BMC Genomics

RESEARCH Open Access

Bioinformatics methods in biomarkers of preeclampsia and associated potential drug applications

Ying Peng[1,2†], Hui Hong[2†], Na Gao[3], An Wan[2] and Yuyan Ma[3*]

Abstract

Background: Preeclampsia is a pregnancy-related condition that causes high blood pressure and proteinuria after 20 weeks of pregnancy. It is linked to increased maternal mortality, organ malfunction, and foetal development limitation. In this view, there is a need critical to identify biomarkers for the early detection of preeclampsia. The objective of this study is to discover critical genes and explore medications for preeclampsia treatment that may influence these genes.

Methods: Four datasets, including GSE10588, GSE25906, GSE48424 and GSE60438 were retrieved from the Gene Expression Omnibus database. The GSE10588, GSE25906, and GSE48424 datasets were then removed the batch effect using the "sva" R package and merged into a complete dataset. The differentially expressed genes (DEGs) were identified using the "limma" R package. The potential small-molecule agents for the treatment of PE was further screened using the Connective Map (CMAP) drug database based on the DEGs. Further, Weight gene Co-expression network (WGNCA) analysis was performed to identified gene module associated with preeclampsia, hub genes were then identified using the logistic regression analysis. Finally, the immune cell infiltration level of genes was evaluated through the single sample gene set enrichment analysis (ssGSEA).

Results: A total of 681 DEGs (376 down-regulated and 305 up-regulated genes) were identified between normal and preeclampsia samples. Then, Dexamethasone, Prednisone, Rimexolone, Piretanide, Trazodone, Buflomedil, Scoulerin, Irinotecan, and Camptothecin drugs were screened based on these DEGs through the CMAP database. Two modules including yellow and brown modules were the most associated with disease through the WGCNA analysis. KEGG analysis revealed that the chemokine signaling pathway, Th1 and Th2 cell differentiation, B cell receptor signalling pathway and oxytocin signalling pathway were significantly enriched in these modules. Moreover, two key genes, PLEK and LEP were evaluated using the univariate and multivariate logistic regression analysis from the hub modules. These two genes were further validated in the external validation cohort GSE60438 and qRT-PCR experiment. Finally, we evaluated the relationship between immune cell and two genes.

Conclusion: In conclusion, the present study investigated key genes associated with PE pathogenesis that may contribute to identifying potential biomarkers, therapeutic agents and developing personalized treatment for PE.

†Ying Peng and Hui Hong are contributed equally to this work.

*Correspondence: yuyanma@sdu.edu.cn

[3] Department of Obstetrics and Gynecology, Qilu Hospital of Shandong University, Jinan Shandong 250012, China
Full list of author information is available at the end of the article

Keywords: Preeclampsia, Bioinformatics analysis, Differentially expressed genes, Immune mechanism, Molecular markers, Potential drugs

Introduction

Preeclampsia (PE) is a pregnancy disorder that causes high blood pressure and proteinuria [1] after 20 weeks of pregnancy [2].It affects 3%-10% of all pregnancies [3–5], and is associated with substantial maternal morbidity, mortality, organ dysfunction, iatrogenic premature delivery [6], and foetal development restriction [7]. Preeclampsia affects the brain development and function of the offspring, increasing the risk of intellectual disability [8], epilepsy [9], autism [10], and schizophrenia [11, 12]. Women with a history of PE are more likely to develop cardiovascular disease or hypertension,[5] resulting in financial and psychological consequences on the family and society. There is no obvious therapeutic intervention for PE today; the only effective treatment is pregnancy, which might result in low birth weight and long-term detrimental health implications for infants [13, 14]. Early assessment of PE risk is critical for high-risk pregnant women because it allows for the implementation of preventive strategies to lower the incidence of PE and improve maternal and infant outcomes.

However, the etiology and pathogenesis of PE are yet unknown. Several hypotheses such as maternal-foetal (paternal) immune maladjustment [15] and inflammatory cytokine disorders [16], all suggest that PE is associated with a number of risk factors, including obesity, hypertension, diabetes, oxidative stress, foetal rejection, genetic polymorphism inheritance [17], and trophoblast insufficiency [18, 19]. The main causes are identified as immunological intolerance [20, 21] and angiogenesis imbalance, and numerous research has demonstrated the immune mechanism of PE development. The latter causes PE by generating an imbalance in immune tolerance at the mother-infant interaction [22, 23]. The pathophysiological basis of PE is shallow placental implantation. The mutual adaptation of villous trophoblast cells and the mother's immune system is required for effective placental development. The poor placental formation will result in acute-like graft rejection disease under the influence of specific immunological factors generating immune intolerance in the mother and child.

Immune system modifications have been widely recognized as the key determinants of PE [24], which is a systemic inflammatory response resulting in an imbalance between placental substances and the corresponding adaptation of the mother[25, 26]. Shah et al. [27] in their study compared the expression of CD66B, nuclear factor NF-κB, and cyclooxygenase-2 (COX-2) in the subcutaneous fat of women with PE ($n=7$), normal pregnant women ($n=6$), and normal non-pregnant women ($n=5$). The percentages of CD66B, NF-κB, and COX-2 in PE patients were significantly higher than in normal non-pregnant or normal pregnant patients. Moreover, Xu et al. [28]. discovered five hub genes that are associated with immune cells in the immunologic microenvironment at the maternal-foetal interface.

An increasing number of researchers are actively pursuing molecular markers using data mining and analysis of databases to the diagnosis and treatment of PE. However, most recent studies using single dataset to identify the key genes and lack of external validation cohort, which may induce a unstable and unreliable result. For example, Liu et al. using the GSE60438 dataset to identify 17 hub genes through the differentially expressed analysis, protein–protein network and svm analysis, while lack of external cohort [29]. Lin et al. using the GSE48424 dataset to identify three key genes (HDC, MS4A2 and SLC18A2) associated with PE through PPI network analysis. Although they applied GSE149437 to verify the predictive value of the above these genes, the result still lack reliability [30]. Moreover, Wang et al. explored the immune cell infiltration in PE, however, the dataset that used in the study is relatively small, and lack of the stability [31].

In this study, we aimed to investigate the potential biomarkers, immune cell infiltration and drug in PE via integrative multiple datasets. These results were further validated using the qRT-PCR experiment. Our findings might provide the novel biomarkers for the prediction and diagnosis of PE.

Methods

Selection of the GEO dataset and data processing

The gene expression microarray datasets GSE10588, GSE25906, GSE48424, and GSE60438 were downloaded from the GEO database (http://www.ncbi.nlm.nih.gov/geo/). GSE10588 contains 26 normal placental tissues and 17 preeclampsia placental tissues [32], GSE25906 contains 37 normal placental tissues and 23 preeclampsia placental tissues [33], GSE48424 contains 18 normal blood samples and 18 preeclampsia blood samples [34], and GSE60438 contains 42 normal samples and 35 PE samples [35]. To verify data reliability, we used the combat function in the 'SVA' R package to eliminate the batch effect from three datasets (GSE10588, GSE25906, and GSE48424) and combined these three datasets (Merge

dataset). The distribution of datasets before and after the merger was then evaluated using principal component analysis. The dataset GSE60438 was selected for external validation [35]. The detail pipeline of this study was showed in Fig. 1.

Differential expression analysis

We used the 'Limma' R package (version 4.0; http://www.bioconductor.org/packages/release/bioc/html/limma.html) to discover differential genes between normal and PE samples in a combined dataset of 81 normal samples and 58 PE samples. For the screening criterion, corrected *P*-values less than 0.05 and $\log_2$ |Fold Change| > 0 were considered statistically significant.

CMAP drug database analysis

To find potential small-molecule agents for the treatment of PE, the top 1000 up-regulated and down-regulated DEGs were uploaded to the Connective Map database (CMAP) (http://www.broadinstitute.org/cmap/).

Co-expression network analysis of weighted genes

The "WGCNA" R package was used for co-expression analysis. The genes with the largest variance (25%) were firstly selected to assure the heterogeneity and then filtered to remove outliers using sample clustering methods. The co-expression similarity matrix included absolute values of correlations between transcription expression levels. A Pearson correlation matrix was constructed for matching genes. We used the power function amn = | key | beta (key = the Pearson correlation between gene m and n; AMN = adjacency between gene M and gene N) to construct a weighted adjacency matrix. The parameter β highlighted genes with significant correlations while penalizing those with weak correlations. Following that, the appropriate β value was selected to perform the similarity matrix and investigate the scale-free co-expression network. The adjacency matrix was then transformed into a topological overlap matrix (TOM), which calculated the network connectedness of the genes in the topology matrix as the sum of the adjacent genes generated by all the other networks. Based on TOM similarity data, average linkage level clustering was performed, and the minimum size of the gene tree (genome) was selected 30. The gene significance (GS) was calculated to quantify the correlation between genes and sample traits to determine the value of each module. Module eigengenes (MEs) were regarded as the major components in the principal component analysis of each gene module, and the expression patterns of all genes could be summed as a single typical expression profile within a specific module. The $\log_{10}$ conversion of

