Am J Transl Res 2023;15(1):336-349
www.ajtr.org /ISSN:1943-8141/AJTR0147122

Original Article

Identification of methylation-driven genes, circulating miRNAs and their potential regulatory mechanisms in gestational diabetes mellitus

Yuejun Ju[1,2], Ting Shen[2], Zhanhong Guo[2], Yinghong Kong[2], Yun Huang[1], Ji Hu[1]

[1]Department of Endocrinology, The Second Affiliated Hospital of Soochow University, Suzhou 215000, Jiangsu, P. R. China; [2]Department of Endocrinology, Changshu No. 2 People's Hospital, Changshu 215500, Jiangsu, P. R. China

Received October 20, 2022; Accepted December 6, 2022; Epub January 15, 2023; Published January 30, 2023

Abstract: Objective: Gestational diabetes mellitus (GDM) is a major pregnancy complication. The purpose of this study is to investigate the molecular regulatory mechanisms of GDM. Methods: RNA-seq and methylation data of GDM were retrieved from the Gene Expression Omnibus database. Following principal component analysis (PCA), differentially expressed mRNAs and microRNAs (miRNAs) in the blood were highlighted between GDM and the control. Then, an abnormally expressed miRNA-mRNA network was constructed, based on which a protein-protein interaction (PPI) network was established to identify hub genes. Differentially expressed and methylated genes were identified for GDM, followed by functional enrichment analysis. Results: According to PCA results, no outlier samples were found. A total of 35 differentially expressed circulating miRNAs were identified for GDM. The miRNA-mRNA regulatory network consisted of 94 miRNA-mRNA pairs. The PPI network contained 10 hub genes, including HIF1A, TLR2, FOS, IL6R, MYLIP, ABCA1, SELL, BCL3, AP1G1 and NECAP1. Furthermore, 22 down-regulated and hypermethylated genes and 8 up-regulated and hypomethylated genes were identified for GDM, which are related to helper T cell (Th) differentiation. Conclusion: We identified methylation-driven genes and circulating miRNAs for GDM, which have the potential to serve as novel diagnostic biomarkers.

Keywords: Gestational diabetes mellitus, miRNA-mRNA relationship, methylation, protein-protein interaction, hub genes

Introduction

Gestational diabetes mellitus (GDM) is a major complication in pregnant women. Epidemiological findings have shown that GDM can significantly increase the risk of adverse perinatal clinical outcomes such as dystocia and shoulder dystocia [1]. Furthermore, women with a history of GDM exhibit higher risks of type 2 diabetes mellitus (T2DM) later in life [2]. For their offsprings, GDM is also associated with an increased risk of obesity [3]. However, due to the lack of useful biological indicators for monitoring GDM during pregnancy, GDM is often not diagnosed in the early stages of pregnancy. Thus, it is important to understand the molecular changes associated with GDM.

DNA methylation regulates gene expressions without affecting gene sequences, and is involved in many biological processes, such as chromosomal instability, cellular physiological processes and genomic imprinting. There is increasing evidence that DNA methylation of cord blood status is closely related to the occurrence of GDM [4]. Cord blood DNA methylation has been linked to adverse results of GDM exposure in the uterus, such as elevated early childhood weight and obesity.

miRNAs are conserved, non-coding small RNAs (20-22 nucleotides) that may inhibit target mRNA translation. Circulating miRNAs, as important mediators of intercellular communication, can participate in myriad biological functions and may serve as excellent biomarkers for restricting circulating degradation and expression in specific tissues [5]. Increasing research has highlighted the potential of using circulating miRNAs as GDM biomarkers. For example,

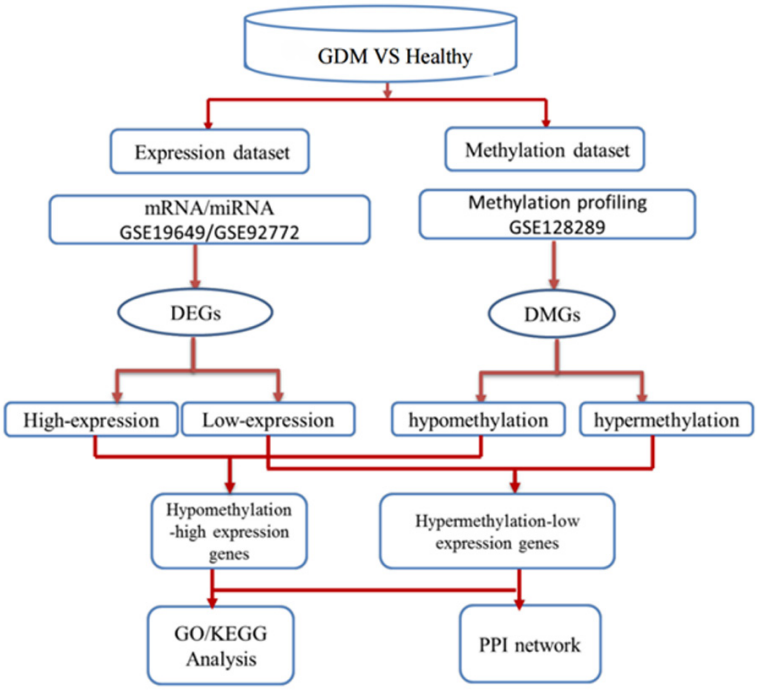


Figure 1. Research flow chart.

in early pregnancy, circulating miR-223 and miR-23a have high sensitivity and accuracy in distinguishing GDM from normal pregnant women [6].

This study comprehensively analyzed methylation and gene expression profile data to investigate DNA methylation pattern changes between GDM and healthy pregnant women and explore potential miRNA-mRNA regulatory relationships. The research process is shown in **Figure 1**.

Materials and methods

Data sources

RNA-seq data of whole blood cells of 8 women with normal glucose tolerance (NGT) and 8 GDM patients were collected from the Gene Expression Omnibus database (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE92772) [7]. The dataset was based on GPL16791 Illumina HiSeq 2500. Genomic expression profiles of blood were obtained from the GSE19649 dataset (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE19649) on GPL16791, with pooled blood of four GDM patients or four healthy women included. Furthermore, methylation profiling data of umbilical cord blood samples from 68 cases of GDM patients treated with insulin and 64 controls without GDM were downloaded from the GSE88929 dataset (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE88929) on the platform of GPL13534 Illumina HumanMethylation450 BeadChip (HumanMethylation450_15017482) [8].

Data preprocessing

For GSE92772 and GSE19649 datasets, expression information was extracted from Series Matrix File (SMF). Probe ID was converted to gene symbol. Duplicate gene symbols or miRNA IDs were averaged to obtain expression value matrices. For the GSE88929 dataset, β value information was extracted from SMF. After annotation of probe ID, methylation β value matrices were established.

Principal component analysis (PCA)

To remove outliers and find samples with high similarity, PCA was performed using genes with significantly different mean values ($P \leq 0.05$) based on mRNA, miRNA and methylation data. The correlation between gene expression levels and samples was researched. The correlation coefficient was calculated to inform potential similarities between samples. The results were visualized as heatmaps.

Differential expression and methylation analysis

The mRNA data expression matrices from the GSE92772 and GSE19649 datasets were analyzed by limma package in R [9]. mRNAs with *p*-value <0.05 and |fold change (FC)| >1.5 or |FC|>2 were selected, respectively. miRNA expression matrices from the GSE92772 dataset were used to distinguish differentially expressed miRNAs with $P<0.05$ and |FC|>2. For β value matrices of the GSE88929 methylation dataset, the ratio of the mean of the GDM group over control group was calculated

as fold changes of the differential methylation. Methylation sites with *p*-value <0.05 were identified to be different.

MiRNA-mRNA regulatory network

mRNAs and miRNAs with P<0.05 and |FC|>2 were used to predict miRNA-mRNA relationships. Down-regulated genes and up-regulated miRNAs were searched through TargetScan (v7.0; targetscan.org) [10], miRTarBase (http://miRTarBase.mbc.nctu.edu.tw/) [11], miRDB [12], miRanda (http://www.microrna.org) [13] and miRMap (http://cegg.unige.ch/mirmap) [14] databases, and used to establish a miRNA-mRNA regulatory network.

Analysis of methylation regulation relationship

Based on the differentially expressed mRNAs with |FC|>1.5 and differentially methylated sites, genes with up-regulation and low methylation and genes with down-regulation and high methylation were identified, respectively, which could be regulated by methylation.

Functional enrichment analysis

Based on the Gene ontology (GO) [15] database and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway [16] database, functional enrichment analysis was conducted on the two types of genes in that network, respectively. F-test was applied to detect significantly enriched GO and KEGG pathways.

Protein-protein interaction (PPI) network construction

Using the STRING (https://string-db.org/) online tool [17], PPI relationships for differentially expressed and methylated genes and genes in the corresponding regulatory networks were predicted. An interaction score over 0.4 was set as the threshold. Based on the obtained PPI relationships, the PPI network was visualized using the Cytoscape software [18]. Connectivity Degree analysis was then performed to identify critical nodes as hub genes in the PPI network.

Statistical methods

R project software (version 4.0.2) was used to conduct all statistical tests and produce the resulting graphics. Princomp function in PCA package was applied to perform a PCA. R package "scatterplot3d", "pheatmap" and "ggplot" were used to plot PCA, heatmaps and Volcano plots. "ClusterProfiler" package was utilized to perform GO analysis. Cytoscape 3.6.1 software was applied to perform the PPI and miRNA-mRNA regulatory network analysis.

