



OPEN ACCESS

EDITED BY

Carlos Galaviz-Hernandez,
National Polytechnic Institute (IPN), Mexico

REVIEWED BY

Alper Uzun,
Brown University, United States
Laura Jazel Barragan-Zuñiga,
Centro Estatal de Cancerología (CECAN),
Mexico

*CORRESPONDENCE

Xueli Gao
✉ 510614019@163.com

RECEIVED 14 February 2024

ACCEPTED 22 May 2024

PUBLISHED 03 June 2024

CITATION

Zhang H, Ma J and Gao X (2024) Identifying
molecular subgroups of patients with
preeclampsia through bioinformatics.
Front. Cardiovasc. Med. 11:1367578.
doi: 10.3389/fcvm.2024.1367578

COPYRIGHT

© 2024 Zhang, Ma and Gao. This is an open-
access article distributed under the terms of
the [Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction
in other forums is permitted, provided the
original author(s) and the copyright owner(s)
are credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Identifying molecular subgroups of patients with preeclampsia through bioinformatics

Huijie Zhang¹, Jianglei Ma² and Xueli Gao^{1*}

¹Department of Obstetrics, The Affiliated Yantai Yuhuangding Hospital of Qingdao University, Yantai, China, ²Department of Infectious Diseases, Yantai Qishan Hospital, Yantai, China

Preeclampsia (PE) is a pregnancy-related disorder associated with serious complications. Its molecular mechanisms remain undefined; hence, we aimed to identify molecular subgroups of patients with PE using bioinformatics to aid treatment strategies. R software was used to analyze gene expression data of 130 patients with PE and 138 healthy individuals from the Gene Expression Omnibus database. Patients with PE were divided into two molecular subgroups using the unsupervised clustering learning method. Clinical feature analysis of subgroups using weighted gene co-expression network analysis showed that the patients in subgroup I were primarily characterized by early onset of PE, severe symptoms at disease onset, and induced labor as the main delivery method. Patients in subgroup II primarily exhibited late PE onset, relatively mild symptoms, and natural delivery as the main delivery method. Gene Ontology and Kyoto Encyclopedia of Genes and Genomes pathway enrichment analyses revealed that the significant enrichment of calcium ion channels in subgroup II indicated the potential efficacy of calcium antagonists and magnesium sulfate therapy. In conclusion, the establishment of PE molecular subgroups can aid in diagnosing and treating PE.

KEYWORDS

bioinformatics, gene, pregnancy, preeclampsia, subgroup

1 Introduction

Preeclampsia (PE) is a pregnancy-related disorder characterized by elevated blood pressure that could be accompanied by proteinuria after 20 weeks of pregnancy (1). PE can cause serious complications, such as cerebral hemorrhage, pulmonary edema, acute kidney injury, liver dysfunction, hemolysis, seizures, and placental abruption (2). Furthermore, PE can result in premature delivery, fetal growth restriction, and an increased risk of fetal morbidity and mortality (1). Globally, PE occurs in 3%–5% of all pregnant women and is a crucial factor in maternal mortality (3). In the United States, 6.6% of pregnancy-related deaths are associated with gestational hypertension (1). Globally, approximately 760,000 maternal and 500,000 infant deaths are associated with PE annually, and the mortality rate of patients is approximately 10 times higher in low-income countries than in high-income countries (4). However, the pathogenesis of PE remains unclear.

Abbreviations

BP, biological process; CC, cellular component; DEG, differentially expressed genes; ERK, extracellular signal-regulated kinase; GAP, GTPase-activating protein; GEO, gene expression omnibus; GO, gene ontology; GSEA, gene set enrichment analysis; KEGG, Kyoto encyclopedia of genes and genomes; MF, molecular function; PCA, principal component analysis; PlGF, placental growth factor; ROS, reactive oxygen species; WGCNA, weighted gene co-expression network analysis.

Owing to the interaction between genetic, environmental, and immune factors, trophoblast invasion and spiral arteriole remodeling in the early stages of placental development can lead to uterine placental blood flow disorders (5). Under placental hypoxia, an imbalance in the expression of soluble factors can cause oxidative stress and inflammatory reactions, leading to systemic endothelial dysfunction and multisystem organ ischemia (1). As the delivery of the fetus is currently the only definitive treatment for PE, it poses a serious threat of increasing the risk of mortality for both pregnant women and newborns. The development of precision medicine has increased interest in identifying biomarkers for the diagnosis and treatment of PE. Over the past few decades, the importance of factors related to oxidative stress, inflammation, angiogenesis, and antiangiogenesis has improved the current understanding of the molecular pathogenesis of PE. Placental growth factor (PlGF) and soluble fibroid tyrosine kinase 1 (sFlt-1) are pro- and anti-angiogenic factors, respectively. In patients with PE, PlGF expression decreases and sFlt-1 expression increases (5). When a patient is <37 weeks pregnant and the sFlt-1:PlGF ratio is ≤ 38 , the possibility of PE occurring within the next 4 weeks can be ruled out, with a predictive value of 99.3%. When the ratio is >38 , the predictive accuracy of PE within 4 weeks is 36.7%, with a prediction sensitivity of 66.2% (6). In a recent study, genetic characteristics analysis using whole-exome sequencing in patients with severe PE identified subsets of protein interaction networks in these patients, suggesting that *LAMB2*, *PTK2*, *RAC1*, *QSOX1*, *FNI*, and *VCAM1* may be related to the pathogenesis of PE (7). Although introducing these biomarkers into clinical practice has yielded promising results, there are still significant unmet clinical needs; this may be because PE, including multiple subtypes, can lead to maternal and infant mortality and morbidity through various pathophysiological pathways. Approaches for correctly identifying and testing the subtype correlations are lacking (8).

Bioinformatics analysis is an important tool for analyzing and evaluating datasets in modern biomedical research. It can be used to predict protein structure, sequence alignment, protein-protein interaction (PPI), and other characteristics, and it is particularly useful for revealing the biochemical and functional roles of proteins (9). This study aimed to use unsupervised computational learning, weighted gene co-expression network analysis (WGCNA), Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses to explore the entry points of the transcriptome dataset of patients with PE in the Gene Expression Omnibus (GEO) database, as well as to conduct bioinformatics analysis on these datasets to construct a molecular subgroup of these patients, providing new insights into the molecular mechanisms of PE. Effective identification of PE subtypes will contribute to the early prevention and diagnosis of PE and guide targeted treatment.

2 Methods

2.1 Downloading and processing of datasets

We searched for “preeclampsia” in the GEO (<https://www.ncbi.nlm.nih.gov/geo/>) of the National Center for Biotechnology

