

# Bioinformatic Analysis of Apoptosis-Related Genes in Preeclampsia Using Public Transcriptomic and Single-Cell RNA Sequencing Datasets

Lingyan Liu<sup>1-3,\*</sup>, Xiuling Li<sup>2,3,\*</sup>, Hongfen Yang<sup>1,2,3</sup>, Fei Xu<sup>3,4</sup>, Xudong Dong<sup>2,3</sup>

<sup>1</sup>Faculty of Life Science and Technology, Kunming University of Science and Technology, Kunming, 650500, People's Republic of China; <sup>2</sup>Department of Obstetrics, The First People's Hospital of Yunnan Province, The Affiliated Hospital of Kunming University of Science and Technology, Kunming, 650500, People's Republic of China; <sup>3</sup>Medical School, Kunming University of Science and Technology, Kunming, 650500, People's Republic of China; <sup>4</sup>Department of Pain Management, The First People's Hospital of Yunnan Province, The Affiliated Hospital of Kunming University of Science and Technology, Kunming, 650500, People's Republic of China

\*These authors contributed equally to this work

Correspondence: Fei Xu; Xudong Dong, The First People's Hospital of Yunnan Province, No. 157 Jinbi Road, Xishan District, Kunming, 650500, People's Republic of China, Tel +861871-63638625, Email 494428530@qq.com; xudong52@outlook.com

**Purpose:** Apoptosis, which is crucial in preeclampsia (PE), affects trophoblast survival and placental function. We used transcriptomics and single-cell RNA sequencing (scRNA-seq) to explore apoptosis-related genes (ARGs) and their cellular mechanisms as potential PE biomarkers.

**Patients and Methods:** All the data included in this study were sourced from public databases. We used scRNA-seq and differential expression analysis, combined with five algorithms from the CytoHubba plugin, to identify ARGs as PE biomarkers. These were integrated into diagnostic nomograms. Mechanistic studies involved enrichment analysis and immune profiling. Biomarker expression was examined at the single-cell level, and verified in clinical samples by RT-qPCR.

**Results:** Cluster of Differentiation 44 (CD44), Macrophage migration inhibitory factor (MIF), PIK3R1, and Toll-like receptor 4 (TLR4) were identified as PE biomarkers. CD44 and TLR4 were down-regulated, while MIF and PIK3R1 were up-regulated. When integrated into the diagnostic nomogram, they showed clinical utility and affected cell functions. In the immune profile of PE, monocytes decreased, resting NK cells increased, and the activities of APC, checkpoint, T-cell co-stimulation, and MHC class I pathways reduced. ScRNA-seq identified 11 cell types, 10 of which were significantly different. Endothelial cell communication with other cell types decreased, while the interaction between common myeloid progenitors (CMP) and villous cytotrophoblasts (VCT) enhanced. The expression levels of CD44, MIF, and PIK3R1 in VCT were significantly different and key to PE. Their decrease in early PE and increase in late PE reflected the placenta's adaptation to adverse pregnancy conditions.

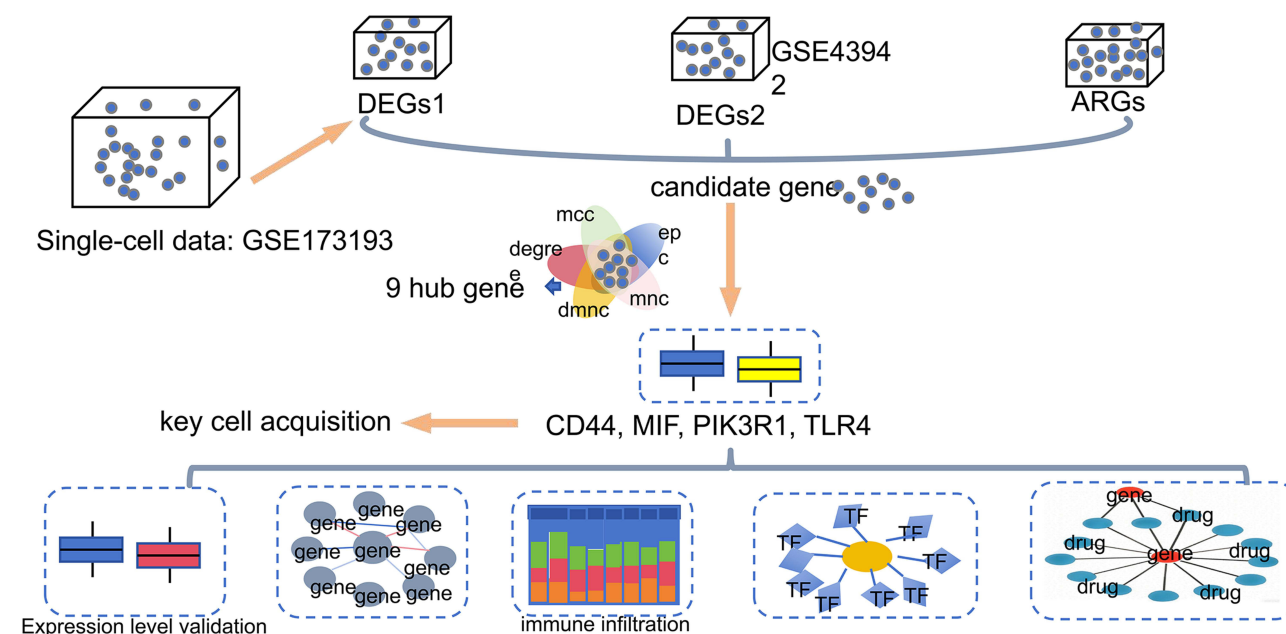
**Conclusion:** Four ARGs, CD44, MIF, PIK3R1, and TLR4, identified through comprehensive analyses, served as significant biomarkers for PE and offered insights into PE's cellular mechanisms of PE, providing valuable references for further research.

**Keywords:** preeclampsia, apoptosis, biomarkers, single-cell RNA sequencing, villous cytotrophoblast cells

## Introduction

Preeclampsia (PE) is a pregnancy-specific complication mainly characterized by high blood pressure and proteinuria after 20 weeks of pregnancy. The global incidence of PE is approximately 2–8%, with higher rates in certain high-risk groups, such as women with first-time pregnancies, older age, multiple pregnancies, or a family history of PE.<sup>1</sup> This disease not only poses a serious threat to the health of pregnant women, but can also lead to intrauterine growth restriction of the fetus, iatrogenic premature delivery, and even endangers the life of the mother and fetus in severe cases.<sup>2</sup> The clinical

## Graphical Abstract



manifestations of PE range from mild hypertension and proteinuria to severe renal dysfunction, liver dysfunction, retinopathy, and neurological complications, which vary in severity and speed of progression, and pose challenges for clinical diagnosis and management. Although previous studies have shown that PE may be related to genetics, immune system disorders, vascular endothelial dysfunction, and other factors, the specific pathogenesis of PE is not fully understood, there is still a lack of effective early prediction methods, and early termination of pregnancy is the only effective treatment is early termination of pregnancy.<sup>3</sup> Therefore, early diagnosis and effective intervention based on pathogenesis are key to improving pregnancy outcomes in PE.

Apoptosis is a form of programmed cell death and is a widespread biological process. It encompasses related mechanisms such as phagocytosis, which is crucial for the clearance of apoptotic cells.<sup>4</sup> Additionally, several other processes are closely related to apoptosis. Programmed necrosis (a regulated form of necrosis), pyroptosis (inflammation-related),<sup>5</sup> autophagy (a self-degradation process that maintains intracellular homeostasis),<sup>6</sup> and ferroptosis (an iron-dependent form of cell death)<sup>7</sup> may all play significant roles in diseases such as preeclampsia.<sup>8</sup> Apoptosis can help remove cells that are damaged or no longer required, thereby protecting the organism from potential pathological changes.<sup>9</sup> Apoptosis plays a key role in the PE<sup>10</sup> cells PE is a pregnancy complication that is closely related to abnormal placental function. The main placental cell types include villous cytotrophoblasts (VCT), syncytiotrophoblasts (SCT), and extravillous trophoblasts (EVT). These cells maintain normal placental functions through continuous differentiation and renewal. Studies have shown that the rate of placental cell apoptosis is abnormal in patients with PE, which may be related to insufficient blood flow and oxygen supply to the placenta.<sup>11,12</sup> In the context of PE, apoptosis of VCT and EVT cells may increase, while apoptosis of SCT cells may decrease, and this unbalanced apoptotic activity may lead to placental dysfunction, which in turn affects the entire course of pregnancy.<sup>13,14</sup>

During a normal pregnancy, the apoptosis of placental cells is in a dynamic balance to maintain the normal development and function of the placenta. However, this balance is disrupted in patients with preeclampsia. Factors such as oxidative stress, inflammatory responses, and placental ischemia and hypoxia can induce abnormal activation of the apoptosis pathways in placental cells.<sup>15</sup> For example, excessive reactive oxygen species (ROS) can damage intracellular DNA, proteins, and lipids, thereby activating endogenous apoptosis pathways, promoting the occurrence of caspase cascade reactions, and ultimately leading to cell apoptosis.<sup>16</sup> Additionally, the abnormal elevation of

inflammatory factors such as tumor necrosis factor- $\alpha$  (TNF- $\alpha$ ) can activate death receptors on the cell surface through exogenous apoptosis pathways, triggering cell apoptosis.<sup>17</sup> This abnormal apoptosis further disrupts the structure and function of the placenta, leading to abnormal placental vascular remodeling and decreased trophoblast invasion ability, ultimately resulting in a series of clinical symptoms of preeclampsia.<sup>17</sup> Current research has identified various potential biomarkers associated with preeclampsia. In addition to the apoptosis-related genes (ARGs) focused on in this study, placental growth factor (PlGF) and soluble fms-like tyrosine kinase-1 (sFlt-1) have also garnered attention.<sup>18</sup> PlGF is a cytokine that promotes angiogenesis, and its levels are typically significantly reduced in patients with preeclampsia, which is closely related to placental vascular dysplasia.<sup>19</sup> sFlt-1 is an anti-angiogenic factor that is highly expressed in patients with preeclampsia; it can bind to angiogenic factors like PlGF, blocking their biological activity and leading to an imbalance in angiogenesis, which is one of the important molecular mechanisms of preeclampsia.<sup>20</sup> In terms of molecular regulatory mechanisms, there exists a complex interaction network among these biomarkers. For instance, certain genes within ARGs may influence the expression of PlGF and sFlt-1 by regulating intracellular signaling pathways, thereby affecting placental vascular generation and function. Currently, commonly used clinical diagnostic methods include measuring the levels of related biomarkers in the blood of pregnant women and ultrasound examinations to assess the condition of the placenta and fetus.<sup>21</sup> However, these methods have certain limitations; therefore, in-depth exploration of new diagnostic biomarkers and molecular regulatory mechanisms is of great significance for improving the accuracy of early diagnosis of preeclampsia.

Although apoptosis plays an important role in the occurrence and development of PE, the specific molecular mechanism remains unclear, prompting researchers to explore the specific biological role of apoptosis in PE and its molecular mechanisms with a view to finding possible therapeutic targets to improve pregnancy outcomes. In summary, the study of apoptosis in PE not only contributes to our understanding of the nature of this complex pathological state but may also provide key information for the development of new prevention and treatment strategies.

Recent studies have identified several biomarkers associated with preeclampsia. For example, CD93, a marker related to phagocytosis, has been shown to be associated with preeclampsia.<sup>22</sup> CD93 is involved in various cell signaling pathways. It can interact with various ligands and receptors on the cell surface, affecting processes such as cell adhesion and migration.<sup>23</sup> In the context of apoptosis and immune response, CD93 may regulate the clearance of apoptotic cells by immune cells, and its abnormal expression in preeclampsia may disrupt the normal balance of placental cell renewal and immune regulation, thereby promoting the development of the disease.<sup>24</sup> In addition to CD93, other potential biomarkers for the diagnosis of preeclampsia have been studied, such as TRA4, PROCR, and TXNIP, which were identified by Li Weiwen et al through bioinformatics and machine learning methods as potential PE biomarkers.<sup>25</sup> These continuously discovered biomarkers deepen our understanding of the pathogenesis of preeclampsia. Furthermore, apoptosis is closely related to oxidative stress, which is also associated with the development of preeclampsia.<sup>26</sup> Xanthine oxidase is an enzyme involved in the generation of reactive oxygen species (ROS) during oxidative stress.<sup>27</sup> The serum uric acid to serum creatinine ratio, as an indirect marker of xanthine oxidase activity, has been shown to be related to the development of preeclampsia.<sup>28</sup> Elevated levels of xanthine oxidase and abnormal serum uric acid to serum creatinine ratios may lead to increased oxidative stress in the placenta, triggering placental cell apoptosis, disrupting normal placental function, and ultimately promoting the occurrence and development of preeclampsia.<sup>29</sup> Thus, research on apoptosis in preeclampsia not only helps us understand the nature of this complex pathological state but may also provide key information for developing new prevention and treatment strategies.

Based on transcriptome and single-cell RNA sequencing (scRNA-seq) data in public databases, this study explored the potential diagnostic biomarkers and molecular regulatory mechanisms of PE through apoptosis-related genes using bioinformatics to provide new ideas for the clinical diagnosis, prevention, and treatment of PE.

## Materials and Methods

### Data Source

Two transcriptome datasets related to PE and a single-cell RNA sequencing (scRNA-seq) dataset were obtained from the Gene Expression Omnibus (GEO) database (<https://www.ncbi.nlm.nih.gov/gds>). The GSE43942 dataset, based on the

GPL10191 platform using a chip sequencing method, served as the training set for this study and included placental tissue samples from 5 PE patients and 7 normal individuals.<sup>30</sup> The clinical information table of the GSE43942 dataset was shown in the [Table S1](#). The GSE25906 dataset, utilized as the validation set, was based on the GPL6102 platform with chip sequencing, featuring placental tissue samples from 23 patients with PE and 37 normal individuals.<sup>31</sup> The clinical information table of the GSE25906 dataset was shown in the [Table S2](#). The scRNA-seq dataset GSE173193, based on the GPL24676 platform with high-throughput sequencing, included placental tissue samples from 2 PE patients and 2 normal individuals.<sup>32,33</sup> The clinical information table of the GSE173193 dataset was shown in the [Table S3](#). In addition, 680 apoptosis-related genes (ARGs) were extracted for analysis from a previously published article for analysis.<sup>34</sup>

## ScRNA-Seq Analysis

To enhance our understanding of PE at the cellular level, we used the Seurat package (v 5.0.1)<sup>35</sup> to perform a comprehensive analysis of the scRNA-seq data in the GSE173193 dataset. Strict quality control (QC) measures were implemented on the scRNA-seq data to exclude low-quality data that might have been caused by cell damage or library preparation defects. To filter the single-cell sequencing data, the data was first integrated. GSM5261695 and GSM5261696 were used as the control group, and GSM5261699 and GSM5261700 were used as the disease group. After being read by read10X, the data was converted into a Seurat object. During the reading process, cells with fewer than 100 genes and genes covered by fewer than 3 cells were filtered out. The PercentageFeatureSet function in the “Seurat” R package was used to calculate the “nCount/nFeature-RNA” and “percent.Mt” of the samples for data quality control. The “Mt” of the samples was calculated by the percentage feature set function. Next, the percentages of the number of genes (nfeature), the number of cells (ncount), and the number of mitochondrial sequencing reads (Mt) in each sample were calculated. Meanwhile, cells with more than 20% mitochondrial genes were filtered out. Cells with the number of intracellular genes less than or equal to 100 and greater than or equal to 10,000 were removed, and genes with count numbers less than or equal to 100 and greater than or equal to 30,000 were also removed, so as to exclude the cases of multi-cell bodies or some cells with poor quality or in a dying state. The 24,100 cells and 22,891 genes that met the screening criteria were used for subsequent analysis. Subsequently, the data were normalized using the NormalizeData function, and the top 2500 highly variable genes were selected for downstream analysis using the FindVariableFeatures function based on the relationship between mean and variance. The results were visualized by applying the LabelPoints function, which highlighted the top ten genes with the greatest variation. The data were further normalized using the ScaleData function, and an elbow plot was created to determine the appropriate principal components (PCs) for further analysis. Utilizing the standard Seurat workflow, Uniform Manifold Approximation and Projection (UMAP) was employed to perform clustering analysis on dimensionality-reduced cells, with the resolution set at 0.5. After obtaining cell clusters through clustering, marker genes were identified from the GSE173193 dataset according to the reference<sup>12</sup> used in this study. Based on these markers, cell annotation was conducted by classifying cells that exhibited high specificity for gene expression within clusters. For clusters lacking marker genes, annotations were supplemented using singleR (v 2.0.0),<sup>36</sup> with the threshold set at a quantile of 0.7. Under these circumstances, cell types specific to this study were annotated and a histogram was created to display the proportions of different cell types in the PE and control samples.