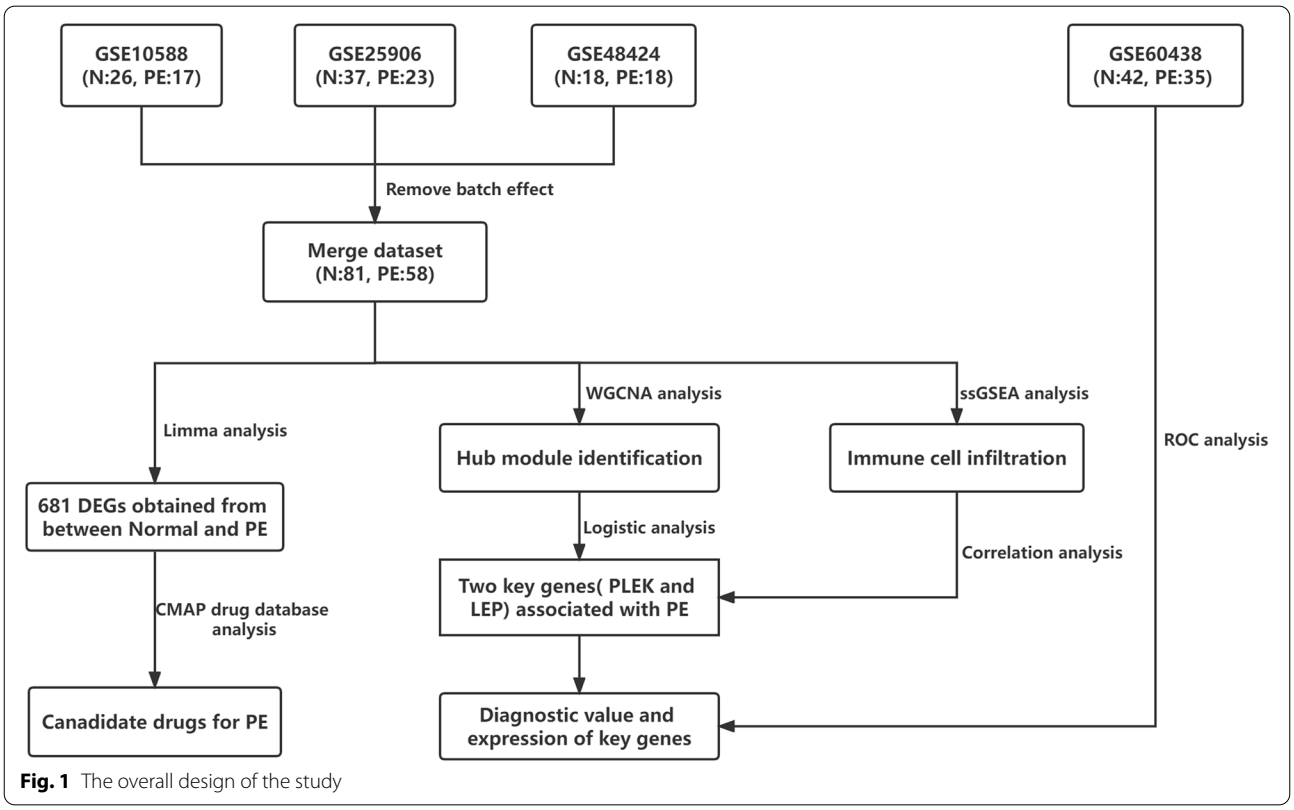

Fig. 1 The overall design of the study

p-value in linear regression between gene expression and clinical data (GS = lgP) was designated as GS. Modular significance (MS) was defined as the average GS within a module, reflecting the correlation between the module and the sample features. We established a threshold value (< 0.25) to increase the capacity of the modules by merging several modules of comparable height. Normal and PE samples were selected as clinical phenotypes. The gene modules were then examined alongside the clinical phenotypes. After identifying the relevant modules, we calculated GS and MM values (module membership [MM] is the correlation between the genes of the module and gene expression profile) and set the threshold values for each key gene.

GO and KEGG enrichment analysis

The R package of 'clusterProfiler' to perform Gene ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) [36–38]enrichment analysis to uncover the function of genes and potential pathways in the module. Furthermore, the screening criteria for significant functions and pathways were adjusted p-value less than 0.05.

Logistic regression

To identified key genes associated with PE, we firstly constructed PPI network based on the hub modules. We then using the MCODE algorithm and logistic regression analysis to screened key genes from the PPI network. The receiver operating characteristic (ROC) curve analysis was applied to the accuracy of the diagnostic model from the result of logistic regression analysis in the merge dataset and external validation dataset.

Immune infiltration analysis

The immune cell gene sets were retrieved from previous related literature, and then used to assess the immune cell infiltration between normal and PE samples [28]. The function of immune-related gene sets was acquired to enrich gene sets and calculate the concentration of each sample score, standardized between 0 and 1, using the single-sample Gene Set Enrichment Analysis (ssGSEA) algorithm [39].

Collection of human tissue specimens

A total of 60 placenta specimens were collected between August 2021 and January 2022, 30 specimens from patients with PE and 30 from healthy pregnant women at the The First Affiliated Hospital of USTC, Division of Life Sciences and Medicine, University of Science and Technology of China (Anhui Provincial Hospital) (Hefei, China). Written informed consent was obtained from all participants. This study was approved by The First Affiliated Hospital of USTC, Division of Life Sciences and Medicine, University of Science and Technology of China (Anhui Provincial Hospital) research ethics board (NO:2021KY161) and is based on the ethical requirements of the Helsinki Declaration. All participants have the right to know.

Detection of the mRNA expression of the hub genes by RT-PCR

Tissue RNA was extracted using TRIzol® (Invitrogen; Thermo Fisher Scientific, Inc.) to assess the expression of important genes in the placentas of PE patients and healthy pregnant women. Total RNA was reverse transcribed into cDNA at 42˚C for 30 min using PrimeScript™ RT Master Mix (Takara Biotechnology Co., Ltd.). Subsequently, qPCR was performed on a CFX96™ Real-Time PCR Detection system (Bio-Rad Laboratories, Inc.) using TB Green™ Fast qPCR mix. The 2-ΔΔCq method was used to quantify the relative mRNA levels for the gene expression analysis [40]. The values were adjusted to endogenous GAPDH expression, and each experiment was repeated three times separately. Critical Technologies: GAPDH sense, 5′-GAAGGTGAAGGTCGGAGT CAA-3′; antisense 5′-CTGGAAGATGGTGATGGG ATTT-3′; PLEK sense, 5′-AGATGCCTGGGTTCGGGA TA-3′; antisense, 5′-GGTTTCTGGCAGTCGAATGGA-3′; LEP sense, 5′-GCTGTGCCCATCCAAAAAGT-3′; antisense, 5′-CCAGGAATGAAGTCCAAACCG-3′. The thermocycling conditions were as follows: 95˚C for 30 s, followed by 40 cycles of 95˚C for 5 s and 60˚C for 30 s.

Statistical analysis

All RT-PCR experimental data were analyzed using GraphPad Prism 7.0. The expression level of each hub gene was expressed as a fold change using the $2^{-\Delta\Delta C}$ method. $P < 0.05$ was considered significant.

Results

Identification of differential genes in PE

GSE10588, GSE25906, and GSE48424 datasets were pooled. We obtained a dataset with 81 normal samples and 58 PE samples after removing the batch effect. As shown in Figs. 2A and B, the datasets were discrete before merging. while the batch effect was removed, we discovered that the datasets were merged, showing that the batch effect was successfully erased. By comparing normal and PE samples, we found 376 down-regulated genes and 305 up-regulated genes (Figs. 3A and B). Figure 3A depicts up-down-regulated differential genes; down-regulated genes are denoted by green and the up-regulated genes are denoted by red point. Moreover, a heat map of the differentially regulated genes is shown in Fig. 3B.

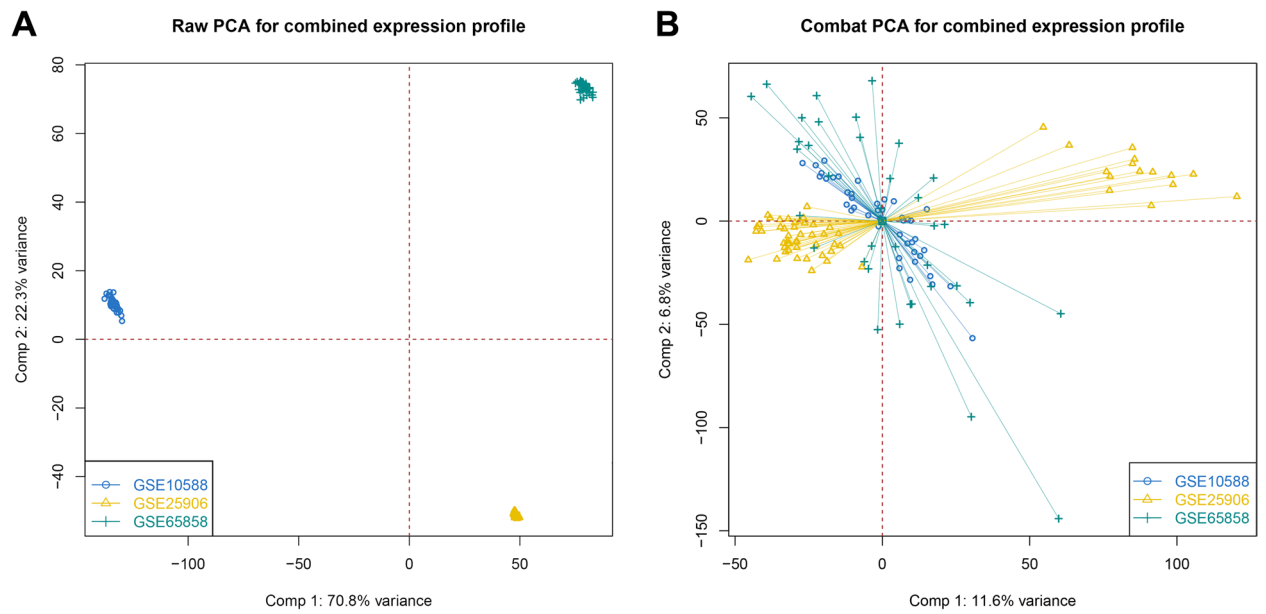

Fig. 2 Evaluation of the batch effect before (**A**) and after (**B**) merging through the principal component analysis

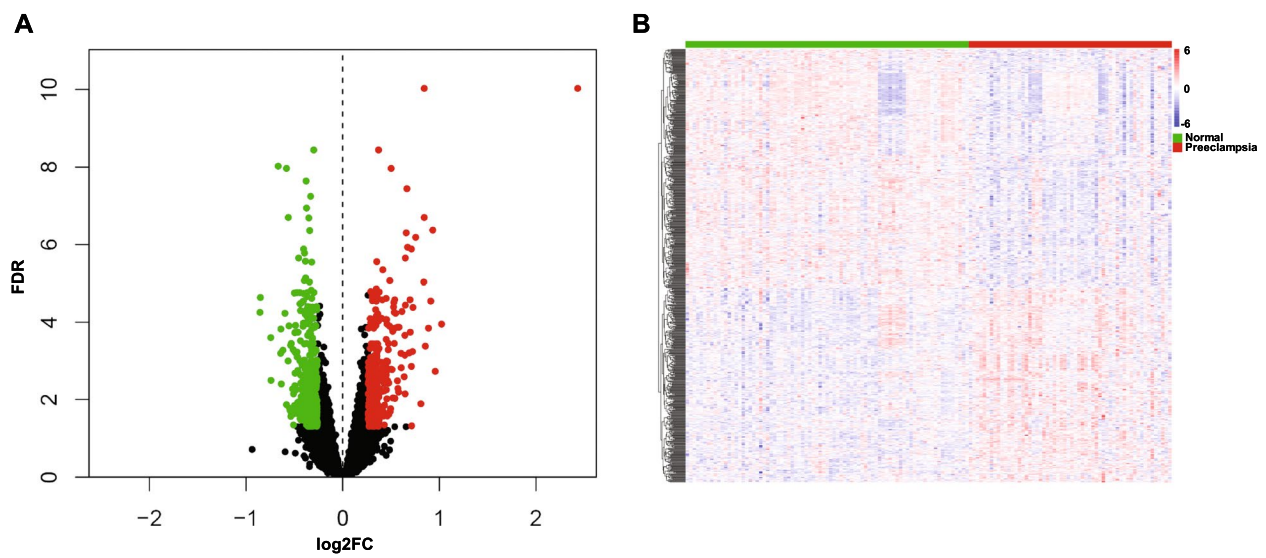

Fig. 3 Identification of differentially expressed genes in the merged dataset. **A** Volcano plot of the genes, the green dots represent the down-regulated genes, red dots represent the up-regulated genes, while the black dots showed genes with no significance. **B** A heatmap plot of the differentially expressed genes

