Medicine®

OPEN

Identification of underlying mechanisms and hub gene-miRNA networks of the genomic subgroups in preeclampsia development

Min Zhang, MD[a], Xiaheng Deng, MD[b], Ziyan Jiang, MD[a], Zhiping Ge, MD[a],*

Abstract

Preeclampsia is a hypertensive disorder of pregnancy that can lead to multiorgan complications in the mother and fetus. Our study aims to uncover the underlying mechanisms and hub genes between genomic subgroups of preeclampsia.

A total of 180 preeclampsia cases from 4 gene profiles were classified into 3 subgroups. Weighted gene coexpression analysis was performed to uncover the genomic characteristics associated with different clinical features. Functional annotation was executed within the significant modules and hub genes were predicted using Cytoscape software. Subsequently, miRNet analysis was performed to identify potential miRNA–mRNA networks.

Three key subgroup-specific modules were identified. Patients in subgroup II were found to develop more severe preeclampsia symptoms. Subgroup II, characterized by classical markers, was considered representative of typical preeclampsia patients. Subgroup I was considered as an early stage of preeclampsia with normal-like gene expression patterns. Moreover, subgroup III was a proinflammatory subgroup, which presented immune-related genomic characteristics. Subsequently, miR-34a-5p and miR-106a-5p were found to be correlated with all 3 significant gene modules.

This study revealed the transcriptome classification of preeclampsia cases with unique gene expression patterns. Potential hub genes and miRNAs may facilitate the identification of therapeutic targets for preeclampsia in future.

Abbreviations: KEGG = Kyoto Encyclopedia of Genes and Genomes, MAPK = mitogen activated protein kinase, PE = preeclampsia, WGCNA = weighted gene coexpression network analysis.

Keywords: hub gene, miRNA–mRNA networks, preeclampsia, subgroup, weighted correlation network analysis

1. Introduction

Preeclampsia (PE) is characterized by hypertension and multiorgan dysfunction (mainly heart, brain, and kidney) after 20 weeks of gestation. The condition may lead to long-term maternal and neonatal complications.[1] The global incidence rate of PE among pregnant women is 5% to 7%, and this pregnancy-specific disorder is responsible for over 70,000 maternal deaths and 500,000 fetal deaths per year worldwide.[2] Several risk factors for progression of PE have been identified, such as previous medical history, environmental factors, oxidative stress, genetic factors, and immunologic dyshomeostasis.[3,4] There are limited clinical interventions for PE, such as antihypertensive drugs and magnesium sulfate therapy; these are used to relieve symptoms and prolong the gestational age. However, these therapies have a minimal effect in preventing the development of PE.[5] A subset of patients experiences severe PE symptoms, leading to termination of pregnancy.[6]

Studies have shown that PE progression can be divided into 2 stages: starting with abnormal placentation and uteroplacental ischemia early in the first trimester and progressing toward excess production of antiangiogenic factors resulting in maternal syndrome.[2,7] In the first stage, inadequate spiral arteriolar remodeling may cause intermittent hypoxia and reoxygenation contributing to oxidative stress.[8] Rajakumar et al[9] found overexpression of HIF-1α and HIF-2α (markers of cellular oxygen deprivation) in proliferative trophoblasts and placental tissue of patients with PE. AP39, a mitochondrial targeting hydrogen sulfide donor, was found to decrease the expression of sFLT1 (a classic antiangiogenic factor) in placental tissues of PE patients to facilitate instability of HIF-1α.[10] In the later stage, there is accumulating evidence to suggest that PE is a proinflammatory disease with insufficient trophoblast invasion, which is characterized by imbalance of interleukin-10 and proinflammatory cytokines, leading to failure of T-cell differentiation.[11–13] The advent of microarray and high-throughput sequencing

Figure 1. The flowchart of the data preparation, processing, and analysis. GO = Gene Ontology, KEGG = Kyoto Encyclopedia of Genes and Genomes, WGCNA = weighted gene coexpression network analysis.

technologies have provided a newfangled method to further understand the genetic and molecular mechanisms of PE. Gong et al[14] conducted high-quality RNA-seq study and found abundant expression of several circ-RNAs, such as hsa-circ-0007444 in PE cases. Moreover, miR-210 was shown to inhibit the proliferation, invasion, and angiogenesis of trophoblasts through targeting mitogen-activated protein kinase (MAPK),[15] Notch homolog 1,[16] and potassium channel modulatory factor 1.[17] However, the precise etiopathogenetic mechanisms of PE are not clear and there is an immense need for identifying high risk groups and discovering novel therapeutic targets.

Here, we acquired data from 4 datasets (GSE10588, GSE25906, GSE60438 and GSE75010) from the Gene Expression Omnibus (GEO). We then divided PE patients into 3 subgroups and constructed a weighted coexpression network analysis of subgroup-specific differentially expressed genes to investigate the genes associated with maternal and fetal complications. Significant gene coexpression modules and genes related to clinical characteristics were identified. Furthermore, the brown, turquoise, and blue module gene network was constructed and hub genes were corroborated using Cytoscape. Finally, miRNAs related to these module hub genes were detected through miRNet, which may serve as biomarkers and therapeutic targets of PE.

2. Materials and Methods

2.1. Data preprocessing

Figure 1 illustrates the flowchart of the data preparation, processing and analysis. As shown in Table 1, 4 datasets (GSE10588,

GSE25906, GSE60438, and GSE75010) were downloaded from the GEO database (https://www.ncbi.nlm.nih.gov/geo/). Batch effects are subgroups of measurements that exhibit qualitatively different behavior across conditions and are unrelated to the biological or scientific variables in a study.[18] The R package "sva" was used to remove the batch effect among these platforms and the principal component analysis was conducted to evaluate the efficiency of batch-effect removal. Given that this study was based on the public database without human or animal experiments, ethical approval and patient consent were not needed.

2.2. Verification of subgroup-specific genes and clinical characteristics among 3 subgroups

Samples were classified into different subgroups using consensus clustering, which was performed by K-means algorithm with the

Table 1

Summary data of GEO database.

Dataset ID	Platform	No preeclampsia	Preeclampsia
GSE10588	GPL2986	50	21
GSE25906	GPL6102	37	23
GSE60438	GPL6884	65	60
	GPL10558		
GSE85307	GPL6244	101	47
	GPL19162		
GSE75010	GPL6244	173	157

GEO = Gene Expression Omnibus.

spearman distance. The maximum cluster number was 10 and the final cluster number was decided by the consensus matrix and cluster consensus score. Then, Wilcoxon sum-rank text was used to verify the differential expression of subgroup-specific genes among cases of each subgroup with the thresholds of Benjamini–Hochberg adjusted $P < .05$. The R package "ggpubr" was used to assess differences of maternal age, body mass index (BMI), gestational age, and infant z-score[19] between the subgroups.

2.3. Weighted gene coexpression network analysis construction

We used R package termed "weighted gene coexpression network analysis" ("WGCNA")[20] to construct the coexpression network of the filtered genes, which could characterize the biological function of each group. In this study, the soft-thresholding power was set to β = 5 (scale free R^2 = 0.9). The topological overlap matrix[21] was used to cluster the genes by analyzing the average linkage and Pearson correlation. Next, we constructed a cluster tree with the following major parameters (cutHeight = 0.4; minimum module size = 55; deepSplit value = 2) and identified functional modules using hierarchical clustering analysis. Finally, the correlation between genes in functional modules and clinical characteristics in the WGCNA package were assessed using Spearman correlation coefficient; P values of < .05 were considered indicative of statistical significance.

2.4. GO and KEGG analysis of Module genes

Gene set enrichment analysis (GSEA) was performed using GSEA software (version 4.1.0). The Gene Ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were performed with the module genes. The R package "clusterProfiler" was used to analyze GO and KEGG enrichment.[22] Terms with false discovery rate of <0.05 and P value of <.05 were considered for subsequent analysis and the dotplot was processed using the R package "ggplot2." The R package "pheatmap" was performed to create the heatmap of subgroups with KEGG significant terms.

