network

STRING is an online database used to predict PPI, which is essential for explaining the molecular mechanism of key cellular activities in GDM.[21] The STRING database was used to construct a PPI network of MeDEGs. PPI network was visually displayed through Cytoscape Version 3.7.2 software[22] (http://www.cytoscape.org/). Subsequently, the GO biological process and KEGG signal pathway analysis were performed based on DAVID.

2.5. Construction of miRNA–mRNA network

The miRNAs related to GDM were obtained from the Human MicroRNA Disease Database database[23] (Human MicroRNA Disease Database, http://www.cuilab.cn/hmdd), which contains a number of miRNA–disease association entries from literature. Then, GDM directly related miRNA target genes are predicted using the starBase database[24] (http://starbase.sysu.edu.cn/), hereby obtaining miRNA–mRNA interaction pairs. Subsequently, the miRNA–mRNA network was constructed and visualized using Cytoscape v 3.6.0 software. The target genes in the network were analyzed based on DAVID for GO biological process and KEGG pathway.

2.6. Construction of miRNA–mRNA and GDM-related pathway regulatory network

The "asthmatic" as the keyword was used to search for KEGG pathways and genes related to GDM in the Comparative Toxicogenomics Database (CTD) 2019 update database[25] (http://ctd.mdibl.org/). The GDM-related genes and pathways were regarded as the overlapping genes and pathways between those identified from the CTD and the genes and pathways in the miRNA–mRNA network. Then the overlapping genes and overlapping pathways were used to construct an interaction network of GDM-related pathways and genes.

2.7. Real-time quantitative PCR (RT-qPCR)

The expression of MeDEGs was further detected by RT-qPCR. Placenta tissues of 5 GDM samples and 5 normal samples were collected from the department of Obstetrics and Gynecology, Shanxi Bethune Hospital. The inclusion criteria for patients were: the GDM group included the patients who have been diagnosed with GDM and have no previous history of hypertension or diabetes. The normal group included pregnant patients with non-GDM, no gestational hypertension, no gestational heart disease, and no previous history of hypertension or diabetes. This study was approved by the ethic committee of Shanxi Bethune Hospital.

Total RNA was extracted using RNAiso Plus (TaKaRa, Tokyo, Japan). According to the manufacturer's instructions, mRNAs were reverse transcribed using a PrimeScript II 1st Strand cDNA Synthesis Kit (TAKARA, catalog number 6210A, Japan). The PCR reactions were performed in a total volume of 20 μL, which includes 10 μL SYBR Premix EX Taq, 1 μL forward primer (10 μM), 1 μL reverse primer (10 μM), 2 μL cDNA, and ddH2O (up to 20 μL). The reaction was performed in a ViiA 7 (Applied Biosystems by Life Technologies, Austin, TX) real-time PCR machines. The PCR conditions were at 50°C for 2 minutes, 95°C for 2 minutes, 40 cycles of 95°C for 15 seconds, and 60°C for 60 seconds. Dissociation curve was analyzed from 60°C to 95°C. The relative gene expression was analyzed by the $2^{-\Delta\Delta CT}$ method. GAPDH was used as endogenous controls for gene expression in the analysis. $\underline{P} < .05$ and $P < .01$ were used as the screening criteria for significant differences and extremely significant differences.

3. Results

3.1. Identification of DEGs and DMGs

For the DEGs of the gene expression microarray, 234 upregulated genes and 319 downregulated genes were identified. For the DMGs of the gene methylation microarray, 232 hypomethylated

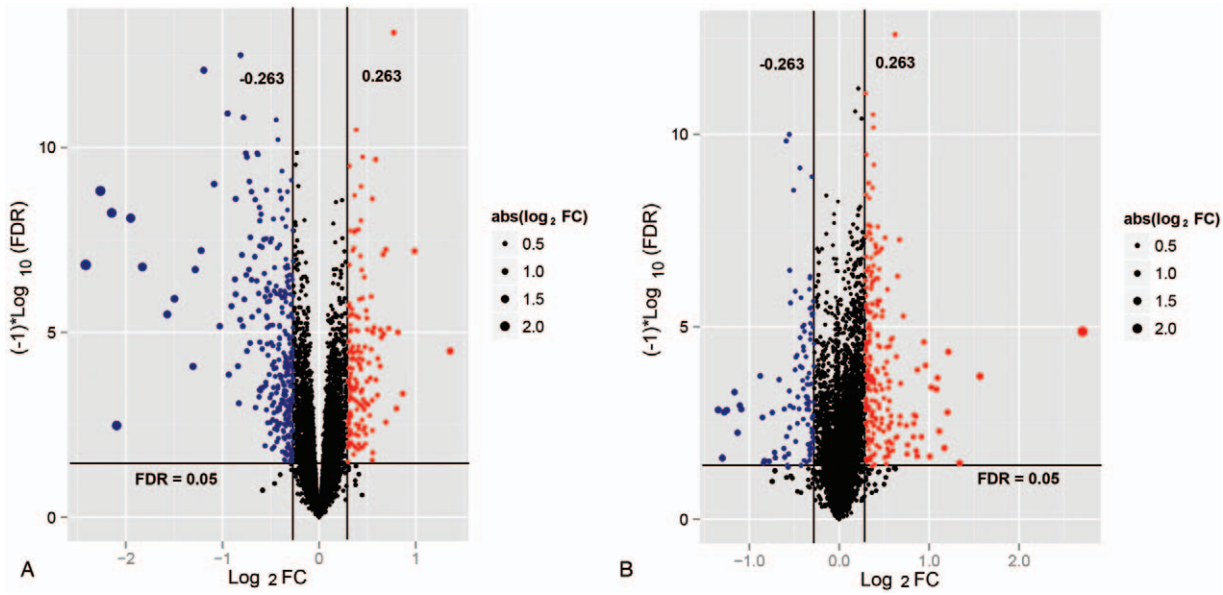

Figure 1. A volcano plot of differentially expressed genes (A) and differentially methylated genes (B). The red plus signs represent upregulated genes, the blue triangles represent downregulated genes, and the black circles represent nondifferentially expressed genes. FC = fold-change.

genes and 513 hypermethylated genes were identified. The volcano map shows the distribution of DEGs and DMGs (Fig. 1). It is found from the heat map that DEGs and DMGs were different between GDM and control samples (Fig. 2).

By overlapping DEGs and DMGs, a total of 245 overlapping genes were obtained (Fig. 3A). Subsequently, 138 MeDEGs were identified from the overlapping genes (Fig. 3B).