Identification of potential drugs for PE

Based on these DEGs, we looked into potential drug targeting pathways in the CMAP database. The analyses of the mechanism of action in the CMAP database revealed 39 mechanisms of action of 51 drugs. As illustrated in Fig. 4, dexamethasone, prednisone, rimexolone, and piretanide interacted via the glucocorticoid receptor agonist, whereas Trazodone, buflomedil, and scoulerine interacted via the adrenergic receptor antagonist. The co-action mechanism of irinotecan, camptothecin, and doxorubicin, on the other hand, was via topoisomerase inhibitor. These results may provide new insights for the treatment of PE.

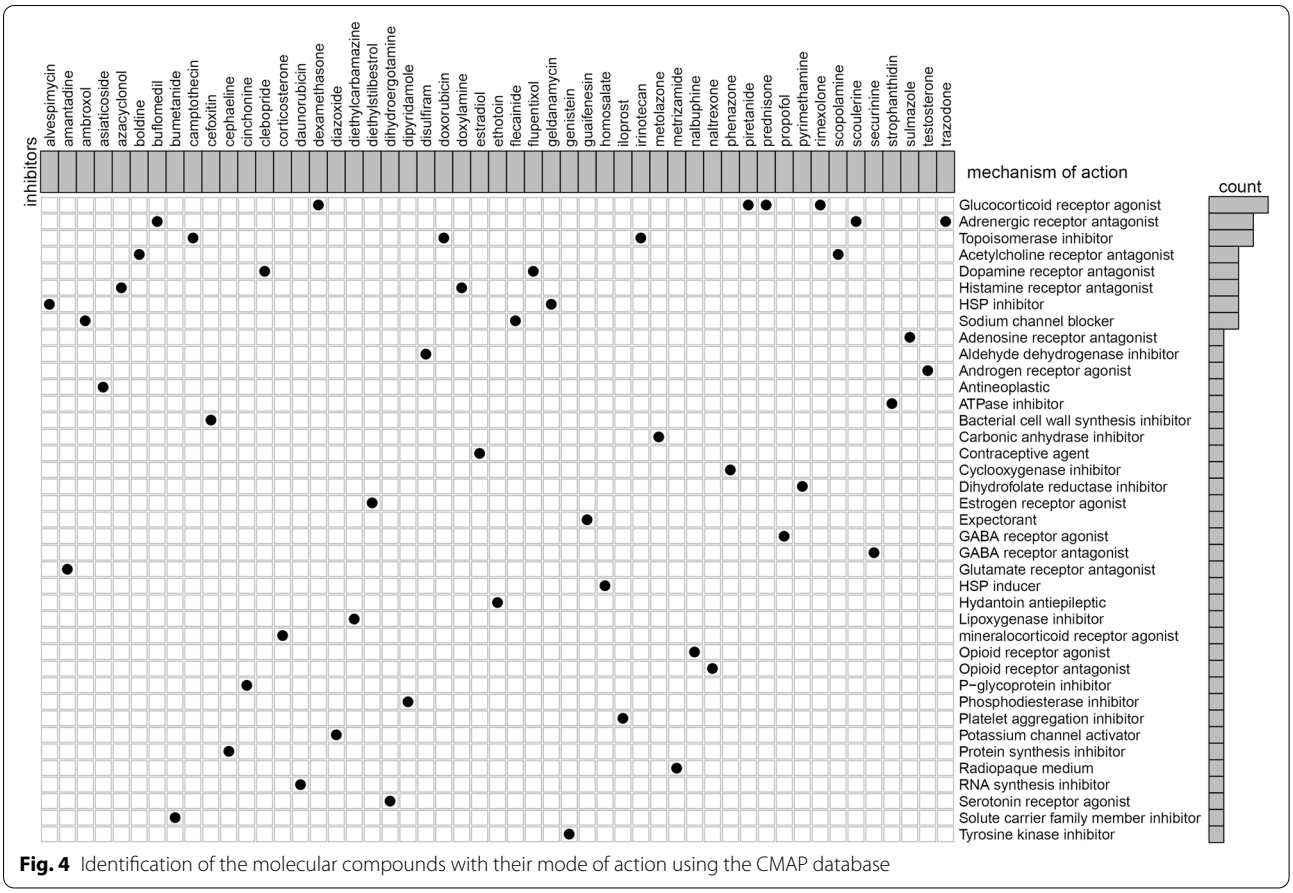

Fig. 4 Identification of the molecular compounds with their mode of action using the CMAP database

Modular identification and key gene screening for PE

WGCNA was used to construct gene co-expression networks to discover biologically essential gene modules and better understand genes linked with clinical traits. We selected the genes with the largest variance (25%) to the WGCNA analysis. In this investigation, $\beta = 5$ (scale-free $R^2 = 0.90$) was selected as the soft threshold to establish a scale-free network. We then applied average-linkage hierarchical clustering method to cluster these genes and six module were obtained, of which yellow and brown modules were the most associated with disease (Fig. 5A-B). We further used MS as the overall gene expression level to assess the correlation with clinical phenotypes. The Fig. 5C and D showed that the correlation between gene significance and clinical trait in yellow and brown modules. According to the GO and KEGG enrichment analyses, the genes in the yellow module were significantly enriched in the NADP metabolic process, cellular aldehyde metabolic process, and other processes (Fig. 6A). Chemokine signalling pathway, Th1 and Th2 cell differentiation, B cell receptor signalling pathway, and oxytocin signalling pathway were all significantly enriched in KEGG enrichment (Fig. 6B). In the brown module, GO enrichment analysis revealed that the genes were significantly enriched in regulation pf lipid metabolic process, aging, cytokine activity, negative regulation of B cell activation, negative regulation of leukocyte activation (Fig. 7A), and cytokine – cytokine receptor interaction, HIF- 1 signalling pathway, TGF—beta signalling pathway, and leukocyte trans-endothelial migration were enriched in KEGG pathway (Fig. 7B). We then constructed a protein–protein interaction network (PPI) from the hub module through the String database and further visualized using the Cytoscape software. Figures 8A and 8B depict the interaction relationship between genes in the yellow and brown modules, respectively. Finally, we used the Maximal Clique Centrality (MCC) algorithm from the cytoscape software to identify 15 hub genes, including LCP1, IGSF6, TLR8, HCK, SLA, DOCK2, RAC2, LCP2, PLEK, CD53, SERPINE1, ENG, ANG, LEP, and FLT1. To restrict the spectrum of core genes, a logistic univariate and multivariate regression analyses were used to construct models associated with PE. Finally, two important genes, PLEK and LEP, were retained with $P < 0.05$ (Table 1). According to ROC analysis, the area under the curve (AUC) value of this model

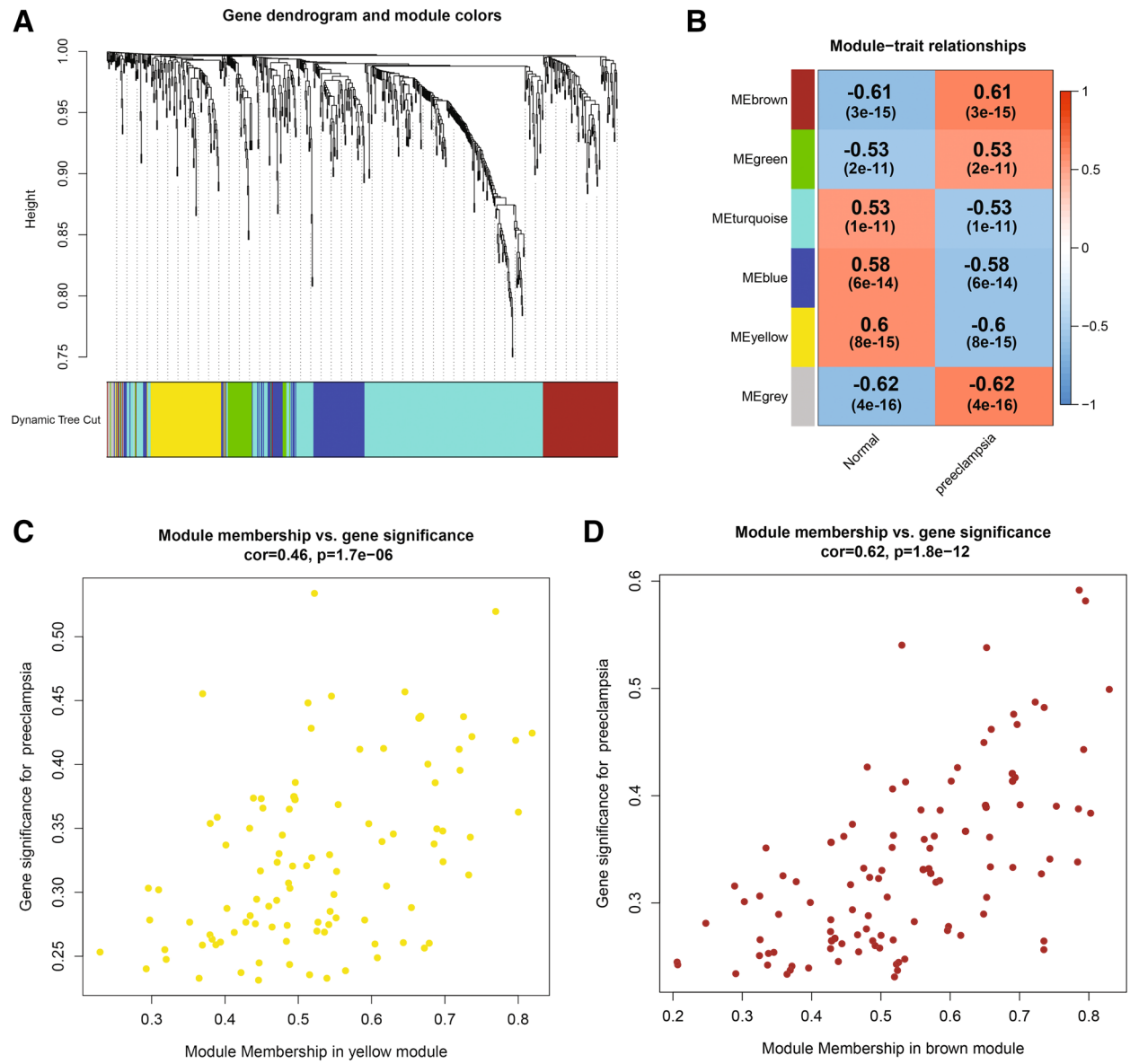

Fig. 5 Identification of hub modules using the WGCNA analysis. **A** The dendrogram of the co-expression network was clustered based on the dissimilarity. **B** The module-trait heatmap showing the correlation between module eigengenes (ME) and traits. **C** Scatter plot of the red module. **D** Scatter plot of the brown module

