ACCSCIENCE PUBLISHING

Gene & Protein in Disease

ORIGINAL RESEARCH ARTICLE

Bioinformatics analysis of the GEO database for the identification of novel biomarkers and potential targeted drugs for pulmonary hypertension

Zhen-Dong Lu[1], Zhi-Liang Jiang[2], Wahab Hussain[2], Xin-Ying Ji[2,3]*, and Umair Ali Khan Saddozai[4]*

[1]Department of Medical Oncology, Beijing Tuberculosis and Thoracic Tumor Research Institute, Beijing Chest Hospital, Capital Medical University, Beijing, China

[2]Henan International Joint Laboratory for Nuclear Protein Regulation, School of Basic Medical Sciences, Henan University, Kaifeng, Henan, China

[3]Henan International Joint Laboratory for Nuclear Protein Regulation, The First Affiliated Hospital of Henan University, Kaifeng, Henan, China

[4]Institute of Translational Medicine, Medical College, Yangzhou University, Yangzhou, Jiangsu, China

***Corresponding authors:**
Xin-Ying Ji
(10190096@vip.henu.edu.cn);
Umair Ali Khan Saddozai
(umairsaddozai2@gmail.com)

Citation: Lu Z, Jiang Z, Hussain W, Ji X, Saddozai UAK. Bioinformatics analysis of the GEO database for the identification of novel biomarkers and potential targeted drugs for pulmonary hypertension. *Gene Protein Dis*. 2025;4(4):025080018. doi: 10.36922/GPD025080018

Received: February 23, 2025

Revised: August 04, 2025

Accepted: August 11, 2025

Published online: August 29, 2025

Abstract

Pulmonary arterial hypertension (PAH) is a progressive and life-threatening cardiopulmonary disorder. This study integrated two PAH datasets (GSE117261 and GSE113439) from the Gene Expression Omnibus database to screen novel PAH biomarkers using bioinformatics methods and explore immune cell infiltration and potential therapeutic drugs. After batch effect correction, 311 differentially expressed genes (DEGs), comprising 182 upregulated and 129 downregulated genes, were identified using the Linear Models for Microarray Data analysis. Weighted gene co-expression network analysis revealed 11 significant modules, and intersecting these with DEGs yielded 24 module DEGs. Gene ontology and Kyoto Encyclopedia of Genes and Genomes enrichment analyses indicated enrichment in extracellular matrix organization, exosome-related functions, and inflammation-related pathways, including the positive regulation of inflammatory response and acute-phase response. A protein-protein interaction network analysis identified 10 hub genes: Upregulated periostin, osteoblast-specific factor (*POSTN*), *OGN*, asporin, and downregulated S100 calcium-binding protein A12, *LCN2*, *SPP1*, S100 calcium-binding protein A9, *CD163*, *S100A8*, and *AQP9*. Immune infiltration analysis (CIBERSORT, dataset GSE169471) revealed markedly altered levels of monocytes, dendritic cells, neutrophils, resting cluster of differentiation 4-positive memory T cells, and macrophages. Single-cell RNA sequencing analysis further confirmed hub gene expression. Notably, macrophage-associated genes (*SPP1*, *S100A8/A9/A12*, *CD163*, *POSTN*, *AQP9*, and *LCN2*) were implicated in vascular inflammation, endothelial dysfunction, and fibrosis, underscoring their role in immune-mediated vascular remodeling in PAH. Finally, drug prediction using the Comparative Toxicogenomic Database identified retinol, arsenic trioxide, and active vitamin D as potential therapeutics with significant regulatory effects on hub genes. This integrated bioinformatics analysis reveals key genes, pathways, and immune cell alterations in

PAH, emphasizing macrophage-associated mechanisms in vascular remodeling and immune dysregulation. The findings provide potential biomarkers and therapeutic targets for improved PAH management.

Keywords: Pulmonary arterial hypertension; Immune infiltration; Single-cell RNA sequencing; Weighted gene co-expression network analysis

1. Introduction

Pulmonary arterial hypertension (PAH) is a progressive and life-threatening cardiopulmonary disorder, primarily defined by sustained elevations in pulmonary vascular resistance and pulmonary arterial pressure.[1] From a pathological perspective, the development of PAH involves pulmonary artery vasoconstriction and remodeling, ultimately leading to right heart failure and pre-mature death of patients. In terms of treatment, current therapies for PAH mainly include supportive care, pharmacological interventions, and cardiopulmonary transplantation.[2] Unfortunately, due to the insidious onset and rapid progression of PAH, despite these treatment options, effective treatments for PAH are still lacking, and the mortality rate remains high.[3] Furthermore, the pathogenesis of PAH remains complex and incompletely elucidated. Therefore, further exploration of the key factors involved in the pathogenesis of PAH and the search for effective drug targets remain valuable for the treatment of PAH.

The regulation of immune cells plays a critical role in the pathogenesis and progression of PAH. Previous research has shown that a reduction in peripheral blood cluster of differentiation 8 (CD8) T cells is linked to increased mortality in PAH patients,[4] while an elevated proportion of peripheral blood T-helper (Th)2 cells contributes to disease progression.[5] In addition, an increased Th17/T-regulatory cell ratio has been strongly correlated with disease severity and poorer prognosis.[6] Given the pivotal role of immune cells in PAH pathogenesis, immune cell-targeted therapies may offer a promising treatment strategy. Increasing attention has been directed toward the functions of various immune cells, including macrophages, monocytes, and neutrophils. This study employed bioinformatics approaches to predict alterations in immune cell populations between PAH patient tissues and normal tissues, providing a foundation for further in-depth investigations into the relationship between PAH and immune regulation.

In recent years, the integration of microarray and bioinformatics technologies has been widely utilized to discover novel biomarkers and elucidate their functional roles across diverse diseases. Among these approaches, weighted gene co-expression network analysis (WGCNA) has emerged as a robust and well-established method.[7] WGCNA assesses correlations between genes and sample traits, clusters genes with similar expression patterns, and identifies key hub genes involved in disease progression. This technique is particularly effective in uncovering underlying disease mechanisms, discovering potential biomarkers, and identifying therapeutic targets.

In this study, two datasets (GSE117261 and GSE113439) were integrated into a single comprehensive dataset to eliminate batch differences and expand the sample size. To identify key genes associated with the pathogenesis of PAH, along with their functional pathways and regulatory networks, DEGs analysis, WGCNA, functional enrichment analysis, protein–protein interaction (PPI) network analysis, and the construction of microRNA (miRNA)–gene–transcription factor (TF) networks were performed. Furthermore, immune cell infiltration under PAH conditions was systematically analyzed, and potential therapeutic drugs or compounds targeting the identified key genes were predicted. This study elucidated the pivotal genes underlying the pathogenesis of PAH, characterized the associated immune cell infiltration patterns, and identified potential therapeutic drugs, thereby facilitating the discovery of novel biomarkers and informing the development of effective PAH treatment strategies.

2. Materials and methods

2.1. Gene chip dataset information

The two datasets, GSE117261 and GSE113439, were retrieved from the National Center for Biotechnology Information's Gene Expression Omnibus database (GEO; https://www.ncbi.nlm.nih.gov/geo/) by searching for "Pulmonary Arterial Hypertension" or "PAH", selecting "Series" as the entry type, "Expression profiling by array" as the study type, and "Homo sapiens" as the species. Both datasets were annotated on the GPL6244 platform. After screening the clinical data, a total of 58 PAH samples and 25 normal samples were obtained from the GSE117261 dataset, and 6 PAH samples and 11 normal samples (excluding data from patients with both PAH and connective tissue disease) were obtained from the

GSE113439 dataset. This resulted in a total of 64 PAH samples and 36 normal samples. Single-cell data were obtained from three PAH samples from GSE169471.