Results

PCA results and sample correlation analysis

To remove outlier samples, we performed PCA on all samples included in this study. The results showed that no sample was removed. Correlation coefficients of gene expression levels between samples were calculated to inform potential similarities. The correlation coefficient has a positive correlation with sample similarity. The PCA and sample correlation analysis results can be found in Supplementary File 1.

Identification of differentially expressed mRNAs for GDM

Differentially expressed mRNAs were identified for GDM based on the GSE92772 and GSE19649 datasets using the Limma package. Based on the criteria of P<0.05 and |FC|>1.5, 104 up-regulated and 96 down-regulated mRNAs were identified between GDM and NGT in the GSE92772 dataset. As shown in the volcano plot, the top 5 up-regulated mRNAs (KIAA1024, SHISA7, ABCA13, MMEL1 and MMP8) and the top 5 down-regulated mRNAs (MCU12, MTUS1, SLC35F2, MIR8061 and MIR4646) were marked (**Figure 2A**). Furthermore, 601 up-regulated and 704 down-regulated genes were found for GDM in the GSE19649 dataset. **Figure 2B** shows the top 5 up-regulated mRNAs (GNLY, EOMES, LOC650263, HBG1 and HBG2) and the top 5 down-regulated mRNAs (C4BPA, SLC2A11, FOSB, FOLR3 and G0S2). These differentially expressed mRNAs enabled reliable identification of GDM blood samples from normal samples (**Figure 2C**, **2D**).

We also analyzed mRNAs with P<0.05 and |FC|>2 for GDM in the GSE92772 and GSE19649 datasets, and identified 28 up-regulated and 9 down-regulated genes between GDM and NGT in the GSE92772 dataset (**Figure 3A**). According to the volcano plot results, there appeared to be 68 up-regulated and 230 down-

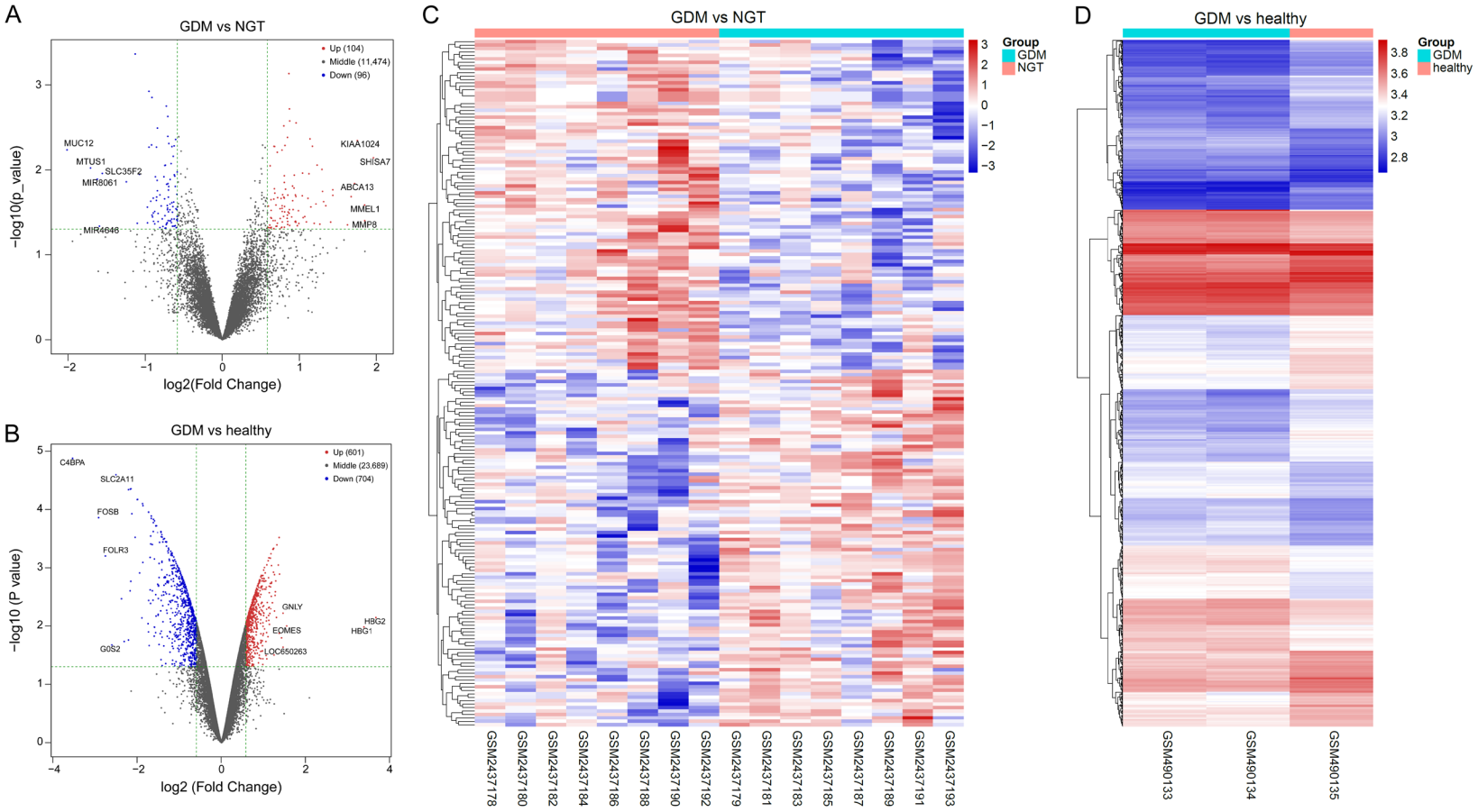


Figure 2. Identification of differentially expressed mRNAs with P<0.05 and |FC|>1.5 for gestational diabetes mellitus. Volcano plot showing all differentially expressed mRNAs in the GSE92772 dataset (A) and GSE19649 dataset (B). Hierarchical clustering analysis results for the GSE92772 dataset (C) and GSE19649 dataset (D). Red represents up-regulation, and blue represents down-regulation.

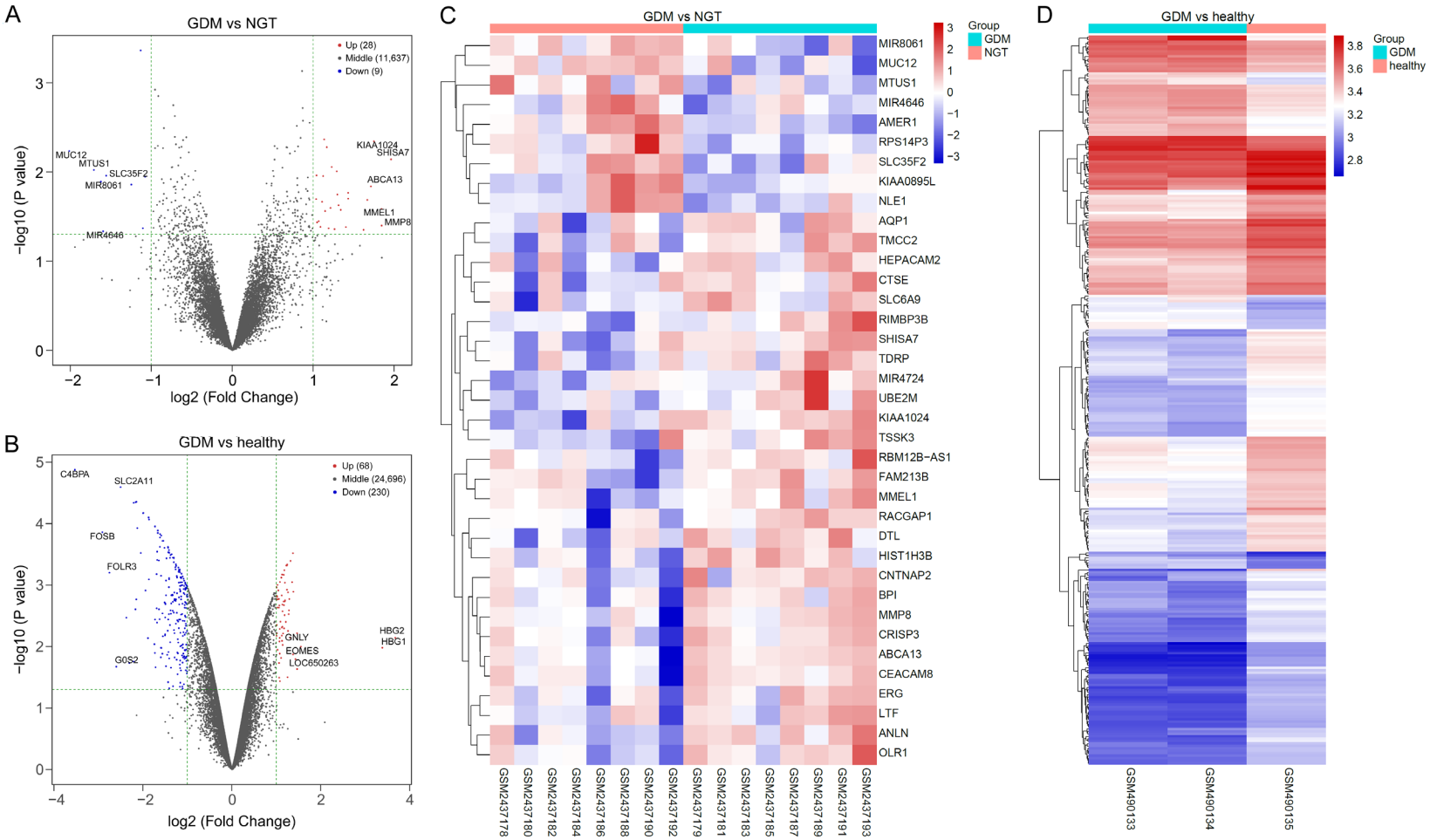


Figure 3. Identification of differentially expressed mRNAs with P<0.05 and |FC|>2 for gestational diabetes mellitus. Volcano plot showing all differentially expressed mRNAs in the GSE92772 dataset (A) and GSE19649 dataset (B). Hierarchical clustering analysis results for the GSE92772 dataset (C) and GSE19649 dataset (D).

regulated genes between GDM and healthy individuals in the GSE19649 dataset (**Figure 3B**). The expression patterns of these mRNAs between GDM and controls in the GSE92772 (**Figure 3C**) and GSE19649 (**Figure 3D**) datasets were visualized.