2.5. Hub gene identification and module miRNA–mRNA network construction

Using Maximal Clique Centrality methods of cytoHubba package, Cytoscape software (version 3.8.0) was utilized to rank the candidate genes and the top 10 nodes were selected as the hub genes. By potentially affecting many biological processes (BPs), miRNAs, with potent posttranscriptional regulatory functions, are closely related to the occurrence and development of diseases. Each miRNA would target hundreds of different genes and multiple miRNAs could regulate their target genes cooperatively. Through miRNet analysis (www.mirnet.ca) in human tissue PLACENTA, based on 11 well-annotated miRNA databases, we could build miRNA-target interaction network and rank the miRNAs by number of interactions, which are most relevant to the hub genes studied.[23] Networks were visualized using Cytoscape software.

2.6. Statistical analysis

R software (v4.0.3) was used for statistical analysis. The correlation between mutant genes and TMB was determined by the Mann–Whitney U test. The enriched functions associated with the differentially expressed genes were determined by the R package "clusterProfiler." Two-tailed $P < .05$ was considered statistically significant for all comparisons.

3. Results

3.1. Clinic characteristics of PE subgroups

A total of 385 samples (180 PE subjects and 205 no-PE cases) from 4 independent GEO datasets (GSE10588,[24] GSE25906,[25] GSE60438,[26] and GSE75010[27]) were analyzed to predict the PE subgroups by evaluation of gene signatures. Principal component analysis method was performed to remove the batch effect among different platforms and datasets (Fig. 2A). The scatterplot revealed successful removal of the batch effect (Fig. 2B). Then, 180 PE cases were classified into 3 subgroups by consensus clustering analysis of gene expression profiles. The minimal cluster consensus score was >0.6 and the cluster count ranged from 2 to 10 (Fig. 2C). There were 77, 88, and 15 cases in subgroups I, II, and III, respectively. The consensus matrix revealed a high similarity of gene expression (Fig. 2D).

Next, we analyzed the differences between the 3 subgroups with respect to maternal age, BMI, gestational age, and infant z-score. We discovered that the patients in subgroup II were older and their gestational age was shorter than those in subgroups I and III (Fig. 3A, B); this indicated that subgroup II may develop more severe PE. However, no significant differences were observed with respect to maternal BMI or infant z-score among the 3 subgroups (Fig. 3C, D). We further clarified the relationship of subgroup with maternal age and gestational age using analysis of variance; the results indicated that the PE subgroups were a maternal age-independent indicator of gestational age (Table 2; $P = .021$).

3.2. Construction of weighted gene coexpression network and modules detection for each subgroup

A total of 2845 subgroup-specific genes (991, 376, and 1478 genes in subgroup I, II, and III, respectively) were obtained by pairwise differential expression analysis across subgroups (Benjamini–Hochberg adjusted $P < .05$; mean difference threshold = 0.2; Table 3; Table S1, Supplemental Digital Content, http://links.lww.com/MD/G938). We further performed differential expression analysis to identify the upregulated genes in each subgroup. GSEA indicated that specific subgroup genes were significantly expressed when compared with normal samples (false discovery rate <0.05; Fig. 4A–C). In addition, subgroup III contained the largest number of subgroup-specific genes and differentially expressed genes compared with normal samples; considering that these patients had moderate PE symptoms, these findings indicated that specific categories of functional genes in subgroup III may have a protective effect against PE.

The R package "WGCNA," a well-known data mining tool, was utilized to construct a weighted coexpression network of these subgroup-specific genes. No outliers or strong clusters were observed; therefore, the selected samples were well clustered in hierarchical cluster analysis. The soft-thresholding power and soft threshold 5 were considered as the optimal power with a scale-free topology fix index 0.90 (Fig. 5A, B). Under specific conditions, the "blue" and "yellow" modules and the "brown" and "red" modules were merged (Fig. 5C, D). Thus, 5 gene coexpression modules were identified in the WGCNA analysis. The correlation of these 5 modules with PE clinical characteristics was assessed to identify the module that showed the most significant level of correlation (Fig. 5E). The module "turquoise" showed the strongest positive correlation with infant z-score and gestational age, while module "brown" showed a significant negative correlation with these characteristics. Module "blue" showed a positive correlation with gestational age and infant z-score, but showed no relationship with maternal age. Other modules, such as green and gray, showed no significant association with maternal age, gestational age, or infant z-score.
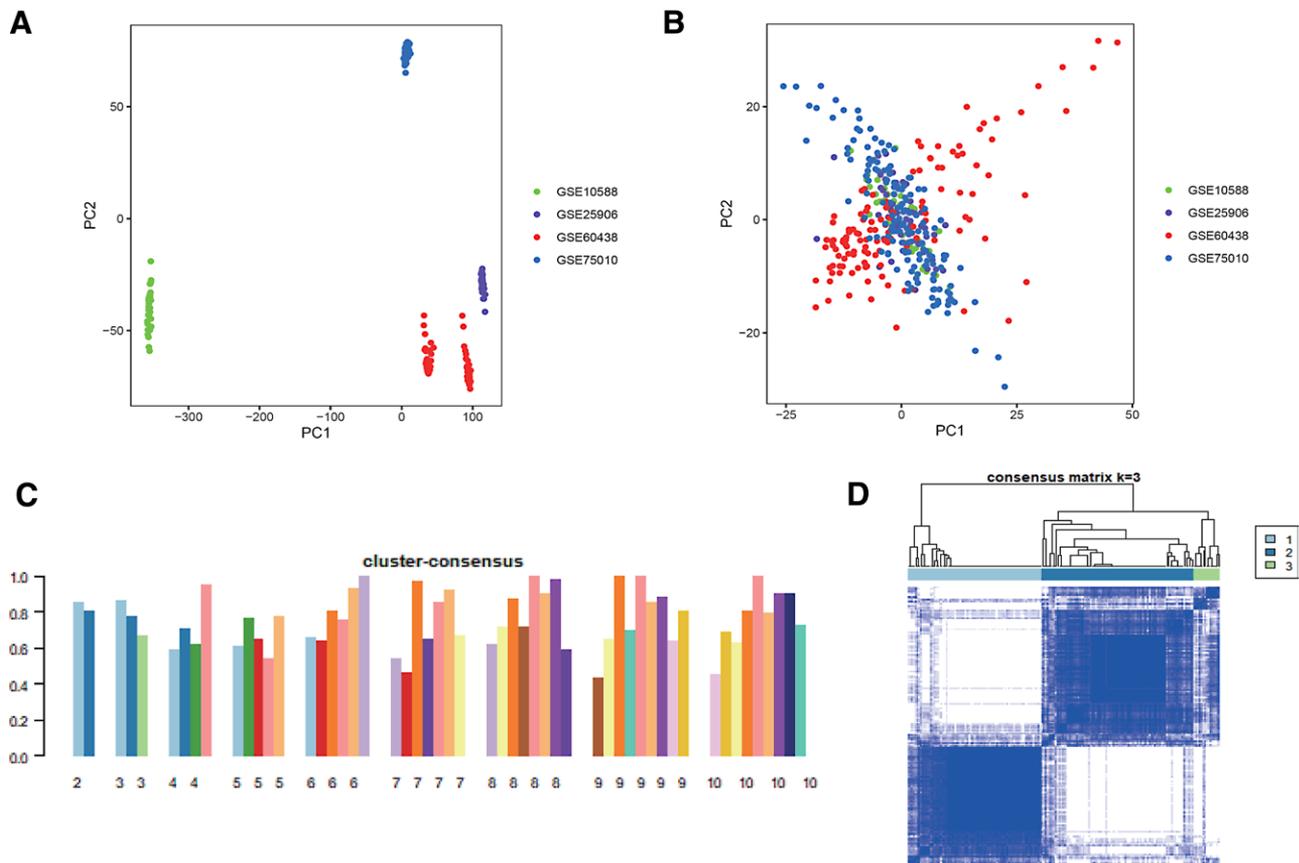
Figure 2. PCA and consensus clustering analysis of the gene expression profiles of PE cases. The colors indicate the samples from 4 different datasets. Each dot in the scatter plots represents the samples on the top 2 principal components (PC1 and PC2) before (A) and after (B) batch-effect removal of gene expression profiles. (C) The bar-plots visualize the consensus scores for subgroup with cluster count ranging from 2 to 10. (D) The heatmap shows the consensus matrix with cluster count of 3, which was determined by the minimal consensus scores for subgroup (>0.6). PC = principal component, PCA = principal component analysis, PE = preeclampsia.