## Differential Expression Analysis

When conducting differential expression analysis, since a large number of genes were simultaneously tested, in order to avoid an excessive number of false-positive results due to multiple testing, we adopted a strict statistical correction method. Specifically, we used the Benjamini-Hochberg method to control the False Discovery Rate (FDR).<sup>37</sup> This method sorted all the P-values of the tests and calculated the corrected q-values based on the sorting results to ensure that the proportion of false discoveries during the entire testing process was maintained at a reasonable level. In actual operation, we took the corrected q-value as one of the important bases for judging the differential expression of genes. Only when the q-value was less than the pre-set threshold (such as 0.05) was the differential expression of the gene considered to be statistically significant.<sup>38</sup> Subsequently, differential expression analysis was performed. Initially, within



the GSE173193 dataset, differences between the PE and control samples across various cell types were compared using the Wilcoxon test, and box plots were used for visualization. Cell types showing significant differences in abundance between samples were defined as differential cells for subsequent analysis ( $P < 0.05$ ). In addition, in the GSE173193 dataset, the FindMarkers function was used to screen for differentially expressed genes (DEGs) within differential cells between PE and control samples, with thresholds set at  $|\log_2\text{Fold Change (FC)}| \geq 0.5$ ,  $\text{adj. } P < 0.05$ , and  $\text{ptc} > 0.1$ , defined as DEGs1, and a Manhattan plot was created for visualization. In the training set, differential expression analysis was performed using the limma package (v 3.54.0)<sup>39</sup> to identify DEGs between PE and control samples, with the thresholds  $|\log_2\text{FC}| > 0.5$  and  $P < 0.05$ , defined as DEGs2. For more intuitive observation of gene differences, volcano plot and heatmap were generated using ggplot2 (v 3.4.1)<sup>40</sup> and ComplexHeatmap (v 2.15.1),<sup>41</sup> respectively, where the volcano plot displayed the top 10 up- and down-regulated genes sorted by  $\log_2\text{FC}$ .

## Identification and Enrichment Analysis of Differentially Expressed ARGs (DE-ARGs)

Previously identified DEGs1, DEGs2, and ARGs were intersected using the VennDiagram package (v 1.7.1),<sup>42</sup> resulting in the identification of DE ARGs. These DE-ARGs were then subjected to Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses to explore their biological functions and signaling pathways. This process was performed using the clusterProfiler package (v. 4.2.2),<sup>43</sup> with a significance level of  $P < 0.05$ . The GO terms included three categories: biological processes (BP), molecular functions (MF), and cellular components (CC). To explore the protein-level interactions of DE-ARGs, these genes were input into the Search Tool for the Retrieval of Interacting Genes (STRING) database (<https://string-db.org>) with a confidence score threshold of  $> 0.15$  for selection. The resulting protein-protein interaction (PPI) network was visualized using the Cytoscape software (v. 3.5.2).<sup>44</sup>

## Recognition and Expression Assessment of Biomarkers

In the training set, using the constructed PPI network, five algorithms from the CytoHubba plugin in Cytoscape software were employed: degree, MCC (Maximal Clique Centrality), DMNC (Density of Maximum Neighborhood Component), MNC (Maximum Neighborhood Component), and EPC (edge-percolated component) were employed. Each algorithm was used to identify the top 9 genes based on their respective scores. The intersection of these genes across all the algorithms yielded a set of hub genes, and Venn diagrams were generated using the VennDiagram package for visual representation. The expression of hub genes in the PE and control samples was verified using both training and validation datasets. Genes exhibiting consistent expression trends and significant differences between the two datasets were defined as biomarkers for subsequent analyses ( $P < 0.05$ ).

## Construction of a Nomogram Through Biomarkers

To maximize the diagnostic value of biomarkers for PE, the rms package (v. 6.5–1)<sup>45</sup> was used to construct the nomogram. Within this nomogram, a scoring system was devised based on the expression levels of biomarkers in the training set, with each factor assigned a specific score. The cumulative sum of the scores for various factors represents the total number of points. Based on the total points, the probability of a PE diagnosis was predicted, with higher scores indicating a greater likelihood of PE. Furthermore, a calibration curve was plotted to perform the Hosmer-Lemeshow (HL) test, a critical diagnostic for the predictive accuracy of the nomogram, and a p-value greater than 0.05, from the HL test, indicated that the predicted values did not significantly deviate from the actual values, thus passing the test. The ggDCA package (v 1.2) (<https://www.rdocumentation.org/packages/ggDCA/versions/1.1>) was used to conduct decision curve analysis (DCA) to assess the clinical utility of the nomogram.

## Gene Set Enrichment Analysis (GSEA) of Biomarkers

After identifying the biomarkers, we conducted a series of analyses to characterize their functionality. To uncover the biological functions and pathways linked to biomarkers, the “h.hallmark.v7.4. symbols.gmt” from the Gene Set Enrichment Analysis (GSEA) database (<http://www.broadinstitute.org/gsea/index.jsp>) was selected as the reference gene set. Spearman correlation coefficients were calculated between each biomarker and all other genes in the training set, ranking the genes in descending order based on these coefficients to identify the related genes for each biomarker.

Subsequent GSEA for each biomarker, conducted via the clusterProfiler package, considered results with  $P < 0.05$  as significantly enriched. The top five enriched pathways were visualized using the enrichment plot package (v. 1.18.3).<sup>46</sup>

## Functional and Annotation Analysis

Variations caused by genetic and non-genetic diseases can be explored through the association between genotypes and phenotypes.<sup>47</sup> The human genome contains hundreds of single-nucleotide variations (SNVs), many of which are known to cause diseases.<sup>48</sup> Approximately 21% of the amino acid variations lead to missense mutations at different protein sites, known as post-translational modifications, which can result in diseases.<sup>49</sup> Consequently, to understand the SNV profiles of biomarkers, the ActiveDriverDB online tool (<https://activedriverdb.org/>) was used for SNV analysis of biomarkers.

Additionally, analysis was performed to evaluate the functional similarity of biomarkers using the GOSemSim package (v.2.27.2).<sup>50</sup> The required GO annotation data for human species across the BP, CC, and MF aspects were constructed using GOSemSim. Semantic similarities were calculated using the mgeneSim function. Thereafter, the geometric mean of these similarities across all categories was computed to derive a final score to evaluate the overall functional impact of the biomarkers across the three GO domains.

GeneMANIA (<http://www.genemania.org/>) was used to predict genes functionally related to biomarkers, and a co-expression network of the top 20 interacting genes and their associated functions was visualized.

Physicochemical properties of proteins reflect the fundamental characteristics of an organism.<sup>51</sup> To understand the physicochemical properties of the biomarkers, the amino acid sequences were retrieved from UniProt (<https://www.uniprot.org/>). Basic information on the protein-coding genes (biomarkers), including amino acid composition, molecular weight, protein half-life, and calculated parameters, such as the protein instability index and theoretical isoelectric point, was analyzed using the ProtParam tool (<https://web.expasy.org/protparam/>).

## Immune Infiltration and Pathway Analyses

Immune infiltration assays were performed to explore variations in the immune environment between patients with PE and healthy controls. Specifically, the CIBERSORT algorithm (v 1.03)<sup>52</sup> was used to estimate the abundance of 22 immune cell types<sup>30</sup> in each sample of the training set. Samples were filtered based on reliability, excluding those with zero expression and removing samples with  $P > 0.05$ . Further analysis was performed to examine the distribution of these immune cells in PE versus control samples, with a significance threshold of  $P < 0.05$ . Spearman correlation analysis was conducted using the psych package (v 2.2.9)<sup>53</sup> to identify potential relationships between significantly different immune cells and biomarkers ( $|\text{cor}| > 0.3$  and  $P < 0.05$ ).

In addition, 13 immune-related pathways previously identified in the literature<sup>54</sup> were explored to assess the differences in pathway activity between patients with PE and control samples in the training set. The single-sample GSEA (ssGSEA) algorithm within the GSVA package (v 1.42.0)<sup>55</sup> was used to calculate single-sample scores for these immune-related pathways, allowing for a comparison of immune-related pathway scores between the PE and control samples ( $P < 0.05$ ). Further analysis was conducted using the psych package to perform Spearman correlation analysis between these differentially immune-related pathways and biomarkers ( $|\text{cor}| > 0.3$  and  $P < 0.05$ ).

## Construction of Regulatory Network

Further analyses were performed to elucidate the molecular regulatory mechanisms that affect these biomarkers. Using the multiMiR package (v 1.18.0),<sup>56</sup> miRNAs associated with biomarkers were predicted using default settings (predicted.cutoff and predicted.cutoff.type="p", selecting the top 20%). Key miRNAs were consistently predicted across five databases: DIANA-microT (<http://www.microrna.gr/microT>), EIMMo (<http://www.mirz.unibas.ch/EIMMo3/>), miRecords (<http://c1.accurascience.com/miRecords/>), TargetScan (<http://www.targetscan.org/>) and miRDB (<http://www.mirdb.org/>). Subsequent predictions of potential lncRNA interactions with these miRNAs were performed using the starBase (<https://starbase.sysu.edu.cn/>) and miRNet (<https://www.mirnet.ca/>) platforms. In starBase, a threshold of pancancerNum  $> 4$  was used to filter the data, retaining only those lncRNAs that were consistently predicted by both databases to have interactive relationships. Using this information, an lncRNA-miRNA-mRNA regulatory network was constructed using the Cytoscape software. Similarly, the ENCODE database (<https://www.encodeproject.org/>) within

miRNet was used to identify transcription factor (TFs)-targeting biomarkers and a TF-mRNA network was constructed using Cytoscape software.

## Drug Prediction and Molecular Docking Analyses

In this study, potential drug-targeting biomarkers were identified using the enrichR package (v. 3.1)<sup>57</sup> from the Drug SIGnatures database (DSigDB) (<http://dsigdb.tanlab.org/DSigDBv1.0/>). Results with an adj.P < 0.05, were considered statistically significant and potential drugs, and the top 10 drugs, based on scoring, were selected for visualization. In addition, the relationship between differential immune cells obtained from immune infiltration analysis and biomarkers was incorporated into an interaction network of drugs, biomarkers, and immune cells using Cytoscape software, which facilitated the understanding of the interactions among drugs, immune cells, and biomarkers.

To understand the stereochemical properties of the biomarkers, AlphaFold v2.0 (<https://alphafold.ebi.ac.uk/>) was employed as an artificial intelligence tool to predict the 3D structures of the biomarkers. The amino acid sequences of these biomarkers were downloaded from the UniProt-KB database (<http://www.uniprot.org/>) and the local distance difference test (LDDT) scores from the AlphaFold database were used to assess the stereochemical performance of these 3D models. To further explore the role of biomarkers in drug therapy, the molecular structures of drugs and protein 3D structures were downloaded from PubChem (<https://pubchem.ncbi.nlm.nih.gov/>) and Protein Data Bank (PDB) (<http://www.rcsb.org/>) databases, respectively. CB-Dock2 was used for the molecular docking of drugs against biomarkers to obtain the binding energies. It is generally accepted that the more stable the ligand-receptor conformation, the lower the required binding energy, with energies  $\leq -5$  kcal/mol considered indicative of strong binding affinity. Drugs with the highest 2D conformation scores were selected for molecular docking, with visualization conducted using PyMOL software (v. 3.1.1).<sup>58</sup>

## Cell Enrichment and Communication Analyses

Among the previously annotated cell types, a series of analyses were conducted to explore the cellular mechanisms of PE across all samples in GSE173193. Applying the ReactomeGSA package (v 1.12.0),<sup>59</sup> functional enrichment was performed on each of the identified differential cells from the prior differential analysis to explore their biological pathways, with the top 20 pathways displayed. Subsequently, with support from the celltalker package (v 0.0.7.9000),<sup>60</sup> PE and control samples were distinguished within the training set to conduct cell communication analysis. This involves detecting the expression and pairing of receptors and ligands within these cell types, with the aim of inferring interactions between different cells.

## Identification of Key Cell and Pseudo-Time Analysis

Biomarkers were examined for their expression differences in differential cells between PE and control samples. Violin plots were used for visualization, identifying cells with significant expression differences in most biomarkers as key cells. Further analysis illustrated the distribution of biomarkers across the different cell types.

To investigate the developmental trajectory and evolution of key cells, the cells were initially subjected to dimensionality reduction and clustering to identify cellular subtypes. Subsequently, the monocle package (v 2.26.0)<sup>61</sup> was employed to conduct a pseudo-time series analysis using 2500 highly variable genes across both the PE and control samples. In addition, temporal changes in biomarker expression within key cells were monitored.

## Reverse Transcription-Quantitative Polymerase Chain Reaction (RT-qPCR)

Placental tissues from six patients (three in the control group and three in the preeclampsia group) were collected after delivery at the First People's Hospital of Yunnan Province. Informed consent was obtained from all participants. This study was approved by the Ethics Committee of First People's Hospital of Yunnan Province (approval number: 221). The RNA extracted from each sample was quantified using 1 mL of triazole reagent and 300  $\mu$ L of chloroform, the concentration of which was quantified using a nanophotometer N50. cDNA was reverse-transcribed using the first-strand cDNA of SweScript. Synthetic Kit (Service Biology, Wuhan, China). RT-qPCR amplification was performed for 40 cycles using a CFX96 real-time fluorescent quantitative PCR instrument (Bio-Rad, Hercules, California, USA). The PCR cycle consisted of 1 min of pre-denaturation at 95°C, 20s of denaturation at 95°C, 20s of annealing at 58°C, and 30s

**Table 1** Sequence Information Table for the Four Gene Primers

Gene Primers	Sequence Information
CD44-F	CCAGAAGGAACAGTGGTTTGGC
CD44-R	ACTGTCCTCTGGGCTTGGTGT
TLR4-F	CCCTGAGGCATTTAGGCAGCTA
TLR4-R	AGGTAGAGAGGTGGCTTAGGCT
MIF-F	AGAACCGCTCCTACAGCAAGCT
MIF-R	GGAGTTGTTCCAGCCCACATTG
PIK3RI-F	CGCCTCTTCTTATCAAGCTCGTG
PIK3RI-R	GAAGCTGTCGTAATTCTGCCAGG

of extension at 72°C. During RT-qPCR analysis, each reaction mixture contained 3 µL CDN A, 5 microliters µL 2x Universal Blue SYBR GreengPCR master mixture, 1 µL forward primer, and 1 µL reverse primer. The sequence information for the four gene primers is in [Table 1](#).