2.2. Pre-processing and differential gene expression analysis

The affy package in R (version 4.2.2) was employed for pre-processing the two original PAH datasets, including background correction, normalization, $\log_2$ transformation, and imputation of missing values using the k-nearest neighbor method.[8] For genes with multiple corresponding probes, the average expression value was calculated. To mitigate batch effects, the SVA package was applied. In addition, principal component analysis (PCA) plots were included before and after correction to demonstrate the successful removal of batch effects. Differentially expressed genes (DEGs) were subsequently identified using the Linear Models for Microarray Data (LIMMA) package, with a significance threshold of $p<0.05$ and $|\log_2 FC| > 0.5$. Finally, the ggplot2 and pheatmap packages were utilized for DEG analysis and visualization, generating heatmaps and volcano plots.

2.3. Development and characterization of WGCNA modules in PAH

To explore gene interactions, WGCNA was utilized to construct a co-expression network. Initially, the top 25% most variable genes from the integrated dataset were selected. After excluding abnormal samples, the optimal soft threshold (β) was determined by assessing co-expression similarity. The pick-Soft-Threshold function was applied to obtain the adjacency matrix, which was then transformed into a topological overlap matrix.[9] Hierarchical clustering and dynamic tree cutting were performed to identify gene modules by grouping genes with similar expression patterns. The two modules showing the strongest positive and negative correlations with PAH (correlation coefficient > 0.5, $p<0.05$) were selected. Finally, DEGs were intersected with the genes from both positively and negatively correlated modules, and their union was defined as the module DEG (MDEGs) for further analysis.

2.4. Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis of the module DEGs

GO and KEGG pathway analyses are commonly utilized bioinformatics tools for investigating gene functions and regulatory mechanisms. These analyses help elucidate the functional enrichment of genes or proteins and the pivotal roles of biological pathways by comparing gene or protein functional data with the three categories in the GO database and identifying pathways associated with specific biological processes. GO annotation analysis focuses on three key categories: Cellular component, biological process, and molecular function, providing valuable insights into the functions and interactions of genes or proteins. Meanwhile, KEGG pathway enrichment analysis highlights pathways crucial for disease progression or drug response. Both the DAVID (https://david.ncifcrf.gov/)[10] and KOBAS (http://kobas.cbi.pku.edu.cn/kobas)[11] databases serve as platforms for the functional classification and annotation of genes and proteins. In this study, these tools were employed for GO analysis of MDEGs, with KOBAS being specifically used for KEGG pathway analysis of MDEGs. The criteria for pathway selection included a $p<0.05$ and a minimum count of ≥2.

2.5. Construction of protein–protein interaction network and screening of key genes for the module DEGs

An initial analysis of the interactions between MDEGs was conducted using the gene interaction retrieval tool, the STRING database (https://cn.string-db.org/).[12] A minimum interaction score of ≥0.4 was set as the cutoff for screening. Further investigation of key genes associated with the pathogenesis of PAH was conducted using the Cytohubba plugin in Cytoscape (v3.10.1).[13] Twelve distinct algorithms—the degree method, betweenness method, bottleneck method, closeness method, clustering coefficient method, density of maximum neighborhood component method, eccentricity method, edge percolated component method, maximal clique centrality method, maximum neighborhood component method, radiality method, and stress method—were employed to identify the top 10 genes from each method. Subsequently, the counts of each gene's occurrences across these methods were aggregated, and the results were visualized using a bar chart. The top 10 genes with the highest frequency of occurrence were statistically identified and designated as hub genes, which are considered crucial for the development of PAH. The findings were then visualized using the Cytoscape software.

2.6. Establishment of a microRNA–gene–TF regulatory network for hub genes

The NetworkAnalyst database (https://www.networkanalyst.ca/)[14] is a comprehensive transcriptome data analysis platform that facilitates gene expression data comparison, quantitative and differential analysis, enrichment analysis, and gene interaction network analysis. The 10 hub genes were input into the NetworkAnalyst database, and both "Gene-miRNA Interactions" and "TF-gene Interactions" functions were selected. For the gene–miRNA interaction data, the "miRTarBase v8.0" dataset was used, while for

the TF–gene interaction data, the "ENCODE" dataset was selected. The default settings were applied, excluding data with peak intensity signals of <500 and predicted regulatory potential scores of <1, using the "BETA Minus" algorithm. The "union" option was selected to merge networks from multiple databases into a single network. In addition, a "Degree Filter" with a cutoff of 1.0 was applied to refine the regulatory network for greater significance.

2.7. Immune infiltration analysis of PAH

The cell-type identification by estimating relative subsets of RNA transcripts (CIBERSORT) algorithm[15] estimates the proportions of 22 immune cell types from tissue sample data using a reference gene expression matrix and linear support vector regression principles. The integrated expression data were analyzed using the CIBERSORT method to estimate the relative proportions of infiltrating immune cells in PAH samples. A *p*-value threshold of <0.05 was applied to filter for suitable samples. The proportion of each immune cell type was calculated for each sample and represented in a bar chart. The "pheatmap," "corrplot," and "vioplot" packages were utilized to compare and visualize the correlation between the levels of 22 immune cell types and immune cell infiltration between PAH and control samples.[16]

2.8. Evaluation of immune dysregulation by single-cell RNA analysis

The single-cell RNA sequencing analysis systematically identifies and characterizes cell populations through key bioinformatics steps. Using the "Seurat" package, 10X Genomics data were processed into Seurat objects, followed by quality control, mitochondrial gene assessment, and normalization. Highly variable genes were selected, and PCA was performed for dimensionality reduction, with visualization through "VizDimLoadings," "DimPlot," and related functions. Clustering was conducted using "Find Neighbors" and "Find Clusters," with t-distributed stochastic neighbor embedding enabling two-dimensional visualization. Marker genes were identified using "FindAllMarkers" and annotated with the "SingleR" package. Core gene expression was visualized using "VlnPlot," "Feature Plot," and "Dot Plot." Lastly, cell communication analysis was performed using "cell chat" to provide a comprehensive interpretation of single-cell data.

2.9. Screening of candidate drugs for the key genes

The Comparative Toxicogenomic Database (CTD)[17,18] compiles toxicological data on chemicals, genes, phenotypes, diseases, and environmental exposures. Using CTD's rich toxicogenomic relationship data, targeted genes, drugs, and diseases can be tightly integrated, leading to the discovery of potential disease drugs. The 10 identified hub genes were imported into CTD and screened for compounds that could reduce the functional effects of these hub genes. Specifically, compounds that could either reduce the expression of overexpressed genes in PAH patients or increase the expression of underexpressed genes were identified, serving as potential targeted drugs for PAH.

2.10. Statistical analysis

All statistical analyses were conducted using R software (version 4.2.2). To account for unequal sample sizes, non-parametric tests, such as the Wilcoxon rank-sum test and Kruskal–Wallis test, were applied for two-group and multi-group comparisons, respectively. For differential expression and regression analyses, linear models with covariate adjustment were used (e.g., via the "limma" package). All tests were two-sided, and $p<0.05$ were considered statistically significant.