Identification of differentially expressed miRNAs for GDM

Using the GSE92772 dataset, 35 differentially expressed miRNAs with P<0.05 and |FC|>2 were identified between GDM and NGT, including 15 up- and 10 down-regulated miRNAs (**Figure 4A**). The top four up-regulated miRNAs were hsa-miR-19a-3p, hsa-miR-96-5p, hsa-miR-190a-5p and hsa-miR-199a-3p. The top four down-regulated miRNAs included hsa-let-7d-3p, hsa-miR-4685-3p, hsa-miR-4732-3p and hsa-miR-485-3p. As visualized in the heatmap, there were significant differences in these miRNAs between GDM and NGT (**Figure 4B**).

Identification of methylated genes for GDM

The volcano plot depicted that there were 426 hypermethylated genes and 528 hypomethylated genes between GDM and non-GDM (**Figure 5A**). As shown in the results from hierarchical clustering analysis, significant differences in methylation status could be found between GDM and non-GDM (**Figure 5B**).

Construction of miRNA-mRNA regulatory network and PPI network for GDM

Based on all mRNAs with P<0.05 and |FC|>2, down-regulated mRNAs and up-regulated miRNAs were matched in at least three databases among the TargetScan, miTarBase, miRDB, miRanda and miRMap databases. Additionally, up-regulated mRNAs and down-regulated miRNAs were also matched. We found that there were 94 miRNA-mRNA relationships in the relevant regulatory network (**Figure 6A**).

The PPI network was established for GDM, consisting of 29 nodes (**Figure 6B**). The top ten nodes were set as hub genes, including HIF1A, TLR2, FOS, IL6R, MYLIP, ABCA1, SELL, BCL3, AP1G1 and NECAP1.

Functional enrichment analysis (FEA) of mRNAs in the miRNA-mRNA regulatory network

FEA of mRNAs in the miRNA-mRNA regulatory network was carried out. For GO-BP terms, these mRNAs were enriched in response to bacteria, other organisms, external biotic stimulus, B-1 B cell homeostasis, intestinal epithelial cell maturation, protein K33-linked deubiquitination and positive regulation of cell death (**Figure 7A**). For GO-CC, these mRNAs were mainly associated with cytosol, clathrin vesicle coat, nucleus, clathrin coat, intracellular organelle, nuclear lumen, cytoskeleton, nucleoplasm, vesicle coat and RNA polymerase II transcription factor complex (**Figure 7B**). For GO-MF, these mRNAs were mainly enriched in cytoskeletal protein binding, Lys63-specific deubiquitinase activity, RNA polymerase II-specific, DNA-binding transcription activator activity, transcription regulatory specific DNA binding, and beta-tubulin binding (**Figure 7C**). In **Figure 7D**, these mRNAs were significantly associated with Th17 cell differentiation, autophagyanimal, PD-L1 expression in cancer, HIF-1, TNF signaling pathway, osteoclast differentiation, apoptosis, measles, cholesterol metabolism, and microRNAs in cancer.

Identification of epigenetically driven genes for GDM

Based on all mRNAs with P<0.05 and |FC|>1.5, we identified 22 down-regulated and hypermethylated genes and 8 up-regulated and hypomethylated genes (**Table 1**). In addition, we also identified 1 up-regulated but hypermethylated gene and 4 down-regulated but hypomethylated genes. These results are presented in Supplementary File 1.

FEA of epigenetically driven genes

We further explored the functions of epigenetically driven genes for GDM. For GO-BP, these genes were mainly enriched in T cell receptor signaling pathway, negative regulation of leukocyte activation, negative regulation, signal transduction, mesoderm formation, mesoderm morphogenesis, DNA damage response, antigen receptor-mediated signaling pathway, and cell communication in response to DNA damage (**Figure 8A**). For GO-CC, these genes were mainly enriched in membrane raft, membrane microdomain, plasma membrane protein complex, dendritic spine, neuron spine, receptor complex, cytosol, condensed chromosome and centrosome (**Figure 8B**). For GO-MF, these genes were significantly associated with elec-

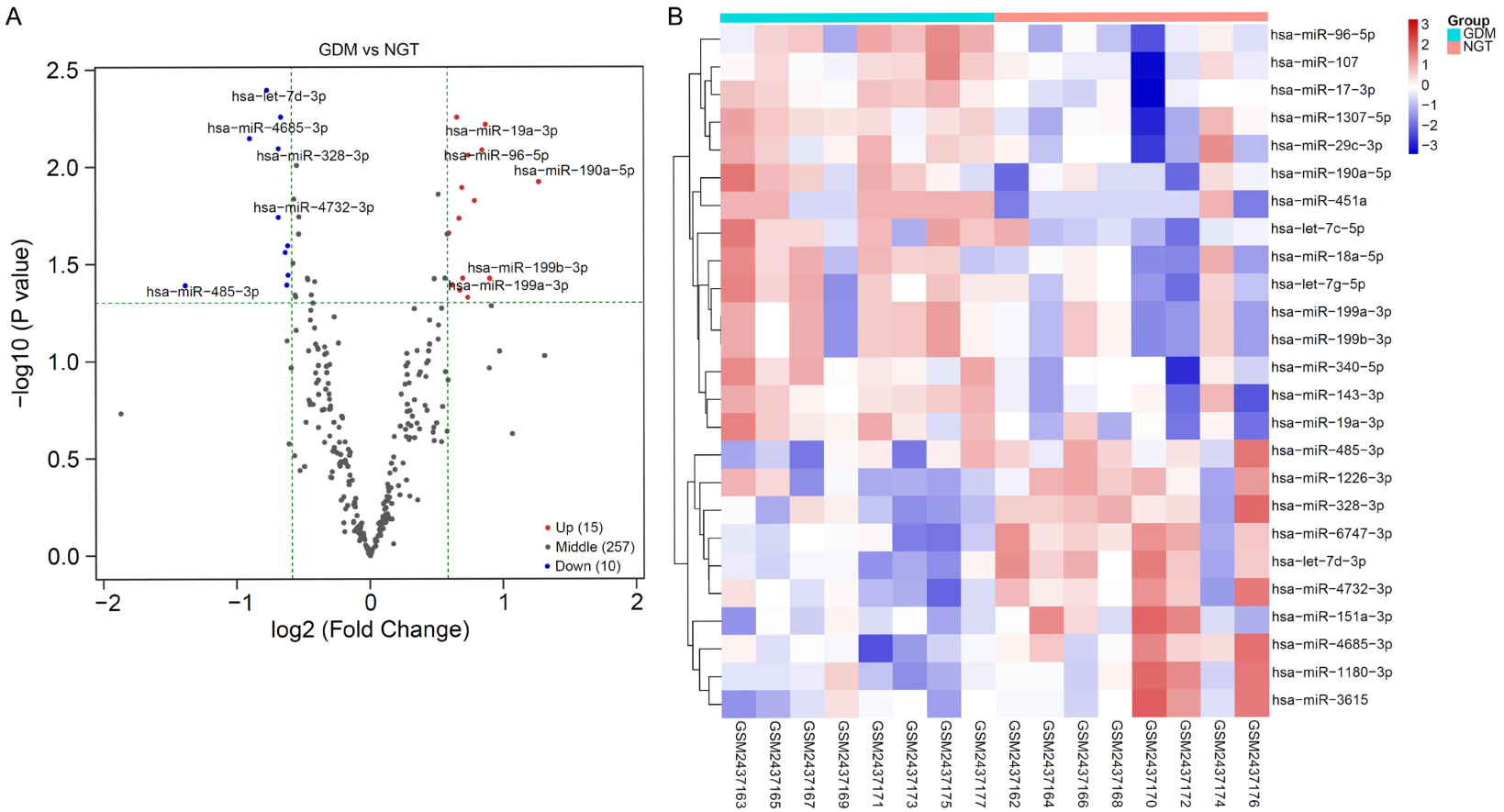


Figure 4. Identification of differentially expressed miRNAs for GDM. A. Volcano plot depicting all differentially expressed miRNAs between GDM and NGT. B. Heat map showing the differences in expression of these miRNAs between GDM and NGT. GDM: Gestational Diabetes Mellitus; NGT: Normal Glucose Tolerance.