Statistical Analysis

All analyses were performed using R software (v 4.2). Differences between the groups were analyzed using the Wilcoxon test. Statistical significance was defined as  $P < 0.05$ , with  $*P < 0.05$ ,  $**P < 0.01$ ,  $***P < 0.001$ ,  $****P < 0.0001$ , indicating increasing levels of significance, and ns denotes non-significant ( $P > 0.05$ ).

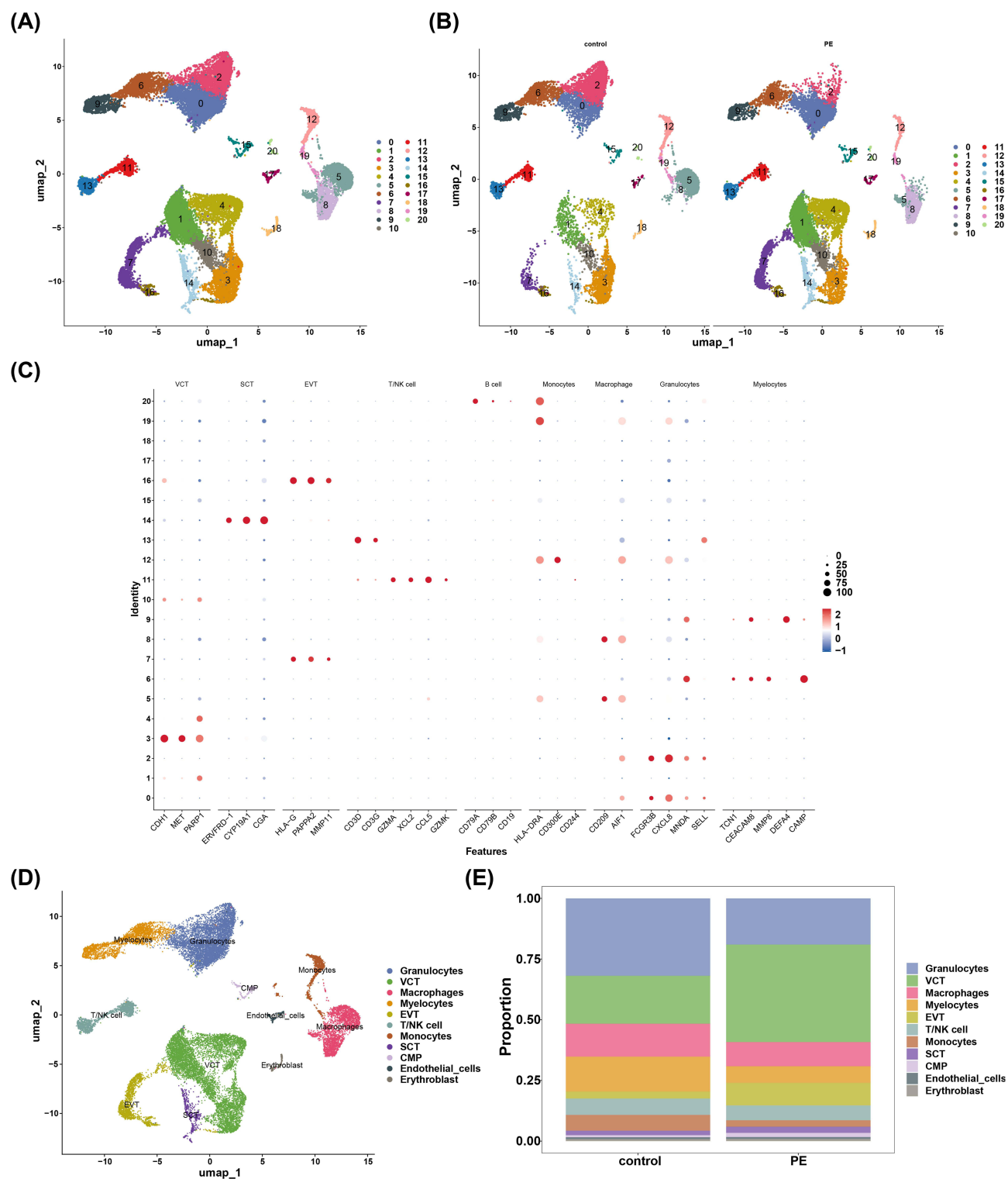
Results

Annotation of 11 highly Specific Cell Types

For the GSE173193 dataset, the results before and after quality control are illustrated in [Figure S1A](#) and [B](#), revealing that 24,100 cells and 22,891 genes met the criteria. From these, 2500 highly variable genes were selected for downstream analysis ([Figure S1C](#)). Subsequently, stability was achieved after the first 20 PCs; hence, the top 20 PCs were selected for further analysis ([Figure S1D](#) and [E](#)). Based on this, UMAP clustering identified 21 cell clusters, and their distribution in PE and control samples is displayed ([Figure 1A](#) and [B](#)). The bubble plot showed a high specificity for marker genes, leading to the cells being named based on these markers ([Figure 1C](#)). Ultimately, 11 cell types were annotated: granulocytes, villous cytotrophoblasts (VCT), macrophages, myelocytes, extravillous trophoblasts (EVT), T/NK cells, monocytes, syncytiotrophoblasts (SCT), common myeloid progenitors (CMP), endothelial cells, and erythroblasts ([Figure 1D](#)). Histograms displayed the proportions of these cell types in PE and control samples, clearly showing that VCT was more prevalent in PE samples, whereas granulocytes were more abundant in the control samples ([Figure 1E](#)).

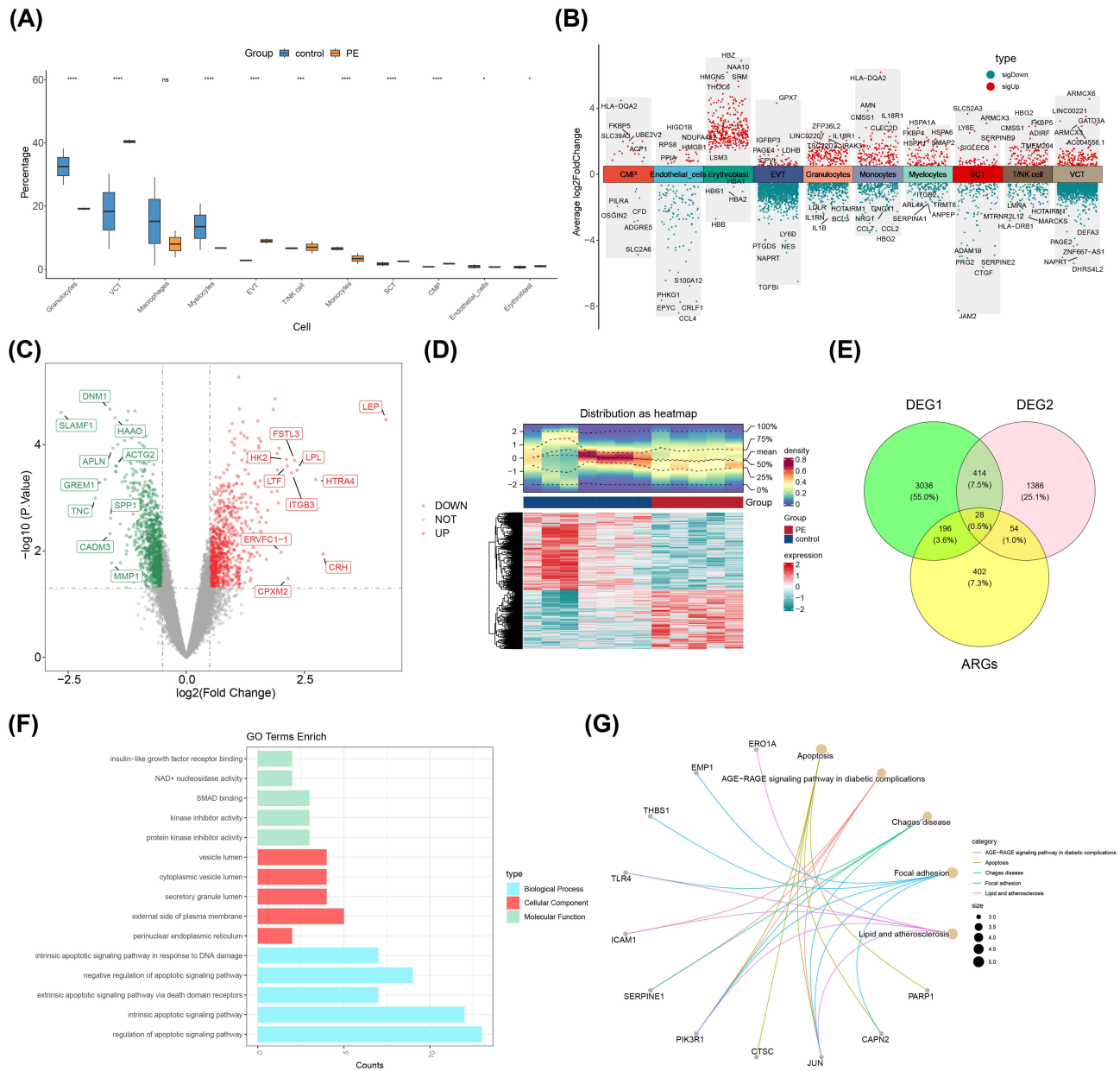
Recognition and Functional Characterization of DE-ARGs

In the GSE173193 dataset, differences between the PE and control samples were compared across 11 cell types, revealing significant variations in 10 cell types ( $P < 0.05$ ), excluding macrophages ( $P > 0.05$ ), which were defined as differential cells ([Figure 2A](#)). Further analysis of 10 differential cells identified 5577 DEGs between PE and control samples. After removing duplicates, 3674 unique DEGs1 were identified, with 790 genes upregulated and 2884 downregulated in the PE samples ([Figure 2B](#)). Similarly, in the transcriptome dataset (training set), 1882 DEGs2 were identified between PE and control samples, with 810 genes upregulated and 1072 downregulated in the PE samples ([Figure 2C](#) and [D](#)). Subsequently, an intersection of 3674 DEGs1, 1882 DEGs2, and 680 ARGs was obtained, yielding 28 DE-ARGs ([Figure 2E](#)), which were further included in the enrichment analysis. This analysis identified 779 GO entries, comprising 673 BP, 26 CC, and 80 MF, including “regulation of the apoptotic signaling pathway”, “intrinsic apoptotic signaling pathway”, “extrinsic apoptotic signaling pathway via death domain receptors”, “negative regulation of the apoptotic signaling pathway”, “intrinsic apoptotic signaling pathway in response to DNA damage”, “NAD+ nucleosidase activity”, and “vesicle lumen” ([Figure 2F](#)). Additionally, 56 KEGG pathways were



**Figure 1** Annotation maps for 11 highly specific cell types. **(A)** 21 samples UMAP cell taxon clustering, Different colors represent different cell populations. **(B)** PE and control grouping UMAP cell taxon clustering. **(C)** Bubble plots of marker genes in different cellular taxa, darker colors represent higher gene expression. **(D)** Clustering map of 11 maker genes. **(E)** Histogram of the percentage of 11 cell types between PE and control groups: different colors represent different cell types.





**Figure 2** Identification and Functional Characterization Map of DE-ARGs. **(A)** Differential boxplots of cell types. “ns” represented not significant, “\*” represented  $p < 0.05$ , “\*\*” represented  $p < 0.01$ , and “\*\*\*” represented  $p < 0.001$ . **(B)** Differential Gene Manhattan Map: Red dots represent up-regulated genes and green dots represent down-regulated genes; the 5 genes with the most significant difference between up- and down-regulation show the gene names. **(C)** Volcano plot of gene expression in PE and control samples, red represents high expression, blue represents low expression, and the shade of the color reflects the amount of expression. **(D)** Heat map distribution of PE and control samples, the top section is a heatmap of the expression density of up- and down-regulated differential genes in the samples, showing the lines of the five quartiles and the mean; the bottom section is a heatmap of the expression of the differential genes. **(E)** Venn diagrams between DEG1, DEG2 and ARGs. **(F)** GO enrichment results: the horizontal axis is the number of genes enriched to any one pathway, and the vertical axis is the name of the GO-enriched entry. **(G)** KEGG enrichment results: gray dots indicate genes, yellow dots indicate different pathway names, and different colored lines represent pathways that genes are enriched into.

identified, such as the “AGE-RAGE signaling pathway in diabetic complications”, “apoptosis”, “chagas disease”, “focal adhesion”, and “lipid and atherosclerosis” (Figure 2G). These enrichment results, involving cell death, signaling pathway regulation, and extracellular interactions, revealed the potential processes through which DE ARGs affected PE.

## Incorporating Biomarkers CD44, MIF, PIK3R1, and TLR4 into a Nomogram

Based on the 28 DE-ARGs, a PPI network with 22 nodes and 38 edges was constructed, where JUN, TLR4, CD44, and ICAM1 had the highest degrees (Figure 3A). Further analysis was conducted using 5 different algorithms from the CytoHubba plugin. As DMNC only scored the top nine genes, we selected the top nine genes from each algorithm and identified their intersection, resulting in nine hub genes: CD44, ICAM1, JUN, MIF, PARP1, PIK3R1, SERPINE1, THBS1, and TLR4 (Figure 3B). These nine hub genes were included in the expression level analyses, which clearly revealed consistent expression trends and significant intergroup differences in both the training and validation datasets for CD44, MIF, PIK3R1, and TLR4 ( $P < 0.05$ ) (Figure 3C and D). Notably, CD44 and TLR4 were expressed at significantly lower levels in the PE samples ( $P < 0.05$ ), whereas MIF and PIK3R1 were expressed at higher levels ( $P < 0.05$ ). These four genes were identified as potential biomarkers in this study.

