# Identification of Lupus-Associated Genes in the Pathogenesis of Pre-eclampsia Via Bioinformatic Analysis

Qianwen Dai[1] (iD), Mengtao Li[2], Xinping Tian[2], Yijun Song[1] and Jiuliang Zhao[2]

[1]Department of Obstetrics and Gynecology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, National Clinical Research Center for Obstetric & Gynecologic Diseases, Beijing, China. [2]Department of Rheumatology and Clinical Immunology, Peking Union Medical College Hospital, Peking Union Medical College and Chinese Academy of Medical Sciences, National Clinical Research Center for Dermatologic and Immunologic Diseases (NCRC-DID), Key Laboratory of Rheumatology and Clinical Immunology, Ministry of Education, Beijing, China.

**ABSTRACT:** Pre-eclampsia (PE) is a severe pregnancy complication that is more common in patients with systemic lupus erythematosus (SLE). Although the exact causes of these conditions are not fully understood, the immune system plays a key role. To investigate the connection between SLE and PE, we analyzed genes associated with SLE that may contribute to the development of PE. We collected 9 microarray data sets from the NCBI GEO database and used Limma to identify the differentially expressed genes (DEGs). In addition, we employed weighted gene co-expression network analysis (WGCNA) to pinpoint the hub genes of SLE and examined immune infiltration using Cibersort. By constructing a protein-protein interaction (PPI) network and using CytoHubba, we identified the top 20 PE hub genes. Subsequently, we created a nomogram and conducted a receiver operating characteristic (ROC) analysis to predict the risk of PE. Our analysis, including gene set enrichment analysis (GSEA) and PE DEGs enrichment analysis, revealed significant involvement in placenta development and immune response. Two pivotal genes, BCL6 and MME, were identified, and their validity was confirmed using 5 data sets. The nomogram demonstrated good diagnostic performance (AUC: 0.82-0.96). Furthermore, we found elevated expression levels of both genes in SLE peripheral blood mononuclear cells (PBMCs) and PE placental specimens within the case group. Analysis of immune infiltration in the SLE data set showed a strong positive correlation between the expression of both genes and neutrophil infiltration. BCL6 and MME emerged as crucial genes in lupus-related pregnancies associated with the development of PE, for which we devised a nomogram. These findings provide potential candidate genes for further research in the diagnosis and understanding of the pathophysiology of PE.

**KEYWORDS:** systemic lupus erythematosus, pre-eclampsia, bioinformatics analysis, nomogram, immune infiltration

**CORRESPONDING AUTHOR:** Yijun Song, Department of Obstetrics and Gynecology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, National Clinical Research Center for Obstetric & Gynecologic Diseases, Shuaifuyuan No. 1, Dongcheng District, Beijing 100730, China. Email: 13681191769@163.com

Jiuliang Zhao, Department of Rheumatology and Clinical Immunology, Peking Union Medical College Hospital, Peking Union Medical College and Chinese Academy of Medical Sciences, National Clinical Research Center for Dermatologic and Immunologic Diseases, Key Laboratory of Rheumatology and Clinical Immunology, Ministry of Education, Shuaifuyuan No. 1, Dongcheng District, Beijing 100730, China. **Email:** zjlpumc@sina.com

## Introduction

Systemic lupus erythematosus (SLE) is an autoimmune illness primarily affecting women, typically occurring during their reproductive years. The development of SLE is influenced by multiple variables, such as the formation of autoantibodies, dysregulation of B cells, abnormalities in T cells, failure of the complement system, and signaling of type I interferons.[1] These variables result in persistent inflammation and the involvement of multiple organs. Individuals with SLE have an increased susceptibility to experiencing complications during pregnancy. These encompass pre-eclampsia (PE), fetal or neonatal death, as well as possible fetal illnesses caused by restricted growth and preterm labor. The incidence of PE, especially early-onset pre-eclampsia (EOPE), in patients with SLE is greater than that in the general population.

Pre-eclampsia is a condition that occurs during pregnancy and is characterized by elevated blood pressure and the possibility of organ damage, particularly in the kidneys. Pre-eclampsia is a condition that occurs during pregnancy and is characterized by elevated blood pressure and the possibility of organ damage, particularly in the kidneys. Pre-eclampsia usually occurs after 20. Pre-eclampsia usually occurs after 20 weeks of pregnancy, affecting around 2% to 8% of pregnancies worldwide.[2] From a clinical perspective, PE can be classified into 2 subgroups according to their onset, each with separate pathophysiological origins. Early-onset pre-eclampsia, which occurs before 34 weeks of gestation, often manifests with more severe symptoms and complications for both the mother and the infant. The prevailing belief is that it originates from inadequate placentation. Conversely, late-onset pre-eclampsia (LOPE), occurring at or after 34 weeks of

pregnancy, is primarily driven by maternal cardiovascular risk factors and placental insufficiency.[3] Initiating low-dose aspirin in early pregnancy (<16 weeks) can effectively reduce the risk of EOPE.[4] Extensive epidemiological research has shown that mothers who have experienced EOPE, and their children have a higher likelihood of developing cardiovascular disease in their later life.[5,6] Our research specifically concentrates on the analysis of EOPE due to its significant implications.

Recent research suggests that the activation of complement, the migration of neutrophils, and the generation of proinflammatory cytokines and antiangiogenic molecules play a significant role in the impaired development of the placenta in both autoimmune illnesses and nonautoimmune models of PE.[7,8] Moreover, evidence indicates a connection between higher levels of neutrophils and the presence of neutrophil extracellular traps (NETs) in the placental intervillous spaces. This is accompanied by inflammatory and vascular alterations in individuals with SLE and PE.[9,10] These data suggest that there may be common pathogenic pathways between SLE and PE.

Nevertheless, the underlying causes of both SLE and EOPE have yet to be fully understood. Only a limited number of studies have investigated the molecular pathways that connect SLE and PE using bioinformatics analysis. Examining the overlapping transcriptional profiles of SLE and PE could elucidate their similar underlying causes and offer a more comprehensive understanding of the pathophysiological mechanism of PE.

We employed various integrative bioinformatics tools in our work to reveal crucial genes and probable pathways that are connected with SLE and PE. We obtained 2 data sets of peripheral blood mononuclear cells (PBMCs) from individuals with SLE and 7 PE data sets of placental and maternal peripheral blood transcriptome expression profiles from the Gene Expression Omnibus (GEO) collection. Our study aimed to identify crucial genes shared by both SLE and PE and investigate their association with immune-infiltrating cells to gain a deeper understanding of the development of PE. In addition, we verified the identified essential genes in multiple data sets to determine their diagnostic potential for PE. Our study employs a thorough bioinformatic method to investigate the pathogenic mechanisms of PE that are associated with SLE for the first time.

## Materials and Methods

### Microarray data

We obtained 2 microarray data sets (GSE81622, GSE50772) of PBMCs from patients with SLE and 7 microarray data sets (GSE10588, GSE25906, GSE190639, GSE48424, GSE14722, GSE75010, and GSE149437) from patients with PE from the NCBI Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/).

The cause of PE is not fully comprehended. Nevertheless, defective placentation is widely acknowledged as a contributing factor in the onset of PE. Consequently, we commenced

our work by comparing the gene expression profiles in the placenta of individuals with EOPE and control groups. Table 1 contained comprehensive information about the data sets.

### Data processing

Microarray expression data sometimes exhibits batch effects and variances, which can make it challenging to compare samples from different batches. It is not appropriate to merge data sets without taking these batch effects and variances into account. To address this issue, we used batch correction on 2 SLE data sets (GSE81622, GSE50772) using the combat function from the "surrogate variable analysis (SVA)" package in R 4.2.1 software (R Foundation for Statistical Computing, Vienna, Austria). Similarly, we applied batch correction on the GSE10588 and GSE25906 data sets for the PE analysis. As a result, we were able to generate combined SLE and PE expression data separately. When combining data sets from multiple platforms, we only selected genes that were shared across all the combined data sets. This ensured that the combined data sets were suitable for further studies.

### Differentially expressed genes analysis

In our analysis of the combined SLE and PE data sets, we used the "limma" package in R to identify differentially expressed genes (DEGs). For the SLE data set, the criteria used were an adjusted $P$ value <.05 and |log2 fold change| >.7, while for the PE data set, the criteria were an adjusted $P$ value <.05 and |log2 fold change| >.5.