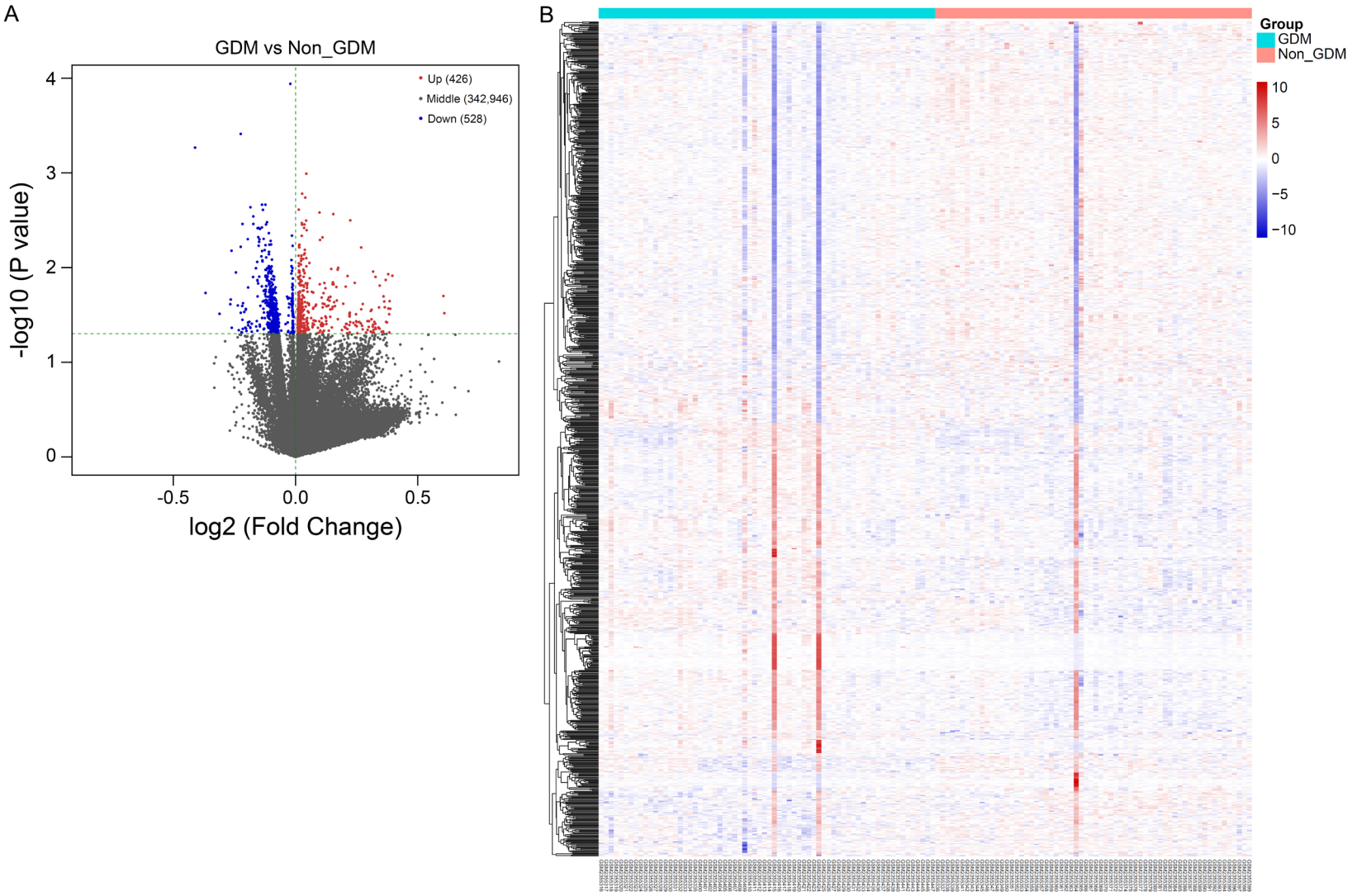


Figure 5. Identification of methylated genes for GDM. A. Volcano plot depicting all methylated genes between GDM and NGT. B. Heat map showing the differences in methylation status between GDM and NGT. Red suggests hypermethylation, and blue represents hypomethylation. GDM: Gestational Diabetes Mellitus; NGT: Normal Glucose Tolerance.

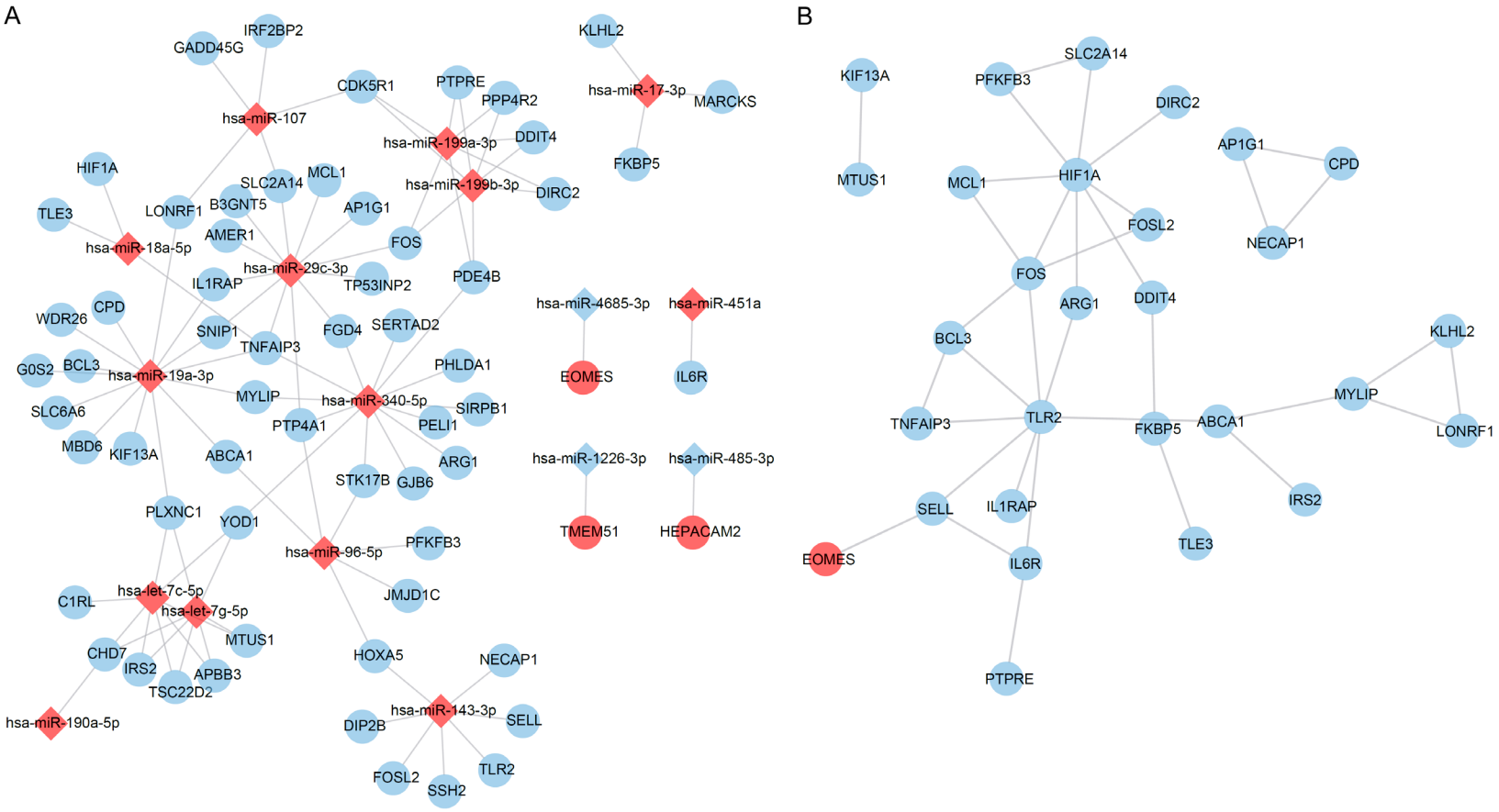


Figure 6. Construction of miRNA-mRNA regulatory network and PPI network for GDM. A. Regulatory network. Diamonds represent miRNAs and circles represent mRNAs. B. PPI network. Red represents up-regulation, and the blue represents down-regulation. GDM: Gestational Diabetes Mellitus; PPI: Protein-Protein Interaction.