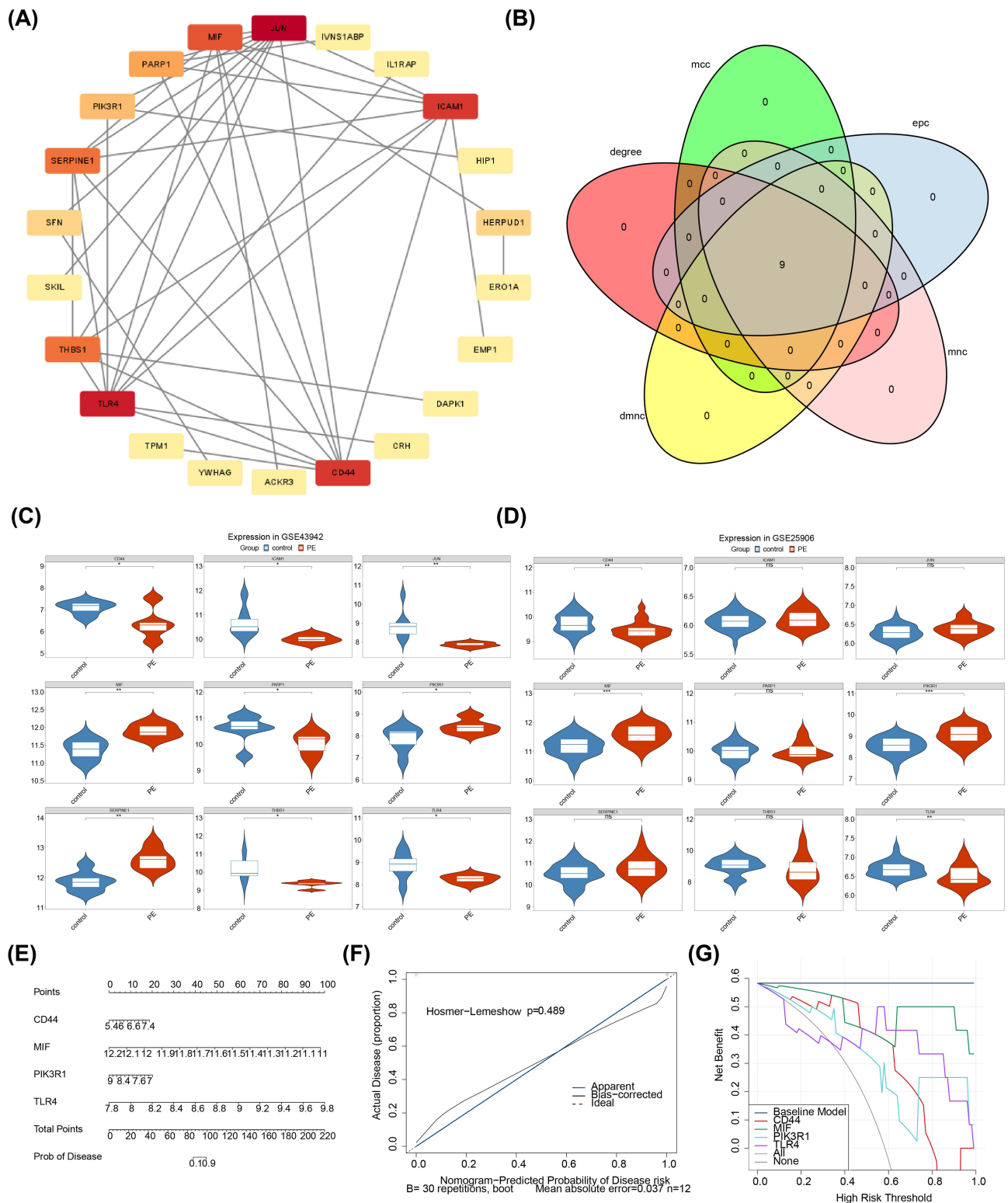
Subsequent analysis integrated these biomarkers into a nomogram to maximize their diagnostic value for PE (Figure 3E). The calibration curve compared the predicted probabilities of the nomogram with the observed actual probabilities (Figure 3F). The P-value of the HL test was 0.489, indicating no significant difference between the predicted and actual values, and the mean absolute error (MAE) was 0.037, suggesting a minimal error between the actual and predicted risk of disease. This demonstrates the high accuracy of the nomogram in predicting the risk of illness in the sample. DCA also indicated that the nomogram's net benefit was higher than those of the positive and negative controls, providing greater net benefits than the use of any single biomarker alone, thereby exhibiting substantial clinical utility (Figure 3G).

## Investigating the Potential Functions and Involved Signaling Pathways of Biomarkers

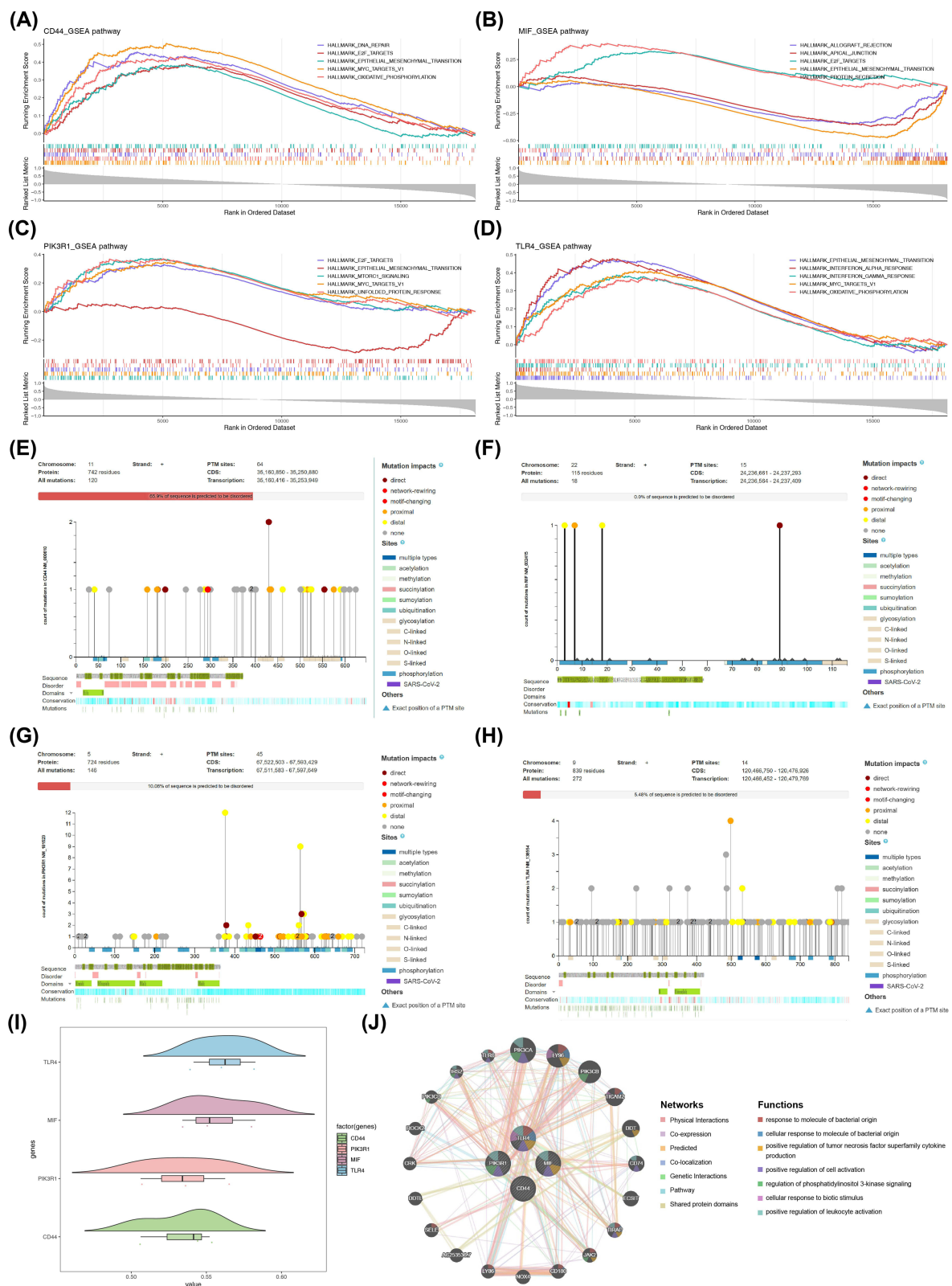
GSEA was performed on biomarkers to explore enriched signaling pathways. The results revealed that CD44, MIF, PIK3R1, and TLR4 were enriched in 24, 14, 30, and 10 hallmark pathways, respectively (Figure 4A–D). In particular, they were commonly enriched in the “epithelial mesenchymal transition” pathway. CD44, MIF, and PIK3R1 were jointly enriched in “E2F targets”; CD44, PIK3R1, and TLR4 were commonly found in “MYC targets V1”; and CD44 and TLR4 were both enriched in “oxidative phosphorylation”. These pathways are closely related to the regulation of cellular functions and pathological states, suggesting the mechanisms by which these biomarkers might influence the development and progression of PE.

Further research on SNVs of biomarkers revealed significant findings. In CD44, 65.9% of the sequences were predicted to be disordered, indicating that these regions lacked a fixed three-dimensional structure and might have played a crucial role in regulating protein function and protein-protein interactions (Figure 4E). The mutations affected various aspects, including direct effects, network rewiring, motif changes, and proximal and distal associations, which could have altered how the protein interacted with other molecules, thereby changing signal transduction pathways and cellular functions. The sites were related to acetylation, methylation, succinylation, SUMOylation, ubiquitination, glycosylation, and phosphorylation, all of which could affect the activity, stability, interaction partners, and cellular localization of the protein. Similar results were observed for other biomarkers, with 0.0%, 10.08%, and 5.48% of the sequences in MIF, PIK3R1, and TLR4, respectively, predicted to be disordered, reflecting potential differences in their cellular functions and stabilities (Figure 4F–H). These results suggest that the sequences and functions of these biomarkers are regulated by various molecular mechanisms and that SNVs can significantly alter their functions.

Friends analysis evaluated the functional similarity among biomarkers, indicating that the four genes shared certain similarities, with TLR4 showing the highest overall functional similarity (scores  $> 0.5$ ) (Figure 4I). GeneMANIA provides a co-expression network for these biomarkers, in which physical interactions are the most prevalent (77.64%), followed by co-expression effects (8.01%). These genes were collectively involved in functions such as “response to molecule of bacterial origin”, “cellular response to molecule of bacterial origin”, “positive regulation of tumor necrosis factor superfamily cytokine production”, “positive regulation of cell activation”, “regulation of phosphatidylinositol 3-kinase signaling”, “cellular response to biotic stimulus”, and “positive regulation of leukocyte activation” (Figure 4J). The physicochemical properties of these biomarkers are listed (Table 2). For example, CD44 has the formula C1719H2701N487O563S18 with a molecular weight of 39,775.36, consisting of 362 amino acids, a theoretical



**Figure 3** Selection, expression and validation of biomarkers. **(A)** PPI network of 28 DE-ARGs, the size of the dots represents the number of cells of that type, and the thickness of the lines represents the strength of communication between the corresponding cell groups. **(B)** 5 Algorithms Top10 Gene Intersection. **(C)** Boxplot of biomarkers expression in GSE43942. The blue and red represented the “control” group and the “PE” group, respectively. “\*” represented  $p < 0.05$ , and “\*\*\*” represented  $p < 0.01$ . **(D)** Boxplot of biomarkers expression in GSE25906. “ns” represented not significant, “\*” represented  $p < 0.01$ , and “\*\*\*” represented  $p < 0.001$ . **(E)** The column-line diagram model of Biomarker. **(F)** Calibration curves for line diagram models: the horizontal coordinates of the calibration curve for the column-line diagram model indicate the probability of illness predicted by the nomogram, and the vertical coordinates indicate the actual probability of illness. **(G)** DCA Decision Curve: the horizontal coordinate is the threshold probability and the vertical coordinate is the net benefit rate after benefits minus drawbacks.



**Figure 4** Potential functions of biomarkers and associated signaling pathways. (A) CD44 enrichment analysis, (B) MIF enrichment analysis, (C) PIK3R1 enrichment analysis, (D) TLR4 enrichment analysis. Each fold represents a pathway with lines marking the genes located in the gene set. Analysis of (E) CD-44, (F) MIF, (G) PIK3R1, (H) TLR4 single nucleotide variants. (I) Functional similarity analysis of biomarkers. (J) Biomarker GGI Network: different colored lines represent different interactions, and different color blocks represent the different functional roles involved.

**Table 2** Physical and Chemical Properties of the Biomarkers

Gene Property	TLR4	PIK3R1	MIF	CD44
Formula	C4334H6734N1122O1246S37	C3715H5840N1018O1143S17	C553H866N152O163S7	C1719H2701N487O563S18
Molecular Weight	95680.13	83598.39	12476.3	39775.36
Number of amino acids	839	724	115	362
Theoretical pI	5.88	5.84	7.73	5.38
Aliphatic index	101.86	81.48	87.39	72.46
Grand average of hydropathicity	0.033	0.716	0.002	0.449
Estimated protein half life	30h	30h	30h	30h
Instability index	43.05	48.89	34.21	32.11

isoelectric point (pI) of 5.38, an aliphatic index of 72.46, a grand average of hydropathicity (GRAVY) of 0.449, an estimated protein half-life of 30 h, and an instability index of 32.11.

Differential Immune Cell Infiltration and Pathway Activation in PE

Subsequent analysis aimed to explore the differences in the immune microenvironment composition between patients with PE and healthy individuals, initially showing the infiltration abundance of 22 immune cells in both types of samples within the training set (Figure 5A), with box plots further visualizing these disparities (Figure 5B). The results revealed significant differences in the two immune cells between the PE and control samples ( $P < 0.05$ ). Monocytes had significantly less infiltration in PE patients ( $P < 0.05$ ), while resting NK cells exhibited a higher infiltration abundance ( $P < 0.05$ ). In addition, biomarkers showed varying degrees of correlation with these differentially infiltrated immune cells, with PIK3R1 displaying the highest negative correlation with monocytes ( $\text{cor} = -0.615$  and  $P < 0.05$ ) and TLR4 showing the highest positive correlation with resting NK cells ( $\text{cor} = 0.650$ ,  $P < 0.05$ ) (Figure 5C). These results suggest differences in the composition of immune cells within the immune microenvironment between PE patients and healthy individuals. The two identified immune cells might play roles in the pathogenesis of PE, and the identified biomarkers could potentially represent the immune microenvironment of patients with PE, which is crucial for their management.

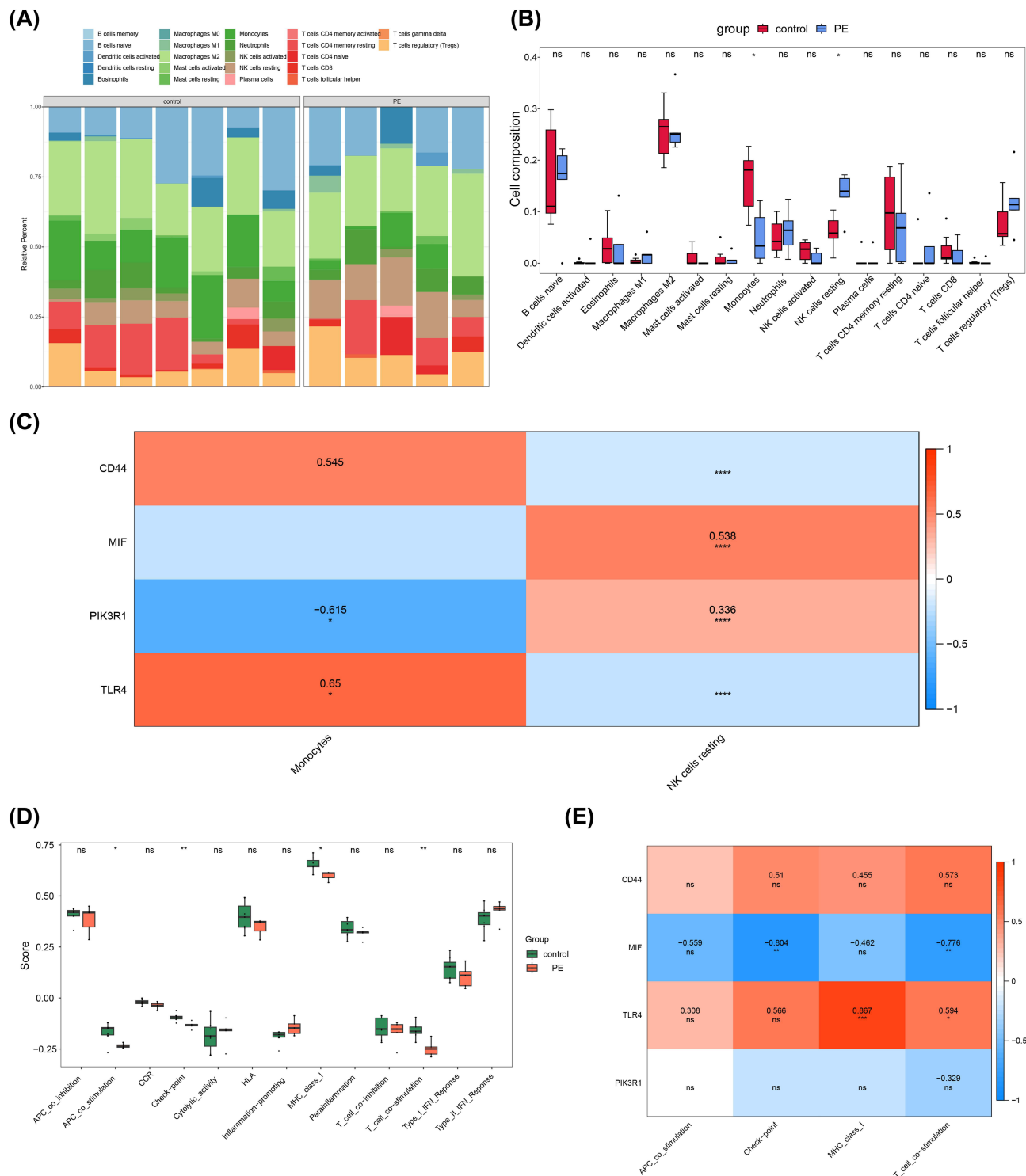
After completing the immune infiltration analysis, we further explored the differences in 13 immune-related pathways between patients with PE and healthy individuals in the training set. This has allowed for a deeper understanding of the pathogenesis of PE, revealing how these cells interact at the molecular level, affect disease progression, and identify potential therapeutic targets. The results showed that all four significantly different immune-related pathways, APC co-stimulation, checkpoint, T cell co-stimulation, and MHC class I, had notably lower scores in the PE samples ( $P < 0.05$ ) (Figure 5D). This suggests that immune regulation in patients might be suppressed in some key pathways that are typically associated with immune activation and antigen presentation. The biomarkers also showed strong correlations with these pathways; for example, the correlation between MIF and checkpoints was as high as  $-0.804$  ( $P < 0.01$ ), and TLR4 had the strongest positive correlation with MHC class I ( $\text{cor} = 0.867$  and  $P < 0.001$ ) (Figure 5E). The reduced activity of these pathways could be a key factor in the development of PE, and the association between these pathways and biomarkers could provide potential targets for the development of targeted therapies for PE.