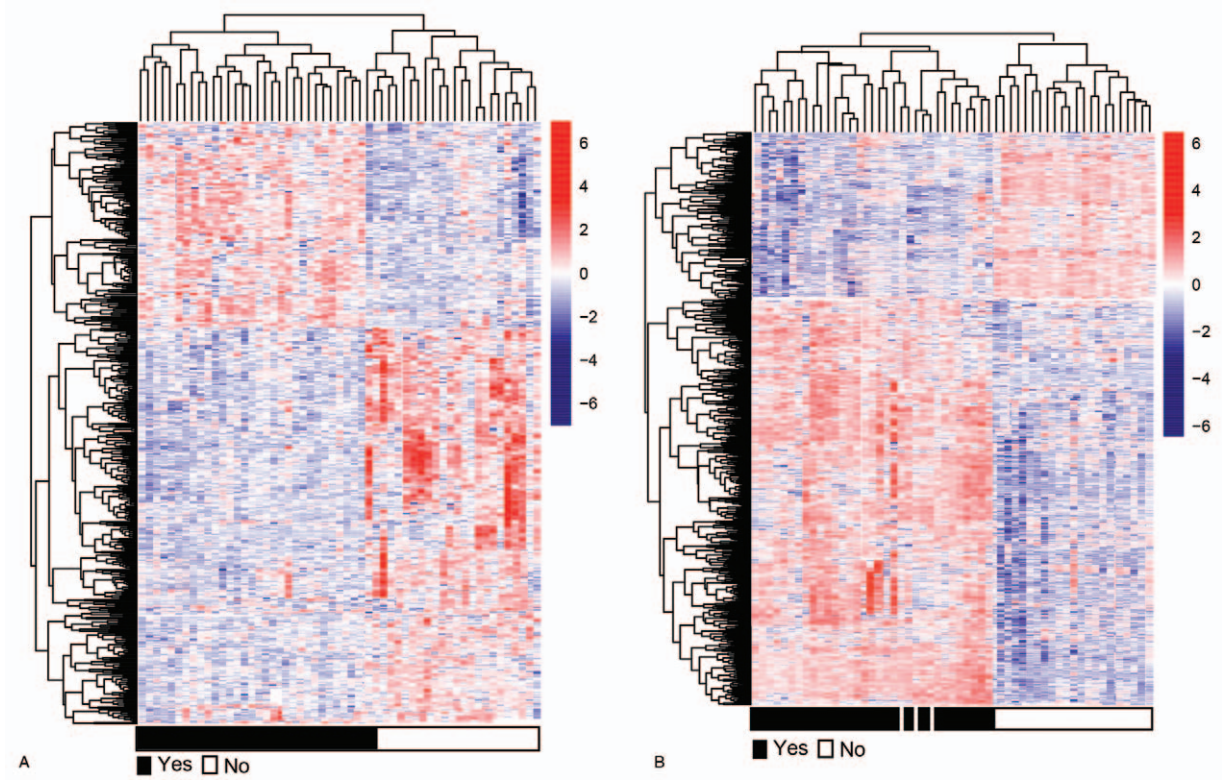

Figure 2. A, Heat map of DEGs. B, Heat map of DMGs. Black represents GDM samples, and white represents control samples. DEGs = differentially expressed genes, DMGs = differentially methylated genes, GDM = gestational diabetes mellitus.

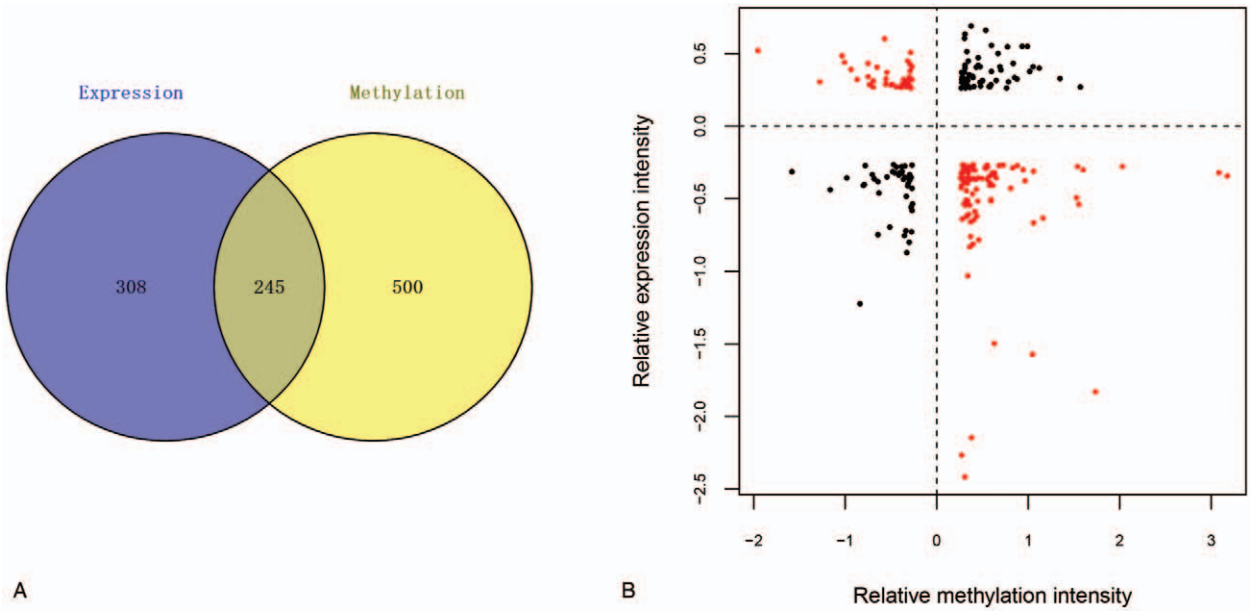

Figure 3. A, Venn diagram for aberrantly methylated-differentially expressed genes by overlapping DEGs and DEMs. B, Scattered distribution diagram of the degree of difference between DEGs and DMGs. DEGs = differentially expressed genes, DMGs = differentially methylated genes.

3.2. Enriched GO terms and KEGG pathways

The GO and KEGG pathway enrichment analysis of MeDEGs was performed. A total of 20 significantly related biological processes and 12 KEGG signaling pathways were obtained (Fig. 4). GO analysis results show that these MeDEGs were rich in 20 biological processes, which are mainly related to "immune response," "defense response," and "response to wounding." In addition, the results of the KEGG pathway enrichment analysis indicated that these MeDEGs were mainly involved in "Allograft rejection," "Graft-versus-host disease," "Type I diabetes mellitus," and "Autoimmune thyroid disease pathways."

3.3. PPI network analysis

STRING was used to build a PPI network. The PPI network contains 91 nodes and 182 edges, among which 73 nodes were established from hypermethylated-downregulated and 18 nodes were established from hypomethylated-upregulated (Fig. 5). Then the GO biological process and KEGG pathway enrichment analysis were carried out on the MeDEGs that constitute the PPI network (Table 1). The results showed that MeDEGs in the PPI network were significantly related to biological processes such as "immune response," "defense response," and "response to wounding." The MeDEGs in the PPI network were mainly enriched in KEGG pathways such as "Allograft rejection," "Graft-versus-host disease," "Type I diabetes mellitus," and "Autoimmune thyroid disease."

3.4. MiRNA–mRNA regulatory network related to GDM

A total of 25 miRNAs related to GDM were identified from the Human MicroRNA Disease Database database. Subsequently, after overlapping the target mRNA and MeDEGs of the GDM-related miRNAs determined by the starBase database, 301 miRNA–mRNA regulatory pairs were obtained (Fig. 6). Together, a miRNA–mRNA network was generated, comprising 95 nodes and 301 connecting edges.