on the merging dataset and external validation dataset achieved 0.923 (Fig. 9A) and 0.670 (Fig. 9B), respectively. PLEK was shown to be significantly expressed in the pooled dataset, although in normal tissues rather than PE tissues (Fig. 10A). However, PLEK was not significantly expressed in the external validation set ($P=0.065$), which could be attributed to the limited PE sample size in the external dataset (Fig. 10C). Interestingly, the LEP gene was significantly expressed in both the merged dataset and the external validation set, but its expression in normal tissues was lower than in PE tissues (Figs. 10B and D).

Immune infiltration analysis

The immune enrichment value of each sample was calculated using the ssGSEA algorithm, and the wilcoxon-rank test was used to evaluated the immune cell infiltration level between normal samples and PE samples (Fig. 11). In addition, the relationship between PLEK, LEP, and immune infiltration was investigated.

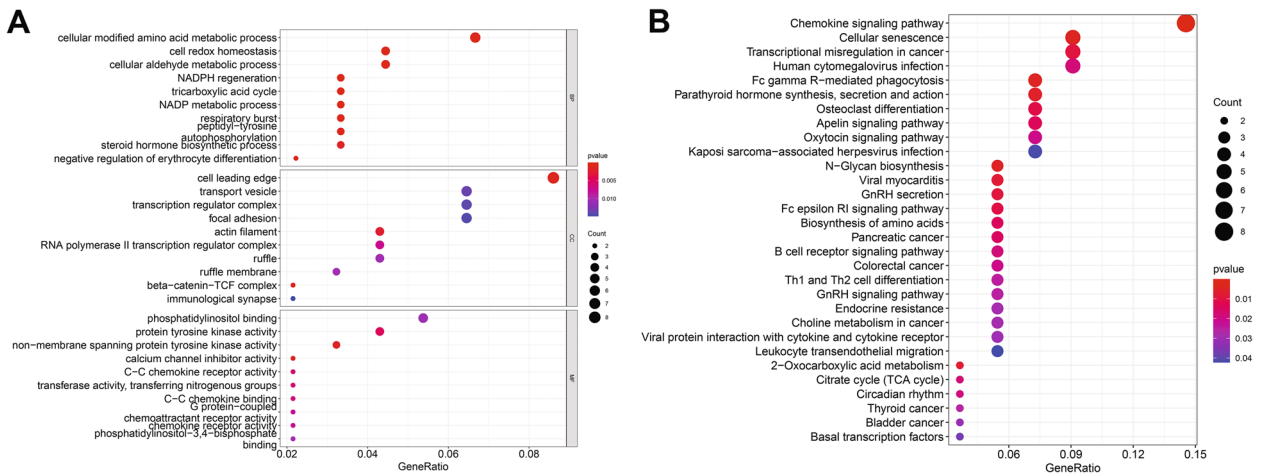

Fig. 6 Gene ontology enrichment (**A**) and KEGG (**B**) enrichment analysis were performed in the yellow modules

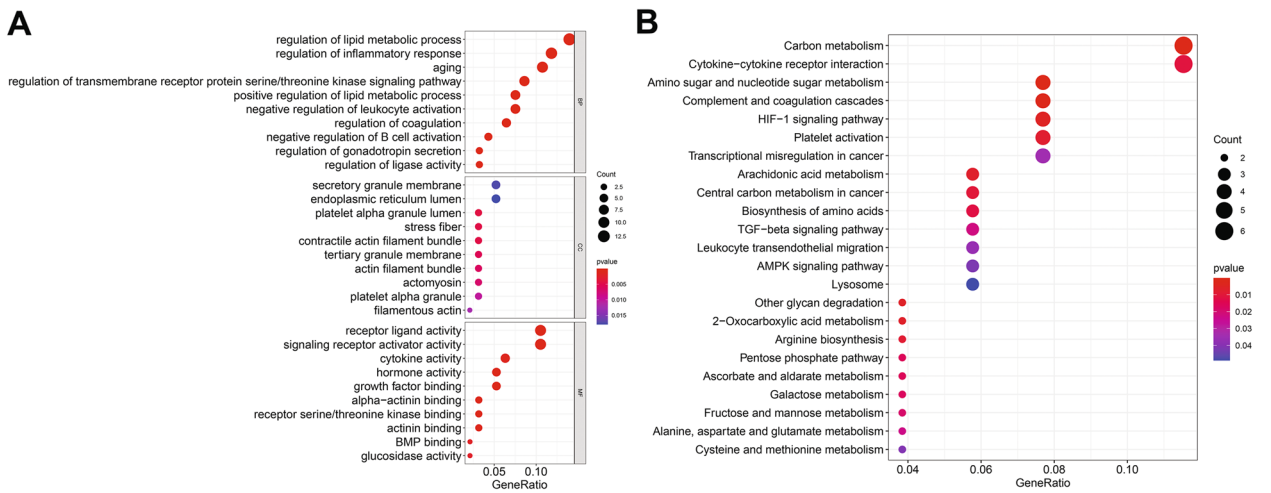

Fig. 7 Gene ontology enrichment (**A**) and KEGG (**B**) enrichment analysis were performed in the brown modules

We discovered a substantial link between PLEK CCR, TIL, and other immune functions, implying that PLEK may play an important role in immunity, whereas LEP was significantly positively correlated with Th2 cells and negatively correlated with Th1 cells. This suggested that LEP was closely linked to T cell function (Figs. 12A and B).

RNA extraction and quantitative real-time PCR

As illustrated in Fig. 13, the PLEK expression levels were significantly lower in the PE samples as compared to normal tissues, while LEP expression levels were significantly higher in normal samples, which was consistent with our analysis results (Figs. 13A and B).

Discussion

This study employed a multi-step integrated bioinformatics analysis of the chip data to identify the hub gene as a biomarker for PE. 376 down-regulated and 305 up-regulated DEGs were firstly identified using the gene expression profiles of the GSE10588, GSE25906 and GSE48424 datasets. According to these DEGs, we further evaluated prospective drugs for PE treatment through the CMAP database. The analysis of the action mode of 51 compounds revealed that the interaction mechanism for prednisone, dexamethasone, rimexolone, and piretanide occurred via the glucocorticoid receptor agonist. Trazodone, buflomedil, and scoulerine interacted via the adrenergic receptor antagonist, whereas the co-action mechanism of irinotecan,

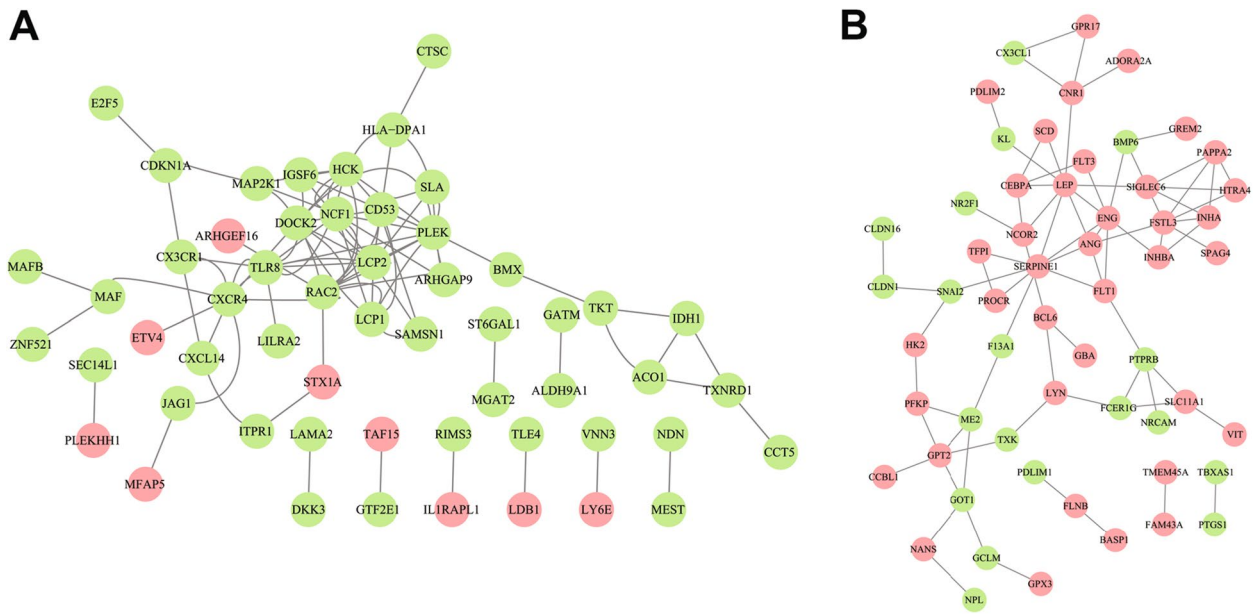

Fig. 8 Protein–protein networks of the hub genes in the red (**A**) and brown (**B**) modules

Table 1 Univariate and multivariate logistic regression

NA	Univariate analysis	NA	Multivariate analysis	NA
	OR (95% CI)	*P* value	*P* value	*P* value
LCP1	0.145 (0.059–0.355)	< 0.001	NA	
IGSF6	0.365 (0.205–0.650)	< 0.001	NA	
TLR8	0.303 (0.141–0.649)	0.002	NA	
HCK	0.257 (0.116–0.567)	< 0.001	NA	
SLA	0.325 (0.158–0.666)	0.002	NA	
DOCK2	0.340 (0.165–0.697)	0.003	NA	
RAC2	0.324 (0.155–0.680)	0.003	NA	
LCP2	0.335 (0.163–0.688)	0.003	NA	
PLEK	0.104 (0.039–0.278)	< 0.001	4.100000e-02 (3.000000e-03–5.150000e-01)	0.013
CD53	0.219 (0.102–0.471)	< 0.001	NA	
SERPINE1	2.040 (1.253–3.321)	0.004	NA	
ENG	4.949 (2.678–9.146)	< 0.001	NA	
ANG	3.438 (1.766–6.691)	< 0.001	NA	
LEP	2.228 (1.698–2.924)	< 0.001	2.524000e + 00 (1.568000e + 00–4.065000e + 00)	0.000
FLT1	2.083 (1.427–3.040)	< 0.001	NA	

Note:

camptothecin, and doxorubicin occurred via topoisomerase inhibitor. Our study identified drugs for PE and may suggest therapeutic targets for future research. However, there is a need to deeply understand whether these treatments are effective in treating PE to provide safe outcomes for the both mother and the foetus.