Revealing the Potential Molecular Regulatory Mechanisms of Biomarkers

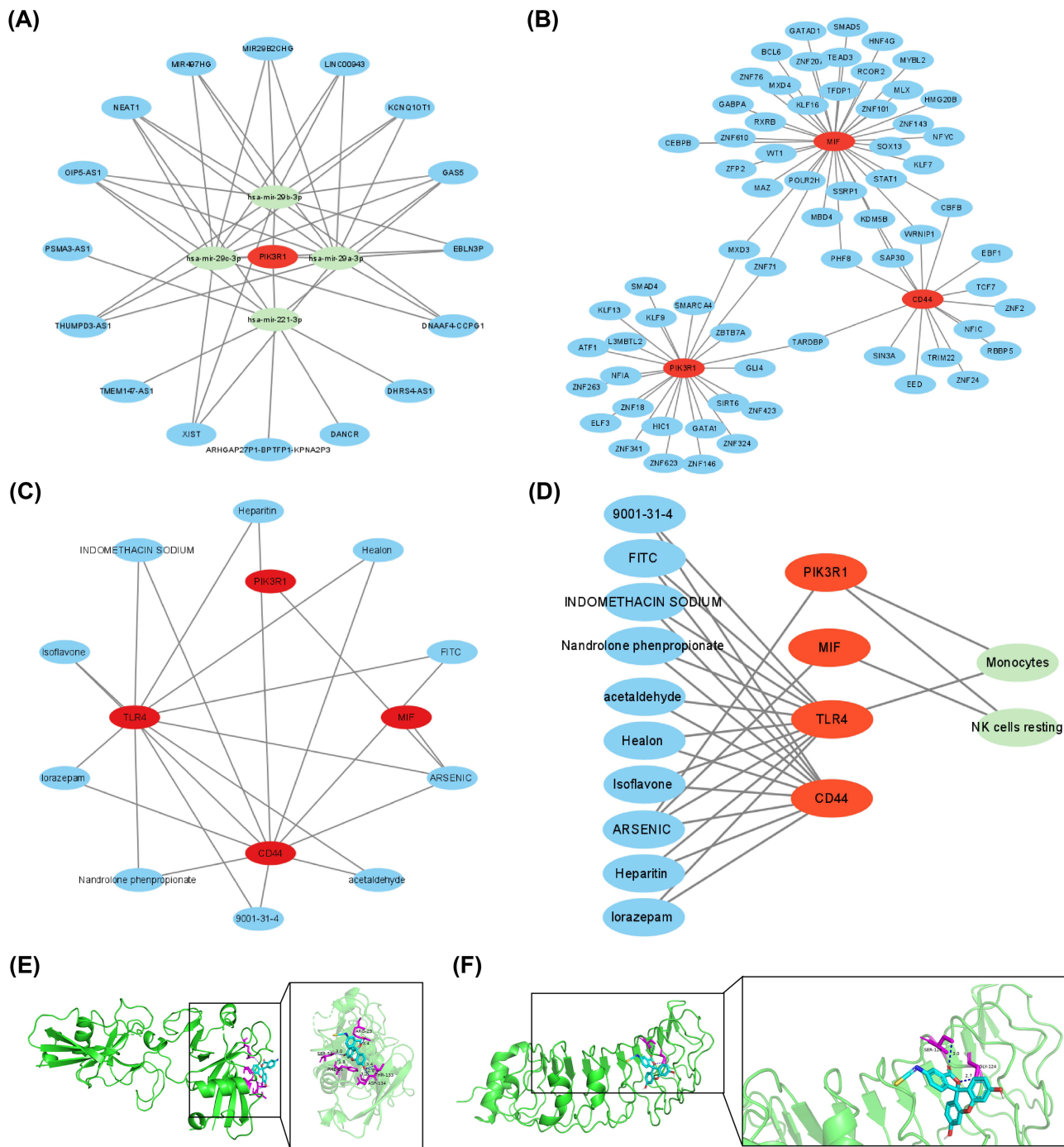
Four miRNAs consistently predicted across five databases (DIANA-microT, EIMMo, miRecords, TargetScan, and miRDB) were identified as key miRNAs, whereas 16 lncRNAs consistently predicted in two other databases (starBase and miRNet) were identified as key lncRNAs. Based on this, an lncRNA-miRNA-mRNA regulatory network was constructed consisting of 21 nodes and 45 edges (Figure 6A). Therefore, PIK3R1 was the only mRNA with intersecting miRNA results, and lncRNAs such as NEAT1, OIP5-AS1, GAS5, and DNAAF4-CCPG1 may simultaneously regulate PIK3R1 through hsa-mir-29a-3p, hsa-mir-29b-3p, hsa-mir-29c-3p, and hsa-mir-221-3p.

We constructed a TF mRNA network to further elucidate the potential regulatory mechanisms of these biomarkers. In this network, 67 TFs were predicted to be associated with three biomarkers, comprising 70 nodes and 75 edges





**Figure 5** Differential immune cell infiltration and pathway activation in PE. **(A)** Map of immune-infiltrating cell abundance, different colors represent different immune cells. **(B)** Box plot of immune cell differences between PE and control groups, horizontal coordinates indicate immune cells and vertical coordinates indicate immune cell immunity scores in the samples; red indicates the control group and blue indicates the PE control group. "ns" represented not significant, and "\*" represented  $p < 0.05$ . **(C)** Heatmap of Differential Immune Cell-Biomarker Correlation: \* denotes significance, numbers represent correlation coefficients, absolute value of correlation coefficients  $< 0.3$  is null. "\*" represented  $p < 0.05$ , and "\*" represented  $p < 0.0001$ . **(D)** Comparative box plots of immune pathway differences. "ns" represented not significant, "\*" represented  $p < 0.05$ , and "\*" represented  $p < 0.01$ . **(E)** Heatmap of biomarkers correlating with immune pathways. "ns" represented not significant, "\*" represented  $p < 0.05$ , "\*" represented  $p < 0.01$ , and "\*" represented  $p < 0.001$ .



**Figure 6** The potential molecular regulatory mechanisms of biomarkers. **(A)** Biomarker-mRNA-lncRNA network: red is biomarker, green is miRNA and blue is lncRNA. **(B)** Biomarker-TF Network, red is biomarkers, blue is TF. **(C)** TOP10 Drug prediction network: biomarkers in red, Top 10 drugs in blue. **(D)** Drug-Biomarker-Immune cell prediction networks: green is an immune cell with differences. **(E)** TLR4 and **(F)** CD44 Molecular Docking: the left image is global, the right image is localized.

(Figure 6B). TFs such as MXD3 and ZNF71 may co-regulate MIF and PIK3R1, whereas PHF8, SAP30, KDM5B, CBFB, and WRNIP1 may co-regulate MIF and CD44, and TARDBP may co-regulate CD44 and PIK3R1.

### Exploring the Structure and Function of Drug Targets Associated with Biomarkers

In DsigDB, 312 potential drug targets identified from biomarkers were visualized, selecting the top 10 ranked drugs, including arsenic, fluorescein isothiocyanate (FITC), isoflavone, heparitin, Healing, acetaldehyde, indomethacin sodium,

nandrolone phenpropionate, lorazepam, and 9001-31-4. A network comprising 14 nodes and 22 edges was constructed with ARSENIC targeting 4r biomarkers (Figure 6C). Moreover, the previously identified differential immune cells (monocytes and resting NK cells) were integrated to construct a network comprising 16 nodes and 26 edges, featuring interaction pairs such as isoflavone-TLR4-monocytes (Figure 6D). This suggests that isoflavone may influence monocyte function through TLR4, potentially affecting immune regulation and inflammatory responses. Subsequent analysis revealed structural models of the proteins corresponding to the biomarkers (Figure S2A–D). Notably, the overall color of the MIF model was blue, with a predicted LDDT score greater than 90, indicating that its predicted structure closely resembled the actual structure. Furthermore, the highest-scoring 2D conformations among the aforementioned drugs were selected for molecular docking, and both CD44 and TLR4 were docked using FITC. Specifically, the binding energy between CD44 and FITC was  $-8.1$  kcal/mol, with residues such as ARG-29, SER-58, PHE-56, THR-133, and ASP-134 forming hydrogen bonds with FITC (Figure 6E). Similarly, the binding energy between TLR4 and FITC was  $-8.0$  kcal/mol, with residues SER-126 and GLY-124 forming hydrogen bonds with FITC (Figure 6F).

## Probing Cellular Communication and Biological Functions Under PE Conditions

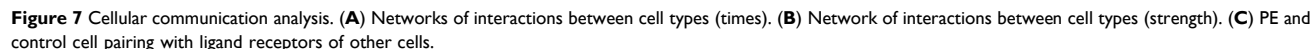
Following transcriptomic analysis, an understanding of the expression patterns and functions of the biomarkers was obtained. To further explore the heterogeneity of these processes at the cellular level, biomarker expression profiles were analyzed from a single-cell perspective to explore unique expression patterns across different cell types. This approach aimed to further understand the mechanisms of PE at the cellular level using the scRNA-seq dataset (GSE173193).

Initially, enrichment analysis of ten different cell types revealed the main biological functions of these cells. These cells were enriched in pathways such as “regulation of thyroid hormone activity”, “hydroxycarboxylic acid-binding receptors”, “mitochondrial uncoupling”, “proton-coupled neutral amino acids”, and “classical antibody-mediated complement activation” (Figure S3A–J). These pathways play crucial roles in maintaining internal homeostasis and in responding to various physiological and pathological conditions.

Further analysis of cellular communication, displaying the number and weight of interactions among the 11 cell types under control and PE conditions, clearly showed a reduction in both metrics under PE conditions (Figure 7A and B). This suggests that signal transmission and interactions between cells may have been suppressed in PE, potentially affecting cellular function and overall tissue coordination. In particular, under PE conditions, endothelial cells, EVT, granulocytes, monocytes, myelocytes, and SCT exhibited reduced communication. These cell types play crucial roles in maintaining vascular stability, immune defense, and placental function, and their impaired communication can be linked to functional disruptions, thus playing a significant role in the pathogenesis of PE. However, the increased communication number and weight observed in CMP and VCT might reflect an activation or compensatory response under PE conditions. This increase in communication could have been a response to the reduction in other cellular interactions, aiming to maintain normal placental and hematological functions. Additionally, the bubble plot under PE conditions illustrated ligand-receptor interactions between these cell types, indicating a level of signaling and interaction and suggesting that these receptors and ligands play key roles in regulating cell behavior, maintaining tissue functions, and responding to environmental changes (Figure 7C).

## Identifying VCT as the Key Cell

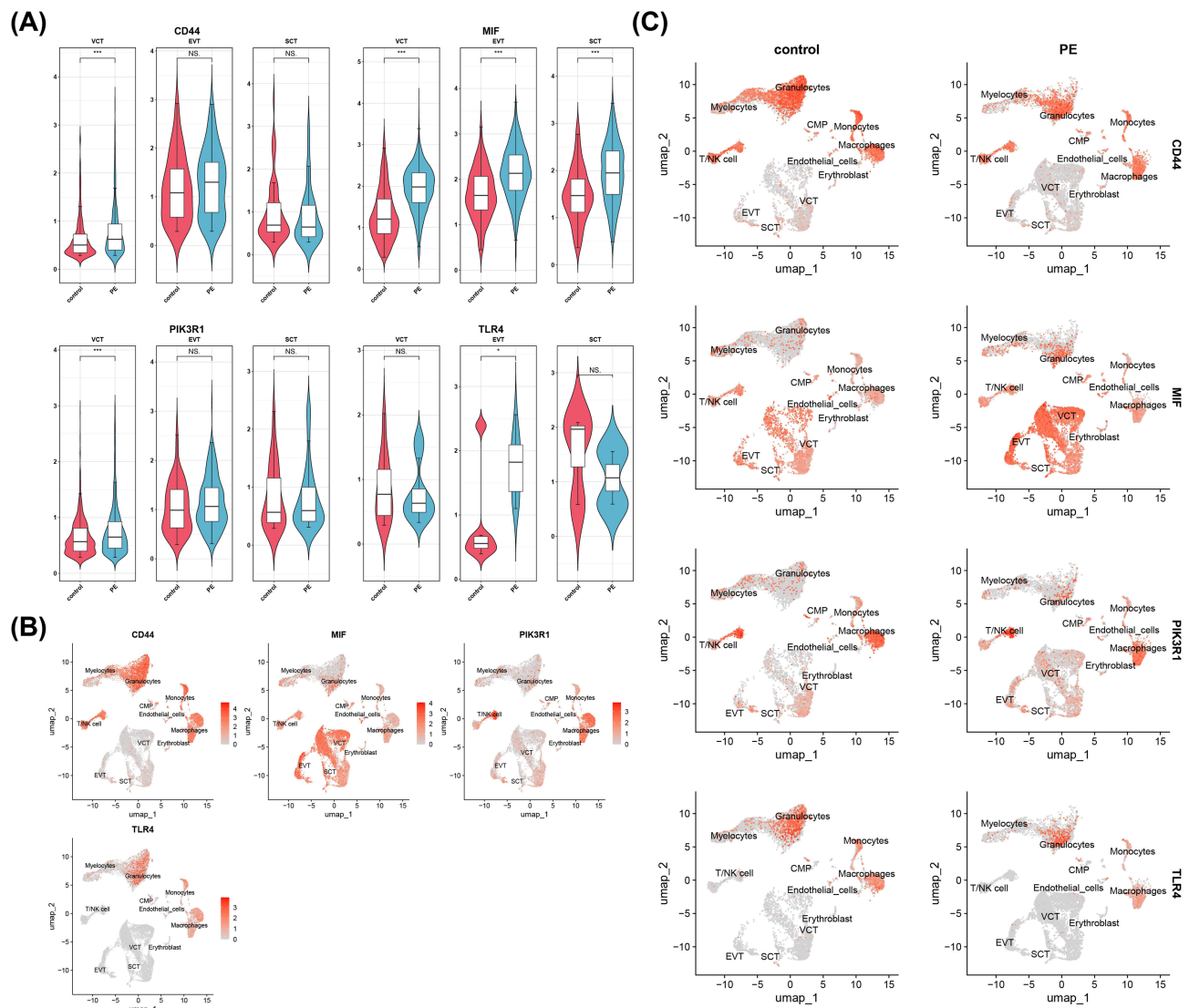
VCT, EVT, and SCT are crucial components of the placenta. Dysfunction of these cells leads to abnormal development and malfunction of the placenta, which are key factors in the development of PE(12). Consequently, this study focused on the expression of biomarkers within these trophoblast layers, selecting cells exhibiting significant differences in biomarker expression as the key cells for further analysis. The results indicated that while only MIF and TLR4 expression differences were observed in EVT ( $P < 0.05$ ), and MIF was the only biomarker differentiated in SCT ( $P < 0.05$ ), significant differences in the expression of CD44, MIF, and PIK3R1 were found in VCT, all of which were also highly expressed under PE conditions ( $P < 0.05$ ) (Figure 8A). Normally, VCT can differentiate into SCT; however, this process may have been disrupted during PE, thereby affecting placental function and maternal health. Thus, VCT was chosen as the key cell type for the subsequent pseudo-time analysis.



