Analysis of micro-RNAs and gene expression profiles in gestational diabetes mellitus: A consensus approach

R. Bhushan¹, P. K. Dubey¹*

¹Department of Genetic Disorders, Institute of Science, Banaras Hindu University Varanasi 221005, Uttar Pradesh, India

Corresponding author: P. K. Dubey
E-mail: pkdubey@bhu.ac.in
Genet. Mol. Res. 20 (2): gmr23678
Received February 5, 2021
Accepted February 19, 2021
Published February 26, 2021
DOI http://dx.doi.org/10.4238/gmr23678

Copyright © 2021 The Authors. This is an open-access article distributed under the terms of the Creative Commons Attribution ShareAlike (CC BY-SA) 4.0 License.

ABSTRACT. Gestational diabetes mellitus (GDM) is a metabolic disorder characterized by insulin resistance. Lack of complete mechanisms involved in its pathophysiology makes its early diagnosis and treatment a difficult task. Recently, micro-RNAs are associated with many diseases including GDM. Its high stability in biological fluids and the ability to modulate genes at large scale makes it potent bio-markers. Here, we analyzed the transcriptomic datasets (GSE98043 and GSE19649) to gain a deeper understanding of the role of miRNAs in GDM. We processed and analyzed the microarray datasets to find differentially expressed miRNAs. Then we used a consensus approach to find the predicted as well as validated GDM related target genes. We then constructed the miRNA-mRNA gene regulatory module to have a better understanding of its regulation. These target genes were further enriched for their functions and pathways. We identified a total of 128 DE miRNAs, of which the top 20 were selected for downstream processing, and 49 validated GDM related target genes among predicted ones that may contribute to the regulatory alterations behind GDM. The micro-RNAs were linked to carbohydrate metabolism, insulin signaling, and cell proliferation and apoptosis. We then focused on miRNAs which were regulating most of
the genes related to GDM, this lead to the identification of four potential GDM miRNAs biomarkers, miR-3065-3p, miR-4650-3p, miR-29b-2-5p and miR-3915 that were significantly altered in GDM. The pathways enrichment analysis shows that they are involved in insulin signaling and pathways related to cancer. We demonstrated the most regulatory and novel miRNAs, miRNA-mRNA interactions, and their related pathways in GDM using Bioinformatics methods. Accordingly, our defined miRNAs and genes could be used for future molecular studies and can be useful in early diagnosis and treatment of GDM.

**Keywords:** Gestational Diabetes Mellitus (GDM); Micro-RNAs; Bio-markers; Genes; Gene ontology; Base population; Genetic distance

**INTRODUCTION**

Gestational diabetes mellitus (GDM) is defined as any degree of carbohydrate intolerance, with onset or first recognition during the second or third trimester of pregnancy. GDM complicates around 7% of all pregnancies while it comprises 90-95% of all cases of diabetes in pregnancy. It is a major cause of perinatal morbidity and mortality, as well as maternal long term morbidity. Major risk factors of GDM are the history of macrosomia, familiarity for type 2 diabetes (T2D), elevated maternal age, and pre-Pregnancy obesity. The pathophysiology of GDM is still not fully characterized. Gestational diabetes mellitus is a metabolic disorder characterized by insulin resistance accompanied by low/absent beta-cell compensatory adaptation to the increased insulin demands. During normal pregnancy, the mother develops insulin resistance during the second or third trimester to properly nourish the fetus. This insulin resistance is compensated by increasing insulin demands associated with hypertrophy and/or hyperplasia of β-cells. The inability of beta-cells to meet this increased demand for insulin leads to glucose intolerance and hyperglycemia that characterizes GDM. It is supposed that a cascade of molecules and several pathways are involved in these adaptive changes, thus contributing to gene expression changes necessary to beta-cells to fulfill the compensatory request. However, the application of molecular mechanisms in the diagnosis of GDM is not clear; therefore deciphering molecular mechanisms and biomarkers related to it should have a high impact on the diagnosis and treatment of GDM.