3.3. Module function enrichment analysis

We used GO analysis (including cell component, BP, and molecular function) to perform enrichment analysis of the biological significance on the 5 WGCNA modules (Fig. 6A–D). In BP group (Fig. 6A, B; Table S2, Supplemental Digital Content, http://links.lww.com/MD/G939), module blue was mainly enriched in neutrophil activation involved in immune response, including neutrophil activation involved in immune response, neutrophil degranulation, T-cell activation, lymphocyte proliferation, and mononuclear cell proliferation. Module brown showed a significant association with multi–multicellular organism process, covering cell–substrate adhesion, response to metal ion, tissue migration, multi–multicellular organism process, female pregnancy, reproductive structure development, and reproductive system development. Module turquoise was mainly associated with extracellular structure organization, extracellular matrix organization and cell-substrate adhesion. In cell component group (Fig. 6A, C; Table S3, Supplemental Digital Content, http://links.lww.com/MD/G940), module blue was enriched in specific granule, secretory granule membrane, tertiary granule, specific granule membrane, and external side of plasma membrane. Module brown was mainly associated with membrane microdomain, membrane region, and membrane raft. Module turquoise showed a significant association with focal adhesion, cell–substrate junction, collagen-containing extracellular matrix, and endoplasmic reticulum lumen. In molecular function group (Fig 6A, D; Table S4, Supplemental Digital Content, http://links.lww.

com/MD/G941), genes of module blue were mainly associated with immune receptor activity, carbohydrate binding, protein serine/threonine kinase activity, and signaling receptor activator activity. Module brown was enriched in protein serine/threonine kinase activity and signaling receptor activator activity. Module turquoise was mainly related to extracellular matrix structural constituent.

The results of KEGG analysis of the 5 WGCNA modules are shown in Figure 6E and F and Table S5 (Supplemental Digital Content, http://links.lww.com/MD/G942). Module blue was enriched in chemokine signaling pathway, osteoclast differentiation, hematopoietic cell lineage, natural killer cell-mediated cytotoxicity, viral protein interaction with cytokine and cytokine receptor, MAPK signaling pathway, and Rap1 signaling pathway. Module brown was enriched in MAPK signaling pathway, Rap1 signaling pathway, proteoglycans in cancer, focal adhesion, cAMP signaling pathway, and breast cancer. Module turquoise was enriched in focal adhesion, MAPK signaling pathway, Rap1 signaling pathway, proteoglycans in cancer, vascular smooth muscle contraction, and cGMP-PKG signaling pathway. Notably, MAPK signaling pathway and Rap1 signaling pathway were both enriched in these 3 key modules.

3.4. Hub gene analysis and prediction of potential miRNA–mRNA networks

According to the relatedness, as assessed by Cytohubb package, Figure 7A to C shows the specific top 10 hub genes in modules blue, brown, and turquoise from strong to weak. The 10
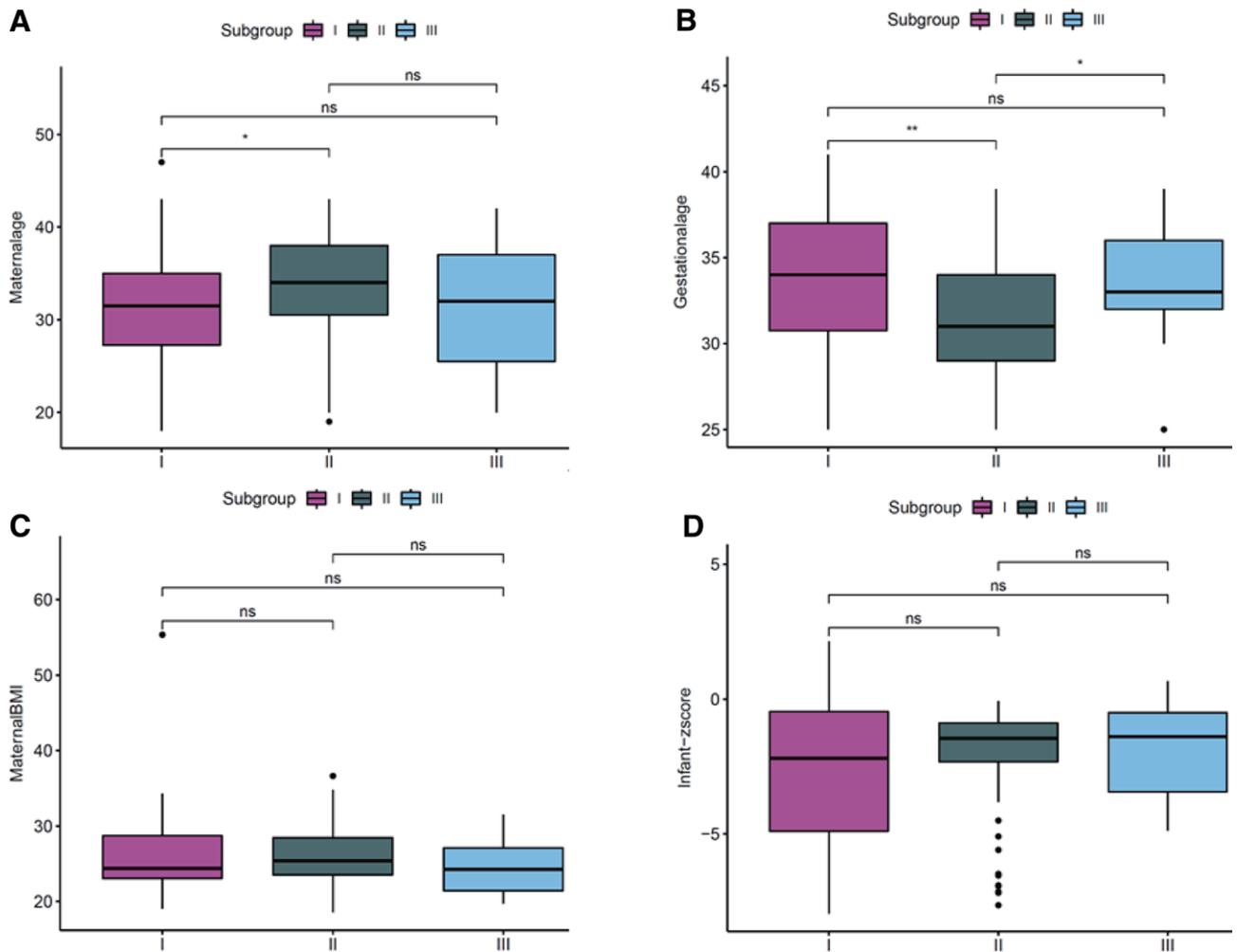
Figure 3. Box-plots representing the pairwise comparison of clinical features, including maternal age (A), gestational age (B), maternal body mass index (C), and infant *z*-score (D) between the subgroups. *$P < .05$, **$P < .01$, ns: no significance. BMI = body mass index.