### Weighted gene co-expression network analysis

Weighted gene co-expression network analysis (WGCNA) is a robust bioinformatics technique that enables the analysis of gene expression data to identify clusters of genes with similar expression patterns. It allows researchers to uncover underlying biological processes (BP) and potential regulatory mechanisms by identifying highly interconnected gene modules. It aids in unraveling complex gene regulatory networks, identifying key driver genes, detecting biomarkers, and predicting gene functions. Its ability to integrate multiple sources of information makes it a valuable tool for exploring gene expression data and gaining a deeper understanding of BP. In this study, we used WGCNA to discover intramodular hub genes in the integrated data set of SLE. First, we calculated the median absolute deviations (MADs) and excluded the lowest 75% of genes with the smallest MADs. Second, the "goodSamplesGenes" function from the "WGCNA" package was used to assess genes and samples that did not match the specified quality standards.[11] In addition, we used hierarchical clustering to detect any abnormal samples. Third, the "WGCNA" package was employed to design a co-expression gene network with a scale-free property using the 1-step network construction function. Fourth, we

**Table 1.** Basic information of GEO data sets in the study.

| ID | GSE | PLATFORM | SAMPLE | | SOURCES | GROUP |
|----|-----|----------|--------|--|---------|-------|
| | | | CONTROL | SLE | | |
| 1 | GSE81622 | GPL10558 | 25 | 30 | PBMCs | **Discovery cohort** |
| 2 | GSE50772 | GPL570 | 20 | 61 | PBMCs | **Discovery cohort** |
| | | | Control | PE | | |
| 3 | GSE10588 | GPL2986 | 26 | 17 | Placenta | **Discovery cohort** |
| 4 | GSE25906 | GPL6102 | 37 | 23 EOPE | Placenta | **Discovery cohort** |
| 5 | GSE190639 | GPL31059 | 13 | 13EOPE and 6 LOPE | Placenta | Validation cohort |
| 6 | GSE48424 | GPL6480 | 19 | 6 nonsevere PE and 13 with severe PE | Whole blood | Validation cohort |
| 7 | GSE14722 | GPL96 | 11 | 12 | Placenta | Validation cohort |
| 8 | GSE75010 | GPL6244 | 77 | 80 | Placenta | Validation cohort |
| 9 | GSE149437 | GPL28460 | 21 | 66 EOPE 168 PPROM 166 sPTD | Whole blood | Validation cohort |

Abbreviations: SLE, systemic lupus erythematosus; PBMCs, peripheral blood mononuclear cells; PE, pre-eclampsia; EOPE, early-onset pre-eclampsia; LOPE, late-onset pre-eclampsia; PPROM, preterm prelabour rupture of membrane; sPTD, spontaneous preterm delivery.

derived module eigengenes (MEs) by calculating each module's primary component of expression. Then, a Pearson correlation analysis was used to identify the relationship between modules and illness phenotypes of interest. Based on the correlations with SLE, we successfully identified crucial modules. Ultimately, we assessed the importance of the module by analyzing the scores of module membership (MM) and gene significance (GS). Genes with a GS value greater than 0.2 and an MM value greater than 0.8 within important modules were identified as potential hub genes in the WGCNA analysis.

*Functional enrichment analysis*

To identify significant gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways associated with the DEGs from the integrated PE data set, we conducted enrichment analysis using the "enrich plot" and "clusterprofiler" packages in R. The significance was determined based on a *P* value <.05, indicating enriched biological functions and pathways.

*Gene set enrichment analysis*

Traditional enrichment analysis uses a threshold to filter genes, which can overlook genes with small yet significant changes. This is particularly the case when only a few DEGs are spread across many pathways. On the contrary, gene set enrichment analysis (GSEA) works at the gene set level instead of

individual genes. It ranks all genes according to relevance (based on factors like fold change or *P* value). Then, it compares how these rankings align with known functional gene sets. The key output from GSEA is the Enrichment Score (ES), which indicates if the gene set under study is related to a particular pathway. Thus, GSEA provides a more comprehensive view as it considers all genes, not just those passing a specific threshold. This study employed both functional enrichment analysis of DEGs and GSEA to identify significant pathways distinguishing PE from the control group. We used 2 reference gene sets obtained from the Molecular Signatures Database (MSigDB), specifically the "c2.cp.kegg_medicus.v2023.2.Hs.symbols.gmt" and "h.all.v2023.2.Hs.symbols.gmt." Pathways with an adjusted *P* value <.05 were considered to be significantly enriched.

*Protein-protein interaction network construction and screening shared hub genes*

To investigate the potential interplay of the DEGs from the PE data set, we mapped them to a protein-protein interaction (PPI) network, which was constructed using the Search Tool for Retrieval of Interacting Genes, commonly known as STRING (https://www.stringdb.org).[12] We set a minimum required interaction score of 0.400 to ensure the reliability of the network. Subsequently, we employed Cytoscape software (version 3.7.2) to visually represent the PPI network. To determine the most significant genes in the PPI network, we

employed CytoHubba, a Cytoscape plug-in. The 20 genes with the highest maximal clique centrality (MCC) scores were identified as hub genes.

### Validation of the expression of the key genes and diagnostic value in PE

We have identified the essential genes shared between the hub genes in the SLE data set and the top 20 hub genes in the integrated PE data set. The mRNA expression levels of these key genes were confirmed by analyzing 3 PE gene expression profiles of the placenta (specifically GSE190639, GSE14722, GSE75010) and 2 PE gene expression profiles of maternal peripheral blood (GSE48424, GSE149437). In addition, a user-friendly nomogram was generated using the "rms" package in R, based on the integrated PE data set. The diagnostic performance was assessed by measuring the sensitivity and specificity of both the nomogram and the key genes using receiver operating characteristic (ROC) curves with the "pROC" package in R.

### Association with immune cell infiltration

In the integrated SLE data set, we used the "Cibersort" package in R to quantify the relative proportions of infiltrating immune cells. Subsequently, a Pearson correlation analysis was performed to investigate the potential relationship linking the expression levels of identified key genes and the degree of immune cell infiltration.

### Statistical analysis

The statistical analyses were conducted in R software (version 4.2.1). A 2-sided $P$ value $< .05$ was considered statistically significant. The normal distribution of the variables was assessed using the Shapiro-Wilk test. We used the $t$-test and 1-way analysis of variance (ANOVA) tests for continuous variables that follow a normal distribution. We employed the Wilcox test and the Kruskal-Wallis test to analyze variables that do not follow a normal distribution.

## Results

### Data preprocessing

The methodology of our bioinformatics analysis was delineated in Figure 1. We obtained 2 data sets of SLE (GSE81622, GSE50772) and 2 data sets of PE (GSE10588, GSE25906) from the GEO database. Following batch correction, we acquired integrated data sets that underwent additional normalization. The combined SLE data set consisted of 91 SLE samples and 45 control samples obtained from PBMCs. The integrated PE data set consisted of 40 samples from patients with PE and 63 samples from control, all were placental tissues. As shown in Supplementary Figure 1, there was a significant decrease in discrepancies among the various data sets after the batch effect removal process.

### DEGs and enrichment analysis

Applying thresholds of $|log2$ fold change$| > .7$ and adjusted $P$ value $< .05$, we identified 180 upregulated and 72 downregulated genes in the integrated SLE data set. Similarly, we discovered 80 upregulated and 24 downregulated genes in the integrated PE data set by adhering to cut-off criteria of adjusted $P$ value $< .05$ and $|log2$ fold change$| > .5$. The volcano plot in Figure 2A and B visualizes all DEGs.

We performed GO and KEGG enrichment analyses on the DEGs obtained from the PE data set. The analyses unveiled significant associations with diverse BP, molecular functions (MFs), and cellular components (CCs). The DEGs were found to be involved in a range of BPs, such as cell adhesion, leukocyte activation, and placenta development. They also played a role in managing leukocyte proliferation and differentiation. Regarding MF, these genes were primarily associated with hormone activity and growth factor receptor binding. Regarding CC, the genes were linked to areas of the cell instrumental in secretion and transport, like vesicles and granules. These DEGs appear to have significant roles in both immune response and placenta development (Figure 2C). The KEGG analysis further indicated that these DEGs contributed substantially to pathways related to cell adhesion, coagulation, complement, and T-cell differentiation (Figure 2D).

Through the GSEA hallmarks gene set, we discovered that the genes in the PE-integrated data set responded actively to UV radiation, estrogen, hypoxia, and glycolysis. However, they were less active in pathways related to specific metabolic targets, fat metabolism, peroxisome, and oxidative phosphorylation (Figure 2E). These alterations could be linked to disrupted cellular homeostasis, altered energy metabolism, or impaired cellular functions, leading to the activation of genes involved in DNA repair and cellular stress response. When we analyzed the data with the KEGG gene set for GSEA, it indicated inhibition of translation and protein degradation pathways (Figure 2F), which implied a potential impairment in protein synthesis and degradation processes.