MicroRNAs (miRNAs) are endogenous ~19-24 small non-coding RNAs that modulate gene expression by inducing the translational arrest and degradation of messenger RNAs. MicroRNA is highly versatile as a single miRNA can potentially modulate multiple genes, whereas a single gene can be regulated by several miRNAs. Such complex nature of miRNAs justifies its role in virtually every cellular process, as well as in development or differentiation, regulation of cell cycle, and immune system homeostasis. Recently, several studies have reported the role of miRNAs in multiple sides of beta-cell function and differentiation, both in normal and diabetic conditions, as well as in beta-cell compensatory processes during pregnancy. Besides their classical role in negative regulation of gene expression, they have also been shown to stimulate gene expression, and also act like hormones. Furthermore, miRNAs have been reported to be present in biological fluids and are highly stable there which can be easily detected and measured. Several studies have reported the expression of microRNAs in the plasma/serum of diabetic patient; their association with the regulation of β cell mass and function and with the immune system homeostasis and certainly represent master players in the progression of this
group of chronic metabolic disorders. Therefore a deep understanding of microRNA functions and genes and pathways related to it could improve the knowledge on the etiology and pathophysiology of GDM and its complications. Furthermore, due to their high stability in body fluids and their accessibility from maternal blood throughout gestation, they could serve as biomarkers for the early diagnosis and treatment of GDM.

This study aims to find altered miRNAs in GDM and their potential validated target genes, the determination of the most important miRNAs, and their related genes and pathways in Gestational Diabetes Mellitus. Here we investigated and identified GDM related differentially expressed miRNAs (DEmiRs), validated GDM related target genes, miRNA-mRNA interactions, and signaling pathways. Our results showed 128 DEmiRNAs of which the top 20 was considered for further analysis. Target genes were predicted and a consensus-based approach leads to the identification of validated GDM related target genes. Of the 128miRNAs, miR-3065-3p, miR-4650-3p, miR-29b-2-5p, and miR-3915 are the most novel promising biomarker. Besides, Functional and pathways enrichment showed that these miRNA and their target genes have important roles in GDM and insulin metabolism.

MATERIAL AND METHODS

We have identified miRNA and their target genes in GDM as well as the Gene ontology and their signaling pathways. First, the top 20 differentially expressed miRNAs (DEmiRNAs) were selected based on their log fold change expression (highest log FC for up-miRNA and lowest logFC for down-miRNA). Target genes of the DEmiRNAs were predicted and GDM specific target genes were selected from gene cards and transcriptomic datasets GSE19649 for GDM blood. Then using the Cytoscape 3.2.1 software miRNA-mRNA regulatory module (MMRM) was constructed for up and down miRNAs separately. Furthermore, Gene Ontology and pathway enrichment analysis were performed using DAVID 7.6, significant functions and pathways involved in GDM were identified.

Microarray data

Two microarray data sets of human GDM have been selected that are available in the public repository: NCBI Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo/)14: GSE98043 and GSE19649, in which GSE98043 was used for its miRNA expression profile while GSE19649 was used for its mRNA expression profile. The dataset GSE98043, submitted by Wang et al., contains 4 chips derived from a study using GPL21575 Agilent-070156 Human miRNA V21.0 Microarray 046064. The samples used for this study were plasma from 4 pregnant women (2 from normal controls and 2 from GDM patients). Another dataset GSE19649, submitted by Zhao et al., contains a total of 5 chips three for blood and two for the placenta, derived from a study using GPL7350 Aalborg University Illumina human-6 v2.0 expression bead chip. Only blood samples namely GSM490133 (blood tissue from GDM woman), GSM490134 (pooled blood tissues from four GDM women), and GSM490135 (pooled blood tissues from four healthy women) were used in this study. The Subjects in both data sets were of Asian (Chinese) ethnicity. Both microarray studies were earlier approved by local Chinese ethics committee and participants; here we are just accessing the data from NCBI and analyzing them.

Data preprocessing and Differentially Expressed miRNA in GDM

The Series matrix file of both datasets was downloaded and processed. The probe-level symbols were converted into gene-level symbols by using GEO2R15. Analyzing GDM has been done in two groups NGT (Normal Glucose Tolerant) control and GDM patients using the GEO2R tool, to detect the differentially expressed miRNAs. Top 20 differentially expressed miRNAs have been selected, of which 10 were highly up-regulated while 10 were highly down-regulated. The selection of miRNA was based on fold change value. All miRNA selected have P-values less than 0.05.
Identification of miRNA-target gene

To identify the target genes of DEmiRNAs, two databases namely Target scan16 and miRDB17 were searched and results were downloaded. Results were tabulated and compared to find out the common target gene in both the data sets. Target score > 90 were taken as criteria for gene selection. Only those target genes which were common in both databases and have a target score > 90 were selected and uploaded in STRING18 for further analysis. Further selection and sorting of target genes were based on combined score (CS), here CS>0.9 were taken as criteria for gene selection. All genes which have TS > 90 and CS > 0.9 were selected. Target prediction and gene sorting were performed separately for up-miRNA and down-miRNA.