Then, the MeDEGs in the miRNA-mRNA network were analyzed for enrichment of biological processes and KEGG pathways, and 7 biological processes and 7 KEGG signaling pathways were obtained (Table 2). The results showed that those MeDEGs were significantly related to biological processes such as "posttranscriptional regulation of gene expression," "oxidation reduction," and "regulation of translation." Those MeDEGs were mainly enriched in KEGG pathways such as "Antigen processing and presentation," "Graft-versus-host disease," "Type I diabetes mellitus," and "Allograft rejection."

3.5. GDM-related miRNA–mRNA–pathway network

In the CTD database, 33 related KEGG signaling pathways and 10 related genes were obtained with "gestational diabetes mellitus" as the key word. An overlapping pathway "hsa04010: MAPK signaling pathway" and an overlapping GDM-related gene "*SOD2*" were identified between the CTD database and the miRNA–mRNA network. The miRNA–mRNA-pathway network consisted of 13 genes and 16 miRNAs (Fig. 7). Three genes *MAP4K3*, *DUSP5*, and *PAK2* were involved in the overlapping pathways (MAPK signaling pathway). Moreover, *SOD2* participates in 3 biological processes: "0055114_oxidation reduction," "0006732_coenzyme metabolic process," and "0000302_response to reactive oxygen species." In addition, it was also found that *SERPINE1*, *PPP1R15B*, and *SOD2* are all involved in the "0055114_oxidation reduction" and "0000302_response to reactive oxygen species" biological process.

3.6. RT-qPCR validation

The expression of *SOD2*, *MAP4K3*, *DUSP5*, *PAK2*, *SERPINE1*, and *PPP1R15B* was evaluated by RT-qPCR in GDM tissues

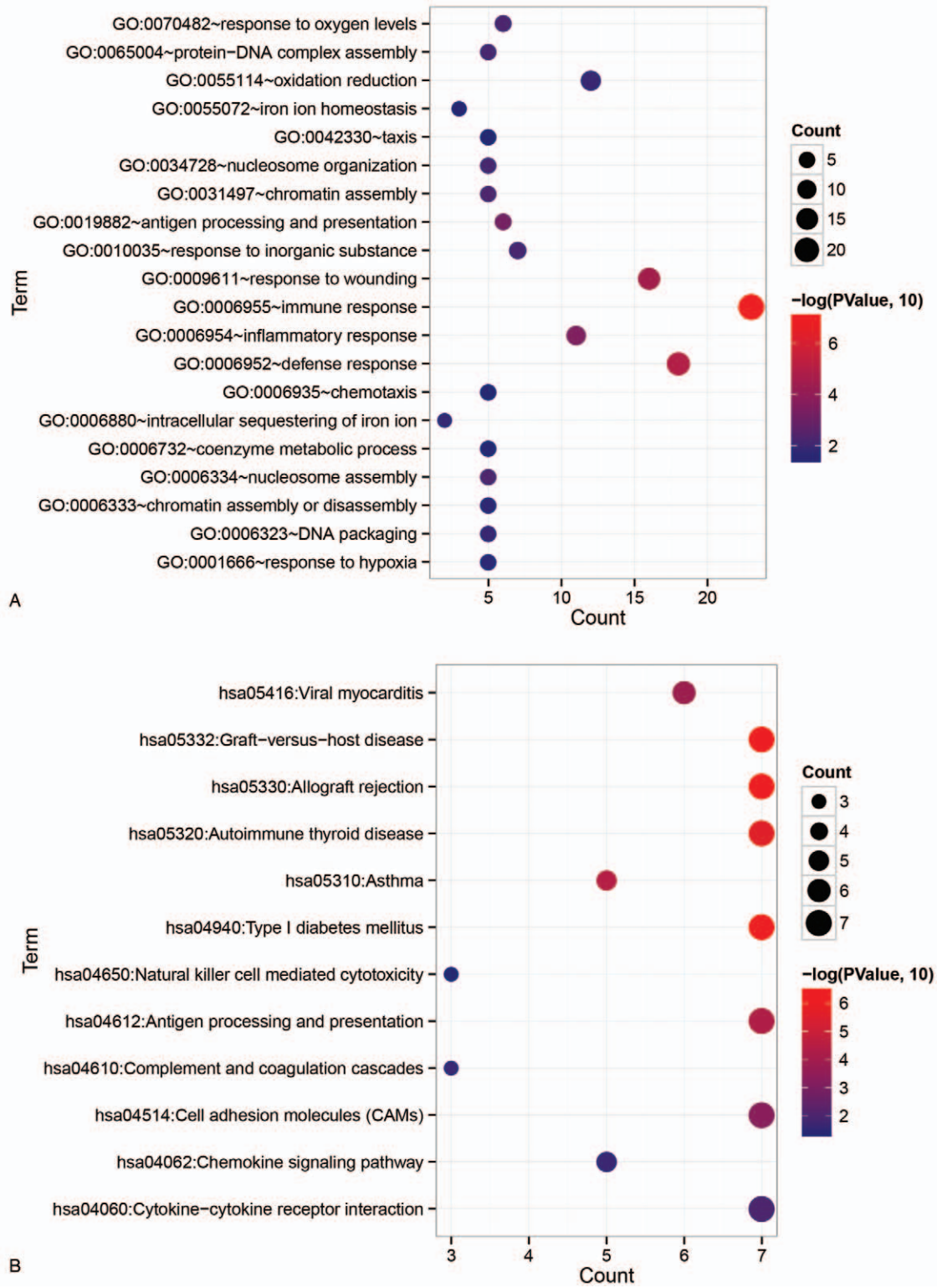

Figure 4. Functional enrichment analysis of methylated-differentially expressed genes. A, Enriched GO terms in the "biological process" category. B, Enriched KEGG biological pathways. The x-axis represents the proportion of genes, and the y-axis represents different categories. The different colors indicate different properties, and the different sizes represent the number of genes. GO = gene ontology, KEGG = Kyoto Encyclopedia of Genes and Genomes.

compared with normal tissues. The characteristics of these patients are displayed in Table 3. There was no difference in prepregnancy BMI, maternal age, gestational age and birth weight between the GDM group and the normal group ($P > .05$). As depicted in Figure 8, the expression level of *DUSP5* in the GDM group was not significantly different from that in the