3. Results

3.1. Screening of DEGs

Analysis of the merged dataset identified a total of 311 DEGs, comprising 182 upregulated and 129 downregulated genes. The top five upregulated genes included hemoglobin beta, periostin, osteoblast-specific factor (*POSTN*), secreted frizzled-related protein 2, asporin (*ASPN*), and WNT inhibitory factor 1. In addition, PCA plots were included before and after correction to demonstrate the successful removal of batch effects (Figure S1A and B). Conversely, the top five downregulated genes were bactericidal/permeability-increasing fold-containing protein beta 1, interleukin 1 receptor type II, S100 calcium-binding protein A12 (*S100A12*), S100 calcium-binding protein A9 (*S100A9*), and ribonuclease A family member 2, mitochondrial. The distribution of DEGs is visualized through a heatmap and a volcano plot in Figure 1A and B, respectively (Figure S1).

3.2. WGCNA for screening key module genes in PAH

Initially, genes were ranked based on variance, and the top 25% (4759 genes) with the highest variability were selected for further analysis. Subsequently, cluster analysis was conducted to identify and remove outlier samples, applying a threshold of 64. This process detected and excluded two abnormal samples, resulting in the selection of Cluster 1, which contained 98 samples.

To determine the optimal power parameter for network construction, the "pick-Soft-Threshold" function in the WGCNA package was used to evaluate power values ranging from 1 to 20. A β of 6 (scale-free R^2 = 0.8) was selected as the optimal parameter. A threshold of 0.4 was applied to merge similar modules within the clustering tree (Figure 2A), leading to

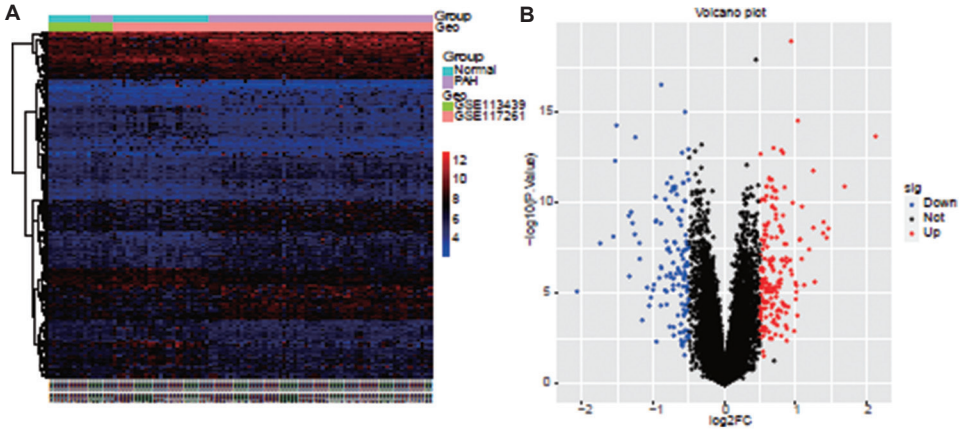

Figure 1. Heatmap and volcano plot of DEGs identified from the integrated dataset. (A) The rows of the heatmap represent DEGs, and the columns represent samples (normal or PAH). The colors reflect upregulation or downregulation of gene expression. (B) Red and blue dots represent upregulated and downregulated DEGs, respectively.
Abbreviations: DEG: Differentially expressed gene; PAH: Pulmonary arterial hypertension.

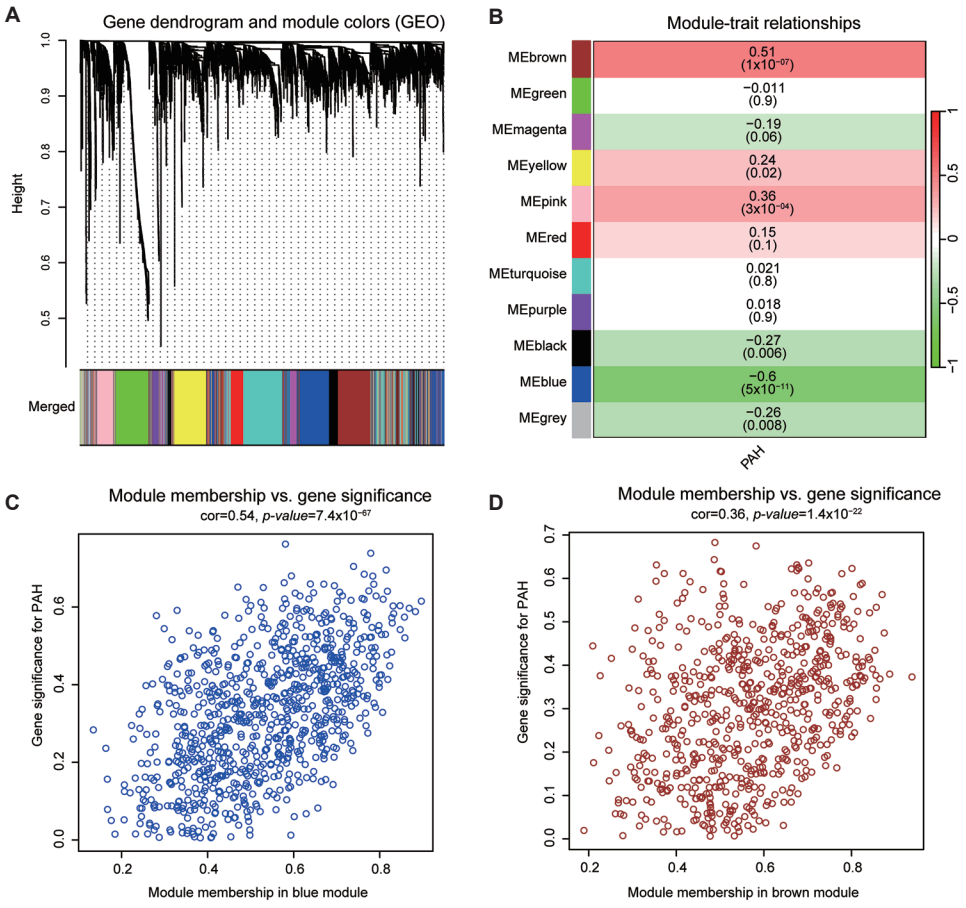

Figure 2. Weighted gene co-expression network analysis. (A) Different colors represent gene modules under the gene tree. (B) Heatmap showing the correlation between modules and PAH. The brown and blue modules are significantly correlated with PAH. The numbers inside and outside the brackets represent *p*-values and correlation coefficients, respectively. (C) Correlation plot between blue module genes and PAH gene significance. (D) Correlation plot between brown module genes and PAH gene significance.
Abbreviations: ME: Module eigengene; PAH: Pulmonary arterial hypertension.

the identification of 11 co-expression modules, each representing a set of genes with similar expression patterns (Figures 2B, S2A and S2B).

Among these, the blue module exhibited the highest negative correlation with PAH (module eigengene [ME]: $r = -0.6$, $p=5\times10^{-11}$), while the brown module showed the highest positive correlation (ME: $r = 0.51$, $p=1\times10^{-7}$). Both modules demonstrated significant associations with PAH compared to other modules (Figure 2B and C). Consequently, the brown module (691 genes) and the blue module (868 genes) were identified as key modules, and their respective genes were retained for further investigation (Tables S2 and S3).