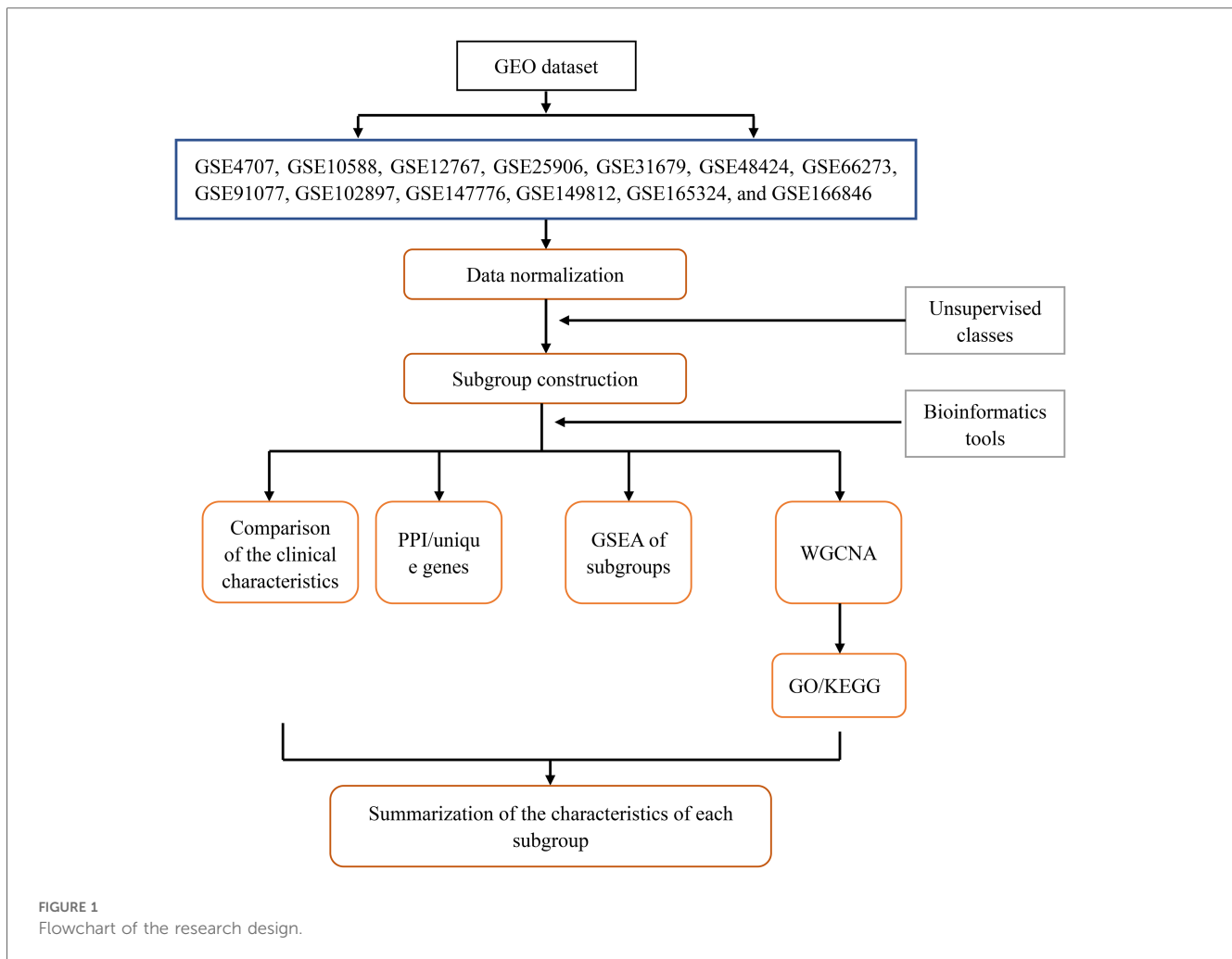
Information. We selected microarray datasets containing patients with PE and healthy controls (GSE4707, GSE10588, GSE12767, GSE25906, GSE31679, GSE48424, GSE66273, GSE91077, GSE102897, GSE147776, GSE149812, GSE165324, and GSE166846) and corresponding platform files (GPL1708, GPL2986, GPL570, GPL1602, GPL6947, GPL6480, GPL4133, GPL13497, GPL22120, GPL20844, GPL212844, GPL22120, and GPL22120) to download. Gene names on the dataset were annotated using Perl version 5.30.0 (<https://www.perl.org/>), and the “limma” and “sva” packages of R software (V 4.2.2) were used for principal component analysis (PCA) clustering batch correction to eliminate batch effects in samples (10). Specifically, the linear modeling method using the “limma” package was used to perform intragroup correction on each microarray dataset. Then, the “sva” package was used to integrate all samples in the datasets. As the data used in this study were sourced from the GEO public database, the contributors to the database obtained ethical approval and informed consent from each patient. Therefore, our study did not require further ethical or any other approval. The flowchart in Figure 1 shows the design for this study.

2.2 Constructing molecular subgroups of patients

Gene expression was analyzed after batch correction using the “limma” package of R software, and consensus clustering was conducted on genes and visual output heatmaps using the “ConsensusClusterPlus” package (11). Specifically, an unsupervised clustering learning method was used to cluster genes with the same genetic characteristics as those expressed in patients with PE. The outputs were visualized by grouping the number of clusters (defined as a parameter κ and set to 2–10), and the internal clustering of each group was evaluated. The higher the evaluation, the better the similarity in gene expression within the group. The subgroup with the highest consistent evaluation score was selected as the study subtype.

2.3 Identification of clinical features of patient subgroups

Based on the continuity of clinical features, the clinical characteristics of patients with PE were divided into discrete variables, including symptom severity (severe PE was defined as one or more of the following conditions: blood pressure above 160/110 mmHg, proteinuria above 5,000 mg/24 h, comorbidities with multiple system diseases, maternal seizures, stroke or fetal intrauterine growth restriction below the third percentile), delivery method, and sex of newborns; and continuous variables, including patient age and gestational week (early onset: gestational age <34 weeks; late onset: gestational age ≥ 34 weeks). The proportion of discrete variables in patients was analyzed using the “rstatix”, “ggplot2”, “ggpubr”, and “reshap2” packages of the R software (12), and the mean and standard deviation values of continuous variables were examined using the



“ggplot2” and “ggpubr” packages. Differences between groups were compared using a one-way analysis of variance.

2.4 Screening of subgroup-unique genes and construction of PPI networks

We used the “limma” package to screen differentially expressed genes (DEGs) between subgroups and control groups and within each subgroup. The intersecting genes of the two genes were the uniquely expressed genes of the subgroup. Unique genes in the subgroups were significantly differentially expressed only in specific subpopulations. The screening criteria were a mean filter of >0.2 and a corrected p -value of <0.05. To better analyze the characteristics of DEGs between subpopulations, we only selected the uniquely upregulated DEGs of each subpopulation. The top 10 unique DEGs from each subgroup were selected and imported into the STRING online tool (<https://string-db.org/>). A confidence level of 0.4 was selected, and a PPI network diagram was created to understand the relationship between protein linkages within each subgroup.

2.5 Gene set enrichment analysis (GSEA) of patient subgroups

To determine the unique DEGs among different subgroups and to evaluate whether the consistency of DEGs between subgroups and normal samples, the comparison files of the subgroups and control groups were converted into gene sets and gene list files using the Perl software. The files converted from each subgroup were inputted into GSEA version 5.32 (13), the running parameters were set to 5,000, and the remaining parameters were obtained from the software by default.

2.6 WGCNA of patient subgroups

We used the “WGCNA” package of R software to construct and visualize the WGCNA for subgroup-unique genes (14). First, we used the “limma” package to filter the corrected abnormal samples. A co-expression network was then constructed using the standard scale-free network analysis function, and the soft threshold power value β was calculated using the power function

“pickSoftThreshold” to enhance the aggregation of gene expression (15). We drew a tree-like gene module diagram for co-expressed gene clustering. Pearson’s correlation analysis was used to analyze the relationship among the delivery mode, symptom severity, sex of newborns, maternal age, and gestational week with the genetic modules to determine the relationship between clinical traits and genetic modules. The expression of gene modules obtained from the WGCNA in subgroups I and II was produced using the “heatmap” package of R software and was presented in the form of a heatmap (14). Finally, a visualized expression heatmap of gene modules in subgroups, which could indirectly reveal the enrichment of subgroups and clinical traits, was drawn using the “limma” and “pheatmap” packages of the R software.

2.7 GO and KEGG pathway analyses of gene modules

We analyzed the GO [biological process (BP), molecular function (MF), and cellular component (CC)] of the gene module using the “org.Hs.eg.db”, “clusterProfiler”, and “enrichplot” packages of R software (16). The KEGG pathway enrichment analysis was performed using the DAVID v6.8 online tool (<https://david.ncifcrf.gov/>) (17). The significantly expressed genes from each gene module were entered into the DAVID website to obtain significantly enriched KEGG pathways.

2.8 Statistical analyses

This study used R software version 4.2.2. Intergroup comparisons were conducted using a one-way analysis of variance. Continuous variables were analyzed using the SPSS 26.0 statistical software (IBM Corp., Armonk, NY, USA), and discrete variables were analyzed using the chi-squared test. The results are expressed as mean \pm standard deviation. A p -value of <0.05 was considered statistically significant.

3 Results

3.1 Data processing and grouping

After organizing all datasets, samples from 130 patients with PE and 138 healthy individuals were obtained. PCA-based visualization images showed that before the batch effects were eliminated, each dataset was located in a different region; there was no intersection between datasets (Figure 2A). After eliminating the batch effects, the PCA results showed that all samples were clustered into a single region, which could be regarded as a unified dataset (Figure 2B). After consensus clustering of genes in patients with PE, based on the set κ , nine clusters were obtained. The visualized clustering results showed that when the cluster was two, the consistency evaluation of each subgroup was at a high level (Figure 2C). The cluster correlation

heatmap of subgroups showed high enrichment within each subgroup and low enrichment between subgroups (Figure 2D).