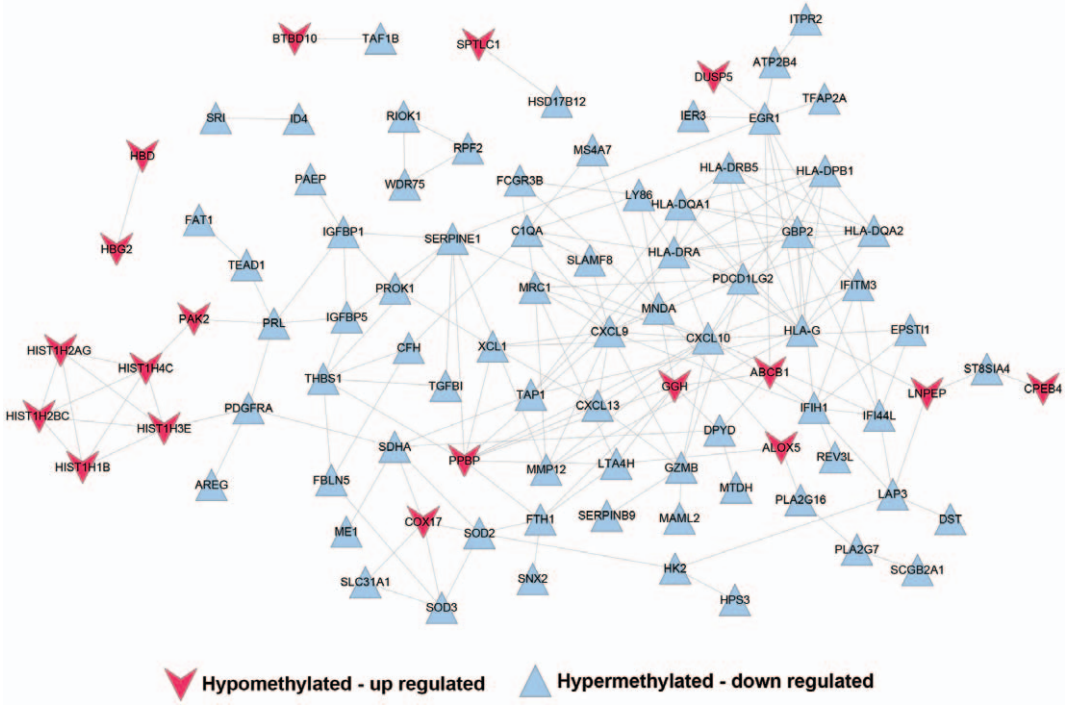

Figure 5. PPI network of methylation-regulated differentially expressed genes. Blue represents the Hypermethylated-downregulated gene, and red represents the Hypomethylated-upregulated gene. PPI = protein–protein interaction.

Table 1

The biological processes and pathways of methylation-regulated differentially expressed genes in the PPI network.

Category	Term	Count	*P* value
Biology process	GO:0006955~immune response	23	4.38E-11
	GO:0006952–defense response	17	4.90E-07
	GO:0009611–response to wounding	15	2.36E-06
	GO:0006954~inflammatory response	10	1.24E-04
	GO:0019882~antigen processing and presentation	6	1.29E-04
	GO:0006334~nucleosome assembly	5	1.53E-03
	GO:0031497~chromatin assembly	5	1.74E-03
	GO:0065004~protein-DNA complex assembly	5	2.05E-03
	GO:0034728~nucleosome organization	5	2.22E-03
	GO:0006323~DNA packaging	5	5.07E-03
	GO:0042127~regulation of cell proliferation	12	6.40E-03
	GO:0006333~chromatin assembly or disassembly	5	6.76E-03
	GO:0010035~response to inorganic substance	6	7.30E-03
	GO:0070482~response to oxygen levels	5	9.71E-03
	GO:0006935~chemotaxis	5	1.49E-02
	GO:0042330~taxis	5	1.49E-02
	GO:0055072~iron ion homeostasis	3	1.91E-02
	GO:0007626~locomotory behavior	6	2.31E-02
	GO:0030198~extracellular matrix organization	4	2.37E-02
	GO:0008284~positive regulation of cell proliferation	7	3.57E-02
	GO:0001666~response to hypoxia	4	4.50E-02
KEGG pathway	hsa05330:Allograft rejection	7	9.14E-07
	hsa05332:Graft-versus-host disease	7	1.50E-06
	hsa04940:Type I diabetes mellitus	7	2.36E-06
	hsa05320:Autoimmune thyroid disease	7	7.58E-06
	hsa04612:Antigen processing and presentation	7	1.26E-04
	hsa05310:Asthma	5	1.51E-04
	hsa05416:Viral myocarditis	6	5.45E-04
	hsa04672:Intestinal immune network for IgA production	5	1.17E-03
	hsa04514:Cell adhesion molecules (CAMs)	7	1.53E-03
	hsa04060:Cytokine-cytokine receptor interaction	7	3.85E-02

GO = gene ontology, PPI = protein–protein interaction.

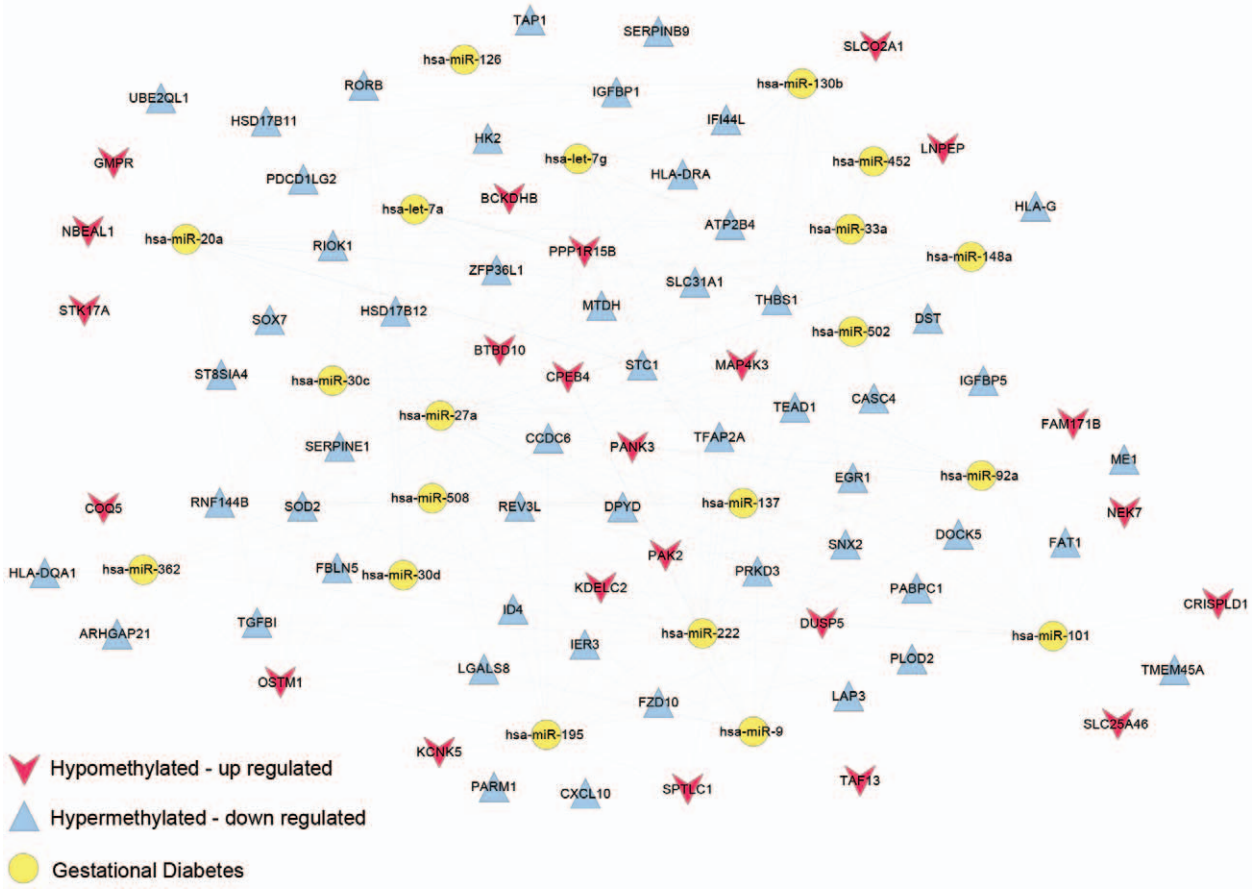

Figure 6. The miRNA-mRNA regulatory network. The blue triangle represents the Hypermethylated-downregulated gene, the red triangle represents the Hypomethylated-upregulated gene, and the yellow circle represents the miRNAs directly related to GDM. GDM = gestational diabetes mellitus.

normal group. The expression levels of *MAP4K3*, *PAK2*, and *PPP1R15B* were significantly higher in GDM group than in control group ($P < .01$). The expression levels of *SOD2* and *SERPINE1* in the GDM group were significantly lower than those in the control group ($P < .01$).