Further analysis assessed the correlation between module membership and gene significance within these key modules. A strong positive correlation was observed in both the brown module (correlation = 0.36, $p=1.4\times10^{-22}$; Figure 2D) and the blue module (correlation = 0.54, $p=7.4\times10^{-67}$; Figure 2C). By intersecting the DEGs from the merged GEO dataset with the brown and blue modules, 30 genes from the brown module and 173 genes from the blue module were identified. Taking the union of both sets, a total of 203 MDEGs were obtained (Figure 3A). Finally, applying the criteria of $p<0.05$ and $|\log_2 FC| > 1$, 24 significantly different MDEGs were identified (Table 1).

3.3. GO and KEGG analysis of the module DEGs

The GO functional enrichment analysis revealed that the 24 MDEGs are predominantly associated with various biological processes, including neutrophil chemotaxis, positive regulation of the inflammatory response, defense response to fungi, acute-phase response, and neutrophil aggregation. In terms of cellular component ontology, these genes were significantly enriched in the extracellular region, extracellular space, extracellular exosome, extracellular matrix, and secretory granule lumen. Molecular function analysis showed that these gene groups are mainly enriched in "RAGE receptor binding," "Toll-like receptor 4 binding," "arachidonic acid binding," "calcium ion binding," and "calcium-dependent protein binding"

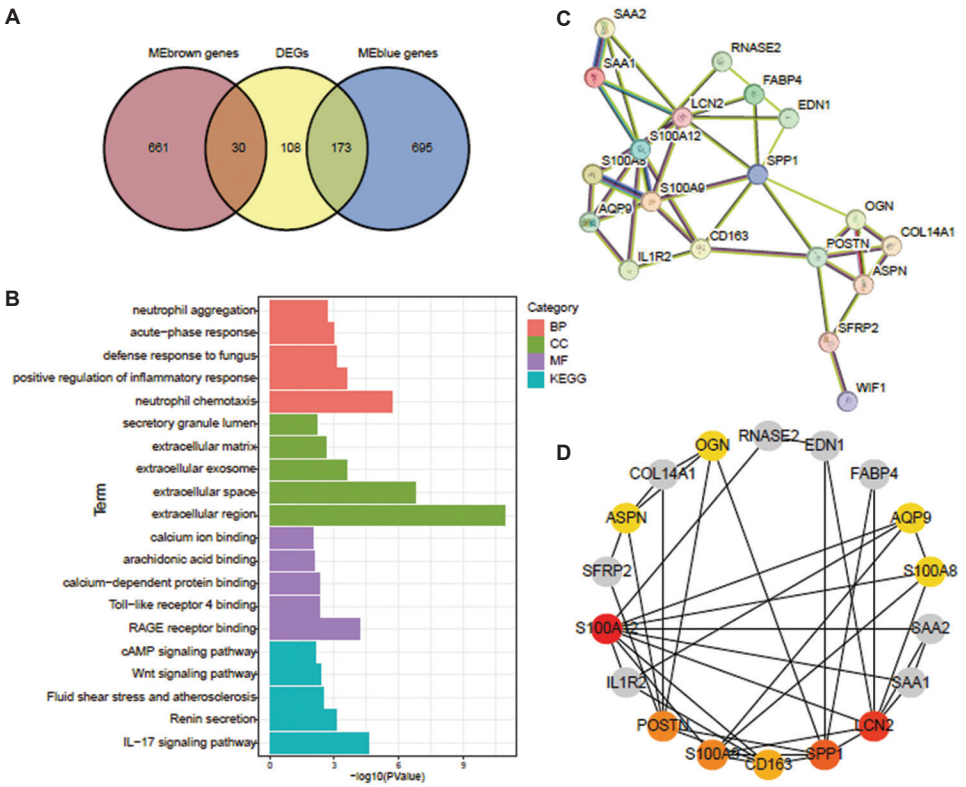

Figure 3. Functional analysis of 24 MDEGs and hub gene selection. (A) Venn diagram showing the overlap between differentially expressed genes (DEGs) and genes in the blue and brown modules. (B) Gene ontology analysis and KEGG pathway analysis of 24 MDEGs. Green bars represent biological processes, red bars represent cellular components, brown lines represent molecular functions, and purple lines represent KEGG pathways. (C) Construction of the MDEG PPI network. (D) Selection of hub genes. The degree of color represents the score of the degree algorithm. The darker the color, the higher the score. Abbreviations: BP: Biological process; CC: Cellular component; KEGG: Kyoto Encyclopedia of Genes and Genomes; MDEG: Module differentially expressed gene; ME: Module eigengene; MF: Molecular function; PAH: Pulmonary arterial hypertension; PPI: Protein–protein interaction.

Table 1. Differentially expressed genes in 24 modules

Gene ID	Log_2FC	Averaged expression	t	*p*-value	Adjusted *p* value	B
IL1R2	−1.747	6.964	−6.138	1.69×10^{-8}	1.31×10^{-6}	9.223
S100A12	−1.561	6.981	−6.327	7.05×10^{-9}	6.73×10^{-7}	10.057
S100A9	−1.535	7.347	−8.318	4.62×10^{-13}	4.88×10^{-10}	19.319
RNASE2	−1.517	5.446	−9.212	5.19×10^{-15}	1.65×10^{-11}	23.637
SAA2	−1.342	4.610	−6.889	5.02×10^{-10}	8.69×10^{-8}	12.594
MGAM	−1.312	5.487	−6.987	3.15×10^{-10}	6.17×10^{-8}	13.043
S100A8	−1.292	9.933	−6.687	1.31×10^{-9}	1.85×10^{-7}	11.670
SAA1	−1.260	4.244	−6.386	5.36×10^{-9}	5.45×10^{-7}	10.321
LOC441081	−1.254	6.931	−8.913	2.34×10^{-14}	5.57×10^{-11}	22.311
LCN2	−1.193	6.230	−6.144	1.64×10^{-8}	1.28×10^{-6}	9.249
AQP9	−1.189	7.745	−5.696	1.23×10^{-7}	6.37×10^{-6}	7.322
SPP1	−1.154	5.325	−3.747	2.99×10^{-4}	2.78×10^{-3}	−0.050
CD163	−1.003	9.106	−4.773	6.20×10^{-6}	1.32×10^{-4}	3.590
EDN1	1.006	7.540	4.475	2.02×10^{-5}	3.36×10^{-4}	2.473
FABP4	1.014	8.153	4.704	8.19×10^{-6}	1.63×10^{-4}	3.327
PDE3A	1.034	7.112	9.330	2.86×10^{-15}	1.09×10^{-11}	24.211
OGN	1.058	7.289	6.237	1.07×10^{-8}	9.53×10^{-7}	9.657
AGBL1	1.067	6.230	6.250	1.01×10^{-8}	9.06×10^{-7}	9.714
ENPP2	1.194	9.377	5.967	3.66×10^{-8}	2.48×10^{-6}	8.479
COL14A1	1.252	7.224	8.063	1.64×10^{-12}	1.49×10^{-9}	18.099
WIF1	1.396	9.565	6.716	1.14×10^{-9}	1.63×10^{-7}	11.805
ASPN	1.440	7.015	6.294	8.20×10^{-9}	7.56×10^{-7}	9.913
SFRP2	1.464	6.555	6.544	2.57×10^{-9}	3.10×10^{-7}	11.024
POSTN	1.693	8.831	7.657	1.22×10^{-11}	6.43×10^{-9}	16.180

Abbreviation: FC: Fold change.