### PPI network construction and hub genes selection

We used the STRING database to construct a protein-protein interaction (PPI) network for 104 DEGs from the integrated PE data set. Subsequently, the results were visualized in Cytoscape (Figure 3A) and analyzed using the cytoHubba plug-in. By implementing the MCC algorithm, we pinpointed the top 20 genes as potential hub genes (Figure 3B).

### The construction of WGCNA and the identification of hub genes in SLE

We then conducted a WGCNA on the integrated SLE data set to explore hub genes in SLE. Upon choosing the highest 25% of variant genes from the data set, we obtained a total of 11,603 genes and 136 samples for analysis. A hierarchical clustering
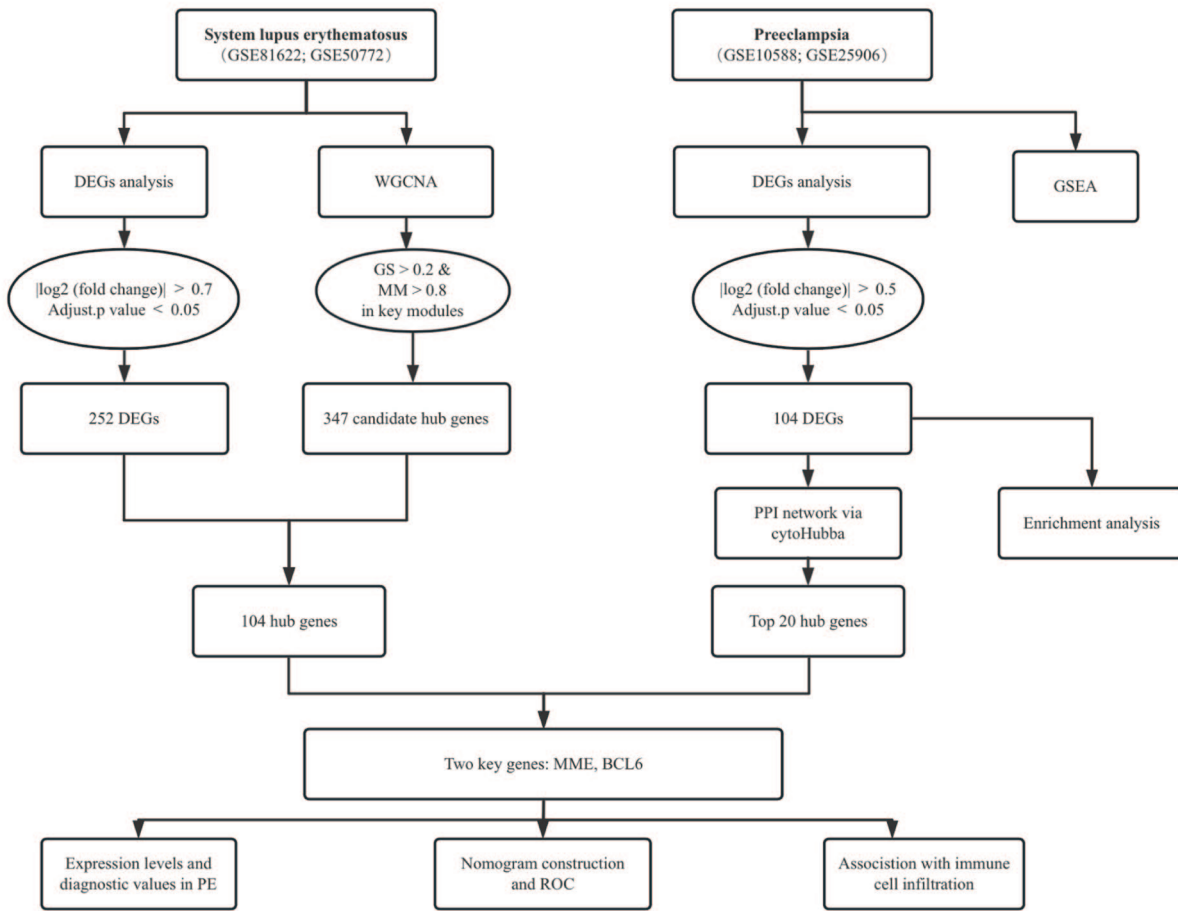
**Figure 1.** Study flowchart. DEGs, differentially expressed genes; GS, gene significance; MM, module membership; WGCNA, weighted gene co-expression network analysis; GSEA, gene set enrichment analysis; PPI, protein-protein interaction; ROC, receiver operating characteristic curve.

analysis was performed with a threshold of 120. An outlier sample (GSM2159873) was identified and excluded from further analyses (Supplementary Figure 3). We used a soft-thresholding power of 6 and established a minimum module size of 50. This led to the identification of 21 modules, which were then subjected to further analysis (Figure 4A and B). Subsequently, we assessed the association between the gene modules and SLE by generating a heatmap using the Pearson correlation coefficient (Figure 4C). Nine modules exhibited significant positive correlations with SLE. Among these, 3 modules (MElightcyan: $r = -.62$, $P = 2e-15$; MEyellow: $r = -.52$, $P = 1e-10$; MEgrey60: $r = -.49$, $P = 1e-09$) displayed strong negative correlations. On the contrary, 6 modules (MEpurple: $r = .52$, $P = 1e-10$; MEgreen: $r = 0.5$, $P = 5e-10$; MElightgreen: $r = .46$, $P = 2e-08$; MEroyalblue: $r = .46$, $P = 2e-08$; MEbrown: $r = .44$, $P = 7e-08$; MEgreenyellow: $r = .43$, $P = 2e-07$) exhibited positive correlations with SLE. To gain a deeper comprehension of the relationship between MM-GS and SLE in the 9 aforementioned modules, we depicted their correlations individually (Supplementary Figure 3). Based on the criterion of $|MM| > 0.8$ and $|GS| > 0.2$, 347 potential hub genes were selected from these modules. Ultimately, through the intersection of the DEGs with

the candidate hub genes acquired from the WGCNA, we successfully identified a total of 104 hub genes in SLE. To investigate a possible link between hub genes associated with SLE and the development of PE, 2 key genes were identified from the intersection of a Venn diagram that included the top 20 hub genes in PE and the hub genes in SLE. These 2 genes were then subjected to further analysis (Figure 4D).

*Validating 2 key gene expression patterns and assessing the diagnostic value of the nomogram model*