## Differential Developmental Dynamics of VCT

4802 <https://doi.org/10.2147/JIR.S507660>





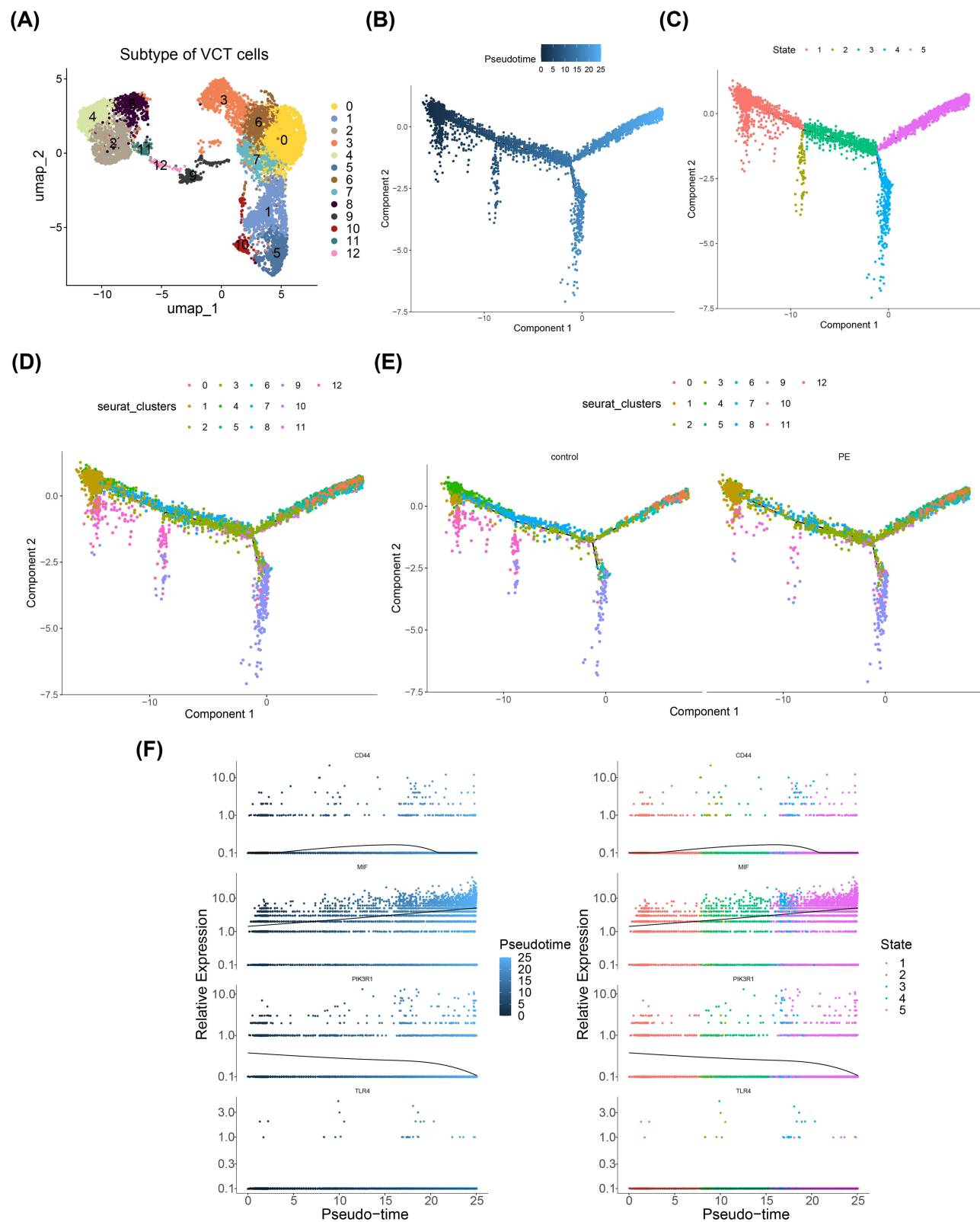
**Figure 8** Identification of biomarkers. (A) The violin map of CD44, MIF, PIK3R1, TLR4 expression in differential cells. "NS" represented not significant, "\*" represented  $p < 0.05$ , and "\*\*\*" represented  $p < 0.001$ . (B) Distribution of biomarkers. (C) Distribution of biomarkers according to subgroups: red color represents the distribution of biomarkers, the left panel is the control group and the right panel is the PE group.

population begins to diverge into different cellular states or fates at this developmental point, highlighting the emergence of heterogeneity within the original cell population, which is substantial enough to steer cells along diverse trajectories.

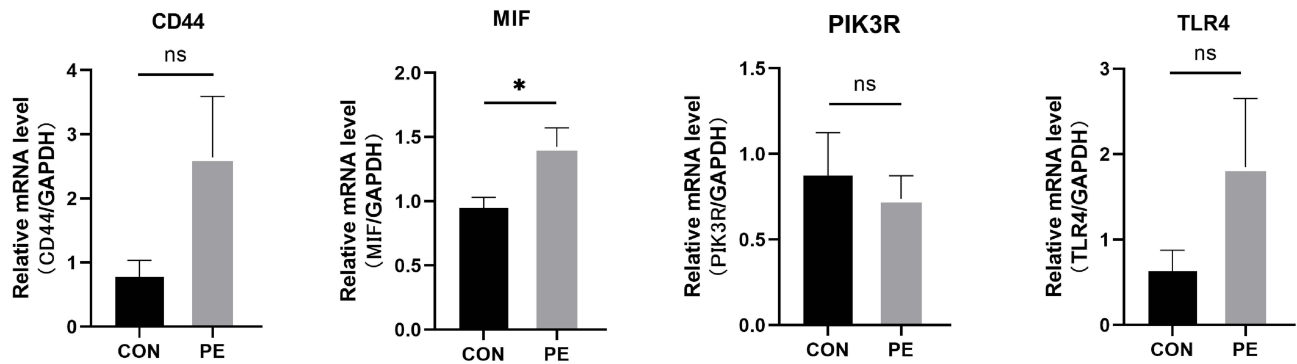
Moreover, compared with the control, there was a notable decrease in the number of VCT in the early stages under PE conditions and an increase in the later stages, suggesting that PE might affect the normal growth and differentiation processes of placental cells (Figure 9E). The change in cell numbers could reflect the adaptive response of the placenta to an adverse pregnancy environment. The reduction in early stage VCT may have been due to insufficient oxygen and nutrient supply associated with PE, affecting the initial establishment and functional development of the placenta. Conversely, the increase in VCT numbers in the later stages could have been a compensatory response by the placenta to enhance its transport capabilities to support ongoing fetal development.

The expression levels of these biomarkers during VCT differentiation were monitored (Figure 9F). As VCT matured, CD44 showed an initial increase, followed by a decrease in expression; MIF exhibited a consistent upward trend, and PIK3R1 consistently decreased. No significant changes in TLR4 expression were observed.





**Figure 9** Proposed time-series trajectory analysis. **(A)** Critical cell clustering results. **(B)** Difference in cell differentiation time (VCT), dark blue indicates an early stage of differentiation, while light blue indicates a later stage of differentiation. This can be used as a starting point for subsequent analysis. **(C)** Classes Occupied by Cell Differentiation (VCT), the different colors represent the taxa occupied by the cells. **(D)** Stages of Cell Differentiation (VCT), different colors represent different groups, 13 groups in total. **(E)** Cell differentiation in different samples. **(F)** Changes in the expression of biomarkers during differentiation of key cells.



**Figure 10** The RT-qPCR results plots for D44, MIF, PIK3R1, and TLR4 genes. “ns” represented not significant, and “\*” represented  $p < 0.05$ .

## Differential Expression of Biomarkers in PE and Controls

RT-qPCR results showed that the expression of CD44, TLR4, and MIF was upregulated in the PE group, while the expression of PIK3R1 was down-regulated (Figure 10).

## Discussion

PE is one of the main causes of maternal death and perinatal mortality; however, the specific mechanism remains unclear, and there is still a lack of effective prediction and intervention measures.<sup>62</sup> Apoptosis is a form of programmed cell death and an important mechanism for maintaining tissue homeostasis. However, abnormal apoptosis of placental trophoblastic cells may lead to the occurrence of PE.<sup>11</sup> Through single-cell transcriptome technology to study apoptosis-related genes, in-depth exploration of potential biomarkers of PE revealed its potential mechanism of action, which has important clinical significance in improving the early diagnosis and effective management of PE.

In this study, all the data were sourced from public databases. Through the application of scRNA-seq analysis and differential expression analysis, coupled with five algorithms from the CytoHubba plugin, along with expression level analysis, four ARGs (CD44, MIF, PIK3R1, and TLR4) were identified as biomarkers for PE at the single-cell level. They were mainly associated with immunomodulatory-related pathways, with CD44 and TLR4 downregulation, and MIF and PIK3R1 showed increased expression in PE.

CD44 is a cell-surface glycoprotein that plays an important role in various biological processes, including inflammatory responses, tumorigenesis, and wound healing. It is considered one of the markers of cancer stem cells and is closely related to the aggressiveness and ability of tumors to metastasize, what's more CD44 leads to the development of colorectal cancer by inducing apoptosis.<sup>63,64</sup> Placental trophoblasts are similar to tumor cells in that they have strong proliferative, migratory, and invasive abilities. Study has shown that increased CD44 positive expression was observed in maternal decidua cells and fibroblast cells close to the root villi.<sup>65</sup> Another study showed that the FKBPL-CD44 pathway appears to play a central role in the pathogenesis of PE and has potential for early diagnostic and therapeutic purposes in PE.<sup>66</sup> Taken together with previous studies, we hypothesized that CD44 may also play a role in PE apoptosis.

Macrophage Migration Inhibitory Factor (MIF) is a widely expressed cytokine that has pro-inflammatory effects, inhibits macrophage migration, and plays an important role in regulating immune responses, inflammatory responses and autoimmune disease.<sup>67</sup> In recent years, MIF has also been found to play a key role in a variety of pathological states such as tumorigenesis, metabolic diseases, and cardiovascular diseases.<sup>68</sup> It has been found that MIF is involved in the apoptotic pathway, and in a lipid metabolism disorder model, inhibition of the MIF pathway can reduce apoptosis of pancreatic islet  $\beta$ -cells and alleviate diabetes mellitus, whereas MIF inhibitors can also have a very good inhibitory effect on apoptosis in nasopharyngeal carcinoma.<sup>69,70</sup> A previous study showed that MIF levels were significantly increased in the maternal peripheral blood of the PE group.<sup>71</sup> Similarly, we found that the expression of MIF in PE placental tissue samples was significantly higher than that in normal samples. It has been suggested that MIF can be used to regulate the inflammatory response of macrophages in the immune environment of PE and influence apoptosis.

The protein encoded by PIK3R1 is a regulatory subunits of phosphatidylinositol 3-kinase (PI3K), that regulates cell growth and differentiation.<sup>72</sup> One study showed that, as an aging-related gene in PE, PIK3R1 may be related to different immune characteristics.<sup>73</sup> PIK3R1 is involved in the regulation of endometrial cell apoptosis.<sup>74</sup> Similar to our findings, another study revealed that PIK3R1 is the hub network and a key gene of PE, and needs more attention in future PE studies.<sup>75</sup>

Toll-like Receptor 4 (TLR4) is a member of the Toll-like receptor family and is mainly expressed in immune cells, such as macrophages and dendritic cells. TLR4 activates the immune response by recognizing foreign pathogens and endogenous signaling molecules.<sup>76</sup> It plays a key role in the recognition of bacterial infections, the initiation of inflammatory responses, and immune regulation. Abnormal activation of TLR4 is associated with the development of multiple inflammatory, autoimmune, and chronic degenerative diseases.<sup>77</sup> TLR4 is directly involved in the regulation of apoptosis.<sup>78</sup> A previous study showed that, in early onset PE, the expression of TLR4 in the placental tissue was significantly higher, and the severity of PE was correlated with the degree of damage to the placental villi,<sup>79</sup> which is similar to our qPCR results.

All the above studies have reflected the key role of these four biomarkers in PE apoptosis. Subsequently, the present study integrated biomarkers to establish a nomogram for predicting PE, which showed good predictive ability and was superior to that of a single biomarker. This demonstrates the great potential of these biomarkers in clinical decision-making and is expected to provide new opportunities for the clinical diagnosis and treatment of PE.

To further explore the potential mechanisms of biomarkers in the development of PE, this study used GSEA enrichment analysis and found that four biomarkers were co-enriched in epithelial mesenchymal transition (EMT). EMT is a biological process in which epithelial cells lose their intercellular adhesion and polarity and gain the ability to migrate, invade, and transform into mesenchymal cells.<sup>80</sup> This process is critical for embryonic development, tissue remodeling, and cancer metastasis.<sup>81</sup> In PE, EMT may be involved in placental formation and remodeling of the cell layer, especially in the invasion and migration of extra-villous trophoblastic cells (EVTs), which are essential for anchoring and functioning of the normal placenta. Abnormal EMT activity may lead to placental insufficiency and is associated with PE development of PE.<sup>82</sup> Therefore, inhibition of EMT can reduce the abnormal activity of PE cells. However, the regulation of EMT by these biomarkers in PE remains unclear and an in-depth understanding of their regulatory mechanisms is of great significance for the treatment of PE.

