Differentially expressed genes identification and bioinformatics analysis of venous blood in patients with mild preeclampsia

Lingling Wang¹, Shoujun Xie², Chunli Ren³, Huachuan Zheng⁴, Xing Zhao⁵

¹²Professor, Department of Laboratory, ³Professor, Department of Obstetrics and Gynecology, ⁴Professor and Head, Department of Central Laboratory, ⁵Professor, Department of Pathology, Affiliated Hospital of Chengde Medical University, Chengde, China

ABSTRACT

Background: Preeclampsia (PE) is a syndrome characterized by hypertension (systolic blood pressure ≥140 mmHg, or diastolic blood pressure ≥90 mmHg) and proteinuria that develops after 20 weeks of gestation. It is classified as mild PE and severe PE. Placental abruption, fetal growth restriction, and fetal death are common complications of PE, which is a serious threat to maternal and infant safety during pregnancy. Bioinformatics analysis can dig into the undiscovered biological information, and have a deeper understanding of the pathogenesis of diseases. At present, the pathogenesis of PE has not been fully studied. There is no information in literature on key gene screening and bioinformatics studies of mild PE. Aims and Objectives: The aim of the study was to explore the differential genes screening and related biological process (BP) in venous blood of patients with mild PE. Materials and Methods: GSE48424 dataset was downloaded from Gene Expression Omnibus database, differential genes were screened. Gene ontology (GO) | Kyoto Encyclopedia of Gene and Genomes (KEGG) enrichment analysis was completed. The protein-protein interaction (PPI) network of differential genes was mapped using STRING and Cytoscape. Identify key modules and genes involved in mild PE. Results: A total of 433 up-regulated and 1242 down-regulated genes were obtained. GO function analysis showed that BP was mainly related to endosome organization, dephosphorylation, and endomembrane system organization. Cellular component is mainly related to promyelocytic leukemia body, nuclear membrane, and nuclear envelope. Molecular function is mainly related to ubiquitin-like protein transferase activity, phosphoric ester hydrolase activity, and phosphatase activity. KEGG showed that differential genes were enriched in sphingolipid, TNF, herpes simplex virus 1 infection, and pancreatic cancer pathway. The PPI network of differential genes was constructed to obtain 10 key genes involved in mild PE: SGMS1, SGPP1, ASAH1, PPAP2C, PPAP2B, PP1R12A, WDR82, PPP2R2A, PPP4R2, and PPP4R1. Conclusion: Using bioinformatics technology to identify the hub genes involved in mild PE can provide a favorable basis for further study of biological markers and pathogenesis of mild PE.

Key words: Mild preeclampsia; Differential genes; Venous blood; Bioinformatics

INTRODUCTION

Preeclampsia (PE) is a placental disease occurring after 20 weeks of gestation. The clinical manifestations are hypertension and proteinuria. Its incidence rate is as high as 8%,¹ which seriously threatens maternal and infant health.² According to the different clinical manifestations, mild PE (MPE) and severe PE (SPE) can be divided into two groups. PE is related to immune, genetic, biological and other factors, such as placental hypoxia, platelet aggregation, endothelial dysfunction and other vascular diseases, immune dysfunction, and oxidative stress injury are all related to PE.³⁴ In recent years, a variety of biological factors, such as soluble endothelial factor, soluble vascular endothelial growth factor receptor and placental growth factor, have been found to be related to PE, but they are not helpful
for the early diagnosis of PE. Therefore, the study of patients’ venous blood-related genes has important clinical significance for further understanding the pathogenic mechanism of PE and the early diagnosis of MPE.

Bioinformatics is an information knowledge integrated by biology, computer science, mathematics, and other disciplines. It clarifies the development mechanism of diseases from multiple perspectives and layers, and provides a new research scheme for medical researchers to conquer diseases. In this study, a Gene chip (GSE48424) containing peripheral blood of healthy pregnant women and patients with mild PE was downloaded from the Gene Expression Omnibus (GEO) database. Bioinformatics technology and the online analysis tool GE02R was used to analyze the differentially expressed genes (DEGs) related to MPE. Subsequently, gene ontology (GO) enrichment analysis and the Kyoto Encyclopedia of Gene and Genomes (KEGG) functional annotation were performed. The results of GO analysis were presented from three aspects: Biological process (BP), cellular component (CC), and molecular function (MF). GO enrichment and KEGG pathway analysis was performed using clusterProfiler (v3.14.3) and org.Hs.eg.db (v3.10.0) software package.