We verified the expression levels of 2 key genes, BCL6, and MME, across 5 different data sets. The data sets GSE190639, GSE14722, and GSE75010 contained placenta expression profiles, whereas GSE48424 and GSE149437 were obtained from maternal peripheral blood samples. Both SLE PBMCs and PE placenta exhibited increased expressions of BCL6 and MME in the diseased group. Nevertheless, there was a significant reduction in the expressions of these markers in peripheral blood samples obtained from patients with PE, as depicted in Figure 5A to G. The samples in data set GSE149437 were
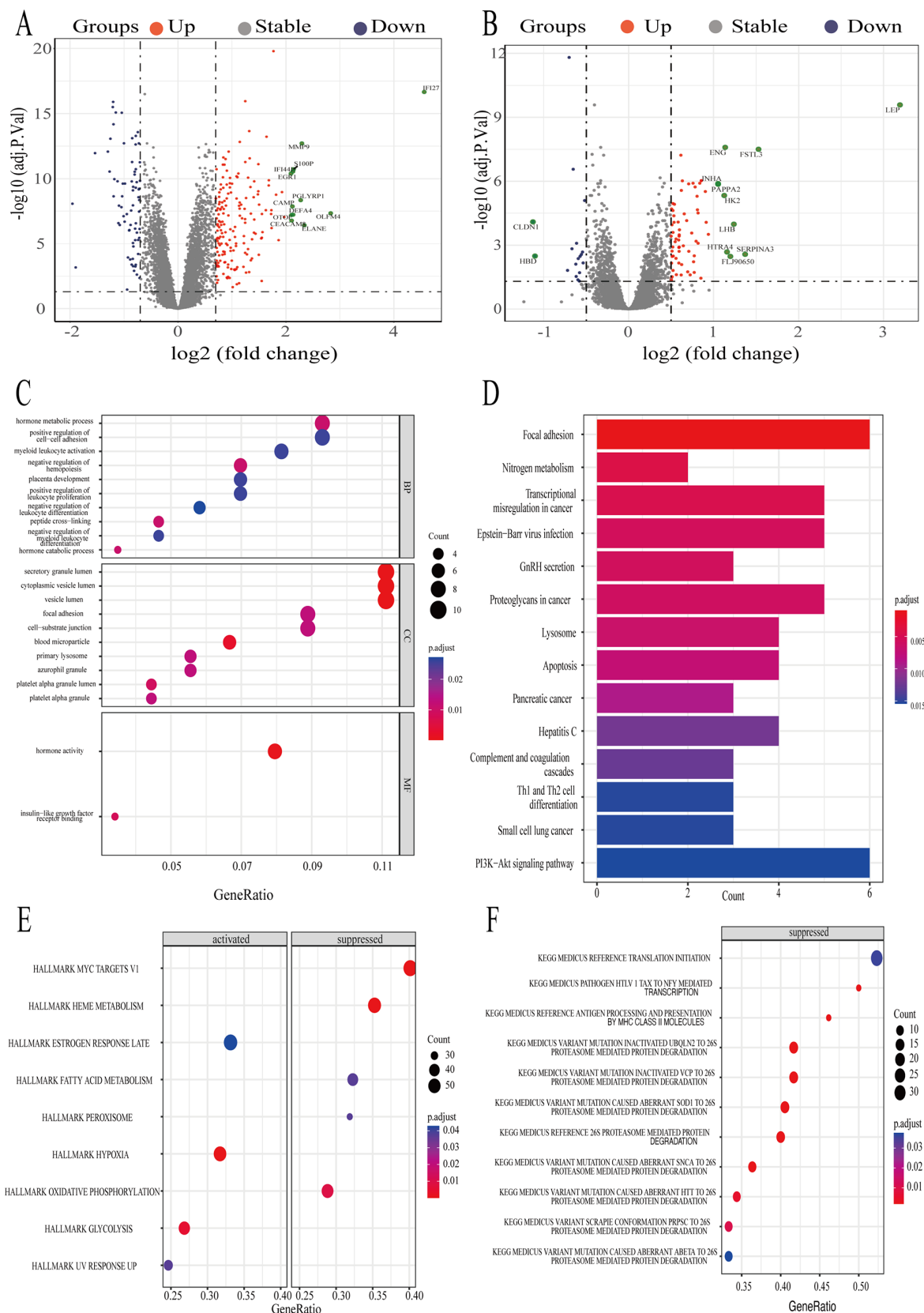
**Figure 2.** Analysis of DEGs in SLE and PE data sets: (A) The DEGs' volcano plot in the SLE data set. Genes with|log2 fold change| >2 are highlighted in green; (B) The DEGs' volcano plot in the PE data set. Genes with|log2 fold change| >1 are highlighted in green; (C) DEGs' GO enrichment results in the PE data set; (D) DEGs' KEGG enrichment pathways in the PE data set; (E) GSEA hallmarks gene set enrichment pathways in the PE data set; (F) GSEA KEGG gene set enrichment in PE data set. DEGs, differentially expressed genes; SLE, systemic lupus erythematosus; PE, pre-eclampsia; GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; GSEA, gene set enrichment analysis.
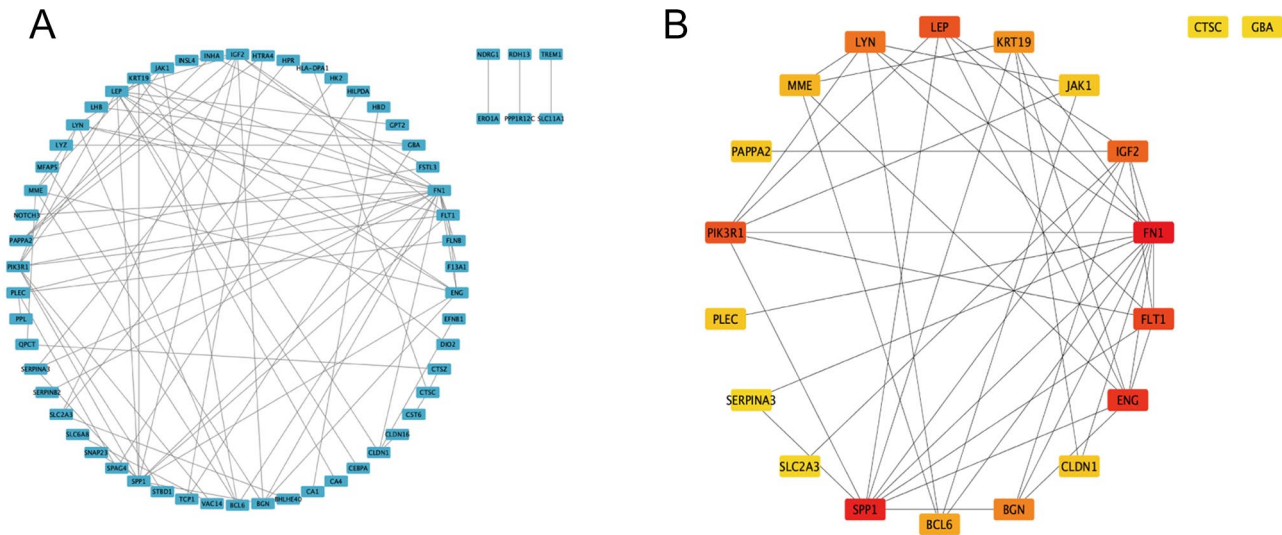
**Figure 3.** PPI network: (A) PPI network of DEGs in the PE data set and (B) Top 20 DEGs extracted by cytoHubba. PPI, protein-protein interaction; DEGs, differentially expressed genes.

classified into 3 groups according to gestational age: <14 weeks, 14 to 28 weeks, and >28 weeks. Our findings indicated that the expression levels of BCL6 and MME exhibited a significant decrease only in the third trimester, as shown in Figure 5G. Subsequently, we developed a prognostic model for PE based on the findings, with particular emphasis on the gene expression levels of BCL6 and MME. The predictive model can be represented as follows: prediction model = −46.05 + 1.24BCL6 + 3.05MME. To accurately assess our model, we constructed a nomogram and evaluated its precision by using a calibration curve (Figure 5 H and I). The diagnostic efficacy of each gene was assessed using ROC curves. The predictive value of our model, as demonstrated in Figure 5J, exceeded that of individual gene predictions in the discovery cohort. The Nomogram achieved an area under the curve (AUC) of 0.88 (95% CI: 0.80-0.95), while BCL6 achieved an AUC of 0.79 (95% CI: 0.70-0.88), and MME achieved an AUC of 0.85 (95% CI: 0.77-0.93). A consistent pattern was observed in all 5 validation cohorts (Supplementary Figure 4), indicating the strong predictive effectiveness of our PE model.

*Infiltrating immune cells in SLE and PE data sets and Correlation between key genes and immune cell infiltrate*

First, we applied the Cibersort algorithm to decipher immune cell profiles, aiming to understand immune regulation and identify links between diagnostic markers and immune cell infiltration in SLE. Figure 6A graphically illustrated the distribution of 22 types of immune cells in each sample. Here we noticed that SLE samples had higher proportions of neutrophils, plasma cells, monocytes, and dendritic cells activated

when compared to the control group (Figure 6B). We then studied the correlation between the expression levels of 2 key genes, BCL6 and MME, and the proportions of infiltrated immune cells. As depicted in Figure 6C, BCL6 and MME were positively correlated with neutrophil accumulation in SLE.

Next, we performed a similar immune infiltration analysis in the integrated PE data set and GSE48424 data set. Notably, PE placenta samples exhibited elevated proportions of Plasma cells and Eosinophils compared to the control group in the integrated PE data set (Figure 7A). While in the GSE48424 data set, higher proportions of CD8+ T cells, T cells memory resting, and Macrophages were observed in the peripheral blood samples of individuals with PE, when compared to the control group (Figure 7C). Furthermore, we found that in the integrated PE data set, there was a weak correlation between the expression levels of BCL6 and MME genes and the proportions of infiltrating immune cells (Figure 7B). However, in the GSE48424 data set, the expression levels of these 2 genes showed a strong positive correlation with neutrophil infiltration and a strong negative correlation with CD8+ T cells (Figure 7D), yet there was no statistically significant difference in neutrophil infiltration between the PE group and the control group in peripheral blood collected after the diagnosis of PE in the GSE48424 data set.