4. Discussion

In this study, 2 types of GDM microarray chips (DNA methylation and gene expression profile data sets) were comprehensively analyzed using bioinformatics analysis to identify GDM-related genes and disease mechanisms. A total

Table 2

Biological processes and pathways that involve targets of GDM-related miRNAs.

Category	Term	Count	*P* value
Biology process	GO:0010608~posttranscriptional regulation of gene expression	5	1.42E-02
	GO:0055114~oxidation reduction	8	2.24E-02
	GO:0006417~regulation of translation	4	2.29E-02
	GO:0006732~coenzyme metabolic process	4	3.04E-02
	GO:0006955~immune response	8	3.23E-02
	GO:0009108~coenzyme biosynthetic process	3	3.75E-02
	GO:0000302~response to reactive oxygen species	3	4.37E-02
KEGG pathway	hsa04612:Antigen processing and presentation	4	1.24E-02
	hsa05330:Allograft rejection	3	1.87E-02
	hsa05332:Graft-versus-host disease	3	2.18E-02
	hsa04940:Type I diabetes mellitus	3	2.50E-02
	hsa05320:Autoimmune thyroid disease	3	3.59E-02
	hsa04514:Cell adhesion molecules (CAMs)	4	4.17E-02
	hsa04010:MAPK signaling pathway	3	4.73E-02

GO = gene ontology, KEGG = Kyoto Encyclopedia of Genes and Genomes.

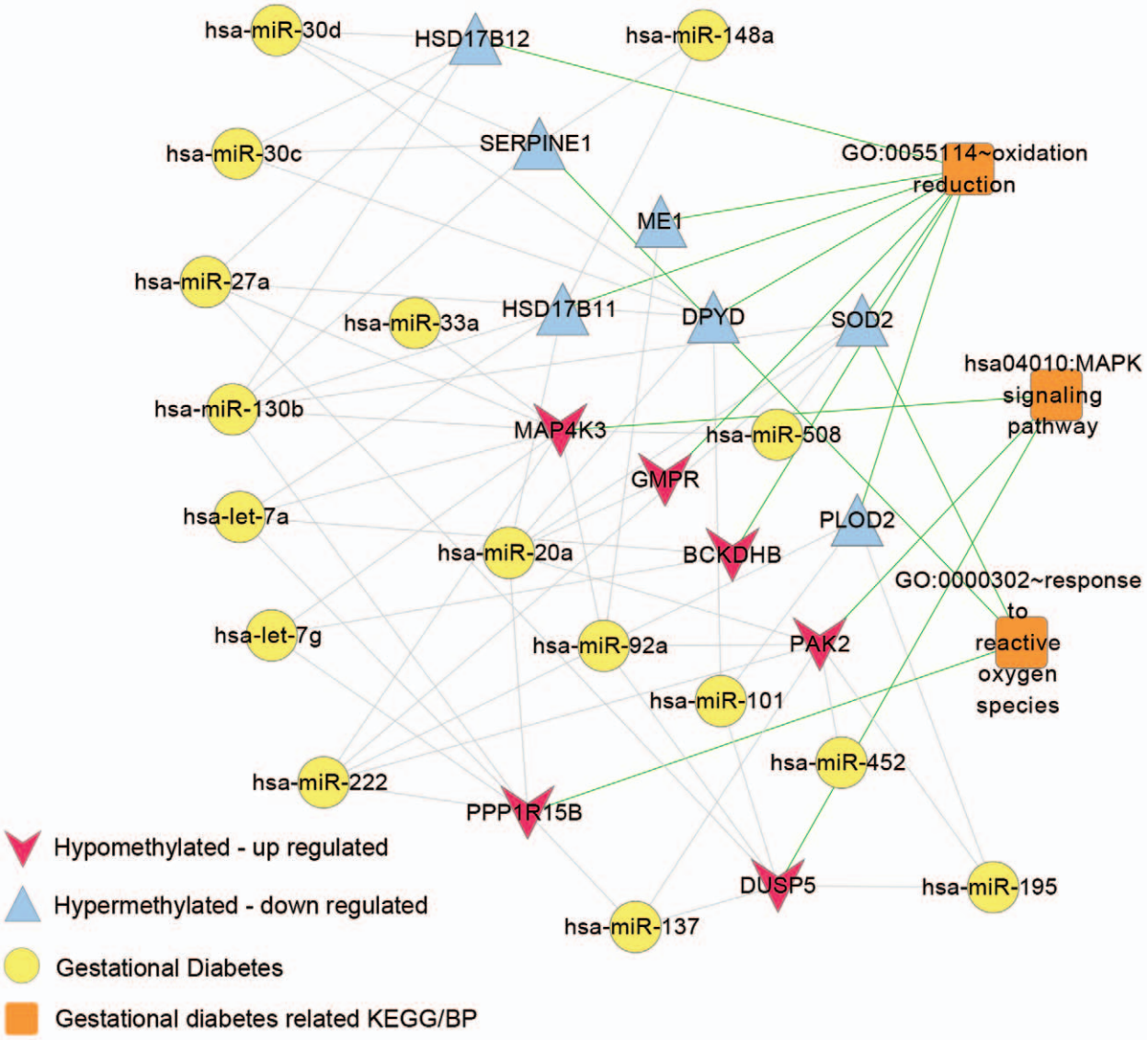

Figure 7. The miRNA–mRNA-KEGG pathway network. The blue triangle represents the Hypermethylated-downregulated gene, the red triangle represents the Hypomethylated-upregulated gene, the yellow circle represents the miRNAs directly related to GDM, and the yellow square represents the KEGG signaling pathway directly related to GDM. GDM = gestational diabetes mellitus, KEGG = Kyoto Encyclopedia of Genes and Genomes.

Table 3

Patient characteristics.