(Table S4). In addition, the KEGG pathways of MDEGs were found to be enriched in "IL-17 signaling pathway," "Renin secretion," "Fluid shear stress and atherosclerosis," "Wnt signaling pathway," and "cAMP signaling pathway" (Figure 3B and Table S5).

3.4. Protein–protein interaction network construction and hub gene identification of the module DEGs

The STRING database screening revealed that the MDEGs' PPI network consisted of 23 nodes and 40 edges, with an average node degree of 3.48. After removing isolated nodes, the final MDEGs' PPI network comprises 19 interconnected genes (Figure 3C and Table S6). The network included eight upregulated genes and 11 downregulated genes. Furthermore, through further analysis using Cytohubba, 10 hub genes were obtained, including the upregulated genes *POSTN*, *OGN*, *ASPN*, and the downregulated genes *S100A12*, *LCN2*, *SPP1*, *S100A9*, *CD163*, *S100A8*, and *AQP9* (Figure 3D and Table S7).

3.5. Regulatory network of hub genes: microRNA–gene–TF interactions

Using the NetworkAnalyst database, a giant miRNA-gene-TF regulatory network with 89 nodes and 91 edges was constructed for the 10 hub genes. In this network, *S100A12* was associated with 45 TFs, and several TFs were simultaneously associated with multiple hub genes, including *S100A12*, *OGN*, *LCN2*, and *S100A8*. In addition, *has-mir-335-5p* was associated with both *ASPN* and *OGN*, while *has-mir-146a-5p* was related to *OGN* and *S100A8*. The findings indicate that these hub genes are regulated by multiple miRNAs and TFs, and there are close connections among multiple hub genes (Figure 4).

3.6. Immune infiltration analysis of PAH

The CIBERSORT algorithm was employed to predict and analyze immune cell infiltration patterns in PAH tissues compared to normal tissues. The distribution of

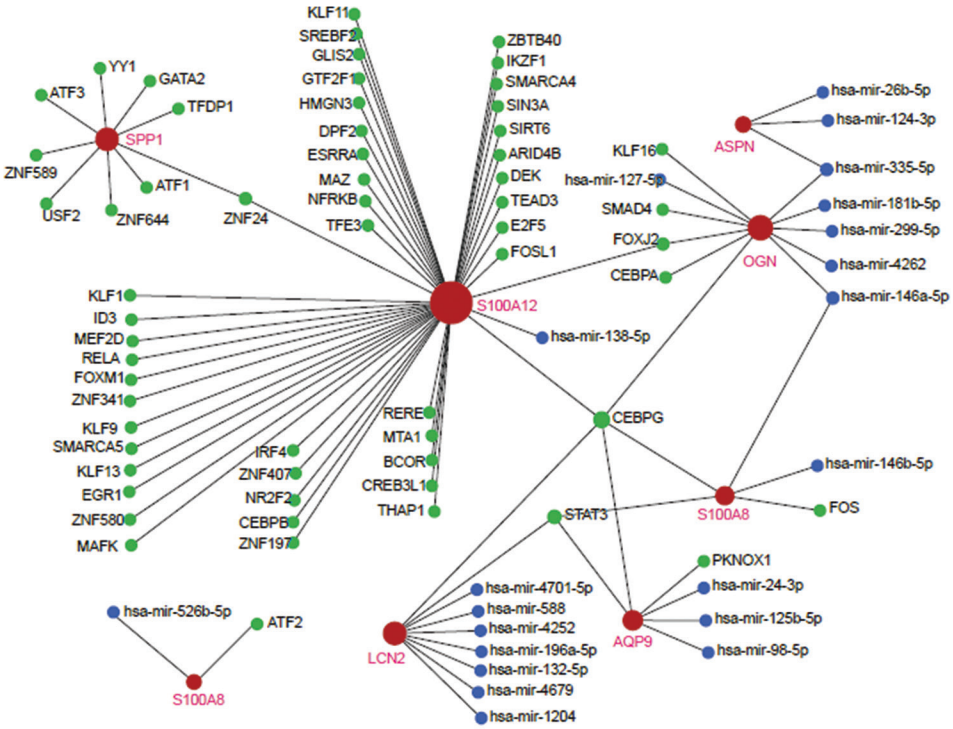

Figure 4. The microRNA–gene–transcription factor regulatory network. Red spheres represent genes, blue spheres represent predicted miRNAs, and green spheres represent transcription factors.

22 immune cell subtypes across samples is illustrated in Figure 5A and B. While the overall composition of immune cells was relatively consistent across samples, macrophage M2 and resting CD4⁺ memory T-cells were found to be predominant subsets. Further analysis using a violin plot revealed distinct differences in immune cell infiltration between PAH and control tissues. Specifically, PAH tissues exhibited elevated levels of resting CD4⁺ memory T-cells, activated dendritic cells, and resting mast cells, whereas naïve CD4⁺ T-cells, resting natural killer (NK) cells, monocytes, and neutrophils were significantly reduced compared to controls.

Correlation analysis of immune cell interactions (Figure 6A and B) demonstrated a negative correlation between M0 macrophages and resting CD4⁺ memory T-cells ($r = -0.43$), as well as between resting mast cells and neutrophils ($r = -0.53$). In contrast, a positive correlation was observed between resting mast cells and resting NK cells ($r = 0.46$). In addition, γδ T-cells showed positive correlations with both M1 macrophages ($r = 0.49$) and CD8⁺ T-cells ($r = 0.51$). These findings highlight significant disparities in immune cell composition between PAH and normal tissues, suggesting that monocytes, dendritic cells, neutrophils, resting CD4⁺ memory T-cells, and macrophages may play pivotal roles in the pathogenesis and progression of PAH.

3.7. Single-cell data of immune dysregulation in PAH

Single-cell RNA analysis of PAH revealed the cellular composition and clustering in three different samples (GSM5206779, GSM5206780, and GSM5206781), indicating a well-mixed integration of datasets (Figure 7A). Clustering of cells into distinct populations, labeled from 0 to 21, suggests heterogeneity within the dataset (Figure 7B). Whereas cell-type annotations revealed diverse immune and stromal populations, including T cells, macrophages, endothelial cells, monocytes, fibroblasts, epithelial cells, NK cells, smooth muscle cells, and common myeloid progenitors (CMPs). Notably, macrophages and monocytes formed large clusters, indicating their significant presence in the dataset, potentially playing key roles in disease pathology (Figure 7C). The cell–cell interaction showed the communication between various immune and stromal cell types. Both networks highlighted strong interactions between macrophages, monocytes, fibroblasts, and endothelial cells, suggesting their central role in the microenvironment. Macrophages exhibited prominent connections with multiple cell types, indicating their key role in immune regulation and inflammation. The presence of dense interactions among smooth muscle cells, T cells, and CMPs further suggests their involvement in disease progression