## Discussion
Our study stands as a pioneer in exploring the potential linkage between SLE and PE, with related research being relatively scarce. Our bioinformatic analysis of SLE and PE data sets aims to identify shared pathogenic genes between these 2 diseases. Interestingly, we identified 2 genes, BCL6 and MME,
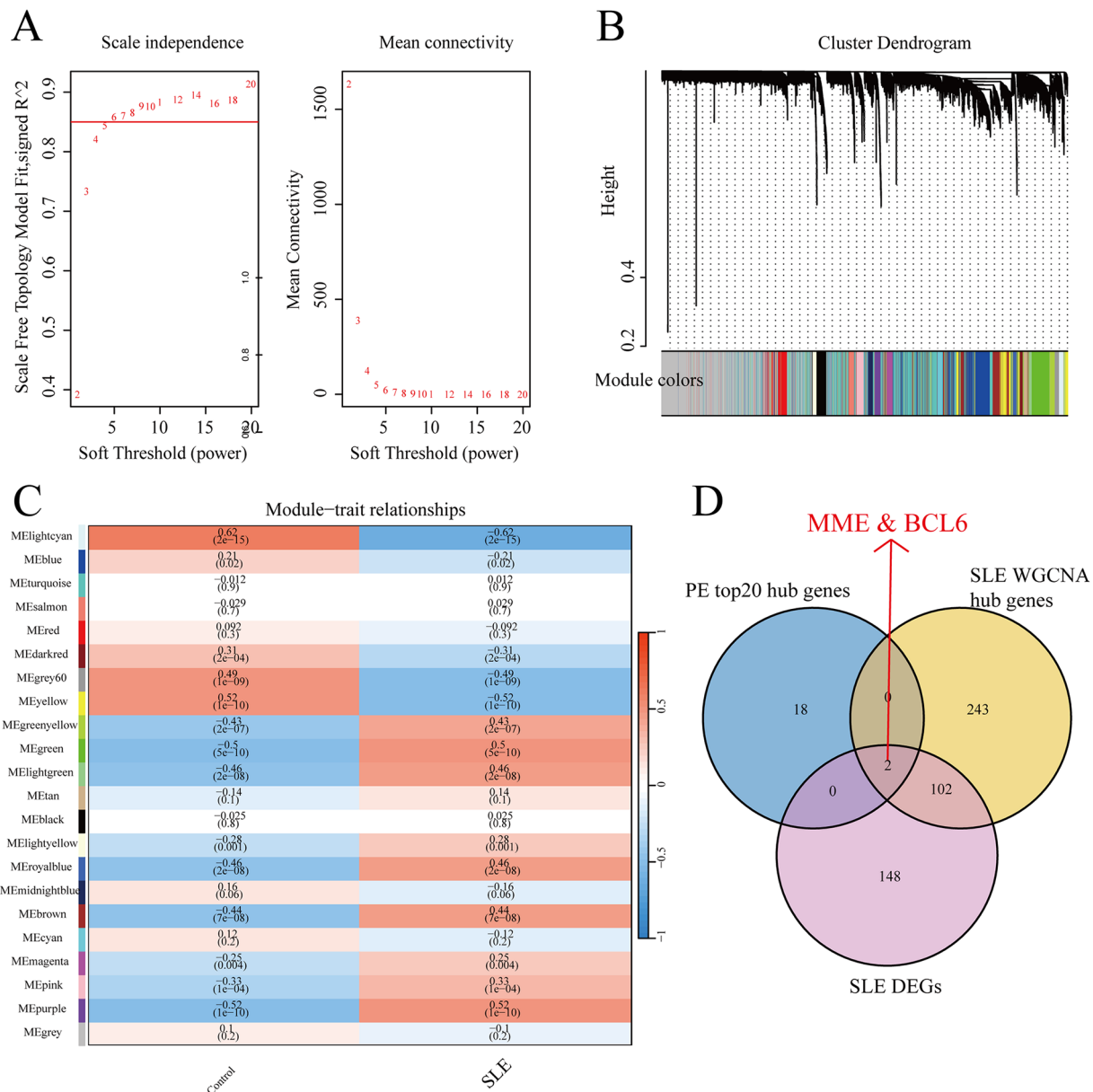
**Figure 4.** Analysis of hub genes in SLE via WGCNA; Identify key genes between SLE and PE: (A) Determination of soft-thresholding power; (B) The dendrogram of coexpression gene clusters in the SLE data set; (C) Heatmap of the correlation between modules and SLE; (D) The Venn diagram shows that 2 key genes are identified in SLE and PE. SLE, systemic lupus erythematosus; WGCNA, weighted gene co-expression network analysis; PE, pre-eclampsia.

which exhibited differential expression patterns across different sample types. Both genes showed upregulation in PBMCs from SLE patients and placental tissues from PE patients. This suggests their potential involvement in the immune dysregulation observed in SLE and the placental dysfunction associated with PE. Even more compelling, we formulated a diagnostic nomogram based on these genes. Patients' total scores, computed by summing each gene's score, hold substantial clinical relevance. High-scoring patients could be screened for PE, paving the way for early interventions to enhance maternal and fetal outcomes. Thus, our novel nomogram offers a valuable tool for managing these patients.