With the development of the high-through sequencing, a large number of disease-related biomarkers have been identified. Non-coding RNAs play a critical role in the development and biological process of disease [41–43]. Chen et al. reviewed and developed a numerous of computational models to predict the potential non coding RNAs–disease associations. For example, a NCMCMDA and DBNMDA model which applied to predict the potential miRNA–disease associations, showed a high performance than many previous computational

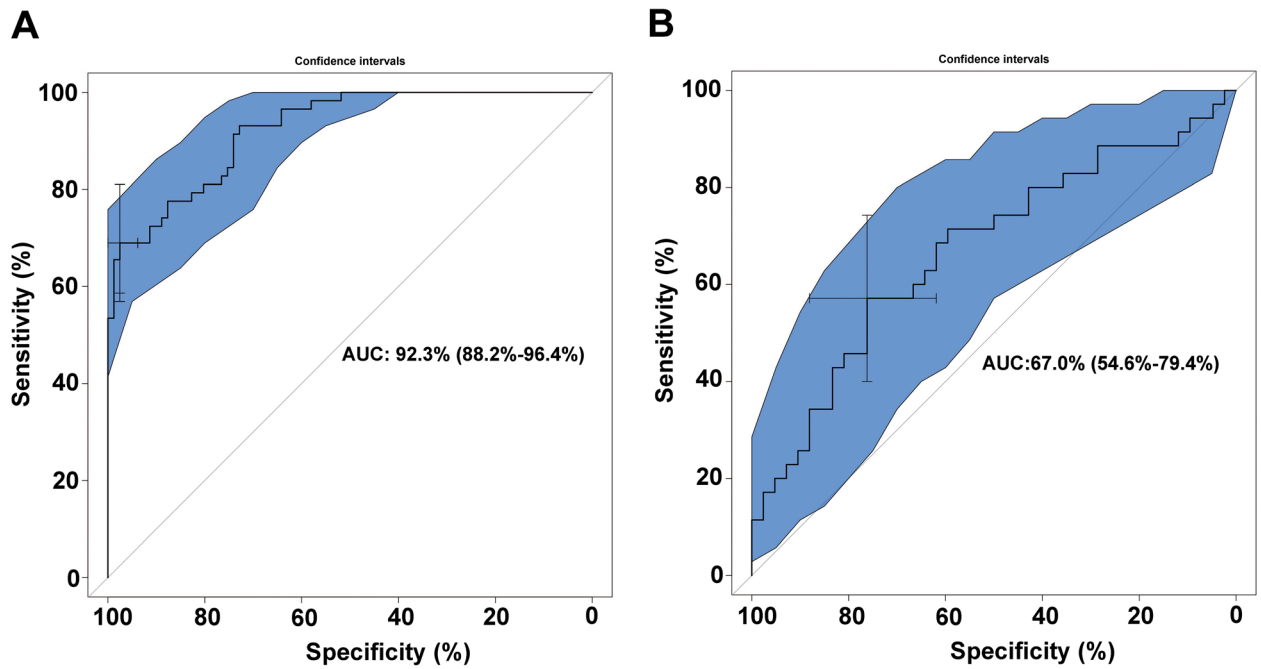

Fig. 9 Evaluation of specificity and sensitivity of the hub genes using the receiver operating characteristic in the training dataset (**A**) and external validation dataset (**B**)

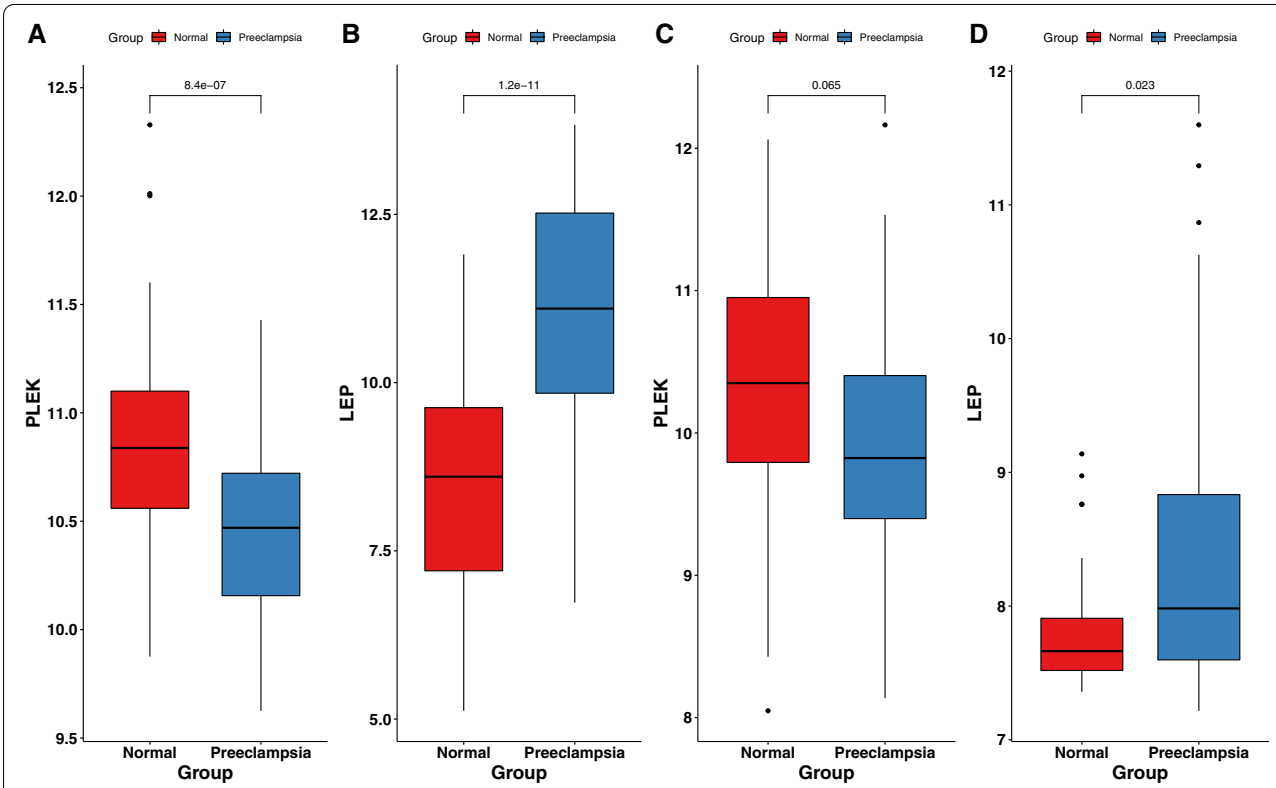

Fig. 10 Validation of expression level of the two key genes (PLEK, LEP) in the training dataset (**A**,**B**) and external validation datasets(**C**,**D**)

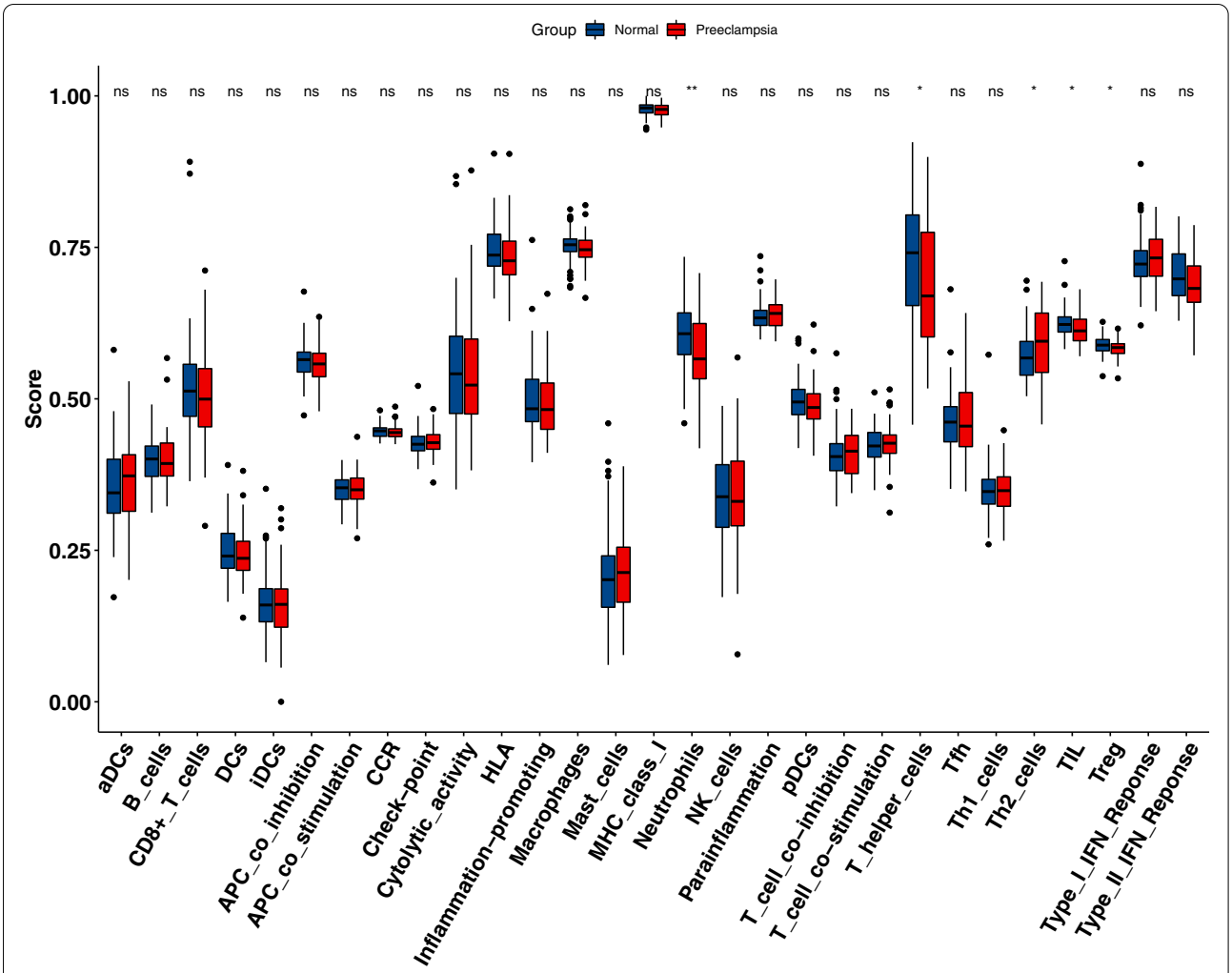

Fig. 11 A comparison of the 29 immune cells and functions between normal and PE samples

methods [44, 45].Cumulative evidence revealed that the unbalanced expression of specific ncRNA is involved in the pathogenesis of PE [46].