Construction of miRNAs and mRNAs regulatory module (MMRM)

Target genes with a combined score > 0.9 were then uploaded in Cytoscape 3.2.1.9. Protein-protein interaction network and the miRNA-mRNA interaction network was then merged to get microRNA-mRNA regulatory module (MMRM). MMRM were constructed separately for up-miRNA and down-miRNA each. The network was analyzed and edge betweenness was taken as criteria for network construction. After that, an organic layout was applied to the network. Furthermore, node size and color have been done on the miRNA-target genes network to identify miRNA, mRNA, and GDM specific genes.

Identification of potential active miRNA-target gene in GDM

To determine the validated GDM related miRNA-target genes as compared to predicted ones that were studied in this analysis; we searched a database Gene Cards and a transcriptomic data sets GSE19649. Those target genes which were common in both were identified as GDM specific miRNA-target genes.

Enriched Gene Ontology and pathway analysis

To identify the biological processes, molecular function, cellular component, and their related pathways in GDM related miRNA-target gene, DAVID 7.6 was used. Based on hypergeometric distribution, DAVID takes the genes with similar or related functions as a whole set. Significant functions were plotted against -log10 of the p-value for up and down miRNA-target genes separately. In this analysis p-value < 0.05 was set as the criterion.

GDM specific target Gene expression profile with RNA-seq

We were then interested to know the mRNA expression of all GDM related miRNA-target genes of up-miRNA and down-miRNA both. To achieve this, transcriptomic data sets GSE19649 of GDM-blood was accessed and expression value was plotted for all genes. All the genes were plotted against average gene expression and log fold change (log FC) values. Graph pad prism 7.0 was used to plot the graph.

Functional analysis of potential GDM specific target genes

The biological process network was created for GDM specific target genes of up-miRNA and down-miRNA separately. All the significant biological processes being regulated by these target genes were selected and uploaded in Cytoscape 3.2.1 for network construction.
RESULTS

Differentially expressed miRNAs in GDM

Both the data sets (one for miRNA and mRNA each) were analyzed according to the workflow Figure 1. After pre-processing and data normalization, expression profiles for both data sets were created. Based on fold change value, exclusive differentially expressed miRNAs were divided into Up- or Down-regulated miRNAs. The p-value < 0.005 and fold change ≥ 1.5 were set as the cut off values of DEmiRs. Our results showed a total of 128 differentially expressed miRNAs, 63 Up-regulated, and 65 Down-regulated miRNAs in GDM. The heat map was constructed for the DEmiRNAs. Out of these 128, we selected the top 20 differentially expressed miRNAs (10 from Up-miRs and 10 from Down-miRs). For visualizing the Differentially Expressed miRNAs, we sorted them and categorized the top 10 Up- and Down-regulated DEmiRNAs in GDM. Further, for better visualization of differential expression, a heat map was constructed for these DEmiRNAs based on their average gene expression (Figure 2A) and bar graph were plotted against its log fold change value.

Figure 1. Work flow and analysis process. Bioinformatics workflow, illustrating the databases and tools employed to reveal the molecules and interactions in the gestational diabetes mellitus (GDM)-associated gene- and microRNA (miRNA)-based regulatory networks.

Figure 2. Heat map and Log fold change expression of the twenties differentially express miRNA sets. Heat map showing the average gene expression of differentially expressed miRNAs (DEmiRNAs) among gestational diabetes mellitus (GDM) and healthy control (HC) The green to red gradation represents the gene expression values change from small to large. Graph pad prism 7 tools were used to draw a heat map.
Target genes of DEmiRNAs

Scanning of target genes from two databases Target scan and miRDB resulted in a total of 41,685 target genes for Up- and 50,831 genes for Down-regulated miRs. Target genes for each miRNA were categorized for Up- and Down-regulated miRNAs. These target genes were further selected and sorted down based on the target score and combined score. Target score > 90 and combined score > 0.9 were set as the criteria for selection. Based on the target score we get a total of 795 target genes, 370 for Up-miRNAs, and 425 for Down-regulated miRNAs. Further selection based on combined score led to the identification of a total of 162 target genes, 72 for Up-miRNAs, and 90 for Down-regulated miRNAs. The next level of sorting was done to identify GDM specific target genes (discussed in other sub-section). Different steps involved in the selection of target genes are depicted diagrammatically in Fig. 3 and number of genes sorted at each steps are depicted.