BCL6, short for B-cell lymphoma 6, is a zinc finger transcription factor located on chromosome 3q27. It plays a crucial role in the development of germinal center (GC) B-cells and follicular helper T (Tfh) cells by specifically targeting a range of genes involved in different cellular functions, including differentiation, survival, DNA damage response, and cell-cycle regulation.[13,14] Prior research has demonstrated the novel function of BCL-6 in the development of lupus by enhancing the production of interleukin (IL)-10 and IL-21, as well as stimulating IgG production by B cells.[15] Moreover, recent findings have highlighted important functions of BCL6 in trophoblastic cells with observations of elevated protein levels of BCL6 in
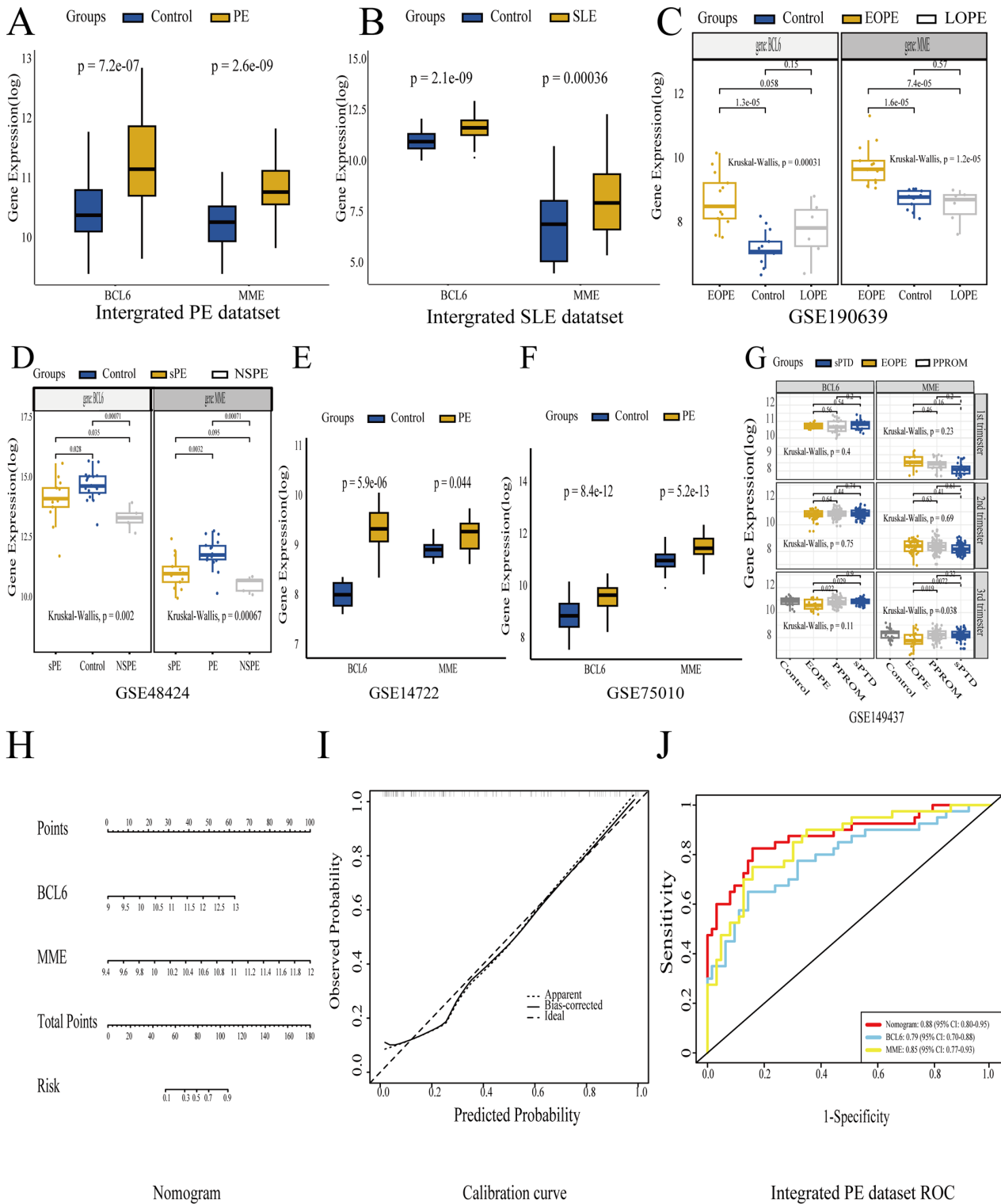
**Figure 5.** The validation of 2 key genes in PE and nomogram construction: (A) The expression of the 2 genes between PE and control in the integrated PE data set; (B) The expression of the 2 genes between SLE and control in the integrated SLE data set; (C) The expression of the 2 genes between EOPE, LOPE, and control in the GSE190639 data set; (D) The expression of the 2 genes in maternal peripheral blood samples from patients with severe PE, nonsevere PE, and control in the GSE48424 data set; (E) The expression of the 2 genes between PE and control in the GSE14722 data set; (F) The expression of the 2 genes between PE and control in the GSE75010 data set; (G) The expression of the 2 genes in maternal peripheral blood samples of EOPE, PPROM, and sPTD in the GSE149437 data set; (H) The nomogram for diagnosing PE; (I)Calibration curve to assess the accuracy of the nomogram; (J)The predictive value of the nomogram and the 2 genes in PE by the ROC. PE, pre-eclampsia; SLE, systemic lupus erythematosus; EOPE, early-onset pre-eclampsia; LOPE, late-onset pre-eclampsia; PPROM, preterm prelabor rupture of membrane; sPTD, spontaneous preterm delivery; ROC, receiver operating characteristic.
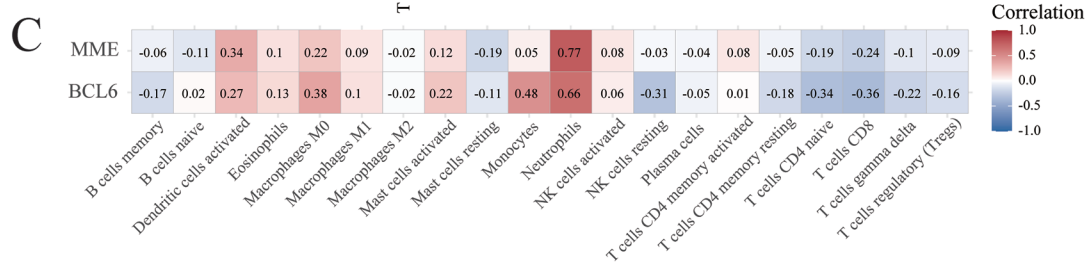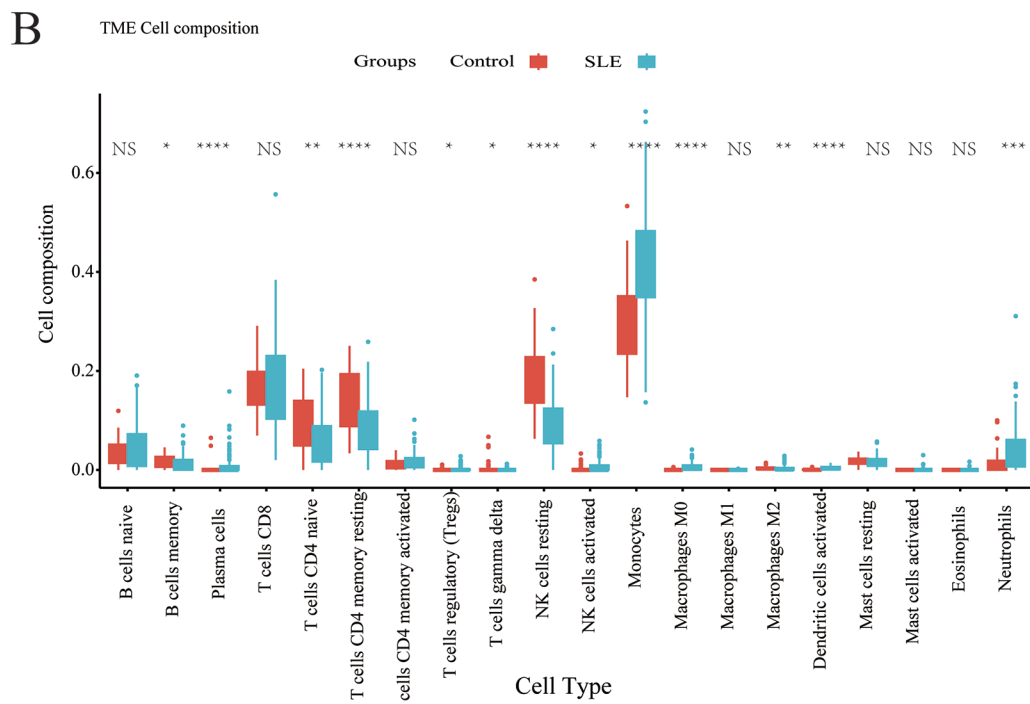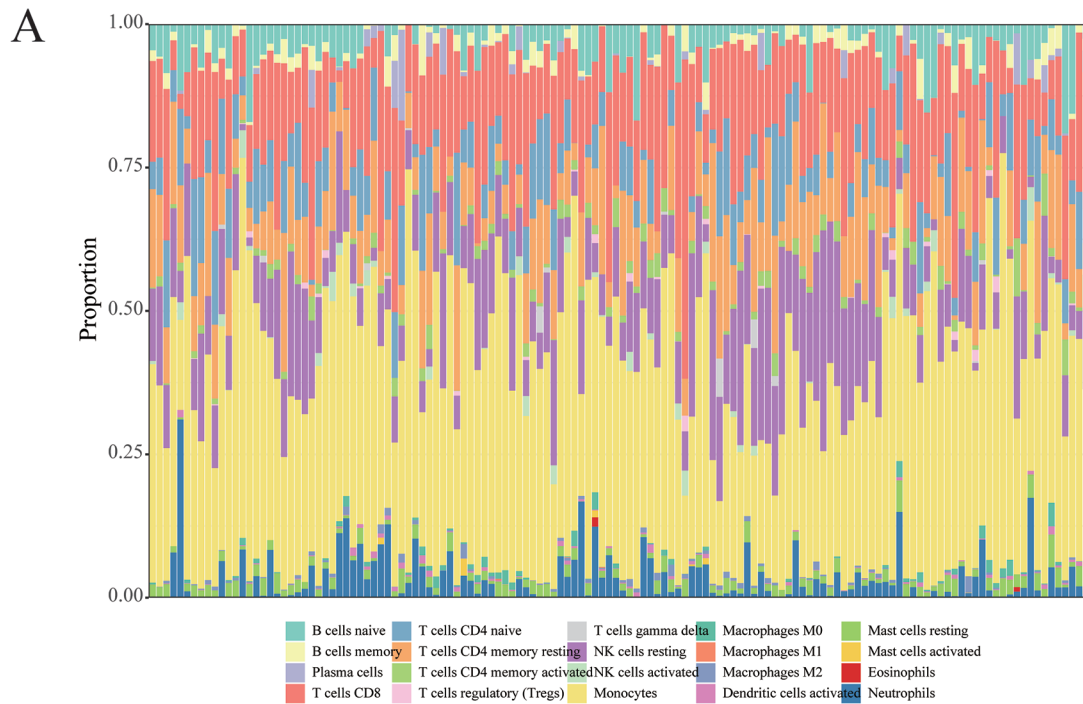
**Figure 6.** Immune cell infiltration analysis between SLE and control in the integrated SLE data set: (A) The bar plot of the proportion of 22 kinds of immune cells in different samples; (B) The boxplot displays the differences in immune cell expression between SLE and control; (C) Correlation analysis of immune cell infiltrations with 2 key genes. SLE, systemic lupus erythematosus.
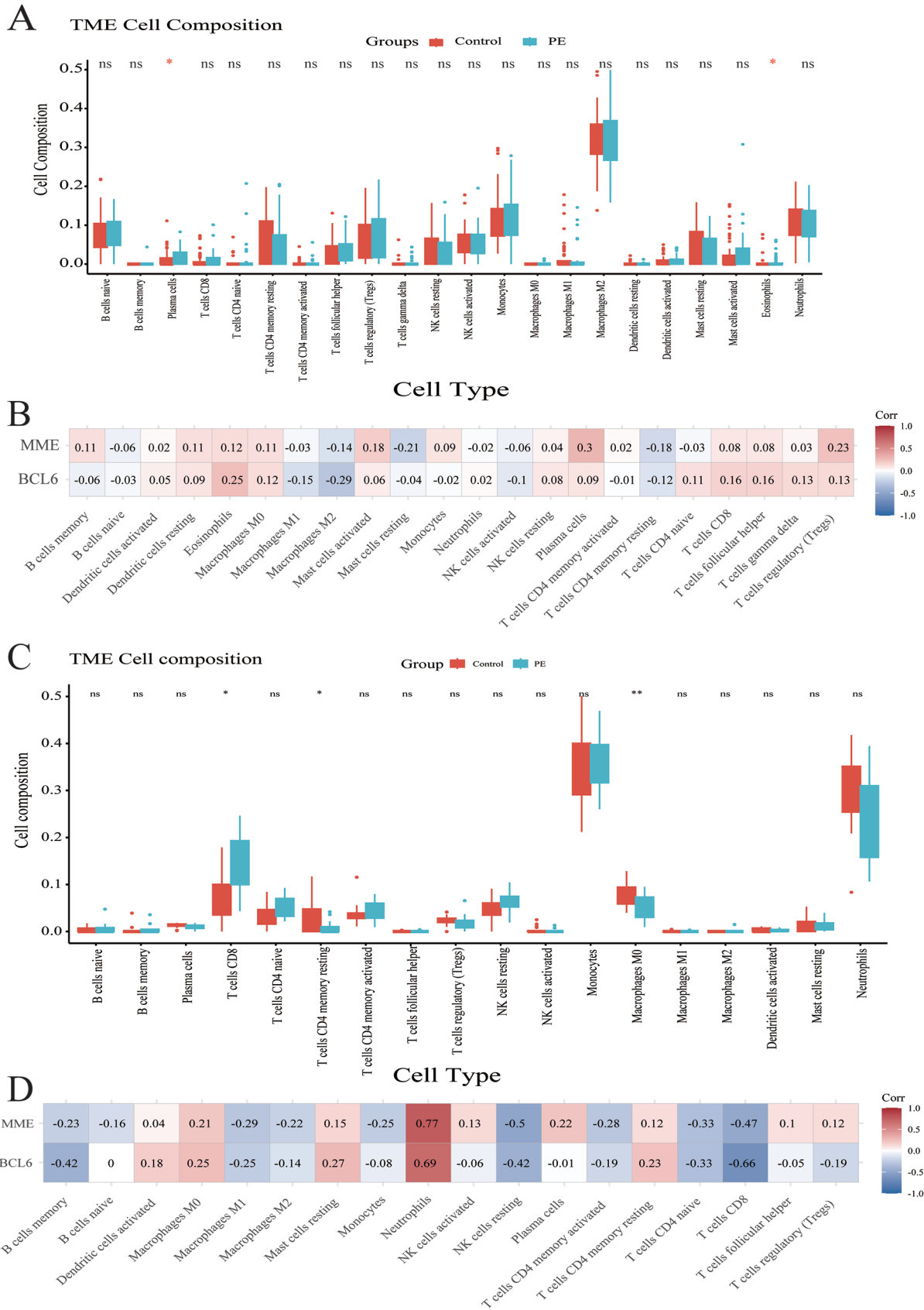
**Figure 7.** Immune cell infiltration analysis between PE and control in the PE data set and GSE48424: (A) The boxplot displays the differences in immune cell expression between PE and control in the integrated PE data set; (B) Correlation analysis of immune cell infiltrations with 2 key genes in the integrated PE data set; (C) The boxplot displays the differences in immune cell expression between PE and control in the GSE48424 data set; (D) Correlation analysis of immune cell infiltrations with 2 key genes in the GSE48424 data set. PE, pre-eclampsia.