However, there is no researches about incorporate the existing predictive model to predict the potential association between non coding RNA and PE presently. Therefore, it is urgent need to applied these model to predict the association between non coding RNA and PE. In our study, trough WGCNA and logistic analysis, we identified two key genes (PLEK and LEP) and constructed a diagnostic model. Leptin (LEP), also known as a placental hormone, is a multifunctional 16 kDa peptide hormone encoded by the LEP gene on chromosome 7 (7q31). It is abundant in adipocytes, the placenta, and tissues such as muscles, liver, brain, and ovaries. LEP is a well-known potential serum marker for PE, particularly in early-onset PE placental tissues [47, 48]. Furthermore, leptin can boost blood supply to the placenta by promoting the formation of new blood vessels [49]. Researchers discovered that leptin can reduce trophoblast cell apoptosis in PE [48], as well as induce endothelial cell proliferation [50], promote immune cell activation, proliferation, and maturation, and prevent monocyte apoptosis. In general, leptin promotes Th-1 response and mediates the production of pro-inflammatory cytokines such as TNF-α, interleukin-2, and interleukin-6. According to Muy-Rivera et al. [51], pregnant women with elevated plasma leptin levels have a 3.8-fold greater risk of developing PE. Leptin has also been shown to raise blood pressure in non-pregnant rabbits [52]. Leptin injection raised ICAM-1 and Eseltin circulation concentrations, resulting in hypertension and proteinuria in pregnant rats [53]. Our findings show that the expression level of this gene was higher in the placenta of the PE group than in the control group, which is consistent with earlier findings. PLEK, another hub gene, was found to be significantly expressed in the

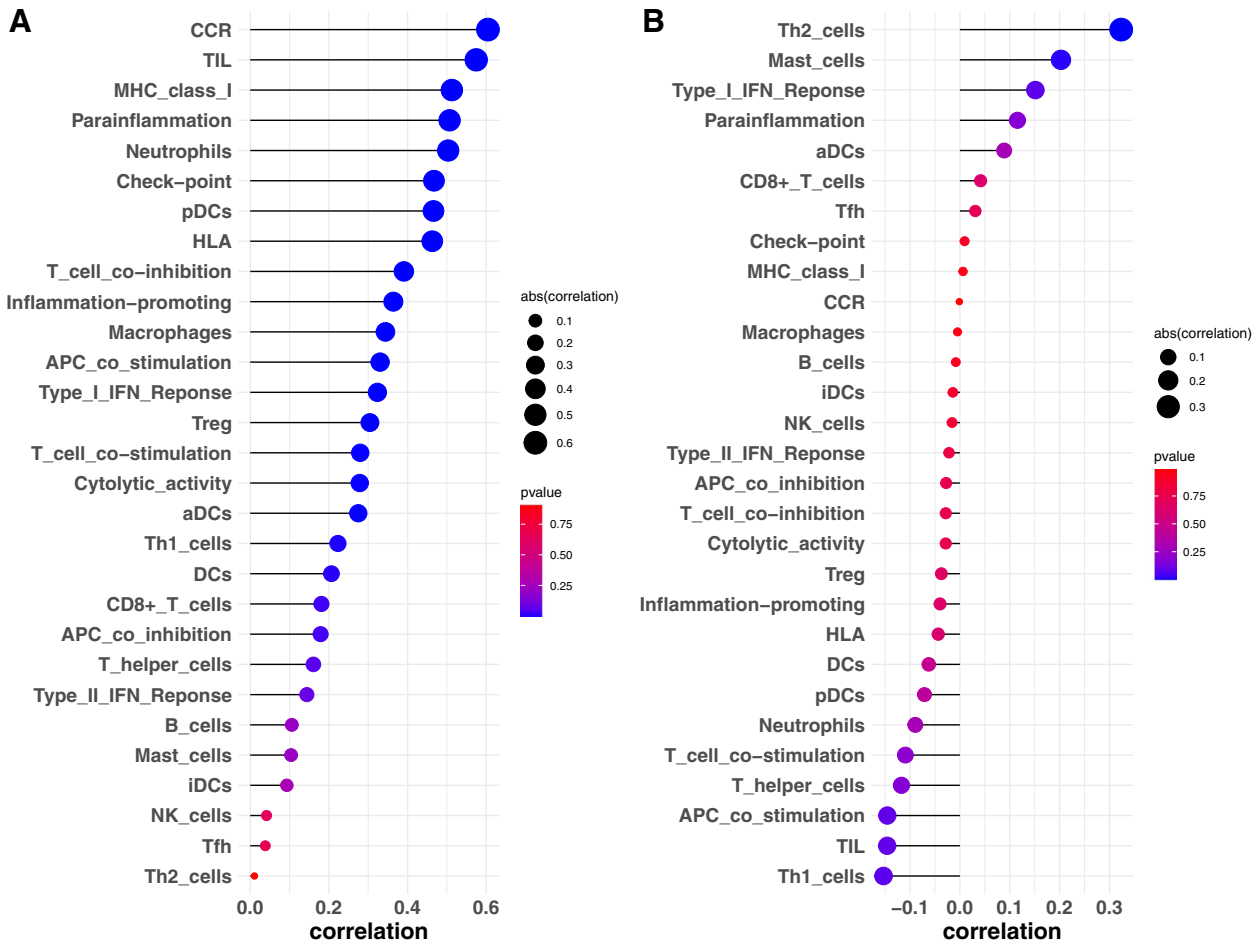

Fig. 12 Correlation of the two key genes, including PLEK (**A**) and LEP (**B**), with the 29 immune cells and functions

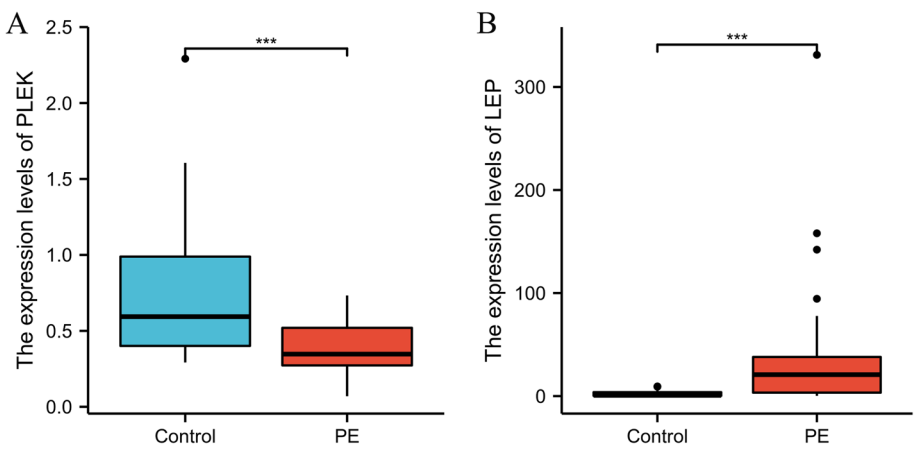

Fig. 13 Two key genes show significantly different expression levels in placenta samples from patients with PE and normal healthy controls through the PCR validation. * $P < 0.05$; ** $P < 0.01$

merged dataset. Its expression was higher in normal tissues than in PE tissues, although it was not significantly higher in the external validation set. Pleckstrin (PLEK), which is found on human chromosome 2, is a protein kinase C target [54] that is involved in signal transduction and hematopoietic cell differentiation [55]. PLEK play an important role in immune and inflammatory responses [56, 57]. Studies have been shown that PLEK is significantly over-expressed in periodontitis, cardiovascular disease, rheumatoid arthritis, and ulcerative colitis [58], and thought to be a crucial mediator in the secretion and activation pathways of pro-inflammatory cytokines TNF-α and IL-1β [59, 60]. While studies have suggested that LILRA2, EVI2A, and PLEK play a role in recurrent miscarriage [61], the underlying mechanism in placental function and regulation has yet to be fully investigated. This is the first study to examine PLEK expression in PE. An examination of its mechanisms would necessitate additional research.

PE is a complex systemic condition, and it has been shown that the immune system plays a significant role in its development [62]. As a result, we evaluated the landscape of 29 immune cell infiltration levels in PE and control samples. The results indicated that the infiltration level of neutrophils, T helper cells, Th2 cells, TIL, and Treg cells were significantly differed between PE and normal tissues. Immune cell infiltration is a new bioinformatics technique that has been used to investigate the diagnosis and prognosis of kidney cancer [63], malignant glioma [64], breast cancer [65], oral squamous cell carcinoma [66], ulcerative colitis[67], osteosarcoma [68], and variety of other diseases. Nonetheless, it has received little attention in the field of PE. The complicated connection between the maternal immune system and its semi-allogeneic foetus is critical in normal pregnancy. Moreover, establishing and maintaining the maternal and foetal immune balance is a prerequisite for normal pregnancy [69]. Abnormal maternal immunological and inflammatory responses to foetal antigens result in increased release of various toxic cytokines, which causes trophoblast cell invasion, vascular remodelling, and placental implantation disorders. Changes in the innate immune system primarily regulate this inflammatory response, with the adaptive immune system possibly playing a supporting role [70].