Table 2

Analysis of variance of gestational age, fetal *z*-score, and their interaction in the subgroups.

	Df	Sum square	Mean square	F value	Pr. (>F)
Subgroup	2	109.7	54.83	4.001	0.0205*
Maternal age	1	8.1	8.14	0.594	0.4421
Subgroup: maternal age	2	39.8	19.90	1.453	0.2376
Residuals	134	1836.2	13.70		

Df = degree of freedom.
Significance codes: *$P < 0.05$.
*0.05.

Table 3

Data of differentially expressed genes in case–control and case–case comparative analysis and the subgroup-specific WGCNA modules.

Subgroup	Genes in case–control comparison	Gene in case–case comparison	WGCNA modules
I	318	991	Green, turquoise
II	1124	376	Brown
III	3476	1478	Blue, gray

WGCNA = weighted gene coexpression network analysis.

hub genes in module blue were *C3AR1, CCR2, CCL5, CCL4, CXCL9, CXCR4, CCR7, CX3CR1, CXCL16*, and *CXCL11*. In module brown, the top 10 hub genes were *EGFR, MAPK3, FGF7, FGF10, FGF9, FLT1, FLT4, FN1, PIK3CB*, and *PPARG*. The 10 hub genes in module turquoise were *COL6A1, COL4A1, COL4A2, COL6A2, COL1A2, COL3A1, COL5A2, COL6A3, COL15A1*, and *COL1A1*.

Moreover, miRNet database was utilized to predict the networks of hub genes and potential miRNA in the 3 key modules. To the hub genes in module brown, the top 5 miRNAs (ranked by degrees) of 36 miRNAs predicted were hsa-mir-34a-5p (degree 7), hsa-mir-19b-3p (degree 4), hsa-mir-522-5p (degree 3), hsa-mir-210-3p (degree 3), and hsa-mir-145-5p (degree 3;

Fig. 7D). In module turquoise, the top 6 miRNAs of the 52 miRNAs predicted by the hub genes were hsa-mir-145-5p (degree 6), hsa-let-7a-5p (degree 6), hsa-mir-34a-5p (degree 4), hsa-mir-106a-5p (degree 4), hsa-mir-19b-3p (degree 3), and hsa-mir-498 (degree 3; Fig. 7E). In module blue, the top 3 miRNAs of the 20 miRNAs predicted by the hub genes were hsa-mir 34a-5p (degree 3), hsa-mir-210-3p (degree 2), and hsa-mir-106a-5p (degree 2; Fig. 7F).

4. Discussion

In the present study, based on the analysis of the gene expression profiles of 205 normal controls and 180 PE cases from 4

A

B

C
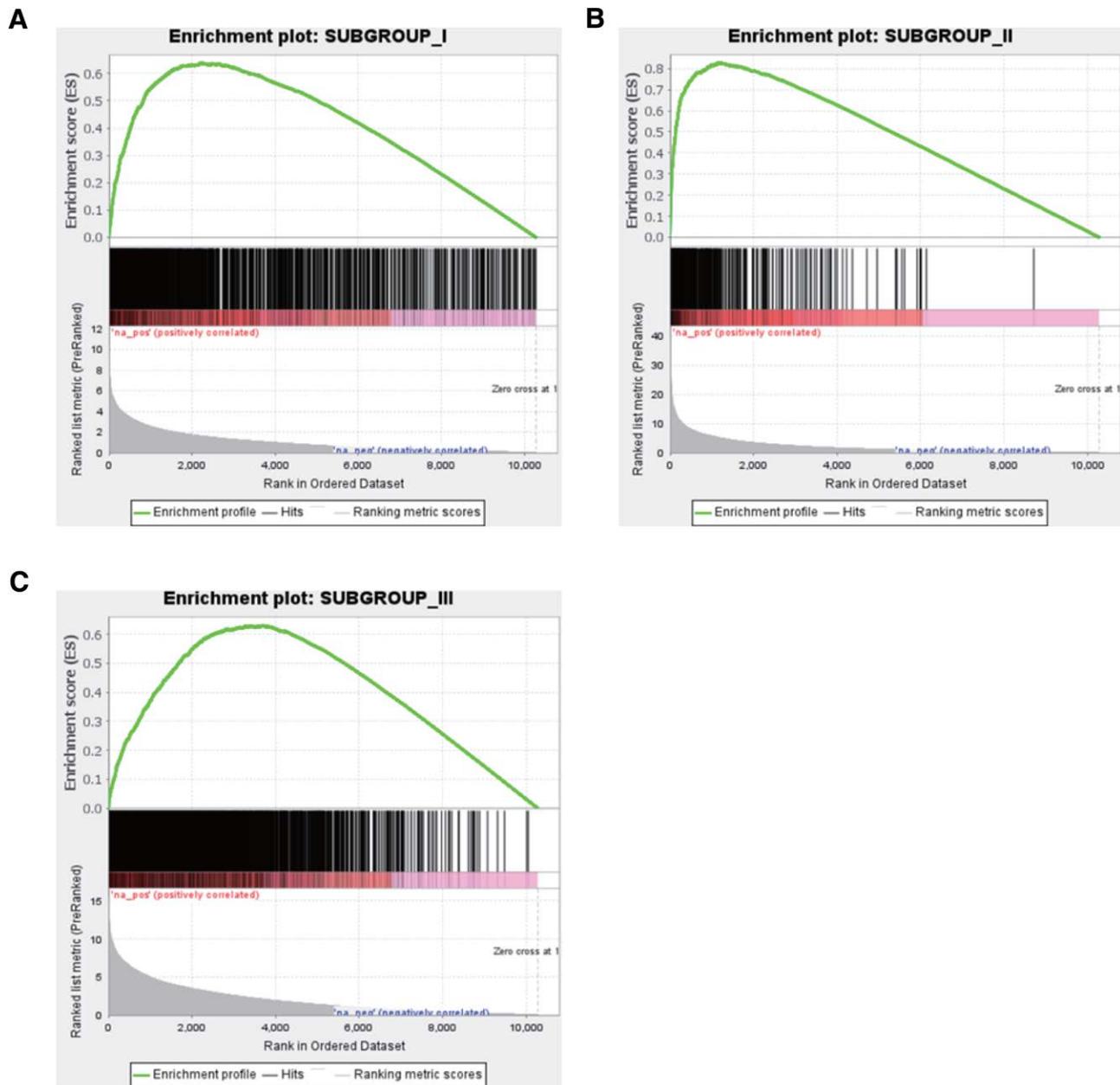
Figure 4. Expression profiles of subgroup-specific genes. Enrichment plots of (A), (B), and (C) indicate that the subgroup-specific genes have different expression levels than the normal controls.

independent GEO datasets, we classified the PE cases into 3 subgroups after eliminating the batch effect and revealed the subgroup-specific gene modules associated with clinical features using WGCNA. By analyzing the correlation between module genes and clinical characteristics, we identified 3 key gene modules and their hub genes. Moreover, we successfully constructed the hub gene-miRNA ceRNA networks.

As a placental disease, the pathogenesis of PE is not well characterized. The suggested pathogenetic mechanisms include reduced invasive capacity of placental trophoblasts, recast dysfunction of the spiral uterine artery caused by an excess of antiangiogenic factors, and inappropriate immune responses to the allogenic fetus.[28,29] Previous microarray studies that have sought to identify early detection markers and candidates for therapeutic intervention have indicated the involvement of oxidative stress and cytokine-associated genes.[30,31] However, there is considerable heterogeneity in the clinical manifestations and therapeutic effect, which cannot be explained by classical

separation of early-onset and late-onset (diagnosis after 34 weeks) PE patients. Leavey et al[27] divided PE patients into 5 clusters. They found high levels of PE markers (such as FLT1) in cluster 2, while cluster 3 showed a maternal-fetal incompatibility presentation; this may represent the considerable interindividual variability observed in clinical settings.