Aims and objectives

The study aims to explore the differential genes screening and related biological process (BP) in venous blood of patients with mild PE.

MATERIALS AND METHODS

Data set sources

“Preeclampsia” and “homo sapiens” were used to retrieve the gene data of PE patients in GEO database (https://www.ncbi.nlm.nih.gov/geo/). Finally, GSE48424 on GPL6480 platform was selected as the analysis data set, including 6 MPE cases, 13 SPE cases and 19 controls. The data of 6 MPE cases and 19 controls were used as the research data, and their general information is shown in Table 1.

Identification of differential expression genes (DEGs)

GEOR2 (https://www.ncbi.nlm.nih.gov/geo/geor2) was used to analyze the DEGs of 6 MPE patients and 19 healthy pregnant women in the GSE48424 dataset. The threshold criteria of | LogFC | >1 and p.adj< 0.05 was employed to screen DEGs in MPE. The ggplot2 software package (v3.3.3) was used to draw the DEGs volcano map.

Function and pathway enrichment of DEGs

Padj<0.0 and qvalue<0.2 were set as the threshold. GO and KEGG pathway enrichment for MPE samples were performed using clusterProfiler (v3.14.3) and org.Hs.eg.db (v3.10.0) software package.

PPI network construction and module analysis

To further understand the interaction relationship between proteins encoded by DEGs, STRING (https://cn.string-db.org/) was used to construct PPI network. The species was “Homo sapiens”. The interaction score was “high confidence (0.7)”. The MCODE plugin in Cytoscape was used to screen hub modules and genes with a score >4.5.

Expression of hub genes

The online tool GEOR2 was used to analyze the hub genes expression in GSE48424 data set in venous blood of MPE patients and healthy pregnant women.

RESULTS

Identification of DEGs

In the GSE48424 dataset, a total of 1675 DEGs of MPE were screened, among which 433 genes were upregulated and 1242 genes were down-regulated. Volcanic map is drawn, as shown in Figure 1. Red dots represent the up-regulated DEGs. Blue dots represent the down-regulated DEGs.

DEGs function enrichment

GO and KEGG enriched analysis showed there were five BP, consisting of endosome organization, dephosphorylation, endomembrane system organization, regulation of phospholipase activity, vesicle organization, three CC, consisting of promyelocytic leukemia body, nuclear membrane, nuclear envelope, four MF, consisting of phosphoric ester hydrolase activity, phosphatase activity, ubiquitin-like protein

<table>
<thead>
<tr>
<th>Table 1: General information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
</tr>
<tr>
<td>Maternal age (years)</td>
</tr>
<tr>
<td>Weight (kg)</td>
</tr>
<tr>
<td>Smoking</td>
</tr>
<tr>
<td>Previous PE</td>
</tr>
<tr>
<td>Caucasian race</td>
</tr>
<tr>
<td>Gestity/Parity</td>
</tr>
<tr>
<td>Gestational age at inclusion</td>
</tr>
<tr>
<td>Gestational age at birth</td>
</tr>
<tr>
<td>Birth weight (g)</td>
</tr>
<tr>
<td>Blood pH at birth</td>
</tr>
</tbody>
</table>

PE: Preeclampsia, MPE: Mild preeclampsia
transferase activity, ubiquitin-protein transferase activity, and 4 KEGG, consisting of Herpes simplex virus 1 infection, pancreatic cancer, sphingolipid signaling pathway, TNF signaling pathway, as shown in Table 2 and Figure 2.

PPI network construction and module analysis
The DEGs PPI network was plotted using STRING and Cytoscape, as shown in Figure 3a. MCODE plug-in was used to screen hub modules, and a total of 9 modules with a score >4.5 were obtained, among which the two modules with the highest score are shown in Figure 3b. Top 10 hub key genes were obtained: SGMS1, SGPP1, ASAH1, PPAP2C, PPAP2B, PPP1R12A, WDR82, PPP2R2A, PPP4R2, and PPP4R1, as shown in Figure 3c. The functions of the above 10 hub genes are shown in Table 3.

Hub genes expression
The differential expression of top 10 hub genes in GSE48424 data set in venous blood of MPE patients and healthy pregnant women is shown in Figure 4.