Enrichment of GDM specific target genes

In this study, we tried to select those target genes which are validated potential players in GDM. For this, we compared our computationally predicted target gene list with genes obtained from the Gene Cards database, containing experimentally validated GDM related gene and was further validated from transcriptomic data set GSE19649. Venn diagram by VENNY 2.1 tool has been drawn for common genes between Gene Cards, GSE19649, and our gene list of Up- and Down-miRs. Thus the result showed a total of 49 target genes, 22 for Up-miRNAs and 27 for Down-miRNAs as potential GDM specific target genes which are experimentally validated.

miRNA-mRNA regulatory module (MMRM) in GDM

All target genes with combined score > 0.9 along with their respective miRNAs were transferred into the Cytoscape 3.2.1, and the network was constructed separately for Up-miRs and Down-miRs. The regulatory miRNA-target genes network (MMRM) for Up-miRs included a total of 380 genes 10 Up-regulated miRNAs and 528 edges, while MMRM for Down-miRs contain total 435 genes, 10 Down-regulated miRNAs and 621 edges. We filtered the unique target genes for each miRNA and GDM specific target genes were given different shapes and colors based on expression value. The node shape (diamond-ellipse) and color (red-blue) represented the expression value of Down- and Up-regulated respectively. The final MMRM for Up-miRs included a total of 114 genes (20 GDM specific target genes), 10 Up-miRs, and 257 edges while MMRM for Down-miRs included total 166 genes (21 GDM specific target genes), 10 Down-miRs and 350 edges. Out of 20 GDM specific target genes in Up-MMRM, 13 were up-regulated while 7 were down-regulated. Similarly, in Down-MMRM, out of 21, 9 were Up-regulated while 12 were down-regulated.

Functional enrichment analysis

Gene ontology enrichment analysis for target genes involved in the MMRM network was performed and significantly enriched functions, processes, and cellular components ( p-value <0.05) were listed in (for target genes of up-regulated miRs) and (for target genes of down-miRs). Response to insulin stimulus, lipid storage, regulation of apoptosis, and cell proliferation are major significant processes being regulated by the Up-regulated miRNAs through regulation of expression of target genes involved in it. Major significant processes regulated by Down-miRNAs are regulation of fatty acid metabolism, regulation of immune response, tyrosine kinase signaling, Wnt receptor signaling, and response to carbohydrate stimulus.
KEGG pathway enrichment analysis

According to the KEGG pathway analysis of DAVID 7.6 software, we demonstrated the significant pathways for Up-miRs and Down-miRs. Significant pathways enriched for target genes of Down-miRs were long term depression and pathways in cancer while JAK-STAT is signaling. Wnt signaling, Insulin signaling, and ErbB signaling are some of the major significant pathways being regulated by target genes of Up-miRs (Fig. 6C). It should be noted that GDM specific target genes of Up-miRs and Down-miRs were enriched in GO functions and/or KEGG pathways together with their other related genes.