Likewise, in our study, we divided PE patients into 3 subgroups with consensus clustering and the more severe subgroup II had greater maternal age and shorter gestational age. Genes in module brown were significantly upregulated in subgroup II and functional enrichment analysis showed that pathogenesis of this subgroup was mainly related to MAPK, Cyclic adenosine monophosphate (cAMP), and Ras-associated protein 1 (rap1) signaling pathways. Studies have demonstrated the participation of these 3 pathways in the regulation of trophoblast invasion and migration in normal pregnancy.[32–34] Padmini et al[35] revealed that aberrant expression of extracellular signal-regulated kinase 1/2 (ERK1/2, which is also known as MAPK3
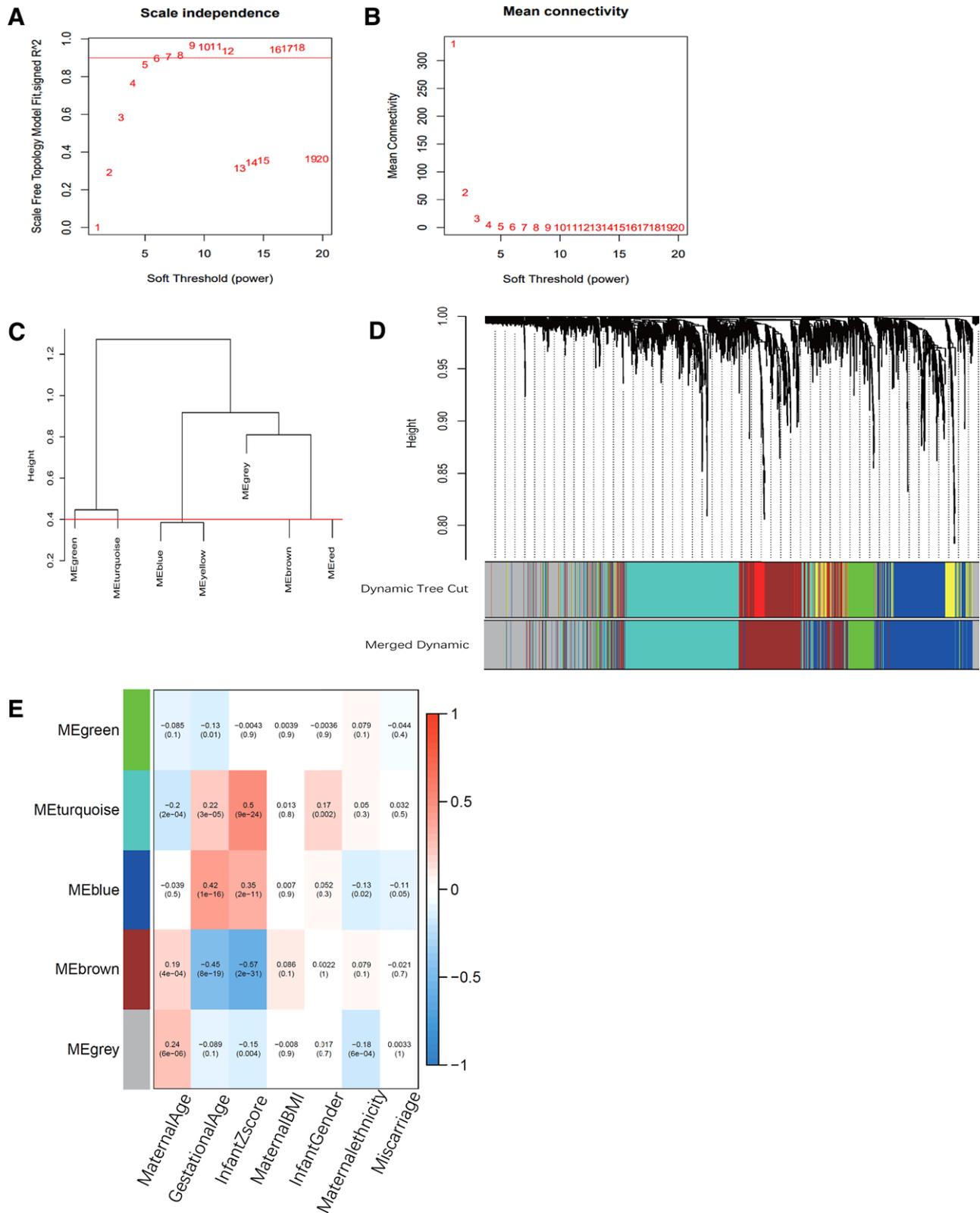
Figure 5. Module detection, gene significance, and module membership of PE samples. (A) Network topology analysis for different soft-thresholding powers. (B) The scale-free fit index and the mean connectivity as a function of the soft-thresholding power. (C) The clusters of the modules, the eigengenes, and the threshold (red line) of modules were merged. (D) Clustering dendrogram of genes with dissimilarity based on the topological overlap, together with assigned module colors. The unmerged colored modules in first row below the dendrogram and the merged colored modules are in the second. (E) The association of the modules and the traits was constructed. Each row refers to a module eigengene and each column refers to a trait. Numbers in each cell represent the correlation and *P* value. The right side of the heatmap reflects the intensity and direction of the correlation (the color depth indicates the strength of the correlation; red, positively correlated; green, positively correlated). BMI = body mass index, PE = preeclampsia.
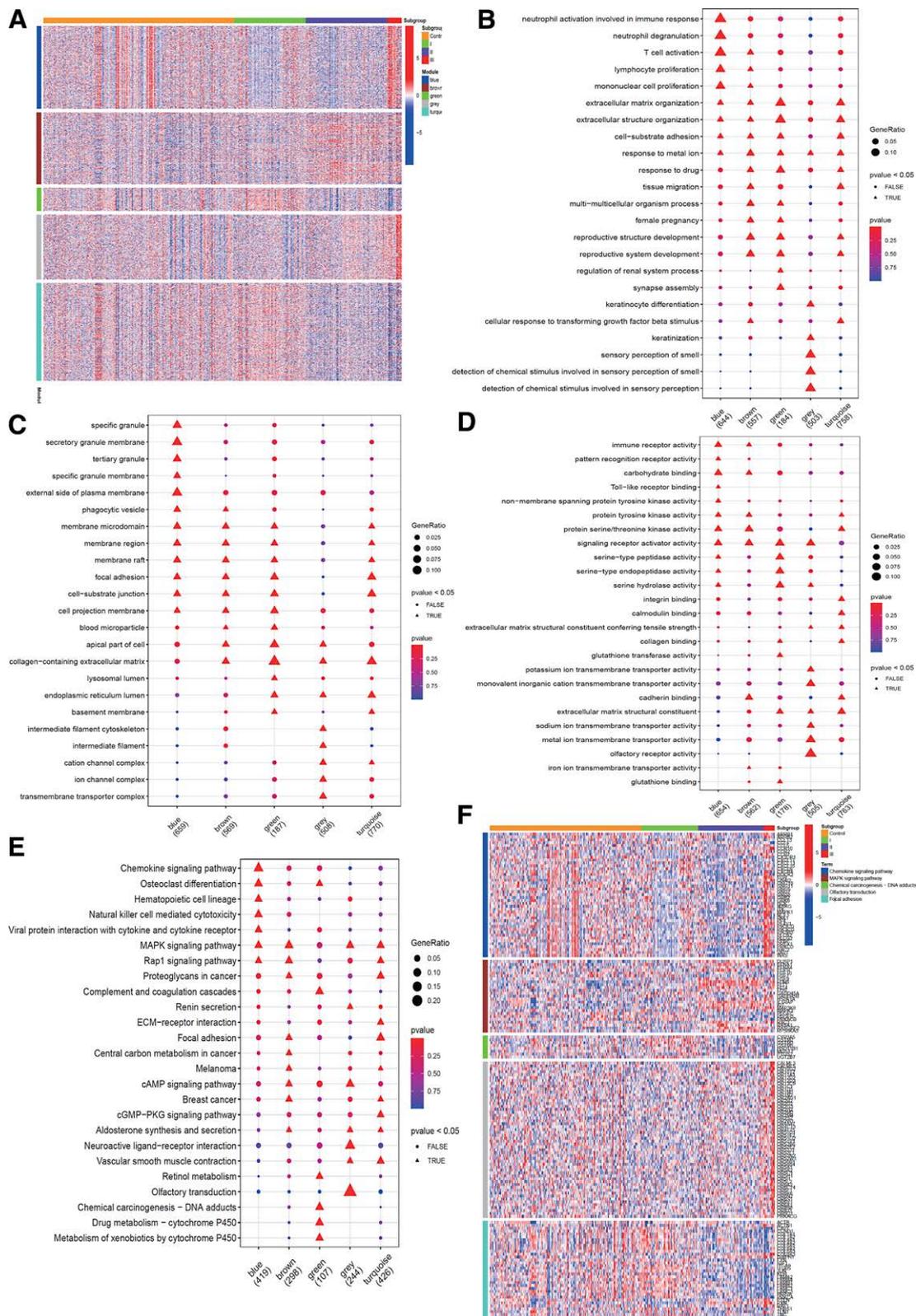
Figure 6. The expression patterns of subgroup-specific genes and functional characterization of WGCNA modules. (A) The scaled expression values of genes that comprise of the 5 gene coexpression network analysis modules are displayed in the heatmap. The bubble plots show the top significantly enriched genes in GO analysis including biological process (B), cellular component (C), and molecular function (D), as well as the KEGG pathways for each WGCNA module (E). (F) The heatmap demonstrates the most significant KEGG pathway in each WGCNA module associated with each subgroup. GO = Gene Ontology, KEGG = Kyoto Encyclopedia of Genes and Genomes, WGCNA = weighted gene coexpression network analysis.