Characteristics	GDM group	Normal group	*P* value
Prepregnancy BMI (kg/m^2)	24.784 ± 2.62	22.096 ± 0.92	.062
Maternal age (y)	35.2 ± 2.77	32.2 ± 2.28	.099
gravidity			1
1	0 (0%)	2 (40%)	
2	3 (60%)	2 (40%)	
≥3	2 (40%)	1 (20%)	
Smoke during pregnancy		1	
Yes	0 (0%)	0 (0%)	
No	5 (100%)	5 (100%)	
Infant sex			1
Males	3 (60%)	3 (60%)	
Females	2 (40%)	2 (40%)	
Gestational age (wks)	38.858 ± 0.706	39.114 ± 0.327	.482
Birth weight (g)	3497 ± 515.40	3616 ± 410.46	.697

BMI = body mass index, GDM = gestational diabetes mellitus.

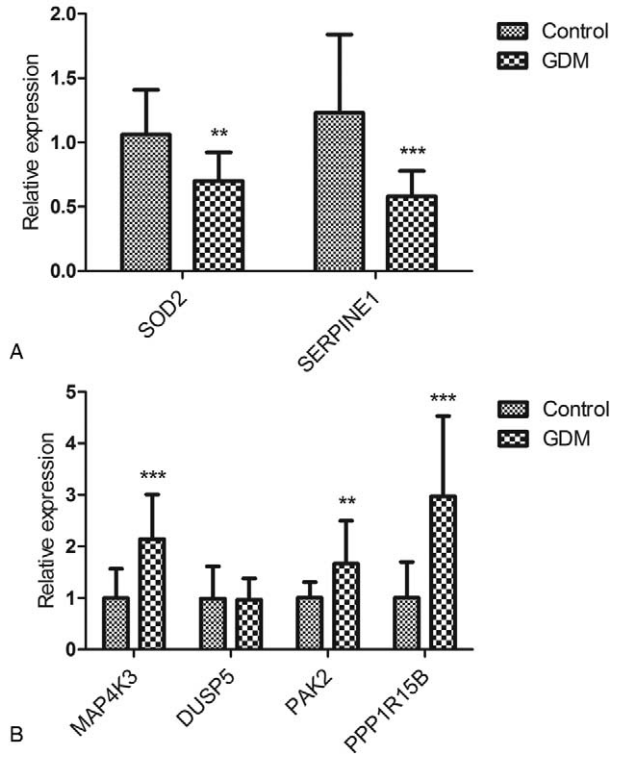

Figure 8. The expression levels of *SOD2*, *MAP4K3*, *DUSP5*, *PAK2*, *SERPINE1,* and *PPP1R15B* in GDM tissues and normal tissues quantified by qPCR. GDM = gestational diabetes mellitus.

of 553 DEGs (234 upregulated and 319 downregulated) and 745 DMGs (232 hypomethylated and 513 hypermethylated) were identified. By overlapping DEG and DMG, 138 MeDEGs were identified. The discovery of the interaction network revealed that related MeDEGs might be involved in the molecular transformation of important pathways related to the occurrence and development of GDM. Gene function and enrichment analysis confirmed the identified pathways and central genes related to methylation, which may provide new insights into the pathogenesis of GDM.

GO analysis shows that the main biological processes of MeDEGs involve "immune response," "defense response," and "response to wounding." These findings are reasonable. It is reported that GDM is associated with impaired maternal immune response.[26] Moreover, the differentially regulated proteins related to GDM identified by Zhao et al[27] also implicated in immune response and defense response. In addition, the enrichment analysis of the KEGG pathway showed that these MeDEGs were mainly enriched in "Allograft rejection," "Graft-versus-host disease," "Type I diabetes mellitus," and "Autoimmune thyroid disease pathways." This indicates that GDM is also involved in the pathway of type I diabetes. This phenomenon was also confirmed by Radaelli et al[28] They found that genes at key steps of fatty acid uptake, transport, and activation pathways were similarly up-regulated in pregnancy with GDM and type I diabetes.

One overlapping gene, *SOD2*, was identified between genes in the miRNA–mRNA network and genes associated to GDM in the CTD. *SOD2* is a mitochondrial enzyme encoded by genomic DNA, which is involved in the detoxification of free radicals produced by mitochondrial respiration. The SOD2 gene is up-regulated under oxidative stress and protects cells from the harmful effects of reactive oxygen species.[29] It has been reported that *SOD2* is associated with an increased risk of reducing gestational age and birth weight.[30] In addition, *SOD2* overexpression inhibits mitochondrial translocation of pro-apoptotic Bcl-2 family members, reduces the number of mitochondrial defects in neuroepithelial cells and reduces mitochondrial membrane potential, thereby eliminating mitochondrial dysfunction caused by maternal diabetes. An in vivo experiment showed that maternal diabetes can cause the inhibition of *SOD2* in the amygdala, leading to autism-like behavior in offspring.[31] Overexpression of *SOD2* was restored, and knockdown of SOD2 mimics this effect, indicating that oxidative stress and *SOD2* expression play an important role in the behavior of the offspring of autism caused by maternal diabetes. Currently, the regulatory mechanism of SOD2 in GDM was not clear. In this study, it was found that the expression level of *SOD2* in patients with GDM was significantly reduced, while the level of methylation was increased. Moreover, the miRNA–mRNA-pathway network suggests that hsa-miR-130b, hsa-miR-20a, hsa-miR-508, and hsa-miR-222 jointly regulate *SOD2*. Therefore, we speculate that the expression of *SOD2* was reduced by regulating these miRNAs to participate in the biological process of the response to reactive oxygen species, which may plays an important role in the pathophysiology of GDM.

In this study, we found that *SERPINE1* and *PPP1R15B* are also involved in the biological process of response to reactive oxygen species. *SERPINE1* is a member of the serine protease inhibitor superfamily. Kohler and Grant[32] found that the *SERPINE1* gene is involved in the pathogenesis of cardiovascular disease. Moreover, some studies have shown that elevated *SERPINE1* levels are associated with an increased risk of type 2 diabetes and its complications (such as diabetic retinopathy and diabetic coronary artery disease).[33–35] This study found that SERPINE1 is hypermethylated and downregulated, and SERPINE1 is simultaneously regulated by hsa-miR-30d, hsa-miR-30c, hsa-miR-130b, and hsa-miR-148a in the miRNA-mRNA-pathway network. *PPP1R15B* is a constitutive repressor of protein phosphatase and eIF2α phosphorylation and is a crucial regulator of translation during cellular stress.[36]*PPP1R15B*-deficient β-cells showed enhanced phosphorylation of eIF2α and were prone to apoptosis.[37] In addition, *PPP1R15B*-deficient mice have low body weight, low survival rate, impaired erythropoiesis, and increased phosphorylation of eIF2α in fibroblasts.[38]*PPP1R15B* is hypomethylated and upregulated in GDM, and it is also affected by hsa-miR-222, hsa-let-7g, hsa-let-7a, hsa-miR-130b, hsa-miR-20a, and hsa-miR-137 regulation. Therefore, we speculate that it may be the combined effect of methylation and miRNA to increase the expression of *PPP1R15B* and decrease the expression of *SERPINE1*, thereby playing a role in the treatment of GDM. These conjectures need further experiments to verify.