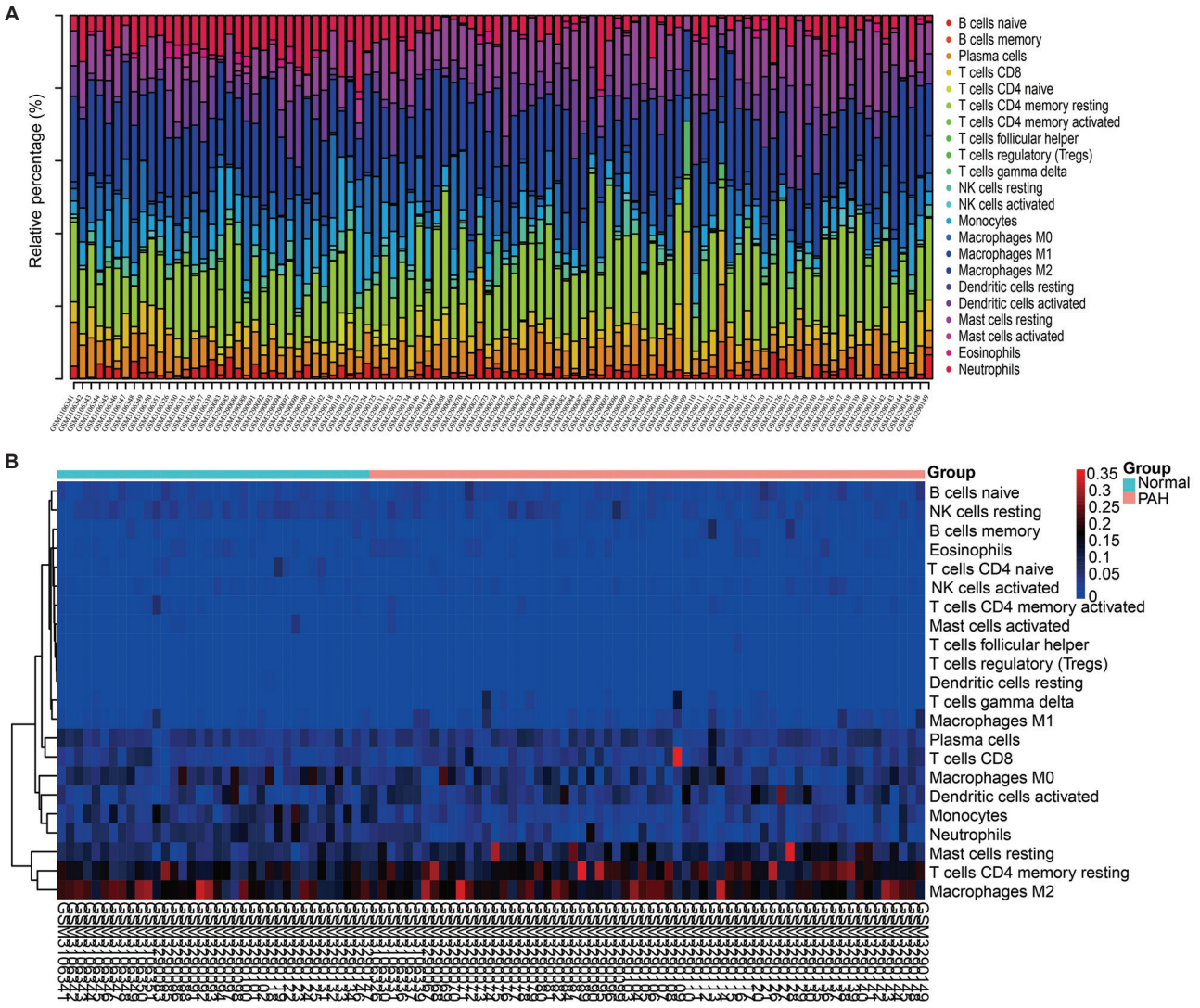

Figure 5. Overall immune cell infiltration between pulmonary arterial hypertension patient tissues and normal tissues. (A) Relative percentages of 22 immune cell types. (B) Heatmap of 22 immune cell types.

(Figure 7D and E). The expression of key macrophage-associated genes, including *SPP1*, *S100A8/A9/A12*, *CD163*, *POSTN*, and *AQP9*, highlights their role in PAH by promoting vascular inflammation, endothelial dysfunction, and fibrosis. Specifically, *SPP1* and *POSTN* contribute to vascular remodeling and extracellular matrix deposition, *S100A8/A9/A12* drive inflammatory responses linked to endothelial injury, *CD163*, a marker of M2 macrophages, is associated with fibrosis and immune regulation, whereas *AQP9* and *LCN2* influence macrophage migration and vascular endothelial activation (Figure 7F). These findings underscore the importance of macrophage-driven immune dysregulation in PAH pathogenesis, providing potential targets for therapeutic intervention.

3.8. Screening of potential candidate drugs for treating PAH

Across the online CTD database, targeted compounds were predicted based on the 10 hub genes screened for PAH. Compounds that could reduce the functional expression of overexpressed hub genes or increase the expression of underexpressed hub genes were selected. After screening, compounds were ranked based on the number of hub genes they affected. Compounds that affected ≥5 hub genes included retinol, sodium arsenite, arsenic trioxide, activated Vitamin D, and nickel. These compounds, which have an impact on the expression of hub genes, may provide assistance in the development of PAH treatment drugs and could potentially be effective candidate drugs for PAH (Table 2).