Figure 7. Interaction network of hub genes and potential miRNA–mRNA regulatory networks. (A–C) Interaction network of the top 10 respective hub genes in the module turquoise, brown, and blue. Using Cytoscape package to obtain a network from DEGs based on WGCNA method and cytoHubba package to search the hub genes in R Studio. The color depth indicates the strength of hub genes. Low to high significance ranked by colors from yellow to red. (D–F) Interaction networks of predicted miRNAs and their target genes; blue squares represent miRNAs, yellow squares represent hub gene mRNAs. Edges between squares denote interactions between miRNAs and hub genes mRNAs. DEG = differentially expressed genes, WGCNA = weighted gene coexpression network analysis.

and MAPK1), a member of MAPK family, plays an important role in PE progression. Studies have also shown that the down-regulation of ERK signaling can aggravate oxidative injury in trophoblasts.[36] Moreover, PE cases showed oxidative stress-induced downregulation of p38 MAPK, which may be caused by elevated phosphorylation of p38 MAPK, leading to defects in growth of placenta.[37,38] In pregnancy, expression of epidermal

growth factor receptor activated MAPK/ERK1/2 signaling pathway in placental tissues to regulate blastocyst implantation may result in occurrence of PE.[39] Consistent with these results, our study demonstrated that these MAPK signaling pathway members including epidermal growth factor receptor and ERK1 (MAPK3) may serve as hub genes in module brown. In addition, Chen et al[34] found that hypoxia, accompanied by dysfunction

of placental multipotent mesenchymal stromal cells, induced cAMP decrease and Rap1 inactivation in trophoblasts, which resulted in decreased trophoblast adhesion and migration during the PE process. Members of the fibroblast growth factor (FGF) family, which were detected in our hub gene analysis as FGF7, FGF9, and FGF10, were shown to participate in the MAPK/ERK/cAMP signaling pathway to modulate uterine–conceptus interactions.[40–42] In a previous study, FGF10 expression was shown to be localized to first trimester extravillous trophoblasts, indicating its potential role in PE pathophysiology.[43] In our study, older maternal age and shorter gestational age were simultaneously observed in subgroup II, suggesting that PE women with older age may manifest more serious disease leading to premature birth. Moreover, we also identified classical markers such as FLT1, FLT4, and FN1 as hub genes, and this suggests that the pathogenesis of PE in subgroup II patients may have a typical placental origin.

In subgroup III, genes in module blue and gray were upregulated and module blue showed a significant positive association with gestational age and infant *z*-score. Our gene enrichment analysis showed that these genes were enriched in chemokine signaling pathway and immune-related BPs such as nature killer (NK) cell-mediated cytotoxicity, neutrophil, T-cell activation, and lymphocyte proliferation. The hypothesis that the immune system is involved in the development of PE has already been confirmed by many reports.[44–46] Autoimmune response may alter the maternal immune pathology, interfering with adaptions for placentation, and increase the incidence of PE as well as intrauterine growth restriction.[47] A typical characteristic of the maternal immunological preparatory response to pregnancy is a T-helper type-2 lymphocyte (Th2) proportion increase simultaneously with Th1 upregulation.[48] A suboptimal interaction between uterine NK cells and the trophoblast is believed to lead to maternal spiral artery modification alteration in patients with PE.[49] Extensive infiltration of activated neutrophils in blood vessels of PE women may lead to systemic vascular inflammation and multiorgan involvement. Walsh et al[50] discovered that cytokine interleukin-17A may increase neutrophil chemokines like *CXCL5* and *CXCL6* gene expression in vascular smooth muscles and result in neutrophil boost in PE pregnancy. The balance expression of chemokines at the maternal-fetal interface affects the immune cell profile and function in the decidua.[51] In our analysis, most of detected hub genes are chemokines associated with multiple immunocytes, such as monocytes (*CCR2*), T cells (*CCL5, CXCL11,* and *CXCL16*), Th1 cells (*CXCL9, CXCR4*), and NK cells (*CCL4*). In addition, decreased expression of another hub gene, *C3AR1*, was also discovered in preeclamptic pregnancies with fetal growth restriction, indicating an inability of preeclamptic placentas to clear the excess of activated complement, promoting accumulation of immune complexes and inflammation.[52] Interestingly, subgroup III showed mild PE clinical features while containing most subgroup-specific genes and differentially expressed genes compared with normal cases; this suggests the existence of a protective mechanism in these patients, which calls for further research. Accordingly, our results indicate that subgroup III may be interpreted as a maternal-fetal incompatibility coupled with a maternal immune response hyperfunction.

Subgroup I had the least number of differentially expressed genes with the normal controls, suggesting that subgroup I may exhibit normal-like expression patterns. Compared to subgroup III, no significant differences were observed with respect to gestational age or maternal age. However, the intrinsic biological characteristics in subgroups I and III showed great differences. Subgroup I showed a significant correlation with module turquoise and genes enrichment analysis showed that most genes were enriched at focal adhesion. Our hub gene analysis showed that all hub genes in module turquoise are collagen family member genes. Collagen is one of the most abundant components of extracellular matrix which could influence and regulate trophoblast invasion and participate in the remodeling of the decidua at the maternal–fetal interface.[53] Studies have shown that excessive collagen deposition may affect smooth muscle growth and decrease the remodeling of spiral arteries, which may be an important factor in the pathogenesis of early-onset PE.[54,55] A previous study identified aberrant expression of collagen fragments in PE; *COL4A1* and COL4A2, which code for collagen type IV α chain 1 and chain 2, respectively, were identified as maternal PE susceptibility genes.[56] Ohmaru-Nakanishi et al found significant upregulation of *COL1A1* expression in fibroblasts from PE placentas compared with normal placentas; in addition, placental fibroblasts were also shown to increase the production of collagen type I and IV in hypoxic conditions in vitro.[57,58] Thus, subgroup I may represent an earlier disease stage based on its normal-like expression patterns and early-onset collagen abnormal expression.