3.2 Clinical characteristics of patient subgroups

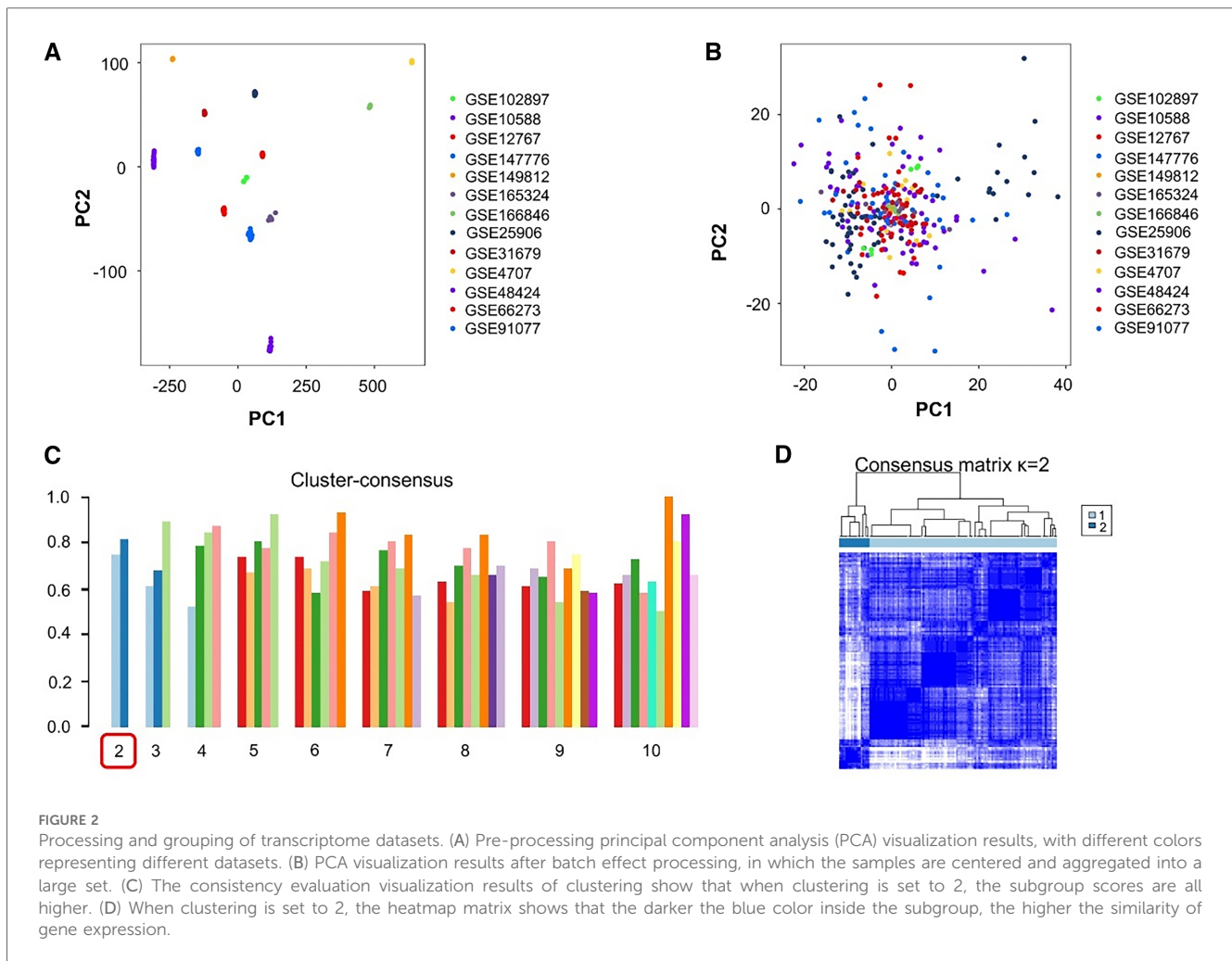
The clinical characteristics of 130 patients with PE (112 in subgroup I and 18 in subgroup II) were analyzed using the downloaded datasets (Supplementary Table S1). The mean and standard deviation values of age (years) and gestational week were 31.14 ± 5.13 and 33.30 ± 3.44 , respectively, in subgroup I and 31.57 ± 5.77 and 35.44 ± 2.28 , respectively, in subgroup II. Intergroup analysis showed no significant difference in age between subgroups I and II (Figure 3A), and the gestational week of subgroup I was significantly lower than that of subgroup II ($p < 0.05$; Figure 3B). The symptoms in subgroup I were more severe than those in subgroup II ($p < 0.05$; Figures 3C,D). Although more patients in subgroup I underwent induced labor based on clinical decision and those in subgroup II experienced spontaneous labor, no significant difference was found in the delivery methods (Figures 3E-F). In addition, no significant difference was observed in the sex of newborns between these two subgroups (Figures 3G-H).

3.3 Unique genes and PPI networks of patient subgroups

According to the screening criteria of a mean filter of >0.2 and corrected p -value of <0.05 , 3,536 upregulated DEGs and 3,413 downregulated DEGs were identified between subgroup I and the control group (Supplementary Table S2); 3,162 upregulated DEGs and 3,787 downregulated DEGs were identified between subgroup II and the control group (Supplementary Table S3); and 3,168 upregulated DEGs and 3,781 downregulated DEGs were identified between subgroups I and II (Supplementary Table S4). After screening for upregulated DEGs and intersections, 1,099 and 1,043 unique DEGs were identified in subgroups I and II, respectively (Supplementary Table S5). The intersection results indicated no common unique DEGs among the subgroups (Figure 4A). The PPI network results showed 20 nodes, 9 protein pairs, and close interactions of DEGs within each subgroup (Figure 4B). Table 1 lists the number of nodes in the top 10 DEGs for each subgroup.

3.4 GSEA of patient subgroups

On the GSEA image, the black vertical lines represent the unique DEGs among subgroups, whereas the gray vertical lines represent the DEGs between subgroups and health control samples. Both sets of data were enriched on the left side of the image. Moreover, the p -value and error detection rate value were <0.05 , indicating that the two sets of DEGs exhibited the same enrichment trend (Figure 5). GSEA results indicated that the



unique DEGs among different subgroups, DEGs between subgroups, and health normal samples were consistent.

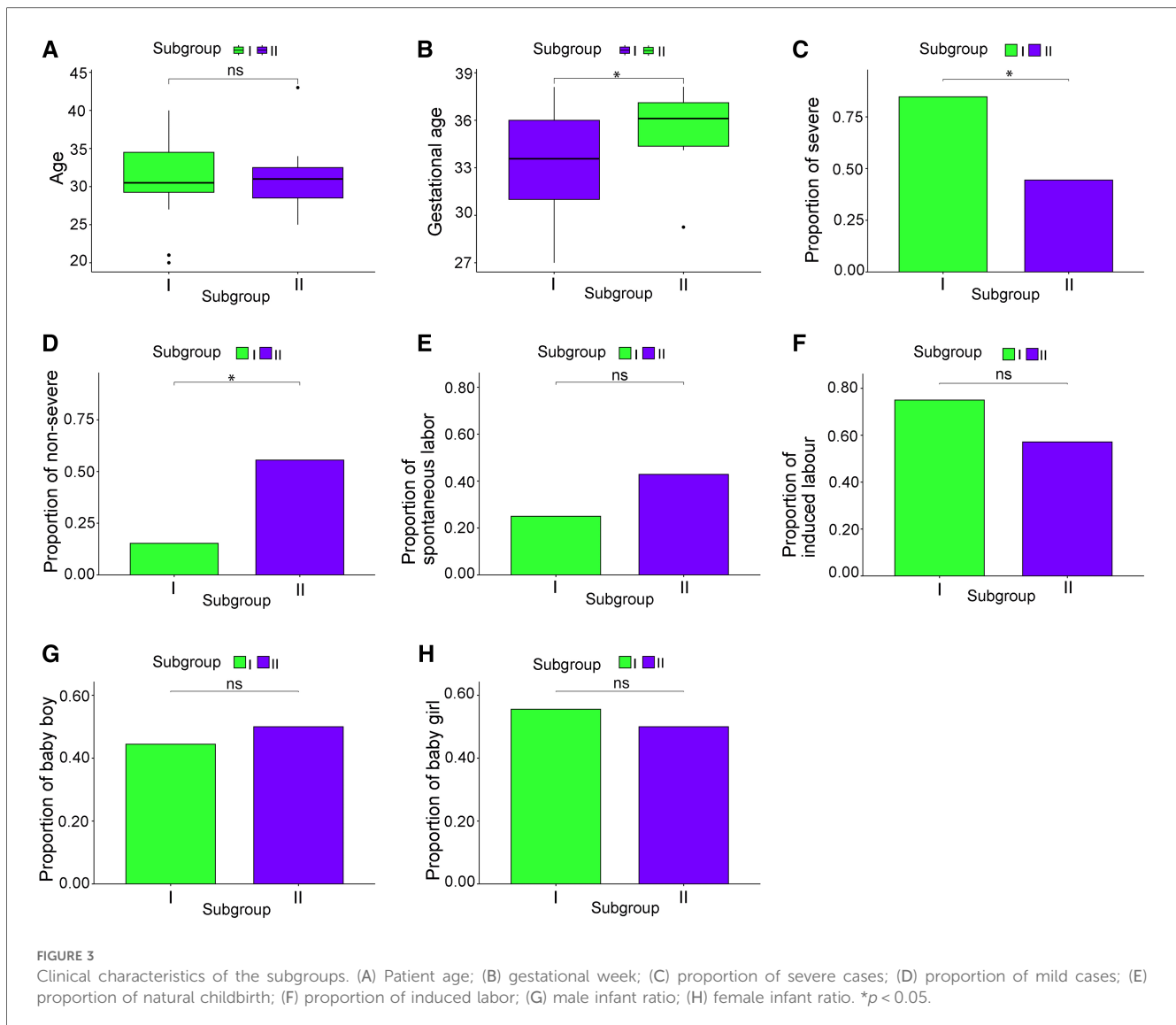
3.5 WGCNA in patients with PE

WGCNA was performed on 6,949 genes in the processed samples. After excluding abnormal samples, the optimal soft threshold power value was selected as 6 (Figure 6A). According to the gene consensus expression tree, three colored gene modules were generated: blue (1,462 genes), brown (242 genes), and gray (438 genes; Figure 6B). The heatmap of gene modules and clinical traits showed a significant negative correlation ($p=0.03$) between spontaneous labor and blue modules and a significant positive correlation ($p<0.001$) with brown modules. Severe symptoms showed a significant positive correlation with the blue module ($p=0.02$) and a significant negative correlation with the gray module ($p<0.001$). Gestational week was significantly negatively correlated with the blue module ($p=0.03$) and positively correlated with the brown ($p=0.02$) and gray ($p=0.001$) modules (Figure 6C). The enrichment heatmap of the gene modules and subgroups showed that the blue modules in subgroup I were highly expressed, whereas brown and gray

modules were poorly expressed. Moreover, blue modules revealed low expression in subgroup II, whereas brown and gray modules showed high expression (Figure 7).