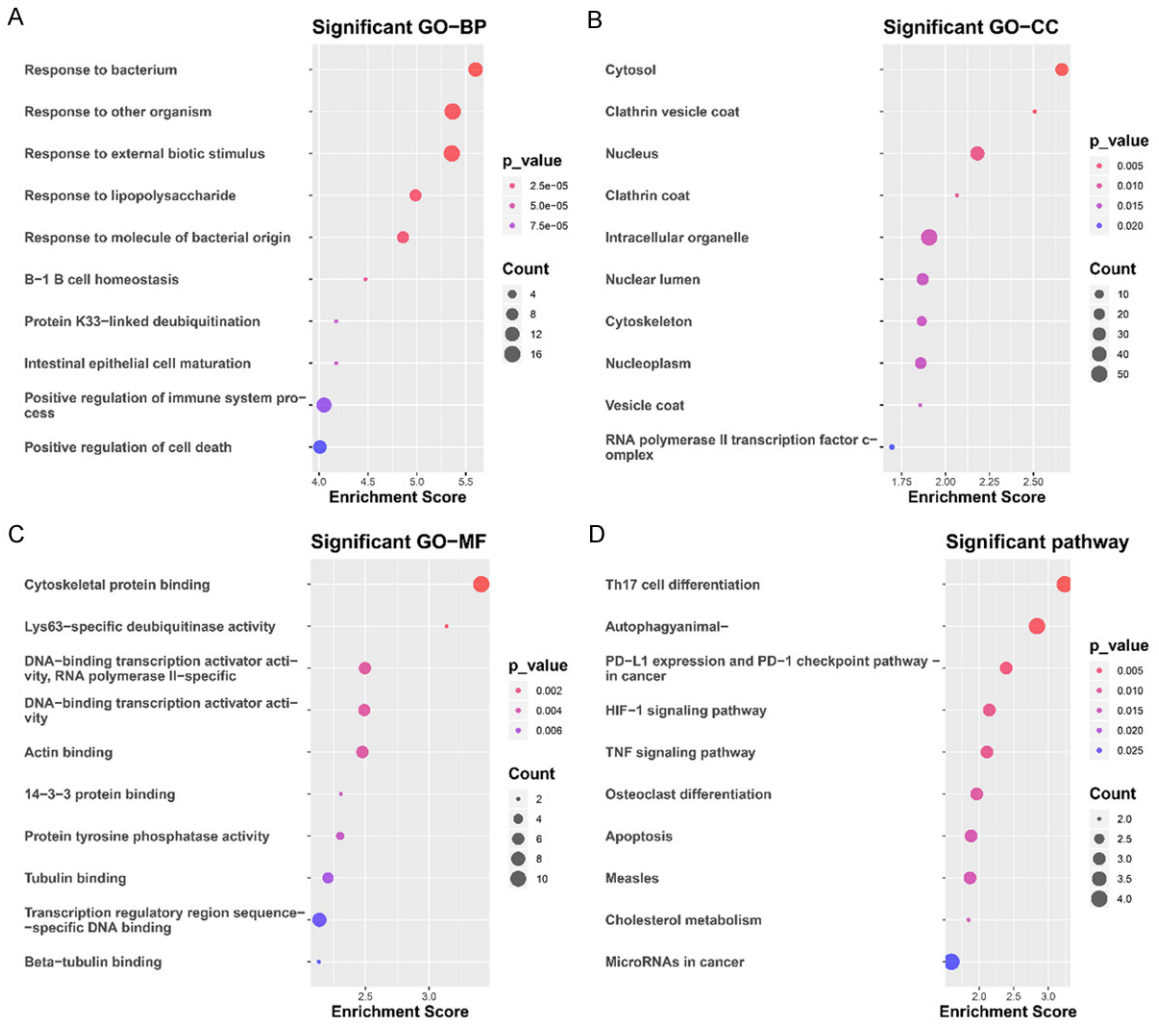


Figure 7. Functional enrichment analysis of mRNAs in the miRNA-mRNA regulatory network. A. The top ten GO-biological processes (BP); B. The top ten GO-cellular components (CC); C. The top ten GO-molecular functions (MF); D. The top ten KEGG pathways.

tron transfer activity, ion channel binding, Ras guanyl-nucleotide exchange factor activity, Ras GTPase binding, guanyl-nucleotide exchange factor activity, GTPase binding, and GTPase activator activity (**Figure 8C**). As shown in **Figure 8D**, four KEGG pathways including Th17 cell differentiation and human T-cell leukemia virus 1 infection were enriched.

Discussion

In this study, we carried out a comprehensive analysis of gene expression and methylation profiles of GDM. We identified novel methylation-driven genes and circulating miRNAs, and examined their potential regulatory mechanisms in GDM.

Based on the results of abnormal gene expression, we constructed a miRNA-mRNA regulatory network for GDM. These miRNAs may be involved in GDM pathogenesis through their target genes. Studies have found that at least 60% of human protein-coding genes are regulated by miRNAs [19]. Intriguingly, we found that hsa-miR-19a-3p, hsa-miR-96-5p and hsa-miR-199a-3p in the network had the highest fold changes of up-regulation in blood samples from GDM pregnant women compared to NGT ones. Several circulating miRNAs have been linked to hyperglycemia in patients with diabetes mellitus [20]. Therefore, our identified circulating miRNAs have the potential to serve as diagnostic biomarkers for GDM. To explore potential functions of mRNAs regulated by miR-

Table 1. Identification of key epigenetics driven genes for gestational diabetes mellitus

Genes	Regulation	Methylation status
ABR	Down-regulation	Hypermethylation
ARID3A	Down-regulation	Hypermethylation
CSNK1D	Down-regulation	Hypermethylation
FNDC3B	Down-regulation	Hypermethylation
GAS7	Down-regulation	Hypermethylation
GLTSCR1	Down-regulation	Hypermethylation
HLA-DMA	Down-regulation	Hypermethylation
IL6R	Down-regulation	Hypermethylation
KLHL2	Down-regulation	Hypermethylation
NPEPL1	Down-regulation	Hypermethylation
PAG1	Down-regulation	Hypermethylation
PDE4B	Down-regulation	Hypermethylation
POFUT2	Down-regulation	Hypermethylation
PPFIA1	Down-regulation	Hypermethylation
RBM33	Down-regulation	Hypermethylation
RERE	Down-regulation	Hypermethylation
RIN2	Down-regulation	Hypermethylation
SLC35F2	Down-regulation	Hypermethylation
TXNRD1	Down-regulation	Hypermethylation
ULK1	Down-regulation	Hypermethylation
USP10	Down-regulation	Hypermethylation
ZAK	Down-regulation	Hypermethylation
C19orf33	Up-regulation	Hypomethylation
CCDC36	Up-regulation	Hypomethylation
ERG	Up-regulation	Hypomethylation
ETFB	Up-regulation	Hypomethylation
MAD1L1	Up-regulation	Hypomethylation
VIPR1	Up-regulation	Hypomethylation
YIF1B	Up-regulation	Hypomethylation
ZAP70	Up-regulation	Hypomethylation

NAs in GDM, we performed FEA and found that these mRNAs were associated with positively regulated immune system processes and cell death [21]. Furthermore, these mRNAs were significantly associated with several GDM-related pathways such as Th17 cell differentiation [22], autophagyanimal [23], HIF-1 signaling pathway [24], TNF signaling pathway [25], apoptosis [26], and cholesterol metabolism [27, 28]. The above results indicate that these mRNAs may participate in GDM pathogenesis.

We identified 10 hub genes in the PPI network, including HIF1A, TLR2, FOS, IL6R, MYLIP, ABCA1, SELL, BCL3, AP1G1 and NECAP1. Hub genes are important mediators of diverse biological activities. For a given pathway, the expression of other genes is usually regulated by the hub gene. Given previous studies, HIF1A methylation might lead to insulin resistance in GDM [29]. Defects in HIF1A could lead to cardiovascular diseases later in life for fetuses exposed to maternal diabetes [24]. Previous studies have linked elevated TLR2 expression in peripheral blood cells and placenta to early metabolic disorders in GDM [30, 31]. GDM is also associated with hyperoxidative stress [28]. ABCA1 is elevated in the cord blood of women with GDM. Lycium barbarum polysaccharide IV may improve the lipid profile of GDM via up-regulating ABCA1 [32]. Furthermore, ABCA1 methylation in both placental and umbilical cord blood can influence the metabolic status of pregnant women [33]. Previous studies support the reliability of our findings regarding hub genes for GDM and the need to further explore the roles of other hub genes in GDM.

Epigenetic changes can occur early in the progression of diseases such as GDM, but the potential consequences of such changes remain unknown. The development of methylation sequencing technologies has facilitated the research to uncover clues related to disease progression. Here, we identified 30 abnormally expressed genes that were driven by DNA methylation. These genes appear to be associated with GDM and may participate in DNA damage. It has been found that DNA damage is associated with maternal blood glucose levels in GDM patients [34]. Furthermore, these genes showed strong correlation with T helper cell subset differentiation, such as Th1, Th2 and Th17. During pregnancy, the immune system needs to adapt in order to maintain immune tolerance to the unborn fetus. Dysregulation of immune adaptive mechanisms is associated with a variety of complications. Increasing evidence suggests that GDM is associated with the upregulation of circulating inflammatory factors [35, 36]. A prospective study reported that women with GDM had higher levels of Th17 than non-GDM women [37]. These findings implicate DNA methylation in inflammatory responses in GDM.

Although we identified potential methylation-driven genes as well as circulating miRNAs and the pathways related to GDM through bioinformatic analysis in several datasets, our study