pre-eclamptic placentas, particularly those that are preterm.[16] These findings were corroborated by our study. The protein seemed to be predominantly localized in the nucleus of villous cytotrophoblasts (vCTBs) beneath the syncytial layer.[17-20] Increased expression of BCL6 seemed to correlate with enhanced ARNT2, LEPTIN, and FLT1, hallmark genes of PE. In the "BCL6-ARNT2" pathway, specific to preterm PE, ischemic stress of trophoblasts coupled with overexpression of BCL6 and ARNT2 led to increased FLT1 expression. This contributed to the antiangiogenic state, hypertension, and EOPE.[21]

Notably, while the BCL6 protein was majorly localized in the nuclei of vCTBs in term placentas, it was also highly expressed in cell column trophoblasts (CCTs) of anchoring villi in first-trimester placentas.[22] Elevated BCL6 might be a molecular mechanism that facilitates the proliferation of cytotrophoblasts under low oxygen conditions in early pregnancy. However, the knockdown of BCL6 seemed to enhance the expression of fusion-related genes in trophoblastic cell lines as well as in primary vCTBs.[23] This suggested that BCL6 acted as a negative regulator for differentiation and fusion of vCTBs into the STB, and increased BCL6 in pre-eclamptic placentas may hinder proper differentiation and successful fusion, contributing to PE pathogenesis. In conclusion, while BCL6 seems to play various cellular roles in trophoblasts, further investigations are needed to define BCL6 expression levels across different stages of pregnancy, understand its functions, identify interaction partners, and ascertain its involvement in molecular networks. These could provide a more comprehensive understanding of the pathogenesis of PE.

Similar to BCL6, the MME gene, also located on human chromosome 3q21-27, is known as a common acute lymphoblastic leukemia antigen, neutral endopeptidase, or cluster of differentiation 10 (CD10). It is part of the CD system and belongs to the peptidase M13 family. Expression of MME is observed in early lymphoid progenitors, neutrophils, and endothelial cells.[24,25] As a zinc-dependent metalloprotease enzyme, MME can degrade various peptides such as amyloid beta-peptide, substance P, brain natriuretic peptide, oxytocin, and bradykinin.[26,27] It is involved in numerous pivotal BP. It has been reported that MME played a role in the hindered development of SLE endothelial progenitor cells into fully mature endothelial cells. This was partially due to the production of IFNα by CD10+ SLE LDNs.[28] In addition, mature CD10+ LDNs have been linked to the severity of noncalcified plaque burden and lower high-density lipoprotein (HDL) cholesterol efflux capacity in SLE patients.[29,30] Substrates of MME can induce migration and aggregation of neutrophils, acting as mediators of neutrophil inflammatory responses. In addition, previous studies have shown an increase in MME expression in the placentas during PE, suggesting its potential involvement in regulating peptide activity at the fetal-maternal interface.[31] Evidence also indicated that active MME released into maternal circulation from the placenta was significantly elevated in PE, potentially leading to complications such as

hypertension and heart failure.[32] The results of KEGG suggested a close relationship between DEGs in PE and protein degradation-related pathways. In conclusion, MME appeared to play critical roles in the pathology of SLE and PE.

Pathological inflammation in early pregnancy causing placental dysfunction, followed by the influx of placental factors into maternal circulation leading to endothelial damage and clinical signs, are generally accepted stages in PE development.[33] Importantly, we found both BCL6 and MME were positively correlated with neutrophils in SLE. We hypothesize that in lupus pregnancies, overexpression of these 2 key genes in the first trimester may result in trophoblast differentiation disorders, maternal vascular malperfusion, placental oxidative stress, and an antiangiogenic state, eventually leading to PE manifestations.

Owing to constraints in collecting placental samples and the late manifestation of PE symptoms, we verified the expression of 2 key genes, BCL6 and MME, in maternal circulation. Interestingly, when exploring their potential as early diagnostic biomarkers for PE in the GSE149437 data set, we found that during early to midpregnancy, the expression levels of BCL6 and MME in peripheral blood were elevated, although not statistically significant. However, in the third trimester, a significant downregulation of these genes was observed in peripheral blood samples from PE patients. In the third trimester after the diagnosis of PE, various physiological changes occur as the disease advances, affecting multiple systems and organs. This downregulation might reflect changes related to disease progression, resolution, or other factors associated with PE. These changes may be influenced by the complex interplay of various BP occurring in the late stages of pregnancy and the pathophysiology of PE. Overall, these findings highlight the complexity of gene expression dynamics in different sample types and stages of pregnancy. The differential expression patterns observed in peripheral blood and placental tissues suggest tissue-specific responses and underline the intricate nature of the underlying pathophysiology involved in PE. As there is not enough evidence, it is important to note that further experimental validations are required and further research is needed to elucidate the functional implications of BCL6 and MME in the context of immune dysregulation, placental dysfunction, and the pathogenesis of PE. In addition, larger-cohort studies and more comprehensive investigations are required to understand the specific roles of these genes and their potential as therapeutic targets.

In our study, we addressed the disparity of smaller placental transcriptomics studies by using advanced bioinformatics techniques to amalgamate microarray data sets across multiple platforms, thus generating large patient sample data sets. Our use of 5 different data sets, including peripheral blood samples, for key gene verification, added considerable strength to the study. However, it is important to note that this is a preliminary analysis and more in-depth research is needed to identify phenotype-specific networks. Further large-scale studies and functional analyses of key genes in PE are required to elucidate the mechanisms involved in PE pathogenesis.

## Conclusion

Our study predicts that key genes differentially expressed in SLE and associated with neutrophils may play significant roles in the pathogenesis of PE, especially EOPE. These findings lay the ideal groundwork for future experimental confirmation.