T cells are thought to be the most important cells in regulating immunological homeostasis [71, 72]. T lymphocytes account for 1%-3% of decidual immune cells [73]. In a normal pregnancy, the mother exhibits Th2 cell-type immunological tolerance, preventing embryo rejection [74]. However, in PE patients, the Th1/Th2 ratio increases. As a result, the Th1/Th2 balance changes towards Th1 [75]. Aside from the Th1/Th2 imbalance that contributes to the onset and progression of the disease, there is also an imbalance of Th17/ regulatory T cells. This imbalance, which was exacerbated by the Th17 immune bias, also contributed to the development of PE. Th17/ Treg cells are balanced at the maternal-foetal interface during normal pregnancy to preserve maternal immunological tolerance and inhibit the inflammatory response [76]. In our study, we identified most of T cells showed a significant difference between normal and PE samples, reveled that the T cells play an important role in PE.

Conclusion

PLEK and LEP were identified as two genes implicated in the development and progression of PE in this investigation. Although more in vivo and in vitro validations are needed, our findings help to understand the pathological process of PE and may serve as a theoretical foundation for future research. The functional annotation and pathway enrichment analysis results show that the immunological mechanism is important in the etiology of PE. Also, because of the maternal and infant complications of PE, it is critical to uncover the aetiology and molecular mechanism, develop molecular biomarkers and investigate effective drugs for the early detection, prevention, and personalized treatment of PE.

Limitations

This study not only points out several benefits of bioinformatics analysis but also highlights some limitations. The dependability of the original microarray dataset is critical to the validity of our conclusions, although the results are constrained due to the small sample size. Similarly, validation results are limited. Second, despite the identification of two hub genes as prospective biomarkers for PE immunotyping, no in vivo and in vitro studies have been conducted. More research on the functions and regulatory mechanisms of key genes in PE is still needed. As a result, this will be the focus of future efforts.

Abbreviations

PE: Preeclampsia; GEO: Gene Expression Omnibus; DEGs: Differentially expressed genes; WGCNA: Weighted gene co-expression network analysis; PPI: Protein–protein interaction network; CMAP: Connectivity map; TOM: Topological overlap matrix; GS: Gene significance; MEs: Module eigengenes; MM: Module membership; GO: Gene ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; ssGSEA: Single-sample Gene Set Enrichment Analysis; MOA: Mechanism of actions; ROC: Receiver operating characteristic; AUC: Area under the curve; CCR: Chemokine receptor; TIL: Tumor infiltrating lymphocyte; MCC: Maximal clique centrality.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-022-08937-3.

Additional file 1:

Acknowledgements
The researchers thank the obstetrics and gynaecology department, surgery department, and laboratory department of Qilu Hospital and The First Affiliated Hospital of USTC for their support during the data collection process.

Authors' contributions
YP wrote the proposal, participated in data collection, analyzed the data, and drafted the paper. NG,HH and AW approved the proposal with some revisions, participated in data analysis, and revised subsequent drafts of the paper. YM designed the entire study, provided administrative, technical, and material support, and gave a critical review of the intellectual content of the article. All authors read and approved the final manuscript.

Funding
No sources of funding exist for this research. All costs were covered by the researchers.

Availability of data and materials
The datasets that support the findings of this study are openly available in the Gene Expression Omnibus (GEO) database under accession ID GSE10588, GSE25906, GSE48424 and GSE60438, [http://www.ncbi.nlm.nih.gov/geo/].

Declarations

Ethics approval and consent to participate
Written informed consent was obtained from all participants. This study was approved by The First Affiliated Hospital of USTC, Division of Life Sciences and Medicine, University of Science and Technology of China (Anhui Provincial Hospital) research ethics board (NO:2021KY161) and is based on the ethical requirements of the Helsinki Declaration. All participants have the right to know.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Department of Obstetrics and Gynecology, Qilu Hospital of Shandong University, Jinan, Shandong, China. [2]Department of Obstetrics and Gynecology, The First Affiliated Hospital of USTC, Division of Life Sciences and Medicine , University of Science and Technology of China, Hefei, Anhui, China. [3]Department of Obstetrics and Gynecology, Qilu Hospital of Shandong University, Jinan Shandong 250012, China.

Received: 26 May 2022 Accepted: 28 September 2022
Published online: 19 October 2022

References

1. Eiland E, Nzerue C, Faulkner M. Preeclampsia 2012. J Pregnancy. 2012;2012:586578.
2. Al-Jameil N, Aziz Khan F, Fareed Khan M, Tabassum H. A brief overview of preeclampsia. J Clin Med Res. 2014;6(1):1–7.
3. Vatten LJ, Skjaerven R. Offspring sex and pregnancy outcome by length of gestation. Early Hum Dev. 2004;76(1):47–54.
4. Hariharan N, Shoemaker A, Wagner S. Pathophysiology of hypertension in preeclampsia. Microvasc Res. 2017;109:34–7.
5. Perry H, Khalil A, Thilaganathan B. Preeclampsia and the cardiovascular system: An update. Trends Cardiovasc Med. 2018;28(8):505–13.
6. Mol BWJ, Roberts CT, Thangaratinam S, Magee LA, de Groot CJM, Hofmeyr GJ. Pre-eclampsia. The Lancet. 2016;387(10022):999–1011.
7. Genest DS, Falcao S, Gutkowska J, Lavoie JL. Impact of exercise training on preeclampsia: potential preventive mechanisms. Hypertens. 2012;60(5):1104–9.
8. Griffith MI, Mann JR, McDermott S. The risk of intellectual disability in children born to mothers with preeclampsia or eclampsia with partial mediation by low birth weight. Hypertens Pregnancy. 2011;30(1):108–15.
9. Wu CS, Sun Y, Vestergaard M, Christensen J, Ness RB, Haggerty CL, Olsen J. Preeclampsia and risk for epilepsy in offspring. Pediatr. 2008;122(5):1072–8.
10. Walker CK, Krakowiak P, Baker A, Hansen RL, Ozonoff S, Hertz-Picciotto I. Preeclampsia, placental insufficiency, and autism spectrum disorder or developmental delay. JAMA Pediatr. 2015;169(2):154–62.
11. Dalman C, Allebeck P, Cullberg J, Grunewald C, Koster M. Obstetric complications and the risk of schizophrenia: a longitudinal study of a national birth cohort. Arch Gen Psychiatry. 1999;56(3):234–40.
12. Ursini G, Punzi G, Chen Q, Marenco S, Robinson JF, Porcelli A, Hamilton EG, Mitjans M, Maddalena G, Begemann M, et al. Convergence of placenta biology and genetic risk for schizophrenia. Nat Med. 2018;24(6):792–801.
13. Coppage KH, Polzin WJ. Severe preeclampsia and delivery outcomes: is immediate cesarean delivery beneficial? Am J Obstet Gynecol. 2002;186(5):921–3.
14. Leavey K, Wilson SL, Bainbridge SA, Robinson WP, Cox BJ. Epigenetic regulation of placental gene expression in transcriptional subtypes of preeclampsia. Clin Epigenetics. 2018;10:28.
15. Saito S, Shiozaki A, Nakashima A, Sakai M, Sasaki Y. The role of the immune system in preeclampsia. Mol Aspects Med. 2007;28(2):192–209.
16. Ramma W, Buhimschi IA, Zhao G, Dulay AT, Nayeri UA, Buhimschi CS, Ahmed A. The elevation in circulating anti-angiogenic factors is independent of markers of neutrophil activation in preeclampsia. Angiogenesis. 2012;15(3):333–40.
17. Phipps EA, Thadhani R, Benzing T, Karumanchi SA. Pre-eclampsia: pathogenesis, novel diagnostics and therapies. Nat Rev Nephrol. 2019;15(5):275–89.
18. Agius A, Sultana R, Camenzuli C, Calleja-Agius J, Balzan R. An update on the genetics of pre-eclampsia. Minerva Ginecol. 2018;70(4):465–79.
19. Spinillo A, Gardella B, Adamo L, Muscettola G, Fiandrino G, Cesari S. Pathologic placental lesions in early and late fetal growth restriction. Acta Obstet Gynecol Scand. 2019;98(12):1585–94.
20. Lynch A, McDuffie R Jr, Murphy J, Faber K, Orleans M. Preeclampsia in multiple gestation: the role of assisted reproductive technologies. Obstet Gynecol. 2002;99(3):445–51.
21. Wimalasundera RC, Larbalestier N, Smith JH, de Ruiter A, Mc GTSA, Hughes AD, Poulter N, Regan L, Taylor GP. Pre-eclampsia, antiretroviral therapy, and immune reconstitution. Lancet. 2002;360(9340):1152–4.
22. Erlebacher A. Immunology of the maternal-fetal interface. Annu Rev Immunol. 2013;31:387–411.
23. Nancy P, Erlebacher A. T cell behavior at the maternal-fetal interface. Int J Dev Biol. 2014;58(2–4):189–98.
24. Gidlof S, Nisell H. Pre-eclampsia. Lakartidningen. 2010;107(51–52):3288–92.
25. LaMarca B, Cornelius D, Wallace K. Elucidating immune mechanisms causing hypertension during pregnancy. Physiology (Bethesda). 2013;28(4):225–33.
26. Ma Y, Ye Y, Zhang J, Ruan CC, Gao PJ. Immune imbalance is associated with the development of preeclampsia. Medicine (Baltimore). 2019;98(14):e15080.
27. Shah TJ, Walsh SW. Activation of NF-kappaB and expression of COX-2 in association with neutrophil infiltration in systemic vascular tissue of women with preeclampsia. Am J Obstet Gynecol. 2007;196(1):48 e41-48.
28. Xu H, Xie Y, Sun Y, Guo R, Lv D, Li X, Li F, He M, Fan Y, Deng D. Integrated analysis of multiple microarray studies to identify potential pathogenic gene modules in preeclampsia. Exp Mol Pathol. 2021;120:104631.
29. Liu K, Fu Q, Liu Y, Wang C. An integrative bioinformatics analysis of microarray data for identifying hub genes as diagnostic biomarkers of preeclampsia. Biosci Rep. 2019;39(9):1-10.
30. Lin J, Meng Y, Song MF, Gu W. Network-Based Analysis Reveals Novel Biomarkers in Peripheral Blood of Patients With Preeclampsia. Front Mol Biosci. 2022;9:757203.
31. Wang Y, Li B, Zhao Y. Inflammation in Preeclampsia: Genetic Biomarkers, Mechanisms, and Therapeutic Strategies. Front Immunol. 2022;13:883404.
32. Sitras V, Paulssen RH, Gronaas H, Leirvik J, Hanssen TA, Vartun A, Acharya G. Differential placental gene expression in severe preeclampsia. Placenta. 2009;30(5):424–33.
33. Tsai S, Hardison NE, James AH, Motsinger-Reif AA, Bischoff SR, Thames BH, Piedrahita JA. Transcriptional profiling of human placentas from