3.6 GO and KEGG

By analyzing the GO and KEGG enrichment pathways of gene modules, we observed that for BP, intrinsic apoptotic signaling pathways, vessel organization, regulation of mitotic cell cycle phase transition, and positive regulation of protein modification by small protein aggregation or removal were significantly enriched in the blue module ($p<0.05$); synapse organization, skin development, regulation of membrane potential, inorganic anion transport, and chloride transport were significantly enriched in the gray module ($p<0.05$; Figure 8A). For CC, focal adhesions, cell-substrate junctions, pharmacological vesicles, organelle subcompartments, and nuclear species were significantly enriched in the blue module ($p<0.05$); the calcium channel, voltage-gated calcium channel, transporter, transmembrane transporter, and ion channel complexes were significantly enriched in the gray module ($p<0.05$; Figure 8B). For MF, ubiquitin-like protein transfer, ubiquitin-protein



transfer, and GTPase activities were significantly enriched in the blue module ($p < 0.05$); oxidoreductase activity with NAD(P)H as a donor and quinone or a similar compound as an acceptor, NADH dehydrogenase activity, and structural constitution of the ribosome were significantly enriched in the brown module ($p < 0.05$); and gated channel, ion channel, channel, passive transmembrane transporter, and high voltage-gated calcium channel activities were significantly enriched in the gray module ($p < 0.05$) (Figure 8C). In the KEGG enrichment pathway, signal pathways, such as those involved in small cell lung cancer, autophagy, legionellosis, shigellosis, glycosylphosphatidylinositol anchor biosynthesis, and the HIV-1 viral life cycle, and the p53 signaling pathway, were significantly enriched in the blue module ($p < 0.05$); ribosome components, metabolic pathways, oxidative physiology, reactive oxygen specifications, peroxisomes, and purine metabolism were significantly enriched in the brown module ($p < 0.05$); and the MAPK signaling pathway, maturity-onset diabetes of the young, calcium signaling pathway, GABAergic synapse, neuroactive ligand-receptor interaction, B-cell receptor signaling pathway, retrograde endocannabinoid

signaling, phototransduction, taste transduction, and oxytocin signaling pathway were significantly enriched in the gray module ($p < 0.05$; Supplementary Table S6).

3.7 Characteristics of aggregated patient subgroups

To obtain the characteristics of each subgroup more intuitively, Table 2 summarizes the clinical characteristics, unique DEGs, and possible effective treatment methods and shows different clinical and molecular characteristics among the subgroups, indicating that patients with PE can be divided into two subgroups.

4 Discussion

Over the past few decades, extensive research has been conducted on PE. However, the subtypes of patients and therapeutic biomarkers for PE remain unclear. Identifying the

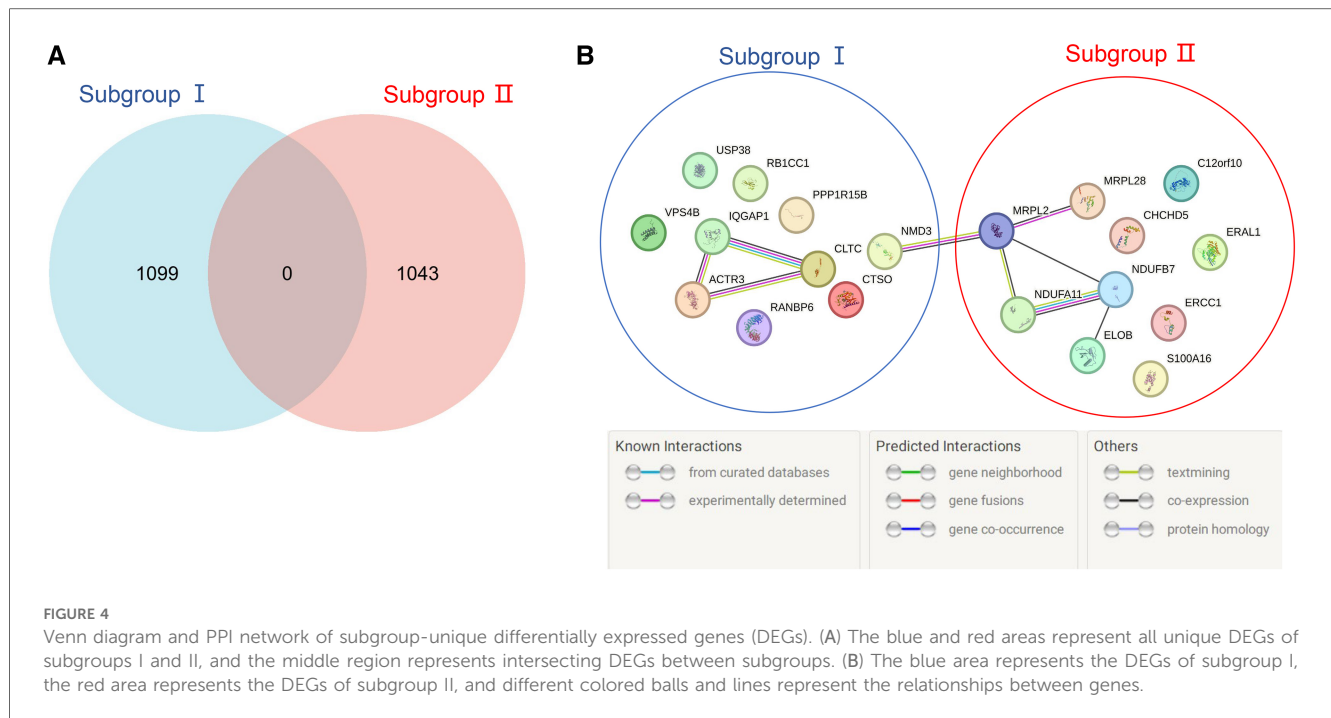


TABLE 1 PPI network nodes of the top 10 DEGs in the subgroups.