## Author Contributions

QD and YS conceived of the original idea for the study, interpreted results, carried out the statistical analysis, edited the paper, and was the overall guarantor. QD obtained ethical approval, contributed to the preparation of the data set, interpreted results, and contributed to drafting the paper. XT, ML, and JZ contributed to editing the paper. All authors have approved the final version of the manuscript.

## Ethics Approval and Consent to Participate

Not applicable.

## Consent for Publication

All authors gave their permission for publication.

## Data Availability Statement

Data used in the study (GSE81622, GSE50772, GSE10588, GSE25906, GSE190639, GSE48424, GSE14722, GSE75010, and GSE149437) are available from the public GEO data sets.

## ORCID iD

Qianwen Dai https://orcid.org/0000-0002-6162-527X

## SUPPLEMENTAL MATERIAL

Supplemental material for this article is available online.

## REFERENCES

1. Lisnevskaia L, Murphy G, Isenberg D. Systemic lupus erythematosus. *Lancet*. 2014;384:1878-1888. doi:10.1016/s0140-6736(14)60128-8
2. Khan KS, Wojdyla D, Say L, Gülmezoglu AM, Van Look PF. WHO analysis of causes of maternal death: a systematic review. *Lancet*. 2006;367:1066-1074. doi:10.1016/s0140-6736(06)68397-9
3. Steegers EA, von Dadelszen P, Duvekot JJ, Pijnenborg R. Pre-eclampsia. *Lancet*. 2010;376:631-644. doi:10.1016/s0140-6736(10)60279-6
4. Rolnik DL, Wright D, Poon LC, et al. Aspirin versus Placebo in pregnancies at high risk for preterm preeclampsia. *N Engl J Med*. 2017;377:613-622. doi:10.1056/NEJMoa1704559
5. Nahum Sacks K, Friger M, Shoham-Vardi I, et al. Prenatal exposure to preeclampsia as an independent risk factor for long-term cardiovascular morbidity of the offspring. *Pregnancy Hypertens*. 2018;13:181-186. doi:10.1016/j.preghy.2018.06.013
6. Bellamy L, Casas JP, Hingorani AD, Williams DJ. Pre-eclampsia and risk of cardiovascular disease and cancer in later life: systematic review and meta-analysis. *BMJ*. 2007;335:974. doi:10.1136/bmj.39335.385301.BE
7. Holers VM. Complement therapeutics are coming of age in rheumatology. *Nat Rev Rheumatol*. 2023;19:470-485. doi:10.1038/s41584-023-00981-x
8. Gelber SE, Brent E, Redecha P, et al. Prevention of defective placentation and pregnancy loss by blocking innate immune pathways in a syngeneic model of placental insufficiency. *J Immunol*. 2015;195:1129-1138. doi:10.4049/jimmunol.1402220
9. Lager S, Sovio U, Eddershaw E, et al. Abnormal placental CD8(+) T-cell infiltration is a feature of fetal growth restriction and pre-eclampsia. *J Physiol*. 2020;598:5555-5571. doi:10.1113/jp279532
10. Wirestam L, Arve S, Linge P, Bengtsson AA. Neutrophils-important communicators in systemic lupus erythematosus and antiphospholipid syndrome. *Front Immunol*. 2019;10:2734. doi:10.3389/fimmu.2019.02734
11. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 2008;9:559. doi:10.1186/1471-2105-9-559
12. Szklarczyk D, Gable AL, Lyon D, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res*. 2019;47:D607-D613. doi:10.1093/nar/gky1131
13. Choi J, Crotty S. Bcl6-mediated transcriptional regulation of follicular helper t cells (T(FH)). *Trends Immunol*. 2021;42:336-349. doi:10.1016/j.it.2021.02.002
14. Mlynarczyk C, Fontán L, Melnick A. Germinal center-derived lymphomas: the darkest side of humoral immunity. *Immunol Rev*. 2019;288:214-239. doi:10.1111/imr.12755
15. Ding S, Zhang Q, Luo S, et al. BCL-6 suppresses miR-142-3p/5p expression in SLE CD4(+) T cells by modulating histone methylation and acetylation of the miR-142 promoter. *Cell Mol Immunol*. 2020;17:474-482. doi:10.1038/s41423-019-0268-3
16. Louwen F, Kreis NN, Ritter A, Friemel A, Solbach C, Yuan J. BCL6, a key oncogene, in the placenta, pre-eclampsia and endometriosis. *Hum Reprod Update*. 2022;28:890-909. doi:10.1093/humupd/dmac027
17. Ren Z, Gao Y, Gao Y, et al. Distinct placental molecular processes associated with early-onset and late-onset preeclampsia. *Theranostics*. 2021;11:5028-5044. doi:10.7150/thno.56141
18. Guo F, Zhang B, Yang H, et al. Systemic transcriptome comparison between early- and late-onset pre-eclampsia shows distinct pathology and novel biomarkers. *Cell Prolif*. 2021;54:e12968. doi:10.1111/cpr.12968
19. Than NG, Romero R, Tarca AL, et al. Integrated systems biology approach identifies novel maternal and placental pathways of preeclampsia. *Front Immunol*. 2018;9:1661. doi:10.3389/fimmu.2018.01661
20. Louwen F, Muschol-Steinmetz C, Friemel A, et al. Targeted gene analysis: increased B-cell lymphoma 6 in preeclamptic placentas. *Hum Pathol*. 2014;45:1234-1242. doi:10.1016/j.humpath.2014.02.002
21. Levine RJ, Lam C, Qian C, et al. Soluble endoglin and other circulating antiangiogenic factors in preeclampsia. *N Engl J Med*. 2006;355:992-1005. doi:10.1056/NEJMoa055352
22. Ritter A, Safdar BK, Jasmer B, et al. The function of oncogene B-cell lymphoma 6 in the regulation of the migration and invasion of trophoblastic cells. *Int J Mol Sci*. 2020;21(21):8393. doi:10.3390/ijms21218393
23. Jasmer B, Muschol-Steinmetz C, Kreis NN, et al. Involvement of the oncogene B-cell lymphoma 6 in the fusion and differentiation process of trophoblastic cells of the placenta. *Oncotarget*. 2017;8:108643-108654. doi:10.18632/oncotarget.20586
24. Marini O, Costa S, Bevilacqua D, et al. Mature CD10(+) and immature CD10(-) neutrophils present in G-CSF-treated donors display opposite effects on T cells. *Blood*. 2017;129:1343-1356. doi:10.1182/blood-2016-04-713206
25. Turner AJ, Isaac RE, Coates D. The neprilysin (NEP) family of zinc metalloendopeptidases: genomics and function. *Bioessays*. 2001;23:261-269. doi:10.1002/1521-1878(200103)23
26. Sankhe R, Pai SRK, Kishore A. Tumour suppression through modulation of neprilysin signaling: a comprehensive review. *Eur J Pharmacol*. 2021;891:173727. doi:10.1016/j.ejphar.2020.173727
27. Bayes-Genis A, Barallat J, Richards AM. A test in context: neprilysin: function, inhibition, and biomarker. *J Am Coll Cardiol*. 2016;68:639-653. doi:10.1016/j.jacc.2016.04.060
28. Denny MF, Yalavarthi S, Zhao W, et al. A distinct subset of proinflammatory neutrophils isolated from patients with systemic lupus erythematosus induces vascular damage and synthesizes type I IFNs. *J Immunol*. 2010;184:3284-3297. doi:10.4049/jimmunol.0902199
29. Mistry P, Nakabo S, O'Neil L, et al. Transcriptomic, epigenetic, and functional analyses implicate neutrophil diversity in the pathogenesis of systemic lupus erythematosus. *Proc Natl Acad Sci U S A*. 2019;116:25222-25228. doi:10.1073/pnas.1908576116
30. Carlucci PM, Purmalek MM, Dey AK, et al. Neutrophil subsets and their gene signature associate with vascular inflammation and coronary atherosclerosis in lupus. *JCI Insight*. 2018;3(8):e99276. doi:10.1172/jci.insight.99276
31. Li XM, Moutquin JM, Deschênes J, Bourque L, Marois M, Forest JC. Increased immunohistochemical expression of neutral metalloendopeptidase (enkephalinase; EC 3.4.24.11) in villi of the human placenta with pre-eclampsia. *Placenta*. 1995;16:435-445. doi:10.1016/0143-4004(95)90101-9
32. Gill M, Motta-Mejia C, Kandzija N, et al. Placental syncytiotrophoblast-derived extracellular vesicles carry active NEP (Neprilysin) and are increased in preeclampsia. *Hypertension*. 2019;73:1112-1119. doi:10.1161/hypertensionaha.119.12707
33. Moufarrej MN, Vorperian SK, Wong RJ, et al. Early prediction of preeclampsia in pregnancy with cell-free RNA. *Nature*. 2022;602:689-694. doi:10.1038/s41586-022-04410-z