DISCUSSION
PE is a multi-system pregnancy-related complication, affecting the health of 1.5%-16.7% of pregnant women, resulting in 60,000 maternal deaths and up to 500,000 premature births globally every year. Although our understanding of PE has been improved in recent years, there are still great difficulties in the development of early biomarkers for PE. The clinical manifestations of PE are diverse, including hypertension, proteinuria, acute kidney injury, thrombocytopenia, liver dysfunction, hemolysis, liver rupture, and seizures. Risk factors include PE history, diabetes mellitus, hypertension, obesity, and multiple pregnancy. At present, the only treatment for PE is the delivery of the placenta, so early prevention and early diagnosis are crucial.

Therefore, the current research is not only focused on the treatment of PE, but also on the search for novel molecular markers for disease prediction and early diagnosis. Some studies have suggested that placental angiogenic factors such as sFlt-1 and sEng are related to PE. They can cause maternal vasoconstriction and hypertension, the activity, and function of proangiogenic placental factors such as PlGF and VEGF. Inflammation is also involved in the occurrence of PE, which activates COX, increases the level of TxA2, reduces the level of PGI2, and causes platelet aggregation and vasoconstriction. CRP rises early in maternal circulation and can be increased before the onset of PE symptoms, which is a potential biomarker for early diagnosis of PE. The placental specific enzyme phosphocholine transferase modifies neurokinin B after translation, and neurokinin B induces activation of neurokinin 3 receptor, ultimately

Table 2: GO and KEGG enrichment results

<table>
<thead>
<tr>
<th>ONTOLOGY</th>
<th>ID</th>
<th>Description</th>
<th>P</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>BP</td>
<td>GO: 0007032</td>
<td>Endosome organization</td>
<td>0.008</td>
<td>19</td>
</tr>
<tr>
<td>BP</td>
<td>GO: 0016311</td>
<td>Dephosphorylation</td>
<td>0.025</td>
<td>59</td>
</tr>
<tr>
<td>BP</td>
<td>GO: 0010256</td>
<td>Endomembrane system organization</td>
<td>0.025</td>
<td>55</td>
</tr>
<tr>
<td>BP</td>
<td>GO: 0010517</td>
<td>Regulation of phospholipase activity</td>
<td>0.025</td>
<td>16</td>
</tr>
<tr>
<td>BP</td>
<td>GO: 0016050</td>
<td>Vesicle organization</td>
<td>0.038</td>
<td>43</td>
</tr>
<tr>
<td>CC</td>
<td>GO: 0031965</td>
<td>Nuclear membrane</td>
<td>0.003</td>
<td>20</td>
</tr>
<tr>
<td>CC</td>
<td>GO: 0005635</td>
<td>Nuclear envelope</td>
<td>&lt;0.001</td>
<td>52</td>
</tr>
<tr>
<td>MF</td>
<td>GO: 0042578</td>
<td>Phosphoric ester hydrolase activity</td>
<td>0.013</td>
<td>50</td>
</tr>
<tr>
<td>MF</td>
<td>GO: 0016791</td>
<td>Phosphatase activity</td>
<td>0.033</td>
<td>38</td>
</tr>
<tr>
<td>MF</td>
<td>GO: 0019787</td>
<td>Ubiquitin-like protein transferase activity</td>
<td>0.033</td>
<td>51</td>
</tr>
<tr>
<td>MF</td>
<td>GO: 0004842</td>
<td>Ubiquitin-protein transferase activity</td>
<td>0.037</td>
<td>48</td>
</tr>
<tr>
<td>KEGG</td>
<td>hsa05168</td>
<td>Herpes simplex virus 1 infection</td>
<td>0.034</td>
<td>59</td>
</tr>
<tr>
<td>KEGG</td>
<td>hsa05212</td>
<td>Pancreatic cancer</td>
<td>0.034</td>
<td>15</td>
</tr>
<tr>
<td>KEGG</td>
<td>hsa04071</td>
<td>Sphingolipid signaling pathway</td>
<td>0.034</td>
<td>20</td>
</tr>
<tr>
<td>KEGG</td>
<td>hsa04668</td>
<td>TNF signaling pathway</td>
<td>0.034</td>
<td>19</td>
</tr>
</tbody>
</table>

PML: Promyelocytic leukemia
leading to increased CRP expression in PE patients. A study using hypersensitive testing showed increased CRP and IL-6 levels in PE patients compared with healthy pregnant women of the same age. Genes such as RPL26, RPL34, RPS15A, RPS26, GPER, and VSIG4 were found to be highly expressed in SPE by gene chip technology.