Breakdown of maternal-fetal immune tolerance is one of the key etiologic factors in PE.<sup>83</sup> The present study also found that there were two different types of immune cells between the two groups: monocytes and resting NK cells. In fetal circulation, dominant monocyte intermediate subsets and increased inflammatory subsets are observed in PE compared to those in normal pregnancy.<sup>84,85</sup> In addition, NK Cell activation has been characterized and is known to play major causative roles in the maternal inflammatory response in PE.<sup>86,87</sup> In addition, the activity of APC co-stimulation, Checkpoint, T cell co-stimulation, and MHC class I pathways was reduced in the PE group, indicating immunosuppression in patients with PE. Immune checkpoints are molecules in the immune system that control and regulate the activity of immune cells by inhibiting signaling and regulating the immune response, thereby helping to maintain the immune system to attack pathogens while avoiding damage to normal tissue.<sup>88</sup> Because of immune dysfunction in PE, immune checkpoint inhibitors are emerging as new therapeutic tools for PE, where CD44 and T cell co-stimulation can effectively improve the immune efficiency of vaccines.<sup>89,90</sup> MHC I and TLR4 play important roles in endogenous antigen presentation and recognition of exogenous pathogens. Both synergistically participate in the immune response while avoiding excessive damage to the host tissues.<sup>91</sup> Based on the differential immune cells and different immune pathways in PE, it has been suggested that targeted biomarkers can be used to regulate the immune infiltration of PE and that T cell co-stimulation or MHC class I expression can be regulated by biomarkers to enhance the immune response of PE. This study aimed to provide new immunotherapy for the clinical treatment of PE and to improve the outcomes of PE.

TLR4 predicts drug isoflavones. Isoflavones are natural plant compounds whose biological activity is mainly related to estrogen receptors and plays an important biological role in cancer, cardiovascular disease, oxidative stress, and other diseases through TLR4.<sup>92</sup> Puerarin (Pue) is an isoflavone that is derived from puerarin. Pue can improve trophoblast movement by inhibiting the PE-related inflammatory response, and is speculated to have a potential therapeutic effect on

PE.<sup>93</sup> This study established a drug-biomarker-immune cell regulatory network that provides a broader vision for PE immunotherapy.

PE is a complex condition that affects pregnancy and involves multiple cell types including VCT, these include VCT, SCT, and EVT. These cell types play key roles in the formation and function of placenta.<sup>94–96</sup> Single-cell analyses confirmed the importance of these three cell types. Finally, in the RT-qPCR experiments, all three biorepresentations showed the opposite trend, except for MIF, which showed the same expression trend as that in the dataset. Combined with single-cell analysis, this study also found that biomarkers showed dynamic expression in VCT cells. Taken together with previous studies, we speculate that this may be due to the period in which the current PE samples were collected, as well as the number of samples, and that more samples will need to be collected in the future for in-depth validation and study of biomarkers.

Although the nomogram constructed in this study shows great potential in predicting preeclampsia (PE), further in-depth exploration is needed on how to integrate it into the existing clinical workflow. Currently, there are many limitations in the early prediction methods for PE in clinical practice, such as poor accuracy and insufficient specificity,<sup>97</sup> which makes the early and accurate diagnosis of PE and timely intervention challenging. From a practical application perspective, integrating the nomogram into the clinical workflow first requires addressing the issue of data acquisition. The use of the nomogram relies on accurate and comprehensive clinical data, including patients' basic information, medical history, and laboratory test indicators. Clinicians need to ensure that they can efficiently and accurately collect these data during daily diagnosis and treatment and input them into the nomogram model.<sup>98</sup> For example, the electronic medical record system can be optimized to automatically extract and integrate relevant data and directly connect to the nomogram prediction model, reducing errors and time costs caused by manual operations.<sup>99</sup> In terms of overcoming the shortcomings of existing early prediction methods, the nomogram has unique advantages. It comprehensively considers the expression levels of multiple biomarkers related to PE (such as CD44, MIF, PIK3R1, and TLR4), providing more comprehensive and accurate predictive information compared to the detection of a single indicator.<sup>100</sup> However, in practical application, the reliability and stability of the nomogram in different populations and clinical environments still need to be further verified. Multi-center, large-sample clinical studies can be conducted to collect data from patients in different regions and races, evaluate the predictive efficacy of the nomogram, and optimize and adjust the model based on the results.<sup>101</sup> In addition, the acceptance of the nomogram by doctors and patients also needs to be considered in clinical application. Doctors should be trained and patients should be well informed.<sup>102</sup> In conclusion, although the nomogram provides a new tool for the early prediction of PE, its wide application in clinical practice still requires in-depth research and practical exploration in aspects such as data acquisition, model validation, doctor training, and patient education.

Through single-cell transcriptome sequencing and bioinformatics analysis, four ARGs (CD44, MIF, PIK3R1, and TLR4) were identified as biomarkers, and the same expression trend was observed in the dataset and PE clinical samples. These biomarkers are involved in immune infiltration, oxidative phosphorylation, and other processes in PE and are expected to become potential markers of PE. However, clinical applications require more data support from samples. The effects of these key genes on PE were further studied in animal and cell experiments.

## Conclusion

PE is a hypertensive multisystem disorder that causes significant fetal-maternal mortality and morbidity worldwide. Apoptosis occurs in PE and is associated with immune cells and pathways. According to the present study, four ARGs, CD44, MIF, PIK3R1, and TLR4, offer new insights into PE's cellular mechanisms of PE and may serve as potential biomarkers for PE. This study opens new avenues for future research on PE's pathogenesis of PE and treatment options. Nevertheless, this study has some limitations. The sample size used in the current research is relatively small, making it difficult to cover the genetic, physiological, and environmental differences among individuals, which limits the generalizability of the research conclusions and affects the accuracy and stability of the nomogram model predictions. Future research will be conducted on a large scale, multi-center, and will include different patient populations; in-depth functional studies will be carried out using in vitro cell experiments and in vivo animal models; the interactions between

these genes and other related factors of PE will be explored in depth to gain a more comprehensive understanding of the disease mechanism, aiding in the development of more effective treatment strategies.

## Abbreviations

PE, preeclampsia; scRNA-seq, transcriptomics and RNA sequencing; ARGs, apoptosis-related genes; RT-qPCR, reverse transcription-quantitative polymerase chain reaction; EVT, extravillous trophoblasts; SCT, syncytiotrophoblasts; CMP, common myeloid progenitors; VCT, villous cytotrophoblasts; GEO, Gene Expression Omnibus; QC, quality control; PCs, principal component; UMAP, uniform manifold approximation and projection.

## Data Sharing Statement

The datasets analysed in this study are available in the Gene Expression Omnibus (GEO) database (<https://www.ncbi.nlm.nih.gov/gds>), including GSE43942, GSE25906, and GSE173193.

## Ethical Approval and Informed Consent

This study followed the ethical requirements of the Declaration of Helsinki. I certify that the research study titled Single-cell RNA sequencing combined with transcriptomic exploration of apoptosis-related genes as biomarkers of preeclampsia and experimental validation has been approved by the Ethics Committee of the First People's Hospital of Yunnan Province. The approval number and date of approval were as follows: [approval number: No.221; approval date: April 22, 2024].

## Consent for Publication

The participants were provided with a clear and understandable explanation of the research objectives, procedures, potential risks, and benefits. They were informed that their participation was voluntary, and that they had the right to withdraw from the study at any time. Participants were given the opportunity to ask questions and provide written informed consent prior to their involvement in the study.

## Acknowledgments

We would like to express our sincere gratitude to all individuals and organizations who supported and assisted us throughout this research. Without your support, this study would not have been feasible.

## Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

## Funding

This study was supported by the following four projects: Key Research and Development Plan of Yunnan Province (grant number: 202403AC100002), Project of National Natural Science Foundation of China (grant number: 42167060), and Social Development Projects of Yunnan Province (202302AA310044).

## Disclosure

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

## References

1. Rana S, Lemoine E, Granger JP, et al. Preeclampsia: pathophysiology, challenges, and perspectives. *Circ Res*. 2019;124:1094–1112. doi:10.1161/CIRCRESAHA.118.313276



2. Phipps EA, Thadhani R, Benzing T, et al. preeclampsia: pathogenesis, novel diagnostics and therapies. *Nat Rev Nephrol.* **2019**;15:275–289. doi:10.1038/s41581-019-0119-6
3. MacDonald TM, Walker SP, Hannan NJ, et al. Clinical tools and biomarkers to predict preeclampsia. *Ebiomedicine.* **2022**;75:103780. doi:10.1016/j.ebiom.2021.103780
4. Hoseinzadeh A, Esmaeili SA, Sahebi R, et al. Fate and long-lasting therapeutic effects of mesenchymal stromal/stem-like cells: mechanistic insights. *Stem Cell Res Ther.* **2025**;16:33.
5. Li J, Wang M, Zhou H, et al. The role of pyroptosis in the occurrence and development of pregnancy-related diseases. *Front Immunol.* **2024**;15:1400977. doi:10.3389/fimmu.2024.1400977
6. Gao Q, Cheng K, Cai L, et al. Aβ1-42 stimulates an increase in autophagic activity through tunicamycin-induced endoplasmic reticulum stress in HTR-8/SVneo cells and late-onset pre-eclampsia. *J Mol Histol.* **2024**;55:513–525. doi:10.1007/s10735-024-10203-7
7. Park C, Alahari S, Ausman J, et al. Placental hypoxia-induced ferroptosis drives vascular damage in preeclampsia. *Circ Res.* **2025**;136:361–378. doi:10.1161/CIRCRESAHA.124.325119
8. You Y, Qian Z, Jiang Y, et al. Insights into the pathogenesis of gestational and hepatic diseases: the impact of ferroptosis. *Front Cell Dev Biol.* **2024**;12:1482838. doi:10.3389/fcell.2024.1482838
9. Liu M, Li Y, Deng Z, et al. Mcm5 mutation leads to silencing of Stat1-bcl2 which accelerating apoptosis of immature T lymphocytes with DNA damage. *Cell Death Dis.* **2025**;16:84. doi:10.1038/s41419-025-07392-8
10. Zhang YC, Qin XL, Ma XL, et al. CLDN1 regulates trophoblast apoptosis and proliferation in preeclampsia. *Reproduction.* **2021**;161:623–632. doi:10.1530/REP-20-0677
11. Mo HQ, Tian FJ, Li X, et al. ANXA7 regulates trophoblast proliferation and apoptosis in preeclampsia. *Am J Reprod Immunol.* **2019**;82:e13183. doi:10.1111/aji.13183
12. Chen Y, Xue F, Han C, et al. Ferulic acid ameliorated placental inflammation and apoptosis in rat with preeclampsia. *Clin Exp Hypertens.* **2019**;41:524–530. doi:10.1080/10641963.2018.1516773
13. Zhou J, Zhao Y, An P, et al. Hsa\_circ\_0002348 regulates trophoblast proliferation and apoptosis through miR-126-3p/BAK1 axis in preeclampsia. *J Transl Med.* **2023**;21:509. doi:10.1186/s12967-023-04240-1
14. Travaglino A, Raffone A, Saccone G, et al. Placental morphology, apoptosis, angiogenesis and epithelial mechanisms in early-onset preeclampsia. *Eur J Obstet Gynecol Reprod Biol.* **2019**;234:200–206. doi:10.1016/j.ejogrb.2018.12.039
15. Gong X, He W, Jin W, et al. Disruption of maternal vascular remodeling by a fetal endoretrovirus-derived gene in preeclampsia. *Genome Biol.* **2024**;25:117. doi:10.1186/s13059-024-03265-z
16. Wang Y, Lv Q, Li J, et al. The protective mechanism of human umbilical cord mesenchymal stem cell-derived exosomes against neutrophil extracellular trap-induced placental damage. *Placenta.* **2024**;153:59–74. doi:10.1016/j.placenta.2024.05.136
17. Smith S, Smith J, Jones K, et al. Placental ischemia during pregnancy induces hypertension, cerebral inflammation, and oxidative stress in dams postpartum. *Hypertens Pregnancy.* **2025**;44:2454597. doi:10.1080/10641955.2025.2454597
18. Xiao F, Li C, Wang X, et al. The predictive value of placental growth factor combined with uterine ultrasound arterial blood flow characteristics for preeclampsia. *Pak J Med Sci.* **2025**;41:426–431. doi:10.12669/pjms.41.2.11349
19. Shu Z, Wang W. Predictive value of prenatal screening markers combined with serum placental growth factor in early pregnancy for preeclampsia. *Pak J Med Sci.* **2025**;41:598–602. doi:10.12669/pjms.41.2.9794
20. Enengl S, Oppelt P, Stelzl P, et al. Risk for imminent delivery in preeclampsia based on the sFlt-1/PIGF ratio: do we need new cut-offs? *Geburtshilfe Frauenheilkd.* **2025**;85:190–199. doi:10.1055/a-2497-8104
21. Alfuraih AM. The emerging role of sonoelastography in pregnancy: applications in assessing maternal and fetal health. *Diagnostics.* **2024**;15:47. doi:10.3390/diagnostics15010047
22. Cristodoro M, Messa M, Tossetta G, et al. First trimester placental biomarkers for pregnancy outcomes. *Int J Mol Sci.* **2024**;25:6136. doi:10.3390/ijms25116136
23. Piani F, Tossetta G, Fantone S, et al. First trimester CD93 as a novel marker of preeclampsia and its complications: a pilot study. *High Blood Press Cardiovasc Prev.* **2023**;30:591–594. doi:10.1007/s40292-023-00608-y
24. Fantone S, Tossetta G, Di Simone N, et al. CD93 a potential player in cytotrophoblast and endothelial cell migration. *Cell Tissue Res.* **2022**;387:123–130. doi:10.1007/s00441-021-03543-3
25. Li W, Zhong L, Zhao K, et al. Identification of critical biomarkers and immune infiltration in preeclampsia through bioinformatics and machine learning methods. *BMC Pregnancy Childbirth.* **2025**;25:136. doi:10.1186/s12884-025-07257-0
26. Li Q, Tong Y, Guo J, et al. Vitamin D receptor regulates oxidative stress and apoptosis via the HIF-1α/HO-1 pathway in cardiomyocytes. *Cell Biochem Biophys.* **2025**. doi:10.1007/s12013-025-01681-x
27. Shuvo AUH, Alimullah M, Jahan I, et al. Evaluation of xanthine oxidase inhibitors febuxostat and allopurinol on kidney dysfunction and histological damage in two-kidney, one-clip (2K1C) rats. *Scientifica.* **2025**;2025:7932075.
28. Annesi L, Tossetta G, Borghi C, Piani F. The role of xanthine oxidase in pregnancy complications: a systematic review. *Antioxidants.* **2024**;13:1234. doi:10.3390/antiox13101234
29. Shenhav S, Harel I, Solt I, et al. Fetoplacental unit involvement in uric acid production in women with severe preeclampsia: a prospective case control pilot study. *J Matern Fetal Neonatal Med.* **2024**;37:2399304. doi:10.1080/14767058.2024.2399304
30. Xiang Y, Cheng Y, Li X, et al. Up-regulated expression and aberrant DNA methylation of LEP and SH3PXD2A in preeclampsia. *PLoS One.* **2013**;8:e59753. doi:10.1371/journal.pone.0059753
31. Tsai S, Hardison NE, James AH, et al. Transcriptional profiling of human placentas from pregnancies complicated by preeclampsia reveals dysregulation of sialic acid acetyltransferase and immune signalling pathways. *Placenta.* **2011**;32:175–182.
32. Zhang B, Zhang F, Lu F, et al. Reduced cell invasion may be a characteristic of placental defects in pregnant women of advanced maternal age at single-cell level. *J Zhejiang Univ Sci B.* **2022**;23:747–759. doi:10.1631/jzus.B2101024
33. Yang Y, Guo F, Peng Y, et al. Transcriptomic profiling of human placenta in gestational diabetes mellitus at the single-cell level. *Front Endocrinol.* **2021**;12:679582.
34. Yang F, Zhang Y. Apoptosis-related genes-based prognostic signature for osteosarcoma. *Aging.* **2022**;14:3813–3825.