One overlapping KEGG pathway, MAPK signaling pathway, was identified between genes in the miRNA–mRNA network and genes associated with GDM in the CTD. The MAPK signal pathway is involved in a variety of cellular activities, including cell proliferation, differentiation, migration, senescence, and apoptosis.[39] There are 3 genes enriched in this pathway, namely *MAP4K3*, *DUSP5,* and *PAK2*. *MAP4K3*, also known as *GLK*, is

a serine/threonine kinase that belongs to the Ste20-like kinase family of mammals.[40,41] Studies have shown that *MAP4K3* is a positive regulator of T cell signaling and T-cell-mediated immune response.[42] Overexpression of *MAP4K3* is associated with human autoimmune diseases such as psoriatic arthritis,[43] rheumatoid arthritis,[44] adult still's disease[45] and systemic lupus erythematosus.[46]*MAP4K3* was significantly Hypomethylated, upregulated in GDM, and it is also affected by hsa-miR-33a, hsa-miR-27a, hsa-miR-130b, hsa-let-7a, hsa-let-7g, hsa-miR-222, hsa-miR-92a, and hsa-miR-508 regulation. *DUSP5* is a member of the dual-specificity protein phosphatase subfamily, which inactivates its target kinases by dephosphorylating phosphoserine/threonine residues and phosphotyrosine residues.[47]*DUSP5* can phosphorylate mitogen-activated protein kinase extracellular signal-regulated kinase (ERK1/2) and play an important role in embryonic vasculature development.[48,49] Moon et al[50] found that during the induction of collagen-induced arthritis, *DUSP5*-overexpressing mice showed reduced pro-inflammatory cytokines in joint tissues. *DUSP5* was significantly Hypomethylated, upregulated in GDM, and the miRNA–mRNA–pathway network indicates that DUSP5 is also regulated by hsa-miR-137, hsa-miR-27a, hsa-miR-92a, hsa-miR-101, and hsa-miR-195. *PAK2* is a serine/threonine kinase that acts as a negative regulator of neuronal glucose uptake and insulin sensitivity.[51] Under basal and insulin-stimulated conditions, GTPase Rac1 activates *PAK2* to inhibit neuronal glucose uptake.[51] In this study, *PAK2* was Hypomethylated and upregulated in GDM, and the miRNA–mRNA–pathway network found that *PAK2* is also regulated by hsa-miR-195, hsa-miR-452, hsa-miR-137, hsa-miR-222, hsa-miR-92a, and hsa-miR-20a. Based on these results, we speculate that MAP4K3, DUSP5, and PAK2 may play a role in the occurrence and development of GDM through the coregulation of methylation and miRNAs, but this speculation still needs to be verified by subsequent experiments.

Moreover, qRT-PCR further validated that *MAP4K3, PAK2,* and *PPP1R15B* were highly expressed in GDM samples compared with the control samples, while *SOD2* and *SERPINE1* were lowly expressed in GDM samples compared with the control samples.

There are some limitations in our study. First, the methylation status of cytosine-phosphate-guanine islands of 6 genes related to GDM was not detected. Second, the experiment did not verify the effect of abnormal methylation and miRNA expression on gene expression. Therefore, further evaluations in clinical trials are needed to verify these genes.

In summary, a comprehensive analysis of DNA methylation and gene expression profiles was conducted to identify genes related to GDM. Six genes (*SOD2*, *MAP4K3*, *DUSP5*, *PAK2*, *SERPINE1,* and *PPP1R15B*) that may be related to the pathogenesis of GDM have been identified, which may provide new methods for the treatment of GDM. In addition, these genes were verified by RT-PCR.

Author contributions

Conceptualization: Jing He, Kang Liu, Xiaohong Hou, Jieqiang Lu.
Formal analysis: Jing He.
Methodology: Jieqiang Lu.
Project administration: Kang Liu, Xiaohong Hou.
Supervision: Kang Liu.
Validation: Jieqiang Lu.
Writing – original draft: Jing He.
Writing – review & editing: Kang Liu, Xiaohong Hou, Jieqiang Lu.

References

[1] Weinert LS. International Association of Diabetes and Pregnancy Study Groups recommendations on the diagnosis and classification of hyperglycemia in pregnancy: comment to the International Association of Diabetes and Pregnancy Study Groups Consensus Panel. Diabetes Care 2010;33:e97.
[2] Zhou SJ, Yelland L, McPhee AJ, Quinlivan J, Gibson RA, Makrides M. Fish-oil supplementation in pregnancy does not reduce the risk of gestational diabetes or preeclampsia. Am J Clin Nutr 2012;95:1378–84.
[3] Ashwal E, Hod M. Gestational diabetes mellitus: where are we now? Clin Chim Acta 2015;451(Pt A):14–20.
[4] Perovic M, Gojnic M, Arsic B, et al. Relationship between mid-trimester ultrasound fetal liver length measurements and gestational diabetes mellitus. J Diabetes 2015;7:497–505.
[5] Perović M, Garalejić E, Gojnić M, et al. Sensitivity and specificity of ultrasonography as a screening tool for gestational diabetes mellitus. J Matern Fetal Neonatal Med 2012;25:1348–53.
[6] Kc K, Shakya S, Zhang H. Gestational diabetes mellitus and macrosomia: a literature review. Ann Nutr Metab 2015;66(suppl 2):14–20.
[7] Nolan CJ. Lipotoxicity, β cell dysfunction, and gestational diabetes. Cell Metab 2014;19:553–4.
[8] Świrska J, Zwolak A, Dudzińska M, Matyjaszek-Matuszek B, Paszkowski T. Gestational diabetes mellitus—literature review on selected cytokines and hormones of confirmed or possible role in its pathogenesis. Ginekol Pol 2018;89:522–7.
[9] Dias S, Adam S, Rheeder P, Louw J, Pheiffer C. Altered Genome-Wide DNA methylation in peripheral blood of South African women with gestational diabetes mellitus. Int J Mol Sci 2019;20:5828.
[10] Hjort L, Martino D, Grunnet LG, et al. Gestational diabetes and maternal obesity are associated with epigenome-wide methylation changes in children. JCI Insight 2018;3:e122572.
[11] Ciechomska M, Roszkowski L, Maslinski W. DNA methylation as a future therapeutic and diagnostic target in rheumatoid arthritis. Cells 2019;8:9.
[12] El Hajj N, Pliushch G, Schneider E, et al. Metabolic programming of MEST DNA methylation by intrauterine exposure to gestational diabetes mellitus. Diabetes 2013;62:1320–8.
[13] Franzago M, Fraticelli F, Stuppia L, Vitacolonna E. Nutrigenetics, epigenetics and gestational diabetes: consequences in mother and child. Epigenetics 2019;14:215–35.
[14] Nazari Z, Shahryari A, Ghafari S, Nabiuni M, Golalipour MJ. In utero exposure to gestational diabetes alters DNA methylation and gene expression of CDKN2A/B in langerhans islets of rat offspring. Cell J 2020;22:203–11.
[15] Strakovsky RS, Zhang X, Zhou D, Pan YX. Gestational high fat diet programs hepatic phosphoenolpyruvate carboxykinase gene expression and histone modification in neonatal offspring rats. J Physiol 2011;589 (Pt 11):2707–17.
[16] Barrett T, Wilhite SE, Ledoux P, et al. NCBI GEO: archive for functional genomics data sets—update. Nucleic Acids Res 2013;41:D991–995.
[17] Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015;43:e47.
[18] Wang L, Cao C, Ma Q, et al. RNA-seq analyses of multiple meristems of soybean: novel and alternative transcripts, evolutionary and functional implications. BMC Plant Biol 2014;14:169.
[19] Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009;4:44–57.
[20] Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res 2009;37:1–13.
[21] Szklarczyk D, Morris JH, Cook H, et al. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. Nucleic Acids Res 2017;45:D362–d368.
[22] Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003;13:2498–504.