Although some biomarkers of PE and SPE have been reported, there are few studies on MPE. Because venous blood collection is relatively easy to implement in clinical practice, it is more important to study the significance of circulating biomarkers in MPE patients. The aim of our study was to identify novel circulating biomarkers for MPE. Based on the analysis of the GSE48424 dataset in GEO database, we found that there were 1675 DEGs in the circulation of MPE patients, 433 up-regulated and 1242 down-regulated compared with healthy pregnant women. DEGs GO function enrichment show: In terms of BP, it is related to inner body organization, phosphoric acid process, membrane systems, and phospholipase activity. In the aspect of CC, it is related to nucleosome, nuclear membrane, and nuclear membrane vesicles. In terms of MF, it is related to phosphate ester hydrolysis enzyme activity, phosphatase, ubiquitin-like protein transferase activity, and ubiquitin protein transferase activity. Multiple signaling pathways have been reported to be involved in the pathogenesis of PE. Low-dose aspirin can reduce hypoxia-induced sFlt-1 release from trophoblast cells and endothelial cells through the JNK/AP-1 pathway. Inhibition of mTOR pathway can reduce the invasion ability of trophoblast cells and induce the occurrence of PE. Kweider et al. found that Nrff-2/HO-1 pathway can induce the transcription of antioxidant protein genes, inhibit oxidative stress, and restore the balance between angiogenic factors and anti-angiogenic factors. In this study, KEGG pathway analysis of MPE DEGs was significantly enriched in herpes simplex virus type 1 infection, pancreatic cancer, sphingolipid pathway, and TNF pathway.

PPI network was constructed with STRING and cytoscape, and key modules were screened. 9 key modules with scores >4.5 and top 10 hub genes were obtained, which were as follows: SGMS1, SGPP1, ASA11, PPAP2C, PPAP2B, PPP1R12A, WDR82, PPP2R2A, PPP4R2, and PPP4R. Among them, SGMS1, SGPP1, ASA11, PPAP2C, and PPAP2B showed the most significant differences in expression in MPE. SGMS1 gene encodes sphingomyelin...
Table 3: Hub genes functions

<table>
<thead>
<tr>
<th>Gene symbol</th>
<th>Gene title</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>SGMS1</td>
<td>Sphingomyelin synthase 1</td>
<td>Expresses in brain predominately. The protein encoded by this gene is predicted to be a five-pass transmembrane protein</td>
</tr>
<tr>
<td>SGPP1</td>
<td>Sphingosine-1-phosphate phosphatase 1</td>
<td>Induces the degradation of sphingosine-1-phosphate</td>
</tr>
<tr>
<td>ASAH1</td>
<td>N-acylsphingosine amidohydrolase 1</td>
<td>Associates with the lysosomal storage disorder, Farber lipogranulomatosis, neuromuscular disorder, spinal muscular atrophy with progressive myoclonic epilepsy and cancer progression. After inhibiting ASAH1 activity in a pregnant mouse model, Megan et al., found that placental ceramide overload, increased trophoblast cell autophagy, and sphingolipid storage disorder led to preeclampsia</td>
</tr>
<tr>
<td>PPAP2C</td>
<td>Phospholipid phosphatase 2</td>
<td>Participates in synthesis of glycerolipids and receptor-activated signal transduction</td>
</tr>
<tr>
<td>PPAP2B</td>
<td>Phospholipid phosphatase 3</td>
<td>Hydrolyze extracellular lysosphatidic acid and short-chain phosphatic acid</td>
</tr>
<tr>
<td>PPP1R12A</td>
<td>Protein phosphatase 1 regulatory subunit 12A</td>
<td>Regulates the interaction of actin and myosin. PPP1R12A related genes were analyzed in 194 pregnant women with HDP and 262 healthy pregnant women, and significant differences were found between PE group and control group, suggesting that PPP1R12A gene may be a PE susceptibility gene</td>
</tr>
<tr>
<td>WDR82</td>
<td>WD repeat domain 82</td>
<td>Is a component of the mammalian SET1A/SET1B histone H3-Lys4 methyltransferase complexes</td>
</tr>
<tr>
<td>PPP2R2A</td>
<td>Protein phosphatase 2 regulatory subunit Balpha</td>
<td>Negative regulation of cell growth and division. PPP2R2A were decreased in mesenchymal stem cells (dMSCs) from patients with PE</td>
</tr>
<tr>
<td>PPP4R2</td>
<td>Protein phosphatase 4 regulatory subunit 2</td>
<td>Participate in organization of microtubules at centrosomes and processing of spliceosomal snRNP</td>
</tr>
<tr>
<td>PPP4R1</td>
<td>Protein phosphatase 4 regulatory subunit 1</td>
<td>Encodes one of several alternate regulatory subunits of serine/threonine protein phosphatase 4 and regulates HDAC3</td>
</tr>
</tbody>
</table>