Finally, miRNA–mRNA network was constructed to detect possible miRNA involvement. miR-34a-5p and miR-106a-5p were found to be correlated with all 3 significant gene modules, indicating their potential involvement in the pathogenesis of PE. Xue et al[59] reported upregulation of miR-34a-5p in placental tissues and serum of patients with severe PE; they found that miR-34a-5p inhibited the invasion and migration of trophoblast cells targeting Smad4. In previous studies, miR-34a elevation was shown to inhibit trophoblast proliferation, migration, and invasion and to simultaneously promote apoptosis.[60,61] However, miR-106 was reported to promote proliferation, migration, and invasion of cytotrophoblast via blocking syncytiotrophoblast differentiation.[62,63] Our results indicated that miR-19b-3b and miR-210-3p participated in the network of module blue and turquoise. Studies have shown downregulation of miR-19b-3p and upregulation of miR-210-3p in the peripheral circulation of women with PE compared with healthy pregnant women.[64,65]

Our findings suggest that PE patients belonging to different subgroups may have different genomic patterns. Further studies are required to validate these findings, especially for the miRNA–mRNA predictions. In conclusion, our study observed heterogeneity between PE subgroups and attempted to explain the heterogeneity according to different gene expression patterns. Our findings suggest that patients in unique subgroups may benefit from more personalized treatment. The predicted hub genes and miRNAs identified in this study may help develop new therapeutic targets in the future.

Author contributions

Conceptualization: Min Zhang, Xiaheng Deng, Ziyan Jiang.
Methodology and Administrative support: Zhiping Ge.
Data curation: Min Zhang, Xiaheng Deng.
Formal analysis and investigation: Min Zhang, Xiaheng Deng.
Data analysis and interpretation: Min Zhang, Xiaheng Deng, Ziyan Jiang.
Writing – review & editing: Min Zhang, Xiaheng Deng, Ziyan Jiang, Zhiping Ge.

References

[1] Ramos JGL, Sass N, Costa SHM. Preeclampsia. Rev Bras Ginecol Obstet. 2017;39:496–512.

[2] Rana S, Lemoine E, Granger JP, et al. Preeclampsia: pathophysiology, challenges, and perspectives. Circ Res. 2019;124:1094–112.

[3] Bartsch E, Medcalf KE, Park AL, et al.; High Risk of Pre-Eclampsia Identification Group. Clinical risk factors for pre-eclampsia determined in early pregnancy: systematic review and meta-analysis of large cohort studies. BMJ. 2016;353:i1753.

[4] Cnattingius S, Reilly M, Pawitan Y, et al. Maternal and fetal genetic factors account for most of familial aggregation of preeclampsia: a population-based Swedish cohort study. Am J Med Genet A. 2004;130A:365–71.

[5] Eddy AC, Bidwell GL 3rd, George EM. Pro-angiogenic therapeutics for preeclampsia. Biol Sex Differ. 2018;9:36.

[6] Covella B, Vinturache AE, Cabiddu G, et al. A systematic review and meta-analysis indicates long-term risk of chronic and end-stage kidney disease after preeclampsia. Kidney Int. 2019;96:711–27.

[7] Redman CW, Sargent IL. Latest advances in understanding preeclampsia. Science. 2005;308:1592–4.

[8] Brosens I, Renaer M. On the pathogenesis of placental infarcts in pre-eclampsia. J Obstet Gynaecol Br Commonw. 1972;79:794–9.

[9] Rajakumar A, Brandon HM, Daftary A, et al. Evidence for the functional activity of hypoxia-inducible transcription factors overexpressed in preeclamptic placentae. Placenta. 2004;25:763–9.

[10] Chandel NS, McClintock DS, Feliciano CE, et al. Reactive oxygen species generated at mitochondrial complex III stabilize hypoxia-inducible factor-1alpha during hypoxia: a mechanism of O_2 sensing. J Biol Chem. 2000;275:25130–8.

[11] Xu J, Gu Y, Sun J, et al. Reduced CD200 expression is associated with altered Th1/Th2 cytokine production in placental trophoblasts from preeclampsia. Am J Reprod Immunol. 2018;79:10.1111.

[12] Peixoto AB, Araujo Junior E, Ribeiro JU, et al. Evaluation of inflammatory mediators in the deciduas of pregnant women with pre-eclampsia/eclampsia. J Matern Fetal Neonatal Med. 2016;29:75–9.

[13] Sowmya S, Sri Manjari K, Ramaiah A, et al. Interleukin 10 gene promoter polymorphisms in women with early-onset pre-eclampsia. Clin Exp Immunol. 2014;178:334–41.

[14] Gong S, Gaccioli F, Dopierala J, et al. The RNA landscape of the human placenta in health and disease. Nat Commun. 2021;12:2639.

[15] Anton L, Olarerin-George AO, Schwartz N, et al. miR-210 inhibits trophoblast invasion and is a serum biomarker for preeclampsia. Am J Pathol. 2013;183:1437–45.

[16] Wang R, Liu W, Liu X, et al. MicroRNA-210 regulates human trophoblast cell line HTR-8/SVneo function by attenuating notch1 expression: implications for the role of microRNA-210 in pre-eclampsia. Mol Reprod Dev. 2019;86:896–907.

[17] Luo R, Shao X, Xu P, et al. MicroRNA-210 contributes to preeclampsia by downregulating potassium channel modulatory factor 1. Hypertension. 2014;64:839–45.

[18] Leek JT, Scharpf RB, Bravo HC, et al. Tackling the widespread and critical impact of batch effects in high-throughput data. Nat Rev Genet. 2010;11:733–9.

[19] UNICEF WHO and World Bank Group. Levels and Trends in Child Malnutrition: Key Findings of the 2019 Edition of the Joint Child Malnutrition Estimates. Geneva, Switzerland: World Health Organization; 2019.

[20] Langfelder P, Horvath S. WGCN. A: an R package for weighted correlation network analysis. BMC Bioinf. 2008;9:559.

[21] Zhang B, Horvath S. A general framework for weighted gene co-expression network analysis. Stat Appl Genet Mol Biol. 2005;4:Article17.

[22] Yu G, Wang LG, Han Y, et al. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012;16:284–7.

[23] Fan Y, Siklenka K, Arora SK, et al. miRNet—dissecting miRNA-target interactions and functional associations through network-based visual analysis. Nucleic Acids Res. 2016;44:W135–41.

[24] Sitras V, Paulssen RH, Gronaas H, et al. Differential placental gene expression in severe preeclampsia. Placenta. 2009;30:424–33.

[25] Tsai S, Hardison NE, James AH, et al. Transcriptional profiling of human placentas from pregnancies complicated by preeclampsia reveals disregulation of sialic acid acetylesterase and immune signalling pathways. Placenta. 2011;32:175–82.

[26] Yong HE, Melton PE, Johnson MP, et al. Genome-wide transcriptome directed pathway analysis of maternal pre-eclampsia susceptibility genes. PLoS One. 2015;10:e0128230.

[27] Leavey K, Benton SJ, Grynspan D, et al. Unsupervised placental gene expression profiling identifies clinically relevant subclasses of human preeclampsia. Hypertension. 2016;68:137–47.

[28] Jena MK, Sharma NR, Petitt M, et al. Pathogenesis of preeclampsia and therapeutic approaches targeting the placenta. Biomolecules. 2020;10:953.

[29] Romero R, Chaiworapongsa T. Preeclampsia: a link between trophoblast dysregulation and an antiangiogenic state. J Clin Invest. 2013;123:2775–7.

[30] Lee GS, Joe YS, Kim SJ, et al. Cytokine-related genes and oxidation-related genes detected in preeclamptic placentas. Arch Gynecol Obstet. 2010;282:363–9.