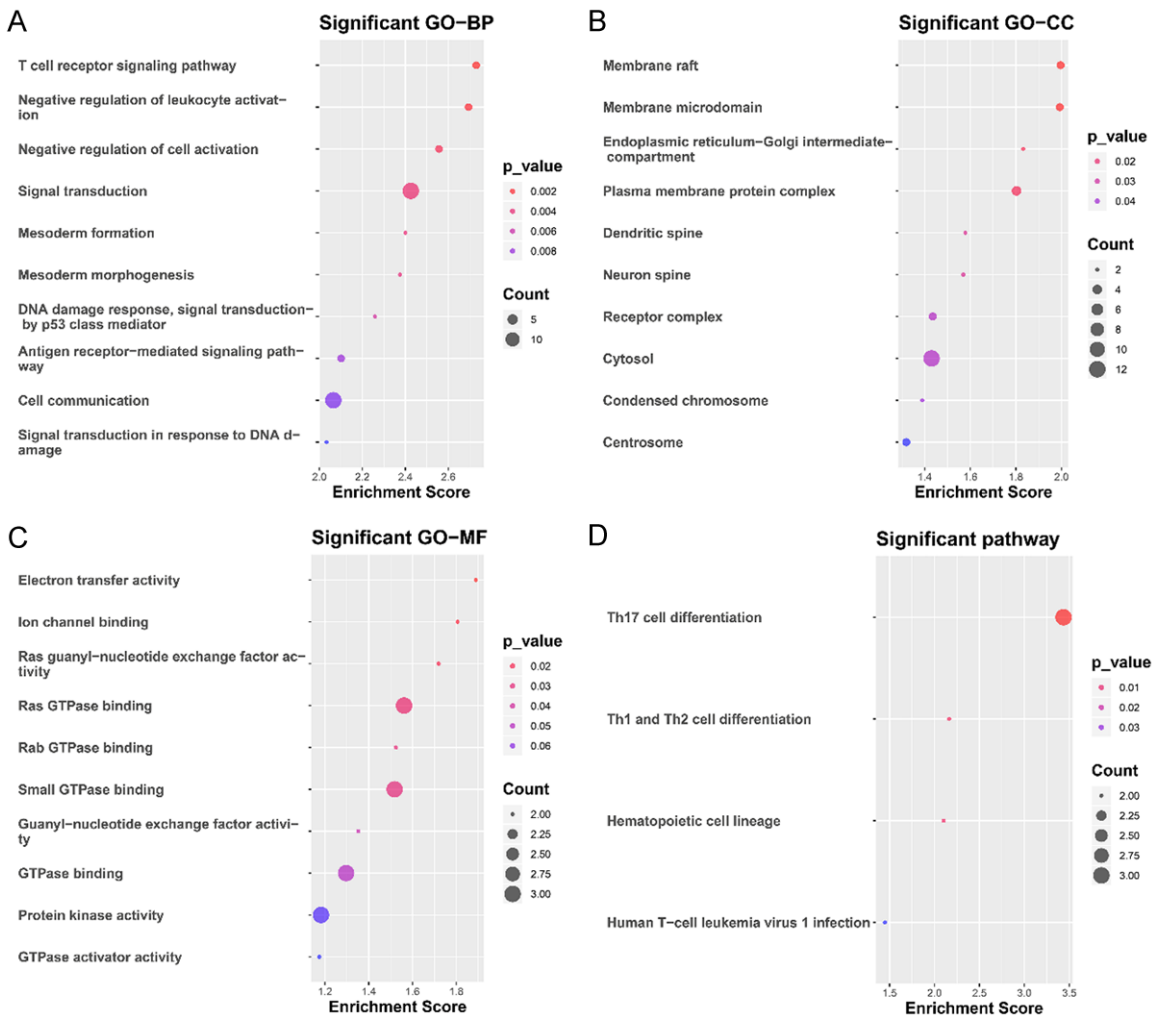


Figure 8. Functional enrichment analysis of epigenetics driven genes. A. The top ten GO-biological processes (BP); B. The top ten GO-cellular components (CC); C. The top ten GO-molecular functions (MF); D. The top ten KEGG pathways.

has several limitations. First, due to the lack of follow-up information in clinical samples, it is difficult to understand the real function of methylation driver genes, circulating miRNAs and the hub genes. We will further collect information from clinical samples to construct models in the first trimester to better evaluate the predictive value of these biomarkers on GDM. Second, we did not carry out confirmation studies to detect the methylation driver genes, circulating miRNAs and hub genes in clinical samples. Finally, because the mRNA sample size was relatively small, we used whole blood cells and blood samples to compensate and increase confidence.

In this work, we comprehensively analyzed transcriptomic gene expression and methylation profiles, identified methylation-driven genes and circulating miRNAs, and explored their potential regulatory mechanisms in GDM. Our results provide insight and reveal novel biomarkers for the diagnosis of GDM.

Acknowledgements

This work was supported by the Changshu Municipal Health Commission Key Project (CSWS202009) and Changshu Municipal Health Commission Guiding Project (CSWSZD-202110).

Disclosure of conflict of interest

None.

Address correspondence to: Ji Hu, Department of Endocrinology, The Second Affiliated Hospital of Soochow University, No. 1055, Sanxiang Road, Suzhou 215000, Jiangsu, P. R. China. E-mail: huji@suda.edu.cn; Yinghong Kong, Department of Endocrinology, Changshu No. 2 People's Hospital, No. 68, Haiyu South Road, Changshu 215500, Jiangsu, P. R. China. Tel: +86-0512-68282030; E-mail: cskyh215500@163.com

References

[1] Elliott HR, Sharp GC, Relton CL and Lawlor DA. Epigenetics and gestational diabetes: a review of epigenetic epidemiology studies and their use to explore epigenetic mediation and improve prediction. Diabetologia 2019; 62: 2171-2178.

[2] Tutino GE, Tam CH, Ozaki R, Yuen LY, So WY, Chan MH, Ko GT, Yang X, Chan JC, Tam WH and Ma RC. Long-term maternal cardiometabolic outcomes 22 years after gestational diabetes mellitus. J Diabetes Investig 2020; 11: 985-993.

[3] Damm P, Houshmand-Oeregaard A, Kelstrup L, Lauenborg J, Mathiesen ER and Clausen TD. Gestational diabetes mellitus and long-term consequences for mother and offspring: a view from Denmark. Diabetologia 2016; 59: 1396-1399.

[4] Howe CG, Cox B, Fore R, Jungius J, Kvist T, Lent S, Miles HE, Salas LA, Rifas-Shiman S, Starling AP, Yousefi P, Ladd-Acosta C, Baccarelli A, Binder EB, Chatzi VL, Czamara D, Dabelea D, DeMeo DL, Ghantous A, Herceg Z, Kajantie E, Lahti JMT, Lawlor DA, Litonjua A, Nawrot TS, Nohr EA, Oken E, Pizzi C, Plusquin M, Raikkonen K, Relton CL, Sharp GC, Sorensen TIA, Sunyer J, Vrijheid M, Zhang W, Hivert MF and Breton CV. Maternal gestational diabetes mellitus and newborn DNA methylation: findings from the pregnancy and childhood epigenetics consortium. Diabetes Care 2020; 43: 98-105.

[5] Vasu S, Kumano K, Darden CM, Rahman I, Lawrence MC and Naziruddin B. MicroRNA signatures as future biomarkers for diagnosis of diabetes states. Cells 2019; 8: 1533.

[6] Yoffe L, Polsky A, Gilam A, Raff C, Mecacci F, Ognibene A, Crispi F, Gratacos E, Kanety H, Mazaki-Tovi S, Shomron N and Hod M. Early diagnosis of gestational diabetes mellitus using circulating microRNAs. Eur J Endocrinol 2019; 181: 565-577.

[7] Stirm L, Huypens P, Sass S, Batra R, Fritsche L, Brucker S, Abele H, Hennige AM, Theis F, Beckers J, Hrabe de Angelis M, Fritsche A, Haring HU and Staiger H. Maternal whole blood cell miRNA-340 is elevated in gestational diabetes and inversely regulated by glucose and insulin. Sci Rep 2018; 8: 1366.

[8] Haertle L, El Hajj N, Dittrich M, Muller T, Nanda I, Lehnen H and Haaf T. Epigenetic signatures of gestational diabetes mellitus on cord blood methylation. Clin Epigenetics 2017; 9: 28.

[9] Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W and Smyth GK. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015; 43: e47.

[10] Agarwal V, Bell GW, Nam JW and Bartel DP. Predicting effective microRNA target sites in mammalian mRNAs. Elife 2015; 4: e05005.

[11] Chou CH, Shrestha S, Yang CD, Chang NW, Lin YL, Liao KW, Huang WC, Sun TH, Tu SJ, Lee WH, Chiew MY, Tai CS, Wei TY, Tsai TR, Huang HT, Wang CY, Wu HY, Ho SY, Chen PR, Chuang CH, Hsieh PJ, Wu YS, Chen WL, Li MJ, Wu YC, Huang XY, Ng FL, Buddhakosai W, Huang PC, Lan KC, Huang CY, Weng SL, Cheng YN, Liang C, Hsu WL and Huang HD. miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions. Nucleic Acids Res 2018; 46: D296-D302.

[12] Wong N and Wang X. miRDB: an online resource for microRNA target prediction and functional annotations. Nucleic Acids Res 2015; 43: D146-152.

[13] Betel D, Wilson M, Gabow A, Marks DS and Sander C. The microRNA.org resource: targets and expression. Nucleic Acids Res 2008; 36: D149-153.

[14] Vejnar CE and Zdobnov EM. MiRmap: comprehensive prediction of microRNA target repression strength. Nucleic Acids Res 2012; 40: 11673-11683.

[15] Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM and Sherlock G. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 2000; 25: 25-29.

[16] Kanehisa M and Goto S. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res 2000; 28: 27-30.

[17] von Mering C, Huynen M, Jaeggi D, Schmidt S, Bork P and Snel B. STRING: a database of predicted functional associations between proteins. Nucleic Acids Res 2003; 31: 258-261.