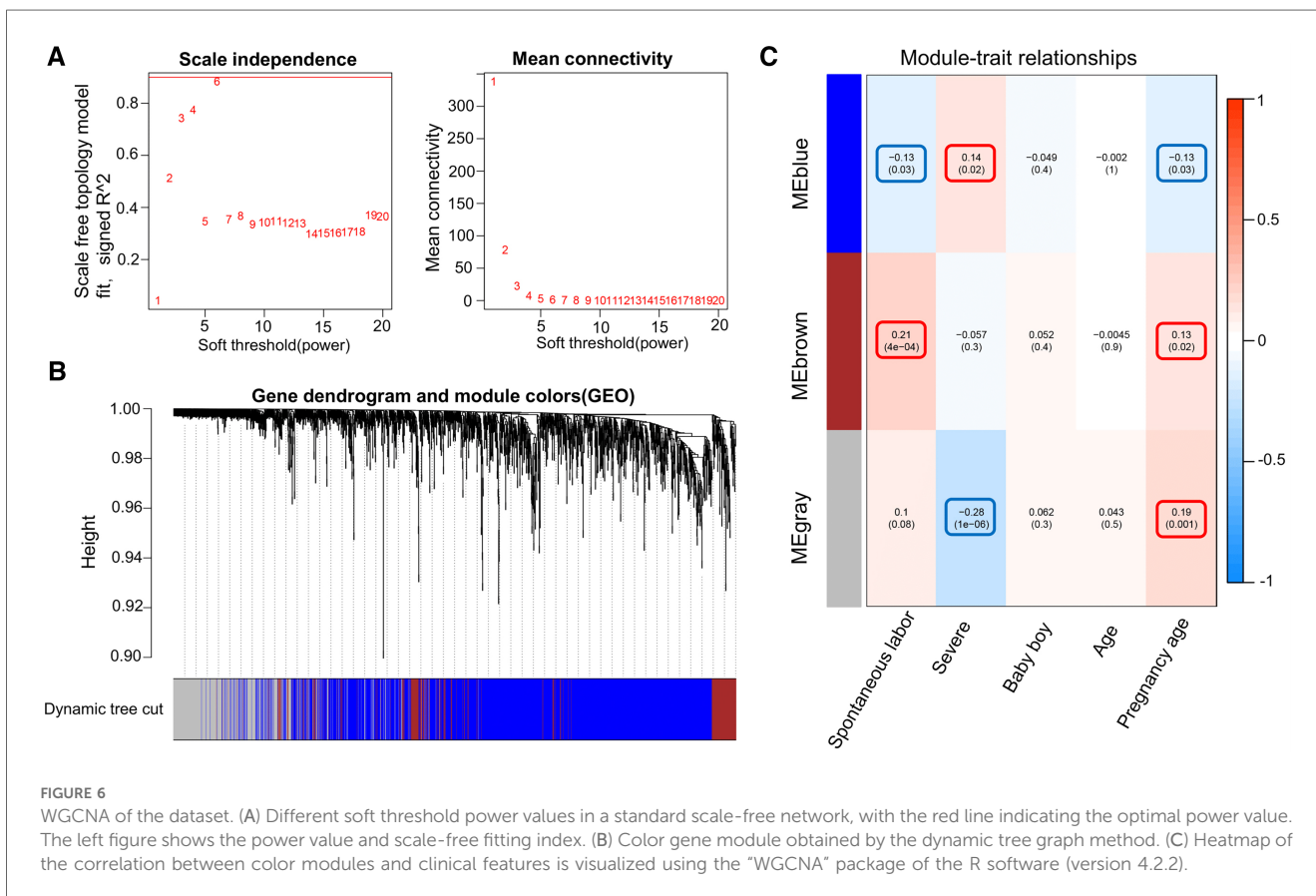
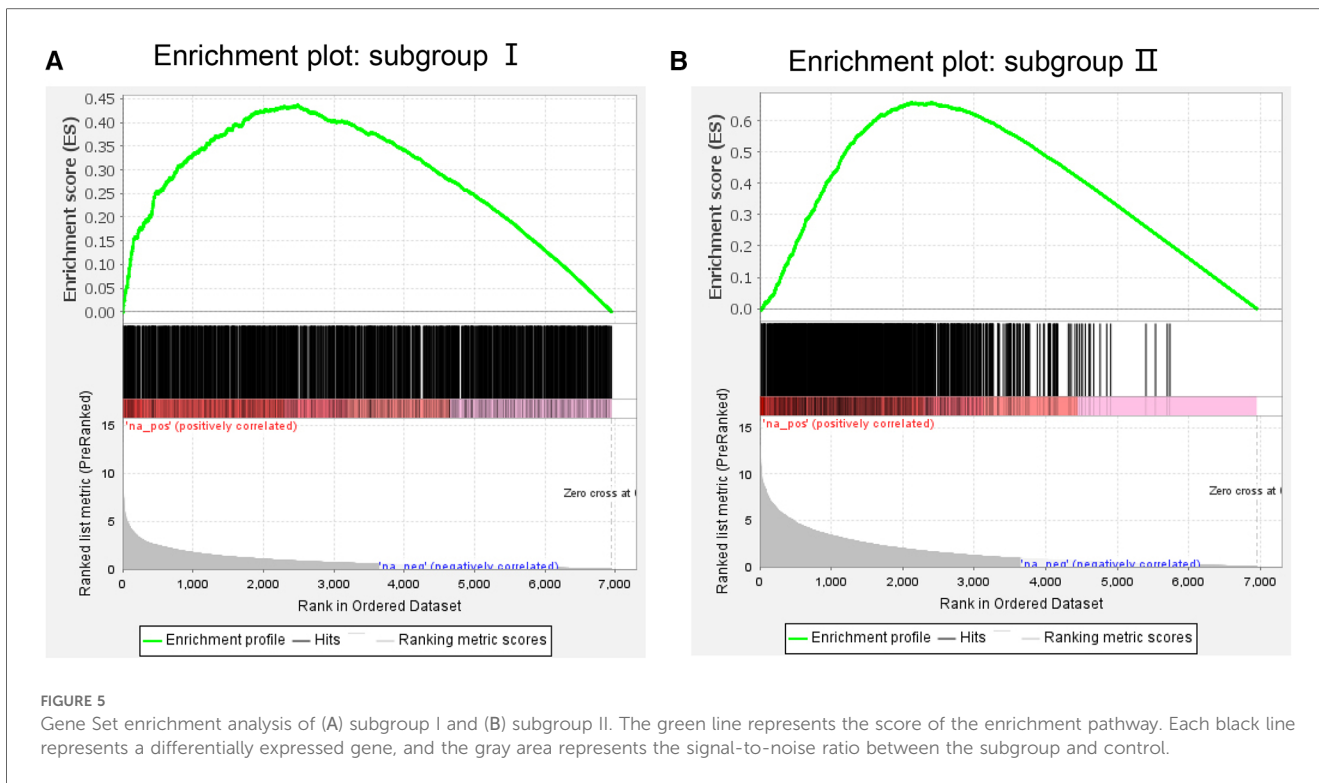
Subgroup	Gene	Nodes
I	<i>IQGAP1</i>	7
	<i>USP38</i>	0
	<i>ACTR3</i>	6
	<i>RB1CC1</i>	0
	<i>NMD3</i>	3
	<i>PPP1R15B</i>	0
	<i>VPS4B</i>	0
	<i>RANBP6</i>	0
	<i>CLTC</i>	7
	<i>CTSO</i>	0
II	<i>C12orf10</i>	0
	<i>CHCHD5</i>	0
	<i>ERAL1</i>	0
	<i>MRPL28</i>	2
	<i>MRPL2</i>	5
	<i>ERCC1</i>	0
	<i>S100A16</i>	0
	<i>TCEB2</i>	1
	<i>NDUFB7</i>	6
	<i>NDUFA11</i>	6

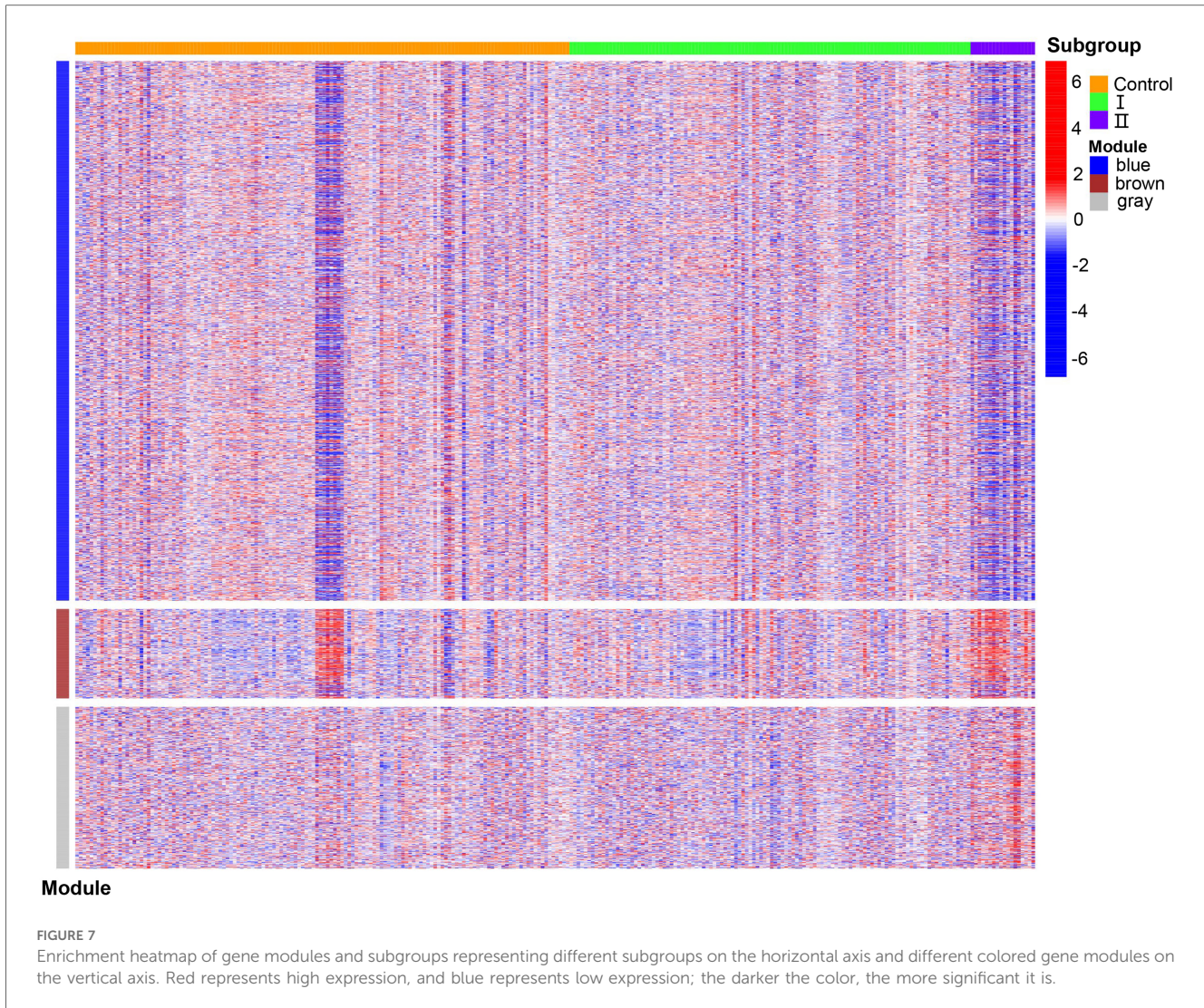
molecular subpopulations of patients with PE is crucial for disease prevention and treatment. To the best of our knowledge, this study is the first to use bioinformatics methods to divide patients with PE into two molecular subgroups and to analyze the subgroups' clinical characteristics, unique DEGs, as well as GO and KEGG enrichment pathways for identifying the feasibility of grouping.