WDR82, which is involved in the synthesis of sphingomyelin and diacylglycerol by phosphatidylecholine and ceramide, and is involved in the regulation of intracellular vesicle trafficking, cholesterol metabolism, cell proliferation and apoptosis, and other important processes. SGPP1 can promote the production of...
long-chain ceramide from sphingosine and catalyze the degradation of bioactive sphingolipid metabolite S1P.\textsuperscript{28} ASAH1 is a member of the acid ceramidase family, which catalyzes ceramide degradation and is overexpressed in some tumors. Mutations in this gene are associated with lysosomal storage disorders, lipogranulomatosis, and neuromuscular disorders.\textsuperscript{29} Phospholipid phosphatase family consists of three enzymes, named LPP1, LPP2, and LPP3, respectively. Their encoding genes are PPAP2A, PPAP2C, and PPAP2B, respectively, which play a crucial role in vascular development. Two genome-wide studies have found that PPAP2B is a new locus related to the susceptibility of coronary artery disease.\textsuperscript{30}

In addition, this study also has some limitations: (1) This study failed to perform the combined analysis of multiple GEO database chips, and only a single dataset was analyzed. However, the dataset GSE48424 covers peripheral blood information of normal pregnant women, mild PE, and severe PE, which is highly representative, and this dataset is still representative to a certain extent. (2) This study only conducted big data analysis through high-throughput sequencing results, without relevant tissue and cytological experimental verification. The relevant content will be further improved in the future to make the study results more convincing.

In summary, this study screened out DEGs, such as SGMS1, SGPP1, ASAH1, PPAP2C, and PPAP2B, and related signaling pathways of MPE, through bioinformatics data mining, which provided a strong basis for further study of the pathogenic mechanism of MPE and biomarkers screening.

**Limitations of the study**

This study only conducted big data analysis through high-throughput sequencing results, without relevant tissue and cytological experimental verification.

**CONCLUSION**

This study screened out DEGs, such as SGMS1, SGPP1, ASAH1, PPAP2C, and PPAP2B, and related signaling pathways of MPE, through bioinformatics data mining, which provided a strong basis for further study of the pathogenic mechanism of MPE and biomarkers screening.

**ACKNOWLEDGMENT**

We thank all the individuals included in this study.

**REFERENCES**

   https://doi.org/10.1016/0002-9378(88)90782-X.

   https://doi.org/10.1371/journal.pone.0082638

   https://doi.org/10.1111/j.1600-0897.2010.00831.x

   https://doi.org/10.1155/2012/586578

   https://doi.org/10.1371/journal.pone.0173711

   https://doi.org/10.1016/j.ajog.2012.08.010

   https://doi.org/10.1056/NEJMoa055352


**Figure 4:** Top 10 hub genes expression. The expression level of top 10 hub genes (SGMS1, SGPP1, ASAH1, PPAP2C, PPAP2B, PP1R12A, WDR82, PPP2R2A, and PPP4R2) in GSE48424 data set in venous blood of MPE patients and healthy pregnant women (a and b).


Authors’ Contribution:
LLW- Concept and design of the study, prepared first draft of manuscript; SJX and CLR- Interpreted the results; reviewed the literature and manuscript preparation; HZC and ZX- Concept, coordination, statistical analysis and interpretation, preparation of manuscript and revision of the manuscript.

Work attributed to:
Affiliated Hospital of Chengde Medical University, Chengde, 067000, China.

Orcid ID:
Dr. Lingjing Wang - https://orcid.org/0000-0002-9072-8948
Shoujun Xie - https://orcid.org/0000-0002-9983-5916
Chunli Ren - https://orcid.org/0000-0002-7187-1298
Huachuan Zheng - https://orcid.org/0000-0003-0965-6818
Dr. Xing Zhao - https://orcid.org/0000-0003-3179-8775

Source of Funding: S and T Program of Chengde (20180A016), Conflicts of Interest: None.