[18] Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B and Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003; 13: 2498-2504.
[19] Friedman RC, Farh KK, Burge CB and Bartel DP. Most mammalian mRNAs are conserved targets of microRNAs. Genome Res 2009; 19: 92-105.
[20] Satake E, Pezzolesi MG, Md Dom ZI, Smiles AM, Niewczas MA and Krolewski AS. Circulating miRNA profiles associated with hyperglycemia in patients with type 1 diabetes. Diabetes 2018; 67: 1013-1023.
[21] Money KM, Barke TL, Serezani A, Gannon M, Garbett KA, Aronoff DM and Mirnics K. Gestational diabetes exacerbates maternal immune activation effects in the developing brain. Mol Psychiatry 2018; 23: 1920-1928.
[22] Sifnaios E, Mastorakos G, Psarra K, Panagopoulos ND, Panoulis K, Vitoratos N, Rizos D and Creatsas G. Gestational diabetes and T-cell (Th1/Th2/Th17/Treg) immune profile. In Vivo 2019; 33: 31-40.
[23] Xu C, Chen X, Reece EA, Lu W and Yang P. The increased activity of a transcription factor inhibits autophagy in diabetic embryopathy. Am J Obstet Gynecol 2019; 220: 108.e1-108.e12.
[24] Cerychova R, Bohuslavova R, Papousek F, Sedmera D, Abaffy P, Benes V, Kolar F and Pavlinkova G. Adverse effects of Hif1a mutation and maternal diabetes on the offspring heart. Cardiovasc Diabetol 2018; 17: 68.
[25] Syngelaki A, Visser GH, Krithinakis K, Wright A and Nicolaides KH. First trimester screening for gestational diabetes mellitus by maternal factors and markers of inflammation. Metabolism 2016; 65: 131-137.
[26] Piao C, Wang X, Peng S, Guo X, Zhao H, He L, Zeng Y, Zhang F, Zhu K and Wang Y. IL-34 causes inflammation and beta cell apoptosis and dysfunction in gestational diabetes mellitus. Endocr Connect 2019; 8: 1503-1512.
[27] Miettinen HE, Rono K, Koivusalo SB, Eriksson JG and Gylling H. Effect of gestational diabetes mellitus on newborn cholesterol metabolism. Atherosclerosis 2018; 275: 346-351.
[28] Sun Y, Kopp S, Strutz J, Gali CC, Zandl-Lang M, Fanaee-Danesh E, Kirsch A, Cvitic S, Frank S, Saffery R, Bjorkhem I, Desoye G, Wadsack C and Panzenboeck U. Gestational diabetes mellitus modulates cholesterol homeostasis in human fetoplacental endothelium. Biochim Biophys Acta Mol Cell Biol Lipids 2018; 1863: 968-979.
[29] Zhang Y, Chen Y, Qu H and Wang Y. Methylation of HIF3A promoter CpG islands contributes to insulin resistance in gestational diabetes mellitus. Mol Genet Genomic Med 2019; 7: e00583.
[30] Kuzmicki M, Telejko B, Wawrusiewicz-Kurylonek N, Lipinska D, Pliszka J, Wilk J, Zielinska A, Skibicka J, Szamatowicz J, Kretowski A and Gorska M. The expression of genes involved in NF-kappaB activation in peripheral blood mononuclear cells of patients with gestational diabetes. Eur J Endocrinol 2013; 168: 419-427.
[31] Stirm L, Kovarova M, Perschbacher S, Michlmaier R, Fritsche L, Siegel-Axel D, Schleicher E, Peter A, Pauluschke-Frohlich J, Brucker S, Abele H, Wallwiener D, Preissl H, Wadsack C, Haring HU, Fritsche A, Ensenauer R, Desoye G and Staiger H. BMI-independent effects of gestational diabetes on human placenta. J Clin Endocrinol Metab 2018; 103: 3299-3309.
[32] Yang S, Si L, Fan L, Jian W, Pei H and Lin R. Polysaccharide IV from Lycium barbarum L. improves lipid profiles of gestational diabetes mellitus of pregnancy by upregulating ABCA1 and downregulating sterol regulatory element-binding transcription 1 via miR-33. Front Endocrinol (Lausanne) 2018; 9: 49.
[33] Houde AA, Guay SP, Desgagne V, Hivert MF, Baillargeon JP, St-Pierre J, Perron P, Gaudet D, Brisson D and Bouchard L. Adaptations of placental and cord blood ABCA1 DNA methylation profile to maternal metabolic status. Epigenetics 2013; 8: 1289-1302.
[34] Dong D, Yu J, Wu Y, Fu N, Villela NA and Yang P. Maternal diabetes triggers DNA damage and DNA damage response in neurulation stage embryos through oxidative stress. Biochem Biophys Res Commun 2015; 467: 407-412.
[35] Baz B, Riveline JP and Gautier JF. Endocrinology of pregnancy: gestational diabetes mellitus: definition, aetiological and clinical aspects. Eur J Endocrinol 2016; 174: R43-51.
[36] Cossu E, Incani M, Pani MG, Gattu G, Serafini C, Strazzera A, Bertoccini L, Cimini FA, Barchetta I, Cavallo MG and Baroni MG. Presence of diabetes-specific autoimmunity in women with gestational diabetes mellitus (GDM) predicts impaired glucose regulation at follow-up. J Endocrinol Invest 2018; 41: 1061-1068.
[37] Sheu A, Chan Y, Ferguson A, Bakhtyari MB, Hawke W, White C, Chan YF, Bertolino PJ, Woon HG, Palendira U, Sierro F and Lau SM. A proinflammatory CD4(+) T cell phenotype in gestational diabetes mellitus. Diabetologia 2018; 61: 1633-1643.

Supplementary File 1

Genes with up-regulation and hypermethylation and genes with down-regulation and hypomethylation

We identified 1 gene with up-regulation and hypermethylation and 3 genes with down-regulation and hypomethylation.

Supplementary Table 1. Gene with up-regulation and hypermethylation and genes with down-regulation and hypomethylation

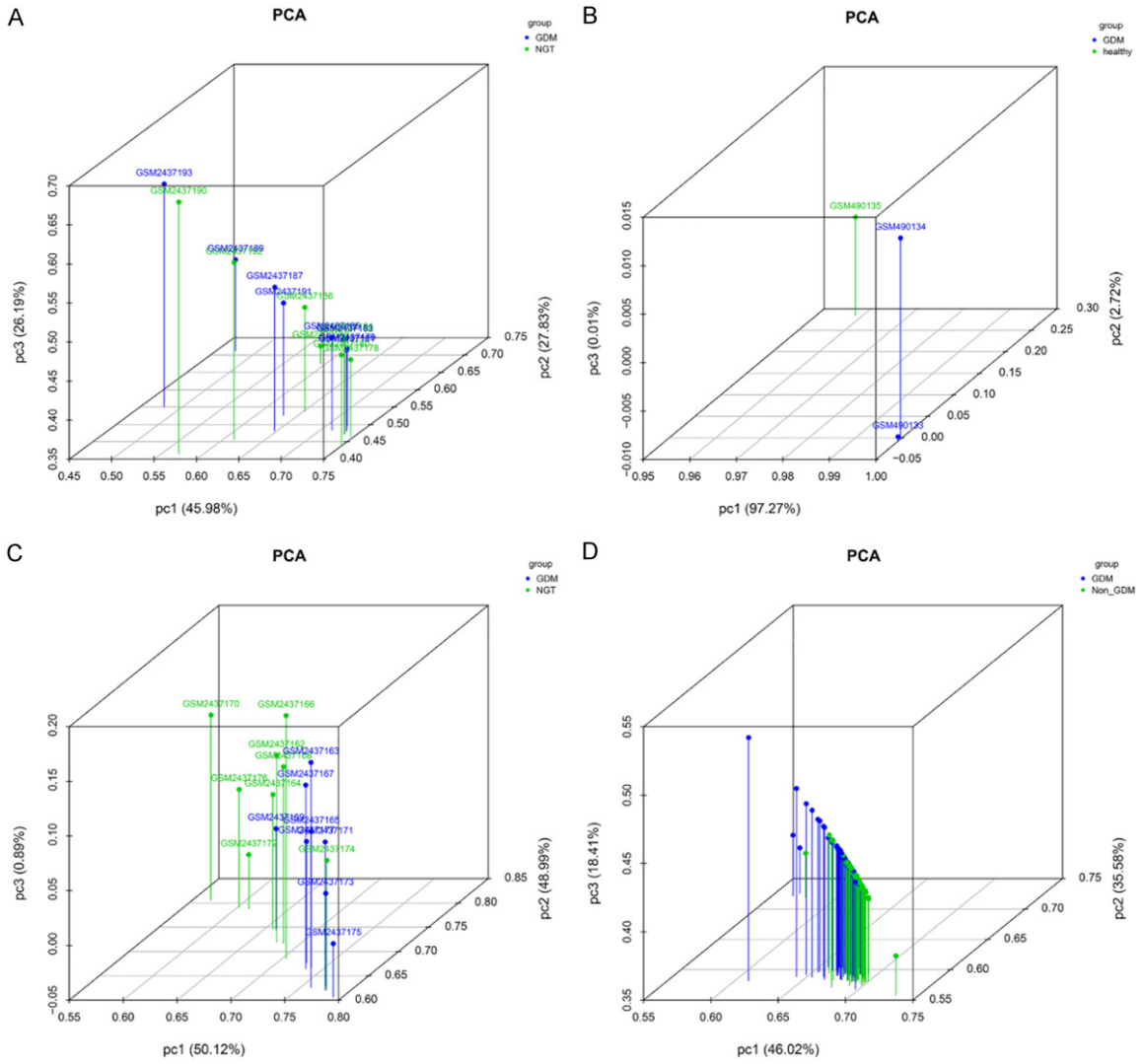
Genes	Regulation	Methylation status
CXXC5	Up-regulation	Hypermethylation
ELL	Down-regulation	Hypomethylation
SLC11A1	Down-regulation	Hypomethylation
YOD1	Down-regulation	Hypomethylation