Investigating GDM specific target Gene expression profile with RNA-seq

In this MicroRNAs study, we also included an mRNA transcriptomic datasets of GDM-Blood to identify validated potential GDM related target genes. The comparative analysis led to the identification of 22 potential GDM specific target genes of Up-miRNAs while 27 potential GDM specific target genes of Down-miRNAs. GDM specific target genes of Up and Down-regulated miRs were categorized and tabulated with their respective miRNAs and expression value. We were then interested to know the mRNA expression of these potential GDM related target genes of Up-miRs and Down-miRs. The expression value for these genes was extracted from the series matrix file of GSE19649. Surprisingly we found that these target genes of Up- and Down-miRNAs were up-regulated as well as down-regulated. Expression values were not available for 2 target genes of Up-miRs while 6 target genes of Down-miRs. The GDM specific target genes were plotted against their fold change values for Up-miRs and Down-miRs. Further, up and down-regulated genes were categorized based on their average gene expression for their respective miRNAs. All validated GDM specific target genes (20 for up- and 21 for down-miRs) were plotted for their average gene expression. For Up-miRs, 13 genes were found to be up-regulated and 7 genes were down-regulated while for Down-miRs, 9 target genes were up-regulated and 12 target genes were down-regulated. Based on their expression, the gene regulatory network of these selected miRNAs and their related target genes were extracted from the main MMRM network. MicroRNA-mRNA gene regulatory network for Up-miR and their validated GDM specific target genes revealed 7 miRNAs namely miR-146a-5p, miR-2110, miR-567, miR-7703, miR-3065-3p, miR-4722-3p and miR-4650-3p regulating total 20 genes of which 13 were up-regulated while 7 were down-regulated. Similarly, MicroRNA-mRNA gene regulatory network for Down-miR and their validated GDM specific target genes revealed 8 miRs namely miR-146a-3p, miR-2467-3p, miR-2682-3p, miR-29b-2-5p, miR-3915, miR-4330, miR-4756-3p, miR-6739-3p regulating total 21 genes of which 9 were up-regulated while 12 were down-regulated.

Study of significant processes being regulated by GDM specific target Genes

The validated target genes for DEmiRNAs were enriched for their biological processes. Biological processes specific to validate GDM related genes of DEmiRNAs were extracted and transferred to Cytoscape, and the network was constructed. Target genes of Up-miRNAs were involved in several significant processes including response to insulin stimulus, regulation of cell-proliferation and cell-death, Tyrosine kinase signaling, oxidative phosphorylation, lipid storage, regulation of gene expression and glycoprotein biosynthesis (Fig. 10A). Similarly, target genes of Down-miRNAs were regulating some of the major processes like regulation of transcription, Wnt signaling pathway, response to carbohydrate stimulus, regulation of lipid metabolism, immune response and metal ion transport. On the basis of our findings, it may suggest that among others, miR-3065-3p, miR-4650-3p, miR-29b-2-5p, and miR-3915 might be the most promising blood-derived miRNA biomarkers in GDM.
DISCUSSION

Gestational Diabetes Mellitus is a general metabolic disorder of glucose metabolism and its pathophysiological process mostly begins before clinical diagnosis. Currently, screening and diagnosis of GDM are accomplished at 24-28 weeks of gestation, this is the period by which diabetes has already been established and presents a high risk of fetal morbidity and mortality. Further, its increasing incidences demand additional biomarkers to predict the onset and to accurately monitor the status of gestational diabetes so that early screening in the first or second trimester of pregnancy could be feasible to promptly set up an adequate therapy which normalizes blood glucose levels, thereby reducing GDM incidences and its associated adverse pregnancy outcomes. Therefore, the molecular study of GDM could have an important role in detecting biomarkers involved in the prognosis of GDM. Recently several studies have reported the diverse role of miRNA in many diseases. Due to its endocrine nature and ability to modulate gene expression, it can be considered as optimal biomarkers and sensors in GDM. However, due to the limited number of public miRNA microarray expression profiles for GDM, candidate miRNAs for GDM has not been identified yet. In the present study, we have tried to identify the blood-derived miRNAs and mRNA biomarkers as well as molecular interactions that clarify biochemical mechanisms involved in GDM. Further, we created a miRNA-mRNA based network based on which we were able to identify some miRNAs biomarkers and pathways potentially involved in GDM.

CONCLUSION

In conclusion, we presented a consensus-based method to analyze and integrate gene and miRNA expression data. Our finding suggests that the different expressions of genes and miRNAs are one of the most important variables in GDM. Hence, for a better understanding of the gene regulatory network, molecular mechanisms of GDM, developing new therapeutic approaches, future studying of miRNA function and regulation, and their potential as diagnostic biomarkers for GDM, bioinformatic analysis is required. We argued that, among others, miR-3065-3p, miR-4650-3p, miR-29b-2-5p, and miR-3915 might be the most promising blood-derived miRNA biomarkers in GDM. Although, our analysis is based on the high throughput data and is not derived in our laboratory, a large number of experimental studies confirm that the pathways and genes which were involved in GDM are supported.

REFERENCES


P. K. Dubey et al.


Genetics and Molecular Research 20 (2): gmr23678