The clinical feature analysis of subgroup I showed that patients in this subgroup had a short gestational week and severe symptoms at the time of onset, which is related to abnormalities in the reactive oxygen species (ROS)-autophagy axis during pregnancy (18). ROS

are products of aerobic energy metabolism and have critical antioxidant activities. Excessive ROS levels can lead to oxidative stress during early pregnancy. Autophagy induces trophoblasts to enter a hypoxic placental environment. Abnormal regulation of the ROS-autophagy axis leads to abnormal autophagy activity, thereby causing PE and intrauterine growth restriction. At the end of pregnancy, the ROS-autophagy interaction changes and is involved in the delivery process. In addition, WGCNA results showed that severe symptoms and lower gestational weeks were significantly positively correlated with the blue module, which was significantly overexpressed in subgroup I; this again indicates that patients in subgroup I had more severe symptoms. DEGs associated with lower gestational weeks and worsening symptoms helped the establishment of subgroup I. A previous study showed no significant correlation between the risk of PE and sex of newborns after analyzing the correlation between them (19). Furthermore, another study on the impact of fetal sex in high-risk premature pregnant women with PE showed no significant relationship between the sex of premature infants and the risk of PE (20). After comparing the sexes of newborns in the subgroups, no significant difference was found between these groups, indicating that the sex of newborns was not an independent factor in distinguishing the subgroups of patients with PE.

GTPase-activating protein 1 containing an IQ motif (*IQGAP1*) is a member of the *IQGAP* protein family and is evolutionarily conserved in eukaryotes. It was first recognized in human osteosarcoma tissues (21). Owing to its similarity in sequence to GTPase-activating protein (*GAP*), it is considered a type of *GAP* that inhibits the inherent GTPase activity of binding partners and stabilizes the active form of the G protein (22). *GAP* also participates in various biological activities, such as cell-cell





adhesion, cytoskeleton dynamics, and cell invasion (23). Our study showed that IQGAP1 was significantly upregulated in subgroup I and had multiple protein pair nodes, indicating its critical role in PE occurrence and development. Placental hypoxia is one of the characteristics of PE. Studies have shown that IQGAP1 is upregulated in macrophages and endothelial cells of hypoxic tissue and can mediate neovascularization by regulating inflammatory cell infiltration and ROS production in hypoxic tissue (24). In hypoxic placental tissues and perivascular areas, the recruitment of endothelial cells or macrophages is impaired, which is related to reduced expression of vascular endothelial growth factor derived from macrophages, affecting angiogenesis and tissue repair in ischemic tissues. IQGAP1-dependent macrophage recruitment plays a crucial role in neovascularization, facilitating post-ischemic and hypoxia blood flow reconstruction and tissue repair. It is also considered a valuable therapeutic target for ischemic vascular diseases (24). In addition, as a scaffold protein for extracellular signal-regulated kinase (ERK), IQGAP1 regulates growth factor-stimulated ERK activity by binding to ERK2, mediates the binding of survival signals to cardiac mast cells, and promotes left ventricular

remodeling after pressure overload (22). These findings suggest that IQGAP1 is a potential biomarker for PE.

In subgroup I, some molecules associated with fetal growth restriction exhibited significant differential upregulation. Phosphorylation of the protein phosphatase 1 regulatory subunit 15B (PPP1R15B) is a conserved cellular response to stress that inhibits overall gene translation and overexpression. This phosphorylation is crucial for embryonic development (25). Cells cannot transition from the G1 phase to the S phase following *PPP1R15B* silencing, leading to decreased proliferation and apoptosis (26). A homozygous mutation in the conserved region of the *PPP1R15B* gene (c.1972G>A; p. Arg658Cys) in twin brothers results in microcephaly, short stature, spinal cord dysplasia, and intellectual disabilities (27). Abdulkarim et al. (28) reported that after the R658C mutation in the *PPP1R15B* gene in two children, a conserved amino acid in the binding region of protein phosphatase-1 was affected, reducing PPP1R15B dephosphorylation and leading to β -cell apoptosis, small head disease, short stature, intellectual disability, and diabetes. Vacuolar protein sorting 4B (VPS4B) is a member of the ATPase protein family and a crucial component of the sorting complex

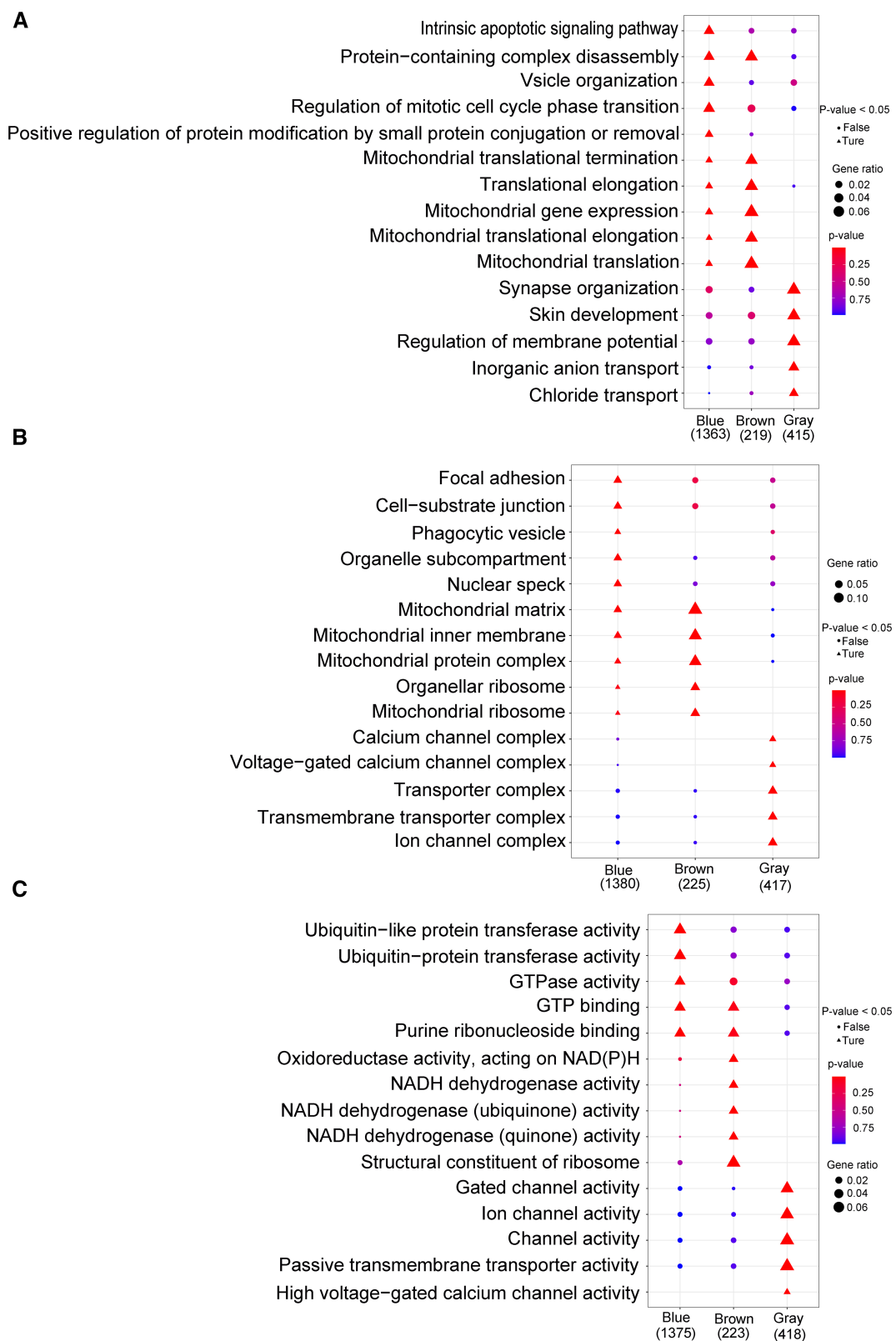


FIGURE 8 Go analysis. (A) Biological process; (B) cellular component; (C) molecular function. The horizontal axis represents different color gene modules, whereas the vertical axis represents terminology. The triangle expression is significantly enriched ($p < 0.05$), whereas the circle expression is not significant ($p > 0.05$); the larger the shape, the more significant the difference.

TABLE 2 Summary of characteristics of the subgroups.

Category	Characteristics
Subgroup I	
Clinical features	Younger gestational age, severe symptoms, induced labor, higher risk of fetal growth restriction or malformation
Unique DEGs	<i>IQGAP1, USP38, ACTR3, RB1CC1, NMD3, PPP1R15B, VPS4B, RANBP6, CLTC, CTSO</i>
Effective therapeutic method	Combined treatment
Subgroup II	
Clinical features	Older gestational age, mild symptoms, spontaneous labor
Unique gene	<i>CL2orf10, CHCHD5, ERAL1, MRPL28, MRPL2, ERCC1, S100A16, TCEB2, NDUFB7, NDUFA11</i>
Effective therapeutic method	Calcium ion blocker, magnesium therapy

that regulates membrane protein internalization and lysosomal degradation (29). VPS4B is closely related to dentin dysplasia pathogenesis (30, 31). VPS4B can participate in cell proliferation and act as a regulator of Wnt- β -catenin signaling in human gingival fibroblasts. The upstream transducer of the catenin signaling pathway promotes cementum formation. Mutations in the β -catenin gene cause a significant decrease in RNA transcription and protein function, leading to the loss of target cell function (30). In mice, homozygous deletion of VPS4B led to embryonic death during early pregnancy (29); this may be because VPS4B knockout, which interferes with RNA transcription in embryos, results in changes in the transcription levels of genes related to apoptosis, cell proliferation, and endocytosis; this affects mouse prenatal death. Clathrin heavy chain (CLTC), a component of clathrin, is a vesicular protein involved in intracellular transport and endocytosis. It is related to human neurological/developmental diseases, with mutations mainly causing phenotypes such as microcephaly, developmental delay, and intellectual impairment (32, 33). Based on these research results, we speculate that patients in subgroup I are more susceptible to fetal growth restriction or developmental abnormalities.

Mitochondria are double-membrane organelles in eukaryotic cells and the main sites of ATP production. They are involved in various cellular processes, including protein and fatty acid synthesis, calcium ion balance regulation, redox reactions, and cell apoptosis (34). Mitochondria are highly sensitive to hypoxic environments and undergo a series of reactions when exposed to adaptation or stress (35). The mother supplies oxygen and nutrition to the fetus via the placenta, which plays a crucial role in pregnancy. The maintenance of placental function depends highly on mitochondria-generated energy. PE originates from placental dysfunction (36). Therefore, mitochondrial abnormalities are the primary cause of placental dysfunction. Mitochondrial damage caused by bioactive factors released by the placenta can lead to endothelial dysfunction and increased maternal blood pressure (37). In subgroup II, NADH: ubiquinone oxidoreductase subunit A11 (NDUFA11) is one of the subunits of mitochondrial respiratory chain complex I, which is crucial for maintaining the balance of mitochondrial energy

metabolism and is involved in regulating mitochondrial respiratory function, cell apoptosis, and oxidative stress response (38). NDUFA11 deficiency is associated with various diseases, including NDUFA11-deficient mitochondrial disease and atherosclerotic disease (39). Another subunit of the mitochondrial respiratory chain complex I, NDUFB7, was significantly upregulated in subgroup II. Correia et al. (40) reported a patient with a dual-allele mutation (c.113-10C>G) in the NDUFB7 intron, which resulted in a significant decrease in NDUFB7 protein function and reduced complex I activity, leading to fetal intrauterine growth restriction, anemia, postpartum hypertrophic cardiomyopathy, and severe lactate acidosis. In addition, another unique gene in subgroup II, *Escherichia coli* Ras protein-like 1 (ERAL1), is also associated with mitochondrial ribosomal components (41). ERAL1 deficiency increases mitochondrial superoxide production, reduces mitochondrial membrane potential, inhibits cell growth, and induces apoptosis. These findings suggest that placental dysfunction is possibly the main target organ of damage in subgroup II patients.

In subgroup II, the GO and KEGG enrichment pathways significantly enriched calcium ion channel activity. Calcium channel blockers dilate small arteries and inhibit coronary vasospasms. The 2021 International Guidelines for the Treatment of Pregnancy-induced Hypertension have identified oral nifedipine (a calcium blocker) as a first-line treatment for PE (42). Magnesium therapy is most widely used to prevent seizures; the underlying mechanisms include blocking calcium ion channels, relaxing blood vessels, and altering neurotransmitter activity, which can significantly reduce maternal blood pressure (1). Therefore, calcium antagonists and magnesium therapy may be more effective for subgroup II patients.

The strengths of our research lie in the analysis of the microarray dataset of PE patients through multiple reliable bioinformatics methods, elucidating possible subtypes of PE patients from a molecular perspective and introducing the characteristics of various subtypes. However, this study has some limitations. First, this study mainly analyzed microarray datasets in public databases; the unique genes of each subgroup need to be validated through cytology, zoology, and even human tissue specimens. Second, the sample size included in this study was limited, and the downloaded patient information lacked clinical features such as blood pressure, urinary protein status, platelet count, and liver and kidney functions; thus, further evaluation of these clinical features in subgroups was impossible. Therefore, before these results are used in clinical practice, they need to be validated in a larger cohort to confirm the robustness of the results. In addition, the construction of molecular subgroups requires a combination of proteomics and metabolomics analysis to improve the accuracy of grouping.

5 Conclusion

In this study, we used transcriptome data to classify PE patients into two subgroups, analyze the differences between

subgroups, and identify the unique DEGs for each subgroup. Our findings provide insights into the diagnosis and personalized treatment of PE.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

Ethics statement

Ethical approval was not required for the study involving humans in accordance with the local legislation and institutional requirements. Written informed consent to participate in this study was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and the institutional requirements. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

HZ: Data curation, Formal Analysis, Investigation, Methodology, Resources, Software, Writing – original draft. JM: Data curation, Formal Analysis, Software, Writing – original draft. XG: Conceptualization, Data curation, Investigation,

Methodology, Project administration, Resources, Software, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcvm.2024.1367578/full#supplementary-material>

References

- Overton E, Tobes D, Lee A. Preeclampsia diagnosis and management. *Best Pract Res Clin Anaesthesiol.* (2022) 36:107–21. doi: 10.1016/j.bpa.2022.02.003
- Macdonald TM, Walker SP, Hannan NJ, Tong S, Kaitu'u-Lino TJ. Clinical tools and biomarkers to predict preeclampsia. *EBiomedicine.* (2022) 75:103780. doi: 10.1016/j.ebiom.2021.103780
- Johnson JD, Louis JM. Does race or ethnicity play a role in the origin, pathophysiology, and outcomes of preeclampsia? An expert review of the literature. *Am J Obstet Gynecol.* (2022) 226:S876–85. doi: 10.1016/j.ajog.2020.07.038
- Goldenberg RL, McClure EM, Macquiere ER, Kamath BD, Jobe AH. Lessons for low-income regions following the reduction in hypertension-related maternal mortality in high-income countries. *Int J Gynaecol Obstet.* (2011) 113:91–5. doi: 10.1016/j.ijgo.2011.01.002
- Shanmugalingam R, Hennessy A, Makris A. Aspirin in the prevention of preeclampsia: the conundrum of how, who and when. *J Hum Hypertens.* (2019) 33:1–9. doi: 10.1038/s41371-018-0113-7
- Zeisler H, Llorba E, Chantraine F, Vatish M, Staff AC, Sennström M, et al. Predictive value of the sFlt-1:plGF ratio in women with suspected preeclampsia. *N Engl J Med.* (2016) 374:13–22. doi: 10.1056/NEJMoa1414838
- Schuster J, Tollefson GA, Zarate V, Agudelo A, Stabila J, Ragavendran A, et al. Protein network analysis of whole exome sequencing of severe preeclampsia. *Front Genet.* (2021) 12:765985. doi: 10.3389/fgene.2021.765985
- Roberts JM, Rich-Edwards JW, McElrath TF, Garmire L, Myatt L, Collaboration GP. Subtypes of preeclampsia: recognition and determining clinical usefulness. *Hypertension.* (2021) 77:1430–41. doi: 10.1161/HYPERTENSIONAHA.120.14781
- Yoodee S, Thongboonkerd V. Bioinformatics and computational analyses of kidney stone modulatory proteins lead to solid experimental evidence and therapeutic potential. *Biomed Pharmacother.* (2023) 159:114217. doi: 10.1016/j.biopha.2023.114217
- Soneson C, Bendel AM, Diss G, Stadler MB. Mutsan—a flexible R package for efficient end-to-end analysis of multiplexed assays of variant effect data. *Genome Biol.* (2023) 24:132. doi: 10.1186/s13059-023-02967-0
- Sun F, Lv H, Feng B, Sun J, Zhang L, Dong B. Identification of natural killer cell-related characteristics to predict the clinical prognosis and immune microenvironment of patients with low-grade glioma. *Aging.* (2023) 15:6264–91. doi: 10.18632/aging.204850
- Whitehead MJ, McCanney GA, Willison HJ, Barnett SC. Myelinj: an ImageJ macro for high throughput analysis of myelinating cultures. *Bioinformatics.* (2019) 35:4528–30. doi: 10.1093/bioinformatics/btz403
- Wang H, Ma X, Li S, Su J, Fan B, Liu B, et al. KIF20A as a potential biomarker of renal and bladder cancers based on bioinformatics and experimental verification. *Aging (Albany NY).* (2023) 15:4714–33. doi: 10.18632/aging.204736
- Sánchez-Baizán N, Ribas L, Piferrer F. Improved biomarker discovery through a plot twist in transcriptomic data analysis. *BMC Biol.* (2022) 20:208. doi: 10.1186/s12915-022-01398-w
- Quan Q, Xiong X, Wu S, Yu M. Identification of immune-related key genes in ovarian cancer based on WGCNA. *Front Genet.* (2021) 12:760225. doi: 10.3389/fgene.2021.760225
- Hein DM, Deng W, Bleile M, Kazmi SA, Rhead B, De La Vega FM, et al. Racial and ethnic differences in genomic profiling of early onset colorectal cancer. *J Natl Cancer Inst.* (2022) 114:775–8. doi: 10.1093/jnci/djac014
- Sherman BT, Hao M, Qiu J, Jiao X, Baseler MW, Lane HC, et al. David: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res.* (2022) 50:W216–21. doi: 10.1093/nar/gkac194

18. De Andrade Ramos BR, Witkin SS. The influence of oxidative stress and autophagy cross regulation on pregnancy outcome. *Cell Stress Chaperones*. (2016) 21:755–62. doi: 10.1007/s12192-016-0715-3
19. Callais NA, Morgan JA, Leggio BA, Wang Y. Association of fetal gender and the onset and severity of hypertensive disorders of pregnancy. *J Matern Fetal Neonatal Med*. (2022) 35:2260–5. doi: 10.1080/14767058.2020.1783654
20. Teoh PJ, Ridout A, Seed P, Tribe RM, Shennan AH. Gender and preterm birth: is male fetal gender a clinically important risk factor for preterm birth in high-risk women? *Eur J Obstet Gynecol Reprod Biol*. (2018) 225:155–9. doi: 10.1016/j.ejogrb.2018.04.025
21. Abel AM, Schuldt KM, Rajasekaran K, Hwang D, Riese MJ, Rao S, et al. IQGAP1: insights into the function of a molecular puppeteer. *Mol Immunol*. (2015) 65:336–49. doi: 10.1016/j.molimm.2015.02.012
22. Wei T, Lambert PF. Role of IQGAP1 in carcinogenesis. *Cancers (Basel)*. (2021) 13:3940. doi: 10.3390/cancers13163940
23. Monteleon CL, McNeal A, Duperret EK, Oh SJ, Schapira E, Ridky TW. IQGAP1 and IQGAP3 serve individually essential roles in normal epidermal homeostasis and tumor progression. *J Invest Dermatol*. (2015) 135:2258–65. doi: 10.1038/jid.2015.140
24. Du Y, Hu Y, Wen N, Fu S, Zhang G, Li L, et al. Abnormal expression of TGFBR2, EGF, LRP10, and IQGAP1 is involved in the pathogenesis of coronary artery disease. *Rev Cardiovasc Med*. (2021) 22:947–58. doi: 10.31083/j.rcm2203103
25. Kloft N, Neukirch C, von Hoven G, Bobkiewicz W, Weis S, Boller K, et al. A subunit of eukaryotic translation initiation factor 2alpha-phosphatase (CreP/PPP1R15B) regulates membrane traffic. *J Biol Chem*. (2012) 287:35299–317. doi: 10.1074/jbc.M112.379883
26. Shahmoradgoli M, Riazalhosseini Y, Haag D, Becker N, Hovestadt V, Heck S, et al. Protein phosphatase 1, regulatory subunit 15B is a survival factor for ERalpha-positive breast cancer. *Int J Cancer*. (2013) 132:2714–9. doi: 10.1002/ijc.27945
27. Kernohan KD, Tétreault M, Liwak-Muir U, Geraghty MT, Qin W, Venkateswaran S, et al. Homozygous mutation in the eukaryotic translation initiation factor 2alpha phosphatase gene, PPP1R15B, is associated with severe microcephaly, short stature and intellectual disability. *Hum Mol Genet*. (2015) 24:6293–300. doi: 10.1093/hmg/ddv337
28. Abdulkarim B, Nicolino M, Igoillo-Estevé M, Daures M, Romero S, Philippi A, et al. A missense mutation in PPP1R15B causes a syndrome including diabetes, short stature, and microcephaly. *Diabetes*. (2015) 64:3951–62. doi: 10.2337/db15-0477
29. Chen D, He F, Lu T, Huang J, Li M, Cai D, et al. VPS4B deficiency causes early embryonic lethality and induces signal transduction disorders of cell endocytosis. *Genesis*. (2021) 59:e23415. doi: 10.1002/dvg.23415
30. Yang Q, Chen D, Xiong F, Chen D, Liu C, Liu Y, et al. A splicing mutation in VPS4B causes dentin dysplasia I. *J Med Genet*. (2016) 53:624–33. doi: 10.1136/jmedgenet-2015-103619
31. Li Q, Lu F, Chen T, Zhang K, Lu Y, Li X, et al. VPS4B mutation impairs the osteogenic differentiation of dental follicle cells derived from a patient with dentin dysplasia type I. *Int J Oral Sci*. (2020) 12:22. doi: 10.1038/s41368-020-00088-z
32. Nabais Sá MJ, Venselaar H, Wiel L, Trimouille A, Lasseaux E, Naudion S, et al. De novo CLTC variants are associated with a variable phenotype from mild to severe intellectual disability, microcephaly, hypoplasia of the corpus callosum, and epilepsy. *Genet Med*. (2020) 22:797–802. doi: 10.1038/s41436-019-0703-y
33. Demari J, Mroske C, Tang S, Nimeh J, Miller R, Lebel RR. CLTC as a clinically novel gene associated with multiple malformations and developmental delay. *Am J Med Genet A*. (2016) 170A:958–66. doi: 10.1002/ajmg.a.37506
34. Roca-Portoles A, Tait SWG. Mitochondrial quality control: from molecule to organelle. *Cell Mol Life Sci*. (2021) 78:3853–66. doi: 10.1007/s00018-021-03775-0
35. Eisner V, Picard M, Hajnóczky G. Mitochondrial dynamics in adaptive and maladaptive cellular stress responses. *Nat Cell Biol*. (2018) 20:755–65. doi: 10.1038/s41556-018-0133-0
36. Hu XQ, Zhang L. Mitochondrial dysfunction in the pathogenesis of preeclampsia. *Curr Hypertens Rep*. (2022) 24:157–72. doi: 10.1007/s11906-022-01184-7
37. Hu XQ, Zhang L. Hypoxia and mitochondrial dysfunction in pregnancy complications. *Antioxidants (Basel)*. (2021) 10:405. doi: 10.3390/antiox10030405
38. Amate-García G, Ballesta-Martínez MJ, Serrano-Lorenzo P, Garrido-Moraga R, González-Quintana A, Blázquez A, et al. A novel mutation associated with neonatal lethal cardiomyopathy leads to an alternative transcript expression in the X-linked complex I NDUFB11 gene. *Int J Mol Sci*. (2023) 24:1743. doi: 10.3390/ijms24021743
39. Yang Y, Li JH, Yao BC, Chen QL, Jiang N, Wang LQ, et al. NDUFB11 and NDUFS3 play a role in atherosclerosis and chronic stress. *Aging (Albany, NY)*. (2023) 15:8026–43. doi: 10.18632/aging.204947
40. Correia SP, Moedas MF, Naess K, Bruhn H, Maffezzini C, Calvo-Garrido J, et al. Severe congenital lactic acidosis and hypertrophic cardiomyopathy caused by an intronic variant in NDUFB7. *Hum Mutat*. (2021) 42:378–84. doi: 10.1002/humu.24173
41. Uchiyama T, Ohgaki K, Yagi M, Aoki Y, Sakai A, Matsumoto S, et al. ERAL1 is associated with mitochondrial ribosome and elimination of ERAL1 leads to mitochondrial dysfunction and growth retardation. *Nucleic Acids Res*. (2010) 38:5554–68. doi: 10.1093/nar/gkq305
42. Magee LA, Brown MA, Hall DR, Gupte S, Hennessy A, Karumanchi SA, et al. The 2021 international society for the study of hypertension in pregnancy classification, diagnosis & management recommendations for international practice. *Pregnancy Hypertens*. (2022) 27:148–69. doi: 10.1016/j.preghy.2021.09.008