[23] Huang Z, Shi J, Gao Y, et al. HMDD v3.0: a database for experimentally supported human microRNA-disease associations. Nucleic Acids Res 2019;47:D1013–d1017.
[24] Li JH, Liu S, Zhou H, Qu LH, Yang JH. starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. Nucleic Acids Res 2014;42:D92–97.
[25] Davis AP, Grondin CJ, Johnson RJ, et al. The comparative toxicogenomics database: update 2019. Nucleic Acids Res 2019;47:D948–d954.
[26] De Luccia TPB, Pendeloski KPT, Ono E, et al. Unveiling the pathophysiology of gestational diabetes: studies on local and peripheral immune cells. Scand J Immunol 2020;91:e12860.
[27] Zhao C, Wang F, Wang P, Ding H, Huang X, Shi Z. Early second-trimester plasma protein profiling using multiplexed isobaric tandem mass tag (TMT) labeling predicts gestational diabetes mellitus. Acta Diabetol 2015;52:1103–12.
[28] Radaelli T, Lepercq J, Varastehpour A, Basu S, Catalano PM, Mouzon SHD. Differential regulation of genes for fetoplacental lipid pathways in pregnancy with gestational and type 1 diabetes mellitus. Am J Obstet Gynecol 2009;201:209e201–10.
[29] Candas D, Li JJ. MnSOD in oxidative stress response-potential regulation via mitochondrial protein influx. Antioxid Redox Signal 2014;20:1599–617.
[30] Poggi C, Giusti B, Vestri A, Pasquini E, Abbate R, Dani C. Genetic polymorphisms of antioxidant enzymes in preterm infants. J Matern Fetal Neonatal Med 2012;25(suppl 4):131–4.
[31] Wang X, Lu J, Xie W, et al. Maternal diabetes induces autism-like behavior by hyperglycemia-mediated persistent oxidative stress and suppression of superoxide dismutase 2. Proc Natl Acad Sci U S A 2019;116:23743–52.
[32] Kohler HP, Grant PJ. Plasminogen-activator inhibitor type 1 and coronary artery disease. N Engl J Med 2000;342:1792–801.
[33] Festa A, Williams K, Tracy RP, Wagenknecht LE, Haffner SM. Progression of plasminogen activator inhibitor-1 and fibrinogen levels in relation to incident type 2 diabetes. Circulation 2006;113:1753–9.
[34] Azad N, Agrawal L, Emanuele NV, et al. Association of PAI-1 and fibrinogen with diabetic retinopathy in the Veterans Affairs Diabetes Trial (VADT). Diabetes Care 2014;37:501–6.
[35] Festa A, D'Agostino RJr, Tracy RP, Haffner SM. Insulin Resistance Atherosclerosis StudyElevated levels of acute-phase proteins and plasminogen activator inhibitor-1 predict the development of type 2 diabetes: the insulin resistance atherosclerosis study. Diabetes 2002;51:1131–7.
[36] Scheuner D, Song B, McEwen E, et al. Translational control is required for the unfolded protein response and in vivo glucose homeostasis. Mol Cell 2001;7:1165–76.
[37] Khan R, Kadamkode V, Kesharwani D, Purkayastha S, Banerjee G, Datta M. Circulatory miR-98-5p levels are deregulated during diabetes and it inhibits proliferation and promotes apoptosis by targeting PPP1R15B in keratinocytes. RNA Biol 2020;17:188–201.
[38] Harding HP, Zhang Y, Scheuner D, Chen JJ, Kaufman RJ, Ron D. Ppp1r15 gene knockout reveals an essential role for translation initiation factor 2 alpha (eIF2alpha) dephosphorylation in mammalian development. Proc Natl Acad Sci U S A 2009;106:1832–7.
[39] Sun Y, Liu WZ, Liu T, Feng X, Yang N, Zhou HF. Signaling pathway of MAPK/ERK in cell proliferation, differentiation, migration, senescence and apoptosis. J Recept Signal Transduct Res 2015;35:600–4.
[40] Chuang HC, Wang X, Tan TH. MAP4K family kinases in immunity and inflammation. Adv Immunol 2016;129:277–314.
[41] Yao Z, Zhou G, Wang XS, et al. A novel human STE20-related protein kinase, HGK, that specifically activates the c-Jun N-terminal kinase signaling pathway. J Biol Chem 1999;274:2118–25.
[42] Chuang HC, Tan TH. MAP4K family kinases and dusp family phosphatases in T-Cell signaling and systemic lupus erythematosus. Cells 2019;8:1433.
[43] Stoeckman AK, Baechler EC, Ortmann WA, Behrens TW, Michet CJ, Peterson EJ. A distinct inflammatory gene expression profile in patients with psoriatic arthritis. Genes Immun 2006;7:583–91.
[44] Chen YM, Chuang HC, Lin WC, et al. Germinal center kinase-like kinase overexpression in T cells as a novel biomarker in rheumatoid arthritis. Arthritis Rheum 2013;65:2573–82.
[45] Chen DY, Chuang HC, Lan JL, et al. Germinal center kinase-like kinase (GLK/MAP4K3) expression is increased in adult-onset Still's disease and may act as an activity marker. BMC Med 2012;10:84.
[46] Zhang Q, Long H, Liao J, et al. Inhibited expression of hematopoietic progenitor kinase 1 associated with loss of jumonji domain containing 3 promoter binding contributes to autoimmunity in systemic lupus erythematosus. J Autoimmun 2011;37:180–9.
[47] Alleboina S, Ayalew D, Peravali R, Chen L, Wong T, Dokun AO. Dual specificity phosphatase 5 regulates perfusion recovery in experimental peripheral artery disease. Vasc Med 2019;24:395–404.
[48] Rushworth LK, Kidger AM, Delavaine L, et al. Dual-specificity phosphatase 5 regulates nuclear ERK activity and suppresses skin cancer by inhibiting mutant Harvey-Ras (HRasQ61L)-driven SerpinB2 expression. Proc Natl Acad Sci U S A 2014;111:18267–72.
[49] Pramanik K, Chun CZ, Garnaas MK, et al. Dusp-5 and Snrk-1 coordinately function during vascular development and disease. Blood 2009;113:1184–91.
[50] Moon SJ, Lim MA, Park JS, et al. Dual-specificity phosphatase 5 attenuates autoimmune arthritis in mice via reciprocal regulation of the Th17/Treg cell balance and inhibition of osteoclastogenesis. Arthritis Rheumatol 2014;66:3083–95.
[51] Varshney P, Dey CS. P21-activated kinase 2 (PAK2) regulates glucose uptake and insulin sensitivity in neuronal cells. Mol Cell Endocrinol 2016;429:50–61.