35. Satija R, Farrell JA, Gennert D, et al. Spatial reconstruction of single-cell gene expression data. *Nat Biotechnol.* **2015**;33:495–502. doi:10.1038/nbt.3192
36. Aran D, Looney AP, Liu L, et al. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat Immunol.* **2019**;20:163–172. doi:10.1038/s41590-018-0276-y
37. Gibbs DC, McCrary MR, Moreno CS, et al. Epidermal growth factor dampens pro-inflammatory gene expression induced by interferon-gamma in global transcriptome analysis of keratinocytes. *BMC Genomics.* **2025**;26:122. doi:10.1186/s12864-025-11237-1
38. Budzinska A, Byl L, Teyssie F, et al. Caloric labels do not influence taste pleasantness and neural responses to erythritol and sucrose. *Neuroimage.* **2025**. doi:10.1016/j.neuroimage.2025.121061
39. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **2015**;43:e47. doi:10.1093/nar/gkv007
40. Gustavsson EK, Zhang D, Reynolds RH, et al. ggtranscript: an R package for the visualization and interpretation of transcript isoforms using ggplot2. *Bioinformatics.* **2022**;38:3844–3846. doi:10.1093/bioinformatics/btac409
41. Gu Z, Hubschmann D. Make Interactive Complex Heatmaps in R. *Bioinformatics.* **2022**;38:1460–1462.
42. Chen H, Boutros PC. VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinf.* **2011**;12:35. doi:10.1186/1471-2105-12-35
43. Yu G, Wang LG, Han Y, et al. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics.* **2012**;16:284–287.
44. Liu P, Xu H, Shi Y, et al. Potential molecular mechanisms of plantain in the treatment of gout and hyperuricemia based on network pharmacology. *Evid Based Complement Alternat Med.* **2020**;2020:3023127. doi:10.1155/2020/3023127
45. Sachs MC. plotROC: a tool for plotting ROC curves. *J Stat Softw.* **2017**;79:1–9.
46. Wang L, Wang D, Yang L, et al. Cuproptosis related genes associated with Jab1 shapes tumor microenvironment and pharmacological profile in nasopharyngeal carcinoma. *Front Immunol.* **2022**;13:989286. doi:10.3389/fimmu.2022.989286
47. Stenson PD, Ball EV, Mort M, et al. Human Gene Mutation Database (HGMD): 2003 update. *Hum Mutat.* **2003**;21:577–581. doi:10.1002/humu.10212
48. Capietto AH, Hoshyar R, Delamarre L. Sources of cancer neoantigens beyond single-nucleotide variants. *Int J Mol Sci.* **2022**;23:10131. doi:10.3390/ijms231710131
49. Krassowski M, Pellegrina D, Mee MW, et al. ActiveDriverDB: interpreting genetic variation in human and cancer genomes using post-translational modification sites and signaling networks (2021 Update). *Front Cell Dev Biol.* **2021**;9:626821. doi:10.3389/fcell.2021.626821
50. Yu G, Li F, Qin Y, et al. GOSemSim: an R package for measuring semantic similarity among GO terms and gene products. *Bioinformatics.* **2010**;26:976–978. doi:10.1093/bioinformatics/btq064
51. Hosseinzadeh F, Ebrahimi M, Goliaei B, et al. Classification of lung cancer tumors based on structural and physicochemical properties of proteins by bioinformatics models. *PLoS One.* **2012**;7:e40017. doi:10.1371/journal.pone.0040017
52. Chen B, Khodadoust MS, Liu CL, et al. Profiling tumor infiltrating immune cells with CIBERSORT. *Methods Mol Biol.* **2018**;1711:243–259.
53. Zein RA, Akhtar H. Getting started with the graded response model: an introduction and tutorial in R. *Int J Psychol.* **2025**;60:e13265. doi:10.1002/ijop.13265
54. Liu B, Li Y, Xiang J, et al. Significance of pyroptosis in immunoregulation and prognosis of patients with acute respiratory distress syndrome: evidence from RNA-Seq of alveolar macrophages. *J Inflamm Res.* **2023**;16:3547–3562. doi:10.2147/JIR.S422585
55. Hanzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinf.* **2013**;14:7. doi:10.1186/1471-2105-14-7
56. Ru Y, Kechris KJ, Tabakoff B, et al. The multiMiR R package and database: integration of microRNA-target interactions along with their disease and drug associations. *Nucleic Acids Res.* **2014**;42:e133. doi:10.1093/nar/gku631
57. Kuleshov MV, Jones MR, Rouillard AD, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **2016**;44:W90–7. doi:10.1093/nar/gkw377
58. Seeliger D, de Groot BL. Ligand docking and binding site analysis with PyMOL and Autodock/Vina. *J Comput Aided Mol Des.* **2010**;24:417–422. doi:10.1007/s10822-010-9352-6
59. Griss J, Viteri G, Sidiropoulos K, et al. ReactomeGSA-efficient multi-omics comparative pathway analysis. *Mol Cell Proteomics.* **2020**;19:2115–2125. doi:10.1074/mcp.TIR120.002155
60. Wang Z, Dai Z, Zheng L, et al. Ferroptosis activation scoring model assists in chemotherapeutic agents' selection and mediates cross-talk with immunocytes in malignant glioblastoma. *Front Immunol.* **2021**;12:747408. doi:10.3389/fimmu.2021.747408
61. Du J, Yuan X, Deng H, et al. Single-cell and spatial heterogeneity landscapes of mature epicardial cells. *J Pharm Anal.* **2023**;13:894–907. doi:10.1016/j.jpha.2023.07.011
62. Dimitriadis E, Rolnik DL, Zhou W, et al. Preeclampsia. *Nat Rev Dis Primers.* **2023**;9:35. doi:10.1038/s41572-023-00451-4
63. Huang YT, Lin YW, Chiu HM, et al. Curcumin induces apoptosis of colorectal cancer stem cells by coupling with CD44 marker. *J Agric Food Chem.* **2016**;64:2247–2253. doi:10.1021/acs.jafc.5b05649
64. Weng X, Maxwell-Warburton S, Hasib A, et al. The membrane receptor CD44: novel insights into metabolism. *Trends Endocrinol Metab.* **2022**;33:318–332. doi:10.1016/j.tem.2022.02.002
65. Obut M, Oglak SC. Expression of CD44 and IL-10 in normotensive and preeclamptic placental tissue. *Ginek Pol.* **2020**;91:334–341. doi:10.5603/GP.2020.0058
66. Todd N, McNally R, Alqudah A, et al. Role of A novel angiogenesis FKBPL-CD44 pathway in preeclampsia risk stratification and mesenchymal stem cell treatment. *J Clin Endocrinol Metab.* **2021**;106:26–41. doi:10.1210/clinem/dgaa403
67. Kang I, Bucala R. The immunobiology of MIF: function, genetics and prospects for precision medicine. *Nat Rev Rheumatol.* **2019**;15:427–437. doi:10.1038/s41584-019-0238-2
68. Jankauskas SS, Wong D, Bucala R, et al. Evolving complexity of MIF signaling. *Cell Signal.* **2019**;57:76–88. doi:10.1016/j.cellsig.2019.01.006
69. Zheng S, Ren X, Han T, et al. Fenofibrate attenuates fatty acid-induced islet beta-cell dysfunction and apoptosis via inhibiting the NF-kappaB/MIF dependent inflammatory pathway. *Metabolism.* **2017**;77:23–38. doi:10.1016/j.metabol.2017.09.001
70. Dong Z, Yang P, Ji Z, et al. MIF inhibition attenuates intervertebral disc degeneration by reducing nucleus pulposus cell apoptosis and inflammation. *Exp Cell Res.* **2024**;439:114089. doi:10.1016/j.yexcr.2024.114089

71. Galbiati S, Inversetti A, Causarano V, et al. HIF1A and MIF as potential predictive mRNA biomarkers of preeclampsia: a longitudinal prospective study in high risk population. *Clin Chem Lab Med*. 2015;53:1339–1347. doi:10.1515/ccm-2014-0745
72. Tsay A, Wang JC. The role of PIK3R1 in metabolic function and insulin sensitivity. *Int J Mol Sci*. 2023;24:12665. doi:10.3390/ijms241612665
73. Wang X, He A, Yip KC, et al. Diagnostic signature and immune characteristic of aging-related genes from placentas in Preeclampsia. *Clin Exp Hypertens*. 2022;1–08. doi:10.1080/10641963.2022.2130930
74. Tan A, Luo R, Ruan P. miR-495 promotes apoptosis and inhibits proliferation in endometrial cells via targeting PIK3R1. *Pathol Res Pract*. 2019;215:594–599. doi:10.1016/j.prp.2019.01.020
75. Xu W, Ru P, Gu Z, et al. Comprehensive analysis of differently expressed and methylated genes in preeclampsia. *Comput Math Methods Med*. 2020;2020:2139270. doi:10.1155/2020/2139270
76. Kim HJ, Kim H, Lee JH, et al. Toll-like receptor 4 (TLR4): new insight immune and aging. *Immun Ageing*. 2023;20:67. doi:10.1186/s12979-023-00383-3
77. Zhang Y, Liang X, Bao X, et al. Toll-like receptor 4 (TLR4) inhibitors: current research and prospective. *Eur J Med Chem*. 2022;235:114291. doi:10.1016/j.ejmech.2022.114291
78. Wei X, Liu H, Li X, et al. Over-expression of MiR-122 promotes apoptosis of hepatocellular carcinoma via targeting TLR4. *Ann Hepatol*. 2019;18:869–878. doi:10.1016/j.aohp.2019.07.005
79. Zhang L, Yang H. Expression and localization of TLR4 and its negative regulator Tollip in the placenta of early-onset and late-onset preeclampsia. *Hypertens Pregnancy*. 2012;31:218–227. doi:10.3109/10641955.2011.642434
80. Manfioletti G, Fedele M. Epithelial-Mesenchymal Transition (EMT). *Int J Mol Sci*. 2023;24:11386. doi:10.3390/ijms241411386
81. Chen T, You Y, Jiang H, et al. Epithelial-mesenchymal transition (EMT): a biological process in the development, stem cell differentiation, and tumorigenesis. *J Cell Physiol*. 2017;232:3261–3272. doi:10.1002/jcp.25797
82. Ge H, Yin N, Han TL, et al. Interleukin-27 inhibits trophoblast cell invasion and migration by affecting the epithelial-mesenchymal transition in preeclampsia. *Reprod Sci*. 2019;26:928–938. doi:10.1177/1933719118799206
83. Jung E, Romero R, Yeo L, et al. The etiology of preeclampsia. *Am J Obstet Gynecol*. 2022;226:S844–S866. doi:10.1016/j.ajog.2021.11.1356
84. Alahakoon TI, Medbury H, Williams H, et al. Characterization of fetal monocytes in preeclampsia and fetal growth restriction. *J Perinat Med*. 2019;47:434–438. doi:10.1515/jpm-2018-0286
85. Faas MM, Spaans F, De Vos P. Monocytes and macrophages in pregnancy and preeclampsia. *Front Immunol*. 2014;5:298. doi:10.3389/fimmu.2014.00298
86. Loewendorf AI, Nguyen TA, Yesayan MN, et al. Preeclampsia is characterized by fetal NK cell activation and a reduction in regulatory T cells. *Am J Reprod Immunol*. 2015;74:258–267. doi:10.1111/aji.12393
87. Deer E, Herrock O, Campbell N, et al. The role of immune cells and mediators in preeclampsia. *Nat Rev Nephrol*. 2023;19:257–270. doi:10.1038/s41581-022-00670-0
88. Rui R, Zhou L, He S. Cancer immunotherapies: advances and bottlenecks. *Front Immunol*. 2023;14:1212476. doi:10.3389/fimmu.2023.1212476
89. Mukherjee I, Singh S, Karmakar A, et al. New immune horizons in therapeutics and diagnostic approaches to Preeclampsia. *Am J Reprod Immunol*. 2023;89:e13670. doi:10.1111/aji.13670
90. Othman AS, Franke-Fayard BM, Imai T, et al. OX40 stimulation enhances protective immune responses induced after vaccination with attenuated malaria parasites. *Front Cell Infect Microbiol*. 2018;8:247. doi:10.3389/fcimb.2018.00247
91. Weimershaus M, Mauvais FX, Saveanu L, et al. Innate immune signals induce anterograde endosome transport promoting MHC class I cross-presentation. *Cell Rep*. 2018;24:3568–3581. doi:10.1016/j.celrep.2018.08.041
92. Krizova L, Dadakova K, Kasparovska J, et al. Isoflavones. *Molecules*. 2019;24:1076. doi:10.3390/molecules24061076
93. Guo J, Bian W, Jiang H. Puerarin attenuates preeclampsia-induced trophoblast mobility loss and inflammation by modulating miR-181b-5p/RBAK axis. *Am J Reprod Immunol*. 2022;87:e13510. doi:10.1111/aji.13510
94. Knofler M, Haider S, Saleh L, et al. Human placenta and trophoblast development: key molecular mechanisms and model systems. *Cell Mol Life Sci*. 2019;76:3479–3496. doi:10.1007/s00018-019-03104-6
95. Redman C, Staff AC, Roberts JM. Syncytiotrophoblast stress in preeclampsia: the convergence point for multiple pathways. *Am J Obstet Gynecol*. 2022;226:S907–S927. doi:10.1016/j.ajog.2020.09.047
96. Zhou W, Wang H, Yang Y, et al. Trophoblast cell subtypes and dysfunction in the placenta of individuals with preeclampsia revealed by single-cell RNA sequencing. *Mol Cells*. 2022;45:317–328. doi:10.14348/molcells.2021.0211
97. Zheng Y, Fang Z, Wu X, et al. Identification of hub genes, diagnostic model, and immune infiltration in preeclampsia by integrated bioinformatics analysis and machine learning. *BMC Pregnancy Childbirth*. 2024;24:847. doi:10.1186/s12884-024-07028-3
98. Liang Q, Sun L. Predictive value of urine misfolded protein in preeclampsia in twin pregnancies. *Arch Gynecol Obstet*. 2024;310:2879–2887. doi:10.1007/s00404-024-07769-8
99. Lin R, Weng X, Lin L, et al. Identification and preliminary validation of biomarkers associated with mitochondrial and programmed cell death in pre-eclampsia. *Front Immunol*. 2025;15:1453633. doi:10.3389/fimmu.2024.1453633
100. Hong DL, Zhu Q, Chen WC, et al. Factors contributing to perioperative blood transfusion during total Hip arthroplasty in patients continuing preoperative aspirin treatment: a nomogram prediction model. *BMC Musculoskelet Disord*. 2025;26:138. doi:10.1186/s12891-025-08399-0
101. Xiao E, Yu R, Cai X, et al. Development and validation of a novel metabolic health-related nomogram to improve predictive performance of cardiovascular disease risk in patients with prediabetes. *Lipids Health Dis*. 2025;24:45. doi:10.1186/s12944-025-02445-5
102. Fente BM, Tesema GA, Gudayu TW, et al. Development and validation of a nomogram for predicting low birth weight among pregnant women who had antenatal care visits at debre markos comprehensive and specialized hospital, Ethiopia. *Front Med*. 2023;10:1253490.

**Journal of Inflammation Research****Dovepress**  
Taylor & Francis Group**Publish your work in this journal**

The Journal of Inflammation Research is an international, peer-reviewed open-access journal that welcomes laboratory and clinical findings on the molecular basis, cell biology and pharmacology of inflammation including original research, reviews, symposium reports, hypothesis formation and commentaries on: acute/chronic inflammation; mediators of inflammation; cellular processes; molecular mechanisms; pharmacology and novel anti-inflammatory drugs; clinical conditions involving inflammation. The manuscript management system is completely online and includes a very quick and fair peer-review system. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/journal-of-inflammation-research-journal>