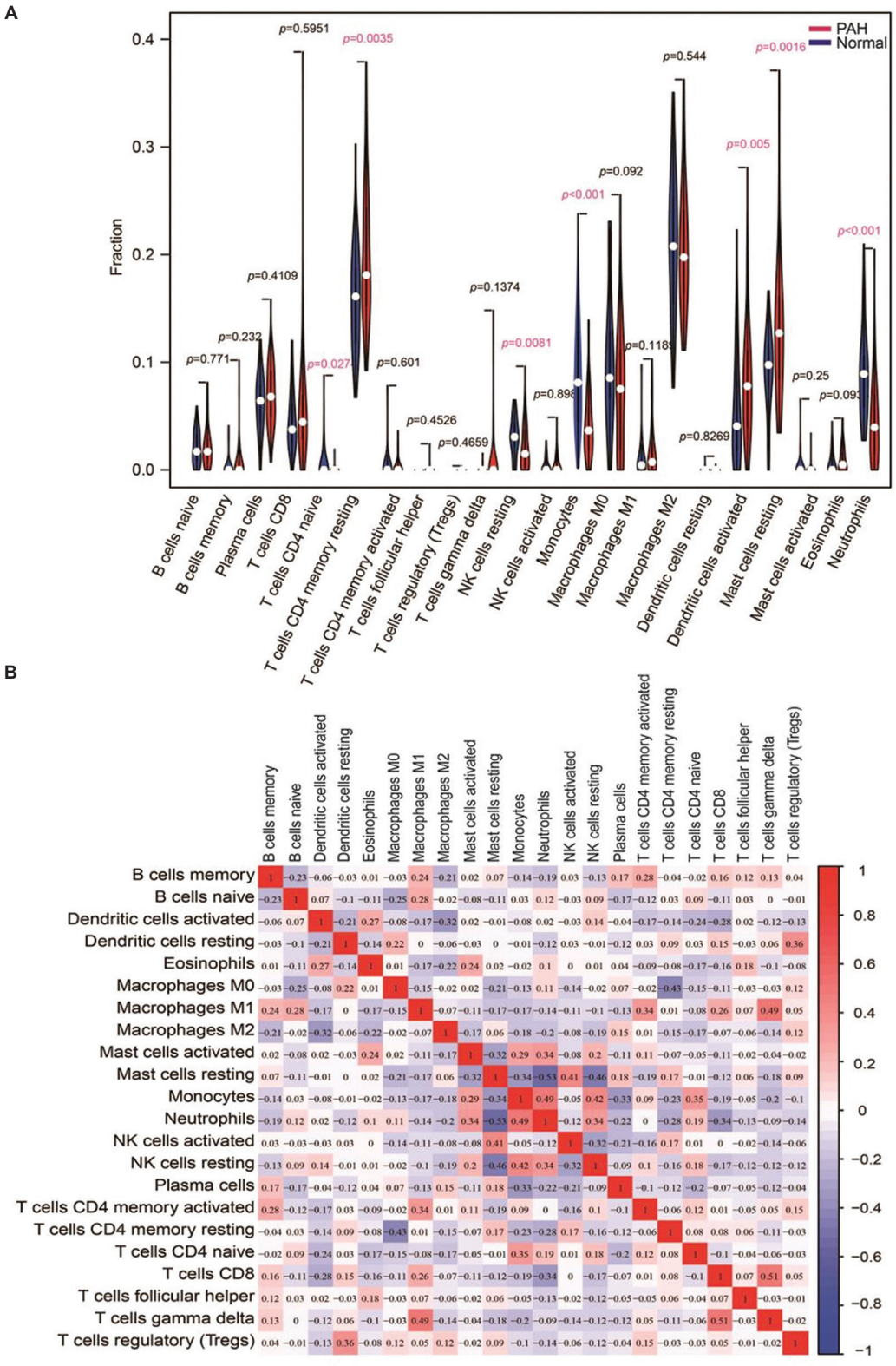

Figure 6. Comparative analysis of immune cell infiltration between pulmonary arterial hypertension (PAH) patient tissues and normal tissues. (A) Comparison of differences in 22 immune cell subtypes between PAH patient tissues and controls. (B) Correlation matrix composed of all 22 immune cell subtypes. Both the horizontal and vertical axes represent immune cell subtypes.

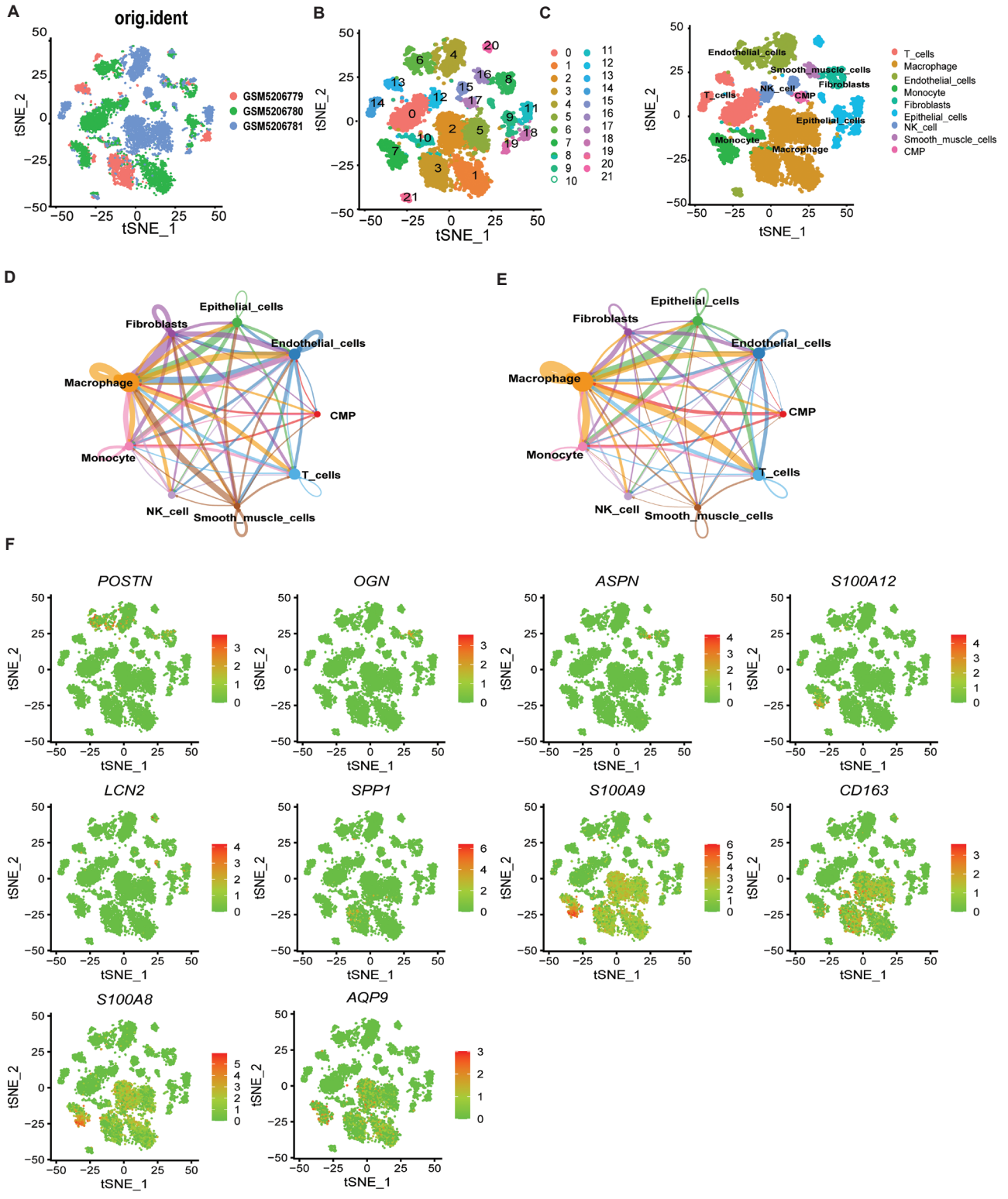

Figure 7. Characterization of cellular composition and marker gene expression in single-cell clusters. (A) The t-distributed stochastic neighbor embedding (t-SNE) plot showing single-cell clustering by sample identity (GSM5206779, GSM5206780, GSM5206781). (B) t-SNE visualization of major cell types: T cells, monocytes, macrophages, epithelial cells, endothelial cells, natural killer cells, fibroblasts, smooth muscle cells, and common myeloid progenitors. (C) t-SNE plot of inferred cell–cell interactions among cell types. (D) t-SNE plot depicting interaction and (E) strength between cell types. (F) Expression patterns of key marker genes (*AQP9*, *POSTN*, *OGN*, *ASPN*, *S100A12*, *LCN2*, *SPP1*, *S100A9*, *CD163*, and *S100A8*) across cell populations, with color intensity representing expression levels.

Table 2. Compounds related to the 10 hub genes, as identified through the Comparative Toxicogenomic Database

Compound name	Compound ID	Compound effect	
		Increase gene expression	Decrease gene expression
Tretinoin	D014212	*AQP9*, *S100A12*, *S100A9*, *S100A8*, *LCN2*, *SPP1*	-
Sodium arsenite	C017947	*S100A12*, *SPP1*, *CD163*, *S100A9*, *S100A8*, *LCN2*	-
Arsenic trioxide	D000077237	*AQP9*, *S100A9*, *S100A8*, *LCN2*	*POSTN*, *OGN*
Calcitriol	D002117	*S100A9*, *S100A8*, *LCN2*, *SPP1*	*ASPN*
Nickel	D009532	*S100A9*, *S100A8*, *LCN2*, *SPP1*, *CD163*	*POSTN*
Estradiol	D004958	*S100A9*, *S100A8*, *LCN2*, *SPP1*	*POSTN*
Ozone	D010126	*S100A12*, *S100A8*, *LCN2*, *SPP1*, *CD163*	-
Tetrachlorodibenzodioxin	D013749	*S100A12*, *S100A9*, *S100A8*, *LCN2*, *SPP1*	-
Tobacco smoke pollution	D014028	*S100A12*, *S100A9*, *S100A8*, *LCN2*	*POSTN*
Valproic acid	D014635	*S100A9*, *S100A8*, *LCN2*, *SPP1*	*ASPN*
Cyclosporine	D016572	*AQP9*, *S100A9*, *LCN2*, *SPP1*	*OGN*

4. Discussion

PAH is a severe and progressive cardiopulmonary disorder; yet, its underlying mechanisms remain poorly understood. Previous gene expression analyses have typically relied on individual datasets, leading to limited sample sizes and a narrow scope for identifying DEGs. To address these limitations, this study integrated two PAH datasets, encompassing a total of 100 samples, thereby enhancing the statistical power of the analysis.