No outlier samples according to PCA

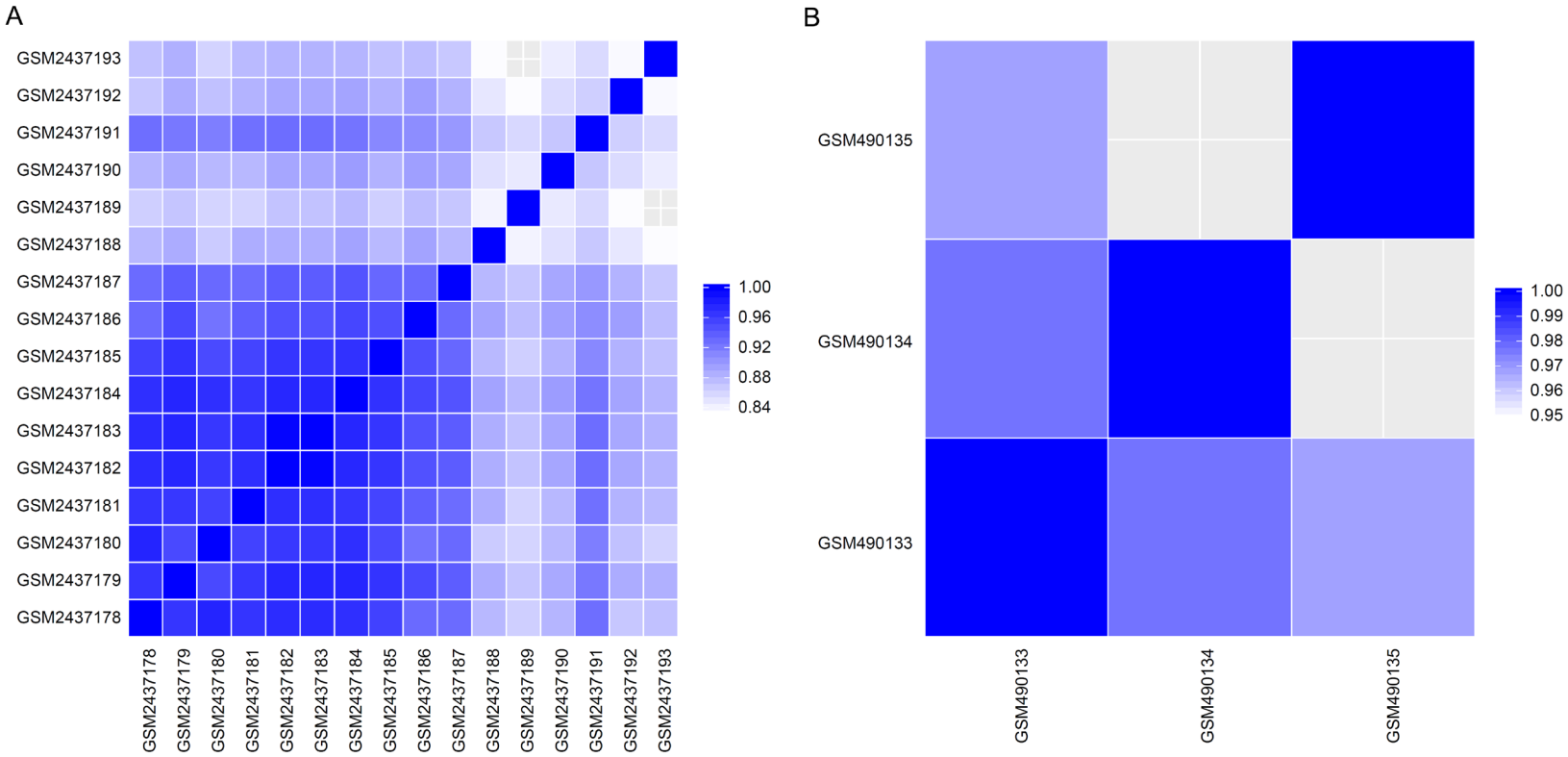
To remove outlier samples, we performed PCA of all samples included in this study. The results showed that no samples were removed. There were no outliers among sixteen whole blood samples from eight NGT and eight GDM patients in the GSE92772 mRNA dataset (Supplementary Figure 1A). PCA results of one blood sample and two pooled blood samples from GDM patients, and four healthy women are shown in Supplementary Figure 1B. Similarly, no outlier samples were found in the GSE92772 miRNA dataset (Supplementary Figure 1C). In Supplementary Figure 1D, there were no outliers among 132 umbilical cord blood samples from 68 GDM patients and 64 controls without GDM in the GSE88929 methylation dataset.

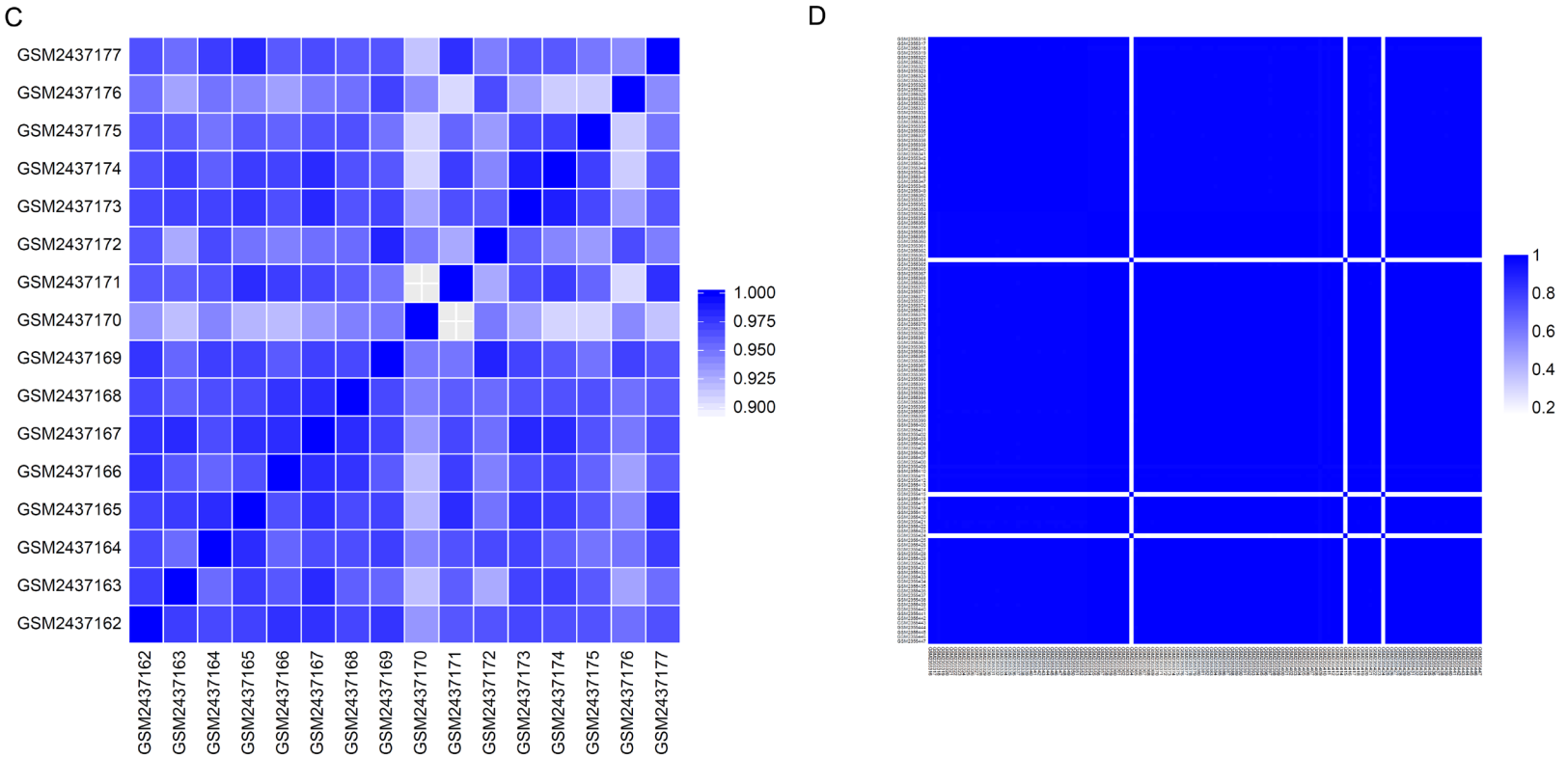
Sample correlation analysis

Correlation coefficients of gene expression levels between samples were calculated, which may represent the similarity between samples. The correlation coefficient has a positive relation with sample similarity. The results reflected that there exists strong correlation between obvious samples in the GSE92772 mRNA dataset (Supplementary Figure 2A), GSE19649 mRNA dataset (Supplementary Figure 2B), GSE92772 miRNA dataset (Supplementary Figure 2C) and GSE88929 methylation dataset (Supplementary Figure 2D).



Supplementary Figure 1. PCA results of samples included in this study. A. GSE92772 mRNA dataset; B. GSE19649 mRNA dataset; C. GSE92772 miRNA dataset; D. GSE88929 methylation dataset. Blue represents GDM samples, and the green represents normal samples.


A
GSM2437193
GSM2437192
GSM2437191
GSM2437190
GSM2437189
GSM2437188
GSM2437187
GSM2437186
GSM2437185
GSM2437184
GSM2437183
GSM2437182
GSM2437181
GSM2437180
GSM2437179
GSM2437178
1.00
0.96
0.92
0.88
0.84
B
GSM490135
GSM490134
GSM490133
1.00
0.99
0.98
0.97
0.96
0.95



Supplementary Figure 2. Heat maps showing the correlation between different samples. A. GSE92772 mRNA dataset; B. GSE19649 mRNA dataset; C. GSE92772 miRNA dataset; D. GSE88929 methylation dataset.