[31] Wang Z, Zhao G, Zibrila AI, et al. Acetylcholine ameliorated hypoxia-induced oxidative stress and apoptosis in trophoblast cells via p38 MAPK/NF-kappaB pathway. Mol Hum Reprod. 2021;27:gaab045.

[32] Alese MO, Moodley J, Naicker TA. Signalling of ERK1/2, P38MAPK and P90RSK in HIV-associated pre-eclampsia. J Obstet Gynaecol. 2019;39:612–8.

[33] Duan FM, Fu LJ, Wang YH, et al. THBS1 regulates trophoblast fusion through a CD36-dependent inhibition of cAMP, and its upregulation participates in preeclampsia. Genes Dis. 2021;8:353–63.

[34] Chen CP, Huang JP, Chu TY, et al. Human placental multipotent mesenchymal stromal cells modulate trophoblast migration via Rap1 activation. Placenta. 2013;34:913–23.

[35] Padmini E, Uthra V, Lavanya S. Effect of HSP70 and 90 in modulation of JNK, ERK expression in preeclamptic placental endothelial cell. Cell Biochem Biophys. 2012;64:187–95.

[36] Nadeau V, Charron J. Essential role of the ERK/MAPK pathway in blood-placental barrier formation. Development. 2014;141:2825–37.

[37] Webster RP, Brockman D, Myatt L. Nitration of p38 MAPK in the placenta: association of nitration with reduced catalytic activity of p38 MAPK in pre-eclampsia. Mol Hum Reprod. 2006;12:677–85.

[38] Adams RH, Porras A, Alonso G, et al. Essential role of p38alpha MAP kinase in placental but not embryonic cardiovascular development. Mol Cell. 2000;6:109–16.

[39] Kupsamy K, Moodley J, Naicker T. Hepatocyte growth factor and epidermal growth factor in HIV infected women with preeclampsia. Eur J Obstet Gynecol Reprod Biol. 2019;240:9–14.

[40] Fischer AJ, Scott MA, Ritchey ER, et al. Mitogen-activated protein kinase-signaling regulates the ability of Muller glia to proliferate and protect retinal neurons against excitotoxicity. Glia. 2009;57:1538–52.

[41] Gupta A, Bazer FW, Jaeger LA. Immunolocalization of acidic and basic fibroblast growth factors in porcine uterine and conceptus tissues. Biol Reprod. 1997;56:1527–36.

[42] Chuang JI, Huang JY, Tsai SJ, et al. FGF9-induced changes in cellular redox status and HO-1 upregulation are FGFR-dependent and proceed through both ERK and AKT to induce CREB and Nrf2 activation. Free Radic Biol Med. 2015;89:274–86.

[43] Anteby EY, Greenfield C, Natanson-Yaron S, et al. Vascular endothelial growth factor, epidermal growth factor and fibroblast growth factor-4 and -10 stimulate trophoblast plasminogen activator system and metalloproteinase-9. Mol Hum Reprod. 2004;10:229–35.

[44] Han X, Ghaemi MS, Ando K, et al. Differential dynamics of the maternal immune system in healthy pregnancy and preeclampsia. Front Immunol. 2019;10:1305.

[45] Wedenoja S, Yoshihara M, Teder H, et al. Fetal HLA-G mediated immune tolerance and interferon response in preeclampsia. EBioMedicine. 2020;59:102872.

[46] Williamson RD, McCarthy FP, Kenny LC, et al. Activation of a TLR9 mediated innate immune response in preeclampsia. Sci Rep. 2019;9:5920.

[47] Barnea ER, Rambaldi M, Paidas MJ, et al. Reproduction and autoimmune disease: important translational implications from embryo-maternal interaction. Immunotherapy. 2013;5:769–80.

[48] Salmon JE, Heuser C, Triebwasser M, et al. Mutations in complement regulatory proteins predispose to preeclampsia: a genetic analysis of the PROMISSE cohort. PLoS Med. 2011;8:e1001013.

[49] Ashkar AA, Di Santo JP, Croy BA. Interferon gamma contributes to initiation of uterine vascular modification, decidual integrity, and uterine natural killer cell maturation during normal murine pregnancy. J Exp Med. 2000;192:259–70.

[50] Walsh SW, Nugent WH, Archer KJ, et al. Epigenetic regulation of interleukin-17-related genes and their potential roles in neutrophil vascular infiltration in preeclampsia. Reprod Sci. 2021;29:154–62.

[51] Rambaldi MP, Weiner E, Mecacci F, et al. Immunomodulation and preeclampsia. Best Pract Res Clin Obstet Gynaecol. 2019;60:87–96.

[52] Lim R, Lappas M. Decreased expression of complement 3a receptor (C3aR) in human placentas from severe preeclamptic pregnancies. Eur J Obstet Gynecol Reprod Biol. 2012;165:194–8.

[53] Guillomot M, Campion E, Prezelin A, et al. Spatial and temporal changes of decorin, type I collagen and fibronectin expression in normal and clone bovine placenta. Placenta. 2014;35:737–47.

[54] Li W, Mata KM, Mazzuca MQ, et al. Altered matrix metalloproteinase-2 and -9 expression/activity links placental ischemia and anti-angiogenic sFlt-1 to uteroplacental and vascular remodeling and collagen deposition in hypertensive pregnancy. Biochem Pharmacol. 2014;89:370–85.

[55] Yong HE, Murthi P, Borg A, et al. Increased decidual mRNA expression levels of candidate maternal pre-eclampsia susceptibility genes are associated with clinical severity. Placenta. 2014;35:117–24.

[56] Johnson MP, Fitzpatrick E, Dyer TD, et al. Identification of two novel quantitative trait loci for pre-eclampsia susceptibility on chromosomes 5q and 13q using a variance components-based linkage approach. Mol Hum Reprod. 2007;13:61–7.

[57] Ohmaru-Nakanishi T, Asanoma K, Fujikawa M, et al. Fibrosis in pre-eclamptic placentas is associated with stromal fibroblasts activated by the transforming growth factor-beta1 signaling pathway. Am J Pathol. 2018;188:683–95.

[58] Semenza GL. Hypoxia-inducible factors in physiology and medicine. Cell. 2012;148:399–408.

[59] Xue F, Yang J, Li Q, et al. Down-regulation of microRNA-34a-5p promotes trophoblast cell migration and invasion via targetting Smad4. Biosci Rep. 2019;39:BSR20181631.

[60] Guo M, Zhao X, Yuan X, et al. Elevated microRNA-34a contributes to trophoblast cell apoptosis in preeclampsia by targeting BCL-2. J Hum Hypertens. 2017;31:815–20.

[61] Sun M, Chen H, Liu J, et al. MicroRNA-34a inhibits human trophoblast cell invasion by targeting MYC. BMC Cell Biol. 2015;16:21.

[62] Zhao YH, Liu YL, Fei KL, et al. Long non-coding RNA HOTAIR modulates the progression of preeclampsia through inhibiting miR-106 in an EZH2-dependent manner. Life Sci. 2020;253:117668.

[63] Kumar P, Luo Y, Tudela C, et al. The c-Myc-regulated microRNA-17~92 (miR-17~92) and miR-106a~363 clusters target hCYP19A1 and hGCM1 to inhibit human trophoblast differentiation. Mol Cell Biol. 2013;33:1782–96.

[64] Sandrim VC, Luizon MR, Palei AC, et al. Circulating microRNA expression profiles in pre-eclampsia: evidence of increased miR-885-5p levels. BJOG. 2016;123:2120–8.

[65] Nejad RMA, Saeidi K, Gharbi S, et al. Quantification of circulating miR-517c-3p and miR-210-3p levels in preeclampsia. Pregnancy Hypertens. 2019;16:75–8.