By combining WGCNA with DEG screening, 24 MDEGs were identified, and they were further refined to uncover 10 key genes associated with PAH. Among these, *POSTN*, *OGN*, and *ASPN* were upregulated, whereas *S100A12*, *LCN2*, *SPP1*, *S100A9*, *CD163*, *S100A8*, and *AQP9* were downregulated. Several of these key genes play significant roles in immune responses and inflammatory pathways. *S100A12*, *S100A8*, and *S100A9* encode calcium-, zinc-, and copper-binding proteins that contribute to inflammatory processes and interact with immune cells, including mast cells and monocytes.[19,20] These proteins, particularly S100A8 and S100A9, are known to drive neutrophil chemotaxis and adhesion.[21] Notably, S100A12 has been identified as a marker for poor cardiac output and mortality in PAH patients, underscoring its prognostic significance, although its precise molecular mechanism remains unclear.[22] In addition, the relationship between S100A8 and T-cell infiltration in PAH tissue warrants further investigation.[23] Interestingly, these proteins are also implicated in the pathophysiology of PAH in patients with SARS-CoV-2 infection, suggesting a potential link between viral infections and PAH progression.[24] Lipocalin 2 (LCN2), also known as neutrophil gelatinase-associated lipocalin, is a member of the lipocalin family involved in apoptosis, innate immunity, and the pathogenesis of kidney disease and heart failure.[25-27] Emerging evidence suggests that LCN2 may serve as a serum biomarker for PAH, particularly in female patients.[28] However, additional studies are required to validate its clinical utility and mechanistic role in PAH.

Among the upregulated genes, *POSTN*, a secreted extracellular matrix protein, has been strongly linked to PAH pathogenesis.[29-31] Increased periostin (POSTN) expression has been observed in lung tissue and serum samples from PAH patients, aligning with previous reports.[32] Recent research highlights its role in vascular remodeling and suggests that targeting POSTN could represent a novel therapeutic strategy for PAH.[33] Further elucidation of its regulatory mechanisms may provide new avenues for disease management. Similarly, osteoglycin (OGN), a small leucine-rich proteoglycan family member, is known to suppress tumor proliferation and invasion in breast and bladder cancers.[34,35] In colorectal cancer, OGN has been shown to modulate immune infiltration by inhibiting vascular endothelial growth factor and enhancing T-cell recruitment.[35] However, its role in PAH-related immune infiltration remains largely unexplored, and further research is needed to determine whether it influences PAH progression through immune modulation.

Another upregulated gene, *ASPN*, encodes an extracellular matrix protein associated with diverse pathological conditions, including osteoarthritis, intervertebral disc disease, and cancer.[36,37] Although its role in PAH remains poorly characterized, studies have reported strong correlations between ASPN expression and immune cell infiltration, particularly in T-cells, B-cells, and fibroblasts.[38] A negative association has also been observed with cytotoxic lymphocytes and NK cells.[39] These findings suggest that

ASPN may contribute to PAH pathogenesis through immune cell regulation, warranting further investigation.

Among the downregulated genes, CD163 is a macrophage-specific scavenger receptor involved in hemoglobin–haptoglobin complex clearance and anti-inflammatory responses. As an acute-phase regulatory receptor, CD163 is considered a marker of selectively activated macrophages.[40,41] Some studies suggest that CD163 may be associated with poor prognosis in PAH, yet additional research is required to confirm its role in disease progression.[42,43]

Secreted phosphoprotein 1 (SPP1), a chemokine-rich matrix phosphoglycerate, plays a critical role in PAH by promoting pulmonary vascular smooth muscle cell proliferation.[44,45] Its expression levels correlate with disease severity, reinforcing its significance in PAH pathology.[46] These findings indicate that SPP1 could serve as a valuable biomarker and therapeutic target. Lastly, aquaporin 9 (AQP9), an aquaporin family member, is implicated in immune cell migration and regulation, particularly in neutrophils, leukocytes, dendritic cells, macrophages, and monocytes.[47] While database analyses suggest a potential role for AQP9 in PAH, experimental validation is necessary to establish its functional relevance in disease pathogenesis.[48]

In addition to gene expression analysis, this study also examined immune cell infiltration in PAH tissues. Significant differences were observed between PAH and normal tissues in various immune cell populations, including monocytes, dendritic cells, neutrophils, resting $CD4^+$ memory T-cells, and macrophages. Given the involvement of key genes in immune regulation, these findings suggest that multiple immune cell types play a pivotal role in PAH progression.

Furthermore, potential therapeutic compounds targeting key PAH-related genes were also explored. The analysis identified retinol, arsenic trioxide, and activated vitamin D as potential regulators of these genes, suggesting their potential therapeutic applications in PAH treatment. However, further experimental studies are necessary to confirm their efficacy and mechanisms of action.

5. Conclusion

This study utilized various bioinformatics analysis methods, including DEG and WGCNA analysis, to identify key genes associated with PAH pathogenesis. Through PPI, miRNA–gene–TF network construction, and functional enrichment analyses, the study further explored the regulatory mechanisms of these genes. Immune infiltration analysis revealed significant differences in monocytes, dendritic cells, neutrophils, resting $CD4^+$ memory T cells, and macrophages between PAH and normal tissues. Notably, macrophage-associated genes *SPP1*, *S100A8/A9/A12*, *CD163*, *POSTN*, *AQP9*, and *LCN2* were linked to vascular inflammation, endothelial dysfunction, and fibrosis, emphasizing their role in immune-driven vascular remodeling in PAH. Finally, targeted drug prediction identified compounds, such as retinol, arsenic trioxide, and active Vitamin D as potential regulators of these key genes. These findings provide novel insights into the molecular mechanisms of PAH, highlighting macrophage-related immune dysregulation as a critical factor in disease progression and offering promising biomarkers and therapeutic targets for future research.

Acknowledgments

None.

Funding

The study was funded by the Cultivation Project for Innovation Team in Teachers' Teaching Proficiency by Zhengzhou Health College (No. 2024jxcxtd01).

Conflict of interest

Xin-Ying Ji is an Associate Editor and Umair Ali Khan Saddozai is an Editorial Board Member of this journal, but were not in any way involved in the editorial and peer-review process conducted for this paper, directly or indirectly. Separately, other authors declared that they have no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

Author contributions

Conceptualization: Zhen-Dong Lu, Umair Ali Khan Saddozai
Data curation: Zhen-Dong Lu, Umair Ali Khan Saddozai
Formal analysis: Zhen-Dong Lu, Umair Ali Khan Saddozai
Funding acquisition: Xin-Ying Ji
Methodology: Zhen-Dong Lu, Xin-Ying Ji, Umair Ali