pregnancies complicated by preeclampsia reveals disregulation of sialic acid acetylesterase and immune signalling pathways. Placenta. 2011;32(2):175–82.
34. Textoris J, Ivorra D, Ben Amara A, Sabatier F, Menard JP, Heckenroth H, Bretelle F, Mege JL. Evaluation of current and new biomarkers in severe preeclampsia: a microarray approach reveals the VSIG4 gene as a potential blood biomarker. PLoS ONE. 2013;8(12):e82638.
35. Yong HE, Melton PE, Johnson MP, Freed KA, Kalionis B, Murthi P, Brennecke SP, Keogh RJ, Moses EK. Genome-wide transcriptome directed pathway analysis of maternal pre-eclampsia susceptibility genes. PLoS ONE. 2015;10(5):e0128230.
36. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000;28(1):27–30.
37. Kanehisa M. Toward understanding the origin and evolution of cellular organisms. Protein Sci. 2019;28(11):1947–51.
38. Kanehisa M, Furumichi M, Sato Y, Ishiguro-Watanabe M, Tanabe M. KEGG: integrating viruses and cellular organisms. Nucleic Acids Res. 2021;49(D1):D545–51.
39. He Y, Jiang Z, Chen C, Wang X. Classification of triple-negative breast cancers based on Immunogenomic profiling. J Exp Clin Cancer Res. 2018;37(1):327.
40. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods. 2001;25(4):402–8.
41. Chen X, Yan CC, Zhang X, You ZH. Long non-coding RNAs and complex diseases: from experimental results to computational models. Brief Bioinform. 2017;18(4):558–76.
42. Wang CC, Han CD, Zhao Q, Chen X. Circular RNAs and complex diseases: from experimental results computational models. Brief Bioinform. 2021;22(6):1-27.
43. Chen X, Xie D, Zhao Q, You ZH. MicroRNAs and complex diseases: from experimental results to computational models. Brief Bioinform. 2019;20(2):515–39.
44. Chen X, Sun LG, Zhao Y. NCMCMDA: miRNA-disease association prediction through neighborhood constraint matrix completion. Brief Bioinform. 2021;22(1):485–96.
45. Chen X, Li TH, Zhao Y, Wang CC, Zhu CC. Deep-belief network for predicting potential miRNA-disease associations. Brief Bioinform. 2021;22(3):bbaa186.
46. Sun N, Qin S, Zhang L, Liu S. Roles of noncoding RNAs in preeclampsia. Reprod Biol Endocrinol. 2021;19(1):100.
47. Rokhafrooz S, Ghadiri A, Ghandil P, Ghafourian M, Hossaini SH, Daraei N, Najafian M, Rouhizadeh A. Association between HLA-G 14bp Gene Polymorphism and Serum sHLA-G Protein Concentrations in Preeclamptic Patients and Normal Pregnant Women. Immunol Invest. 2018;47(6):558–68.
48. Miehle K, Stepan H, Fasshauer M. Leptin, adiponectin and other adipokines in gestational diabetes mellitus and pre-eclampsia. Clin Endocrinol (Oxf). 2012;76(1):2–11.
49. Hauguel-de Mouzon S, Lepercq J, Catalano P. The known and unknown of leptin in pregnancy. Am J Obstet Gynecol. 2006;194(6):1537–45.
50. Cao R, Brakenhielm E, Wahlestedt C, Thyberg J, Cao Y. Leptin induces vascular permeability and synergistically stimulates angiogenesis with FGF-2 and VEGF. Proc Natl Acad Sci U S A. 2001;98(11):6390–5.
51. Muy-Rivera M, Ning Y, Frederic IO, Vadachkoria S, Luthy DA, Williams MA. Leptin, soluble leptin receptor and leptin gene polymorphism in relation to preeclampsia risk. Physiol Res. 2005;54(2):167–74.
52. Prior LJ, Eikelis N, Armitage JA, Davern PJ, Burke SL, Montani JP, Barzel B, Head GA. Exposure to a high-fat diet alters leptin sensitivity and elevates renal sympathetic nerve activity and arterial pressure in rabbits. Hypertension. 2010;55(4):862–8.
53. Ibrahim HS, Omar E, Froemming GR, Singh HJ. Leptin increases blood pressure and markers of endothelial activation during pregnancy in rats. Biomed Res Int. 2013;2013:298401.
54. Tyers M, Haslam RJ, Rachubinski RA, Harley CB. Molecular analysis of pleckstrin: the major protein kinase C substrate of platelets. J Cell Biochem. 1989;40(2):133–45.
55. Kienzle N, Young D, Silins SL, Sculley TB. Induction of pleckstrin by the Epstein-Barr virus nuclear antigen 3 family. Virology. 1996;224(1):167–74.
56. Liu Y, Wang L, Predina J, Han R, Beier UH, Wang LC, Kapoor V, Bhatti TR, Akimova T, Singhal S, et al. Inhibition of p300 impairs Foxp3(+) T regulatory cell function and promotes antitumor immunity. Nat Med. 2013;19(9):1173–7.
57. Ubel C, Sopel N, Graser A, Hildner K, Reinhardt C, Zimmermann T, Rieker RJ, Maier A, Neurath MF, Murphy KM, et al. The activating protein 1 transcription factor basic leucine zipper transcription factor, ATF-like (BATF), regulates lymphocyte- and mast cell-driven immune responses in the setting of allergic asthma. J Allergy Clin Immunol. 2014;133(1):198-206 e191-199.
58. Lundmark A, Davanian H, Bage T, Johannsen G, Koro C, Lundeberg J, Yucel-Lindberg T. Transcriptome analysis reveals mucin 4 to be highly associated with periodontitis and identifies pleckstrin as a link to systemic diseases. Sci Rep. 2015;5:18475.
59. Ding Y, Kantarci A, Badwey JA, Hasturk H, Malabanan A, Van Dyke TE. Phosphorylation of pleckstrin increases proinflammatory cytokine secretion by mononuclear phagocytes in diabetes mellitus. J Immunol. 2007;179(1):647–54.
60. Hasturk H, Kantarci A, Van Dyke TE. Oral inflammatory diseases and systemic inflammation: role of the macrophage. Front Immunol. 2012;3:118.
61. Chen H, Cheng S, Liu C, Fu J, Huang W. Bioinformatics Analysis of Differentially Expressed Genes, Methylated Genes, and miRNAs in Unexplained Recurrent Spontaneous Abortion. J Comput Biol. 2019;26(12):1418–26.
62. Lu HQ, Hu R. The role of immunity in the pathogenesis and development of pre-eclampsia. Scand J Immunol. 2019;90(5):e12756.
63. Geissler K, Fornara P, Lautenschlager C, Holzhausen HJ, Seliger B, Riemann D. Immune signature of tumor infiltrating immune cells in renal cancer. Oncoimmunology. 2015;4(1):e985082.
64. Matsuhisa T. Subset analysis of tumor infiltrating lymphocytes and peripheral blood lymphocytes in malignant glioma patients. No To Shinkei. 1995;47(5):466–73.
65. Yao S, Cheng TD, Elkhanany A, Yan L, Omilian A, Abrams SI, Evans S, Hong CC, Qi Q, Davis W, et al. Breast Tumor Microenvironment in Black Women: A Distinct Signature of CD8+ T-Cell Exhaustion. J Natl Cancer Inst. 2021;113(8):1036–43.
66. Parikh A, Shin J, Faquin W, Lin DT, Tirosh I, Sunwoo JB, Puram SV. Malignant cell-specific CXCL14 promotes tumor lymphocyte infiltration in oral cavity squamous cell carcinoma. J Immunother Cancer. 2020;8(2):e001048.
67. Xue G, Hua L, Zhou N, Li J. Characteristics of immune cell infiltration and associated diagnostic biomarkers in ulcerative colitis: results from bioinformatics analysis. Bioengineered. 2021;12(1):252–65.
68. Fu Y, Bao Q, Liu Z, He G, Wen J, Liu Q, Xu Y, Jin Z, Zhang W. Development and Validation of a Hypoxia-Associated Prognostic Signature Related to Osteosarcoma Metastasis and Immune Infiltration. Front Cell Dev Biol. 2021;9:633607.
69. Perez-Sepulveda A, Torres MJ, Khoury M, Illanes SE. Innate immune system and preeclampsia. Front Immunol. 2014;5:244.
70. Borzychowski AM, Sargent IL, Redman CW. Inflammation and pre-eclampsia. Semin Fetal Neonatal Med. 2006;11(5):309–16.
71. Yang F, Zheng Q, Jin L. Dynamic Function and Composition Changes of Immune Cells During Normal and Pathological Pregnancy at the Maternal-Fetal Interface. Front Immunol. 2019;10:2317.
72. Wang Z, Song K, Zhao W, Zhao Z. Dendritic cells in tumor microenvironment promoted the neuropathic pain via paracrine inflammatory and growth factors. Bioengineered. 2020;11(1):661–78.
73. Lessin DL, Hunt JS, King CR, Wood GW. Antigen expression by cells near the maternal-fetal interface. Am J Reprod Immunol Microbiol. 1988;16(1):1–7.
74. Perricone C, de Carolis C, Perricone R. Pregnancy and autoimmunity: a common problem. Best Pract Res Clin Rheumatol. 2012;26(1):47–60.
75. Vargas-Rojas MI, Solleiro-Villavicencio H, Soto-Vega E. Th1, Th2, Th17 and Treg levels in umbilical cord blood in preeclampsia. J Matern Fetal Neonatal Med. 2016;29(10):1642–5.
76. Eghbal-Fard S, Yousefi M, Heydarlou H, Ahmadi M, Taghavi S, Movasaghpour A, Jadidi-Niaragh F, Yousefi B, Dolati S, Hojjat-Farsangi M, et al. The imbalance of Th17/Treg axis involved in the pathogenesis of preeclampsia. J Cell Physiol. 2019;234(4):5106–16.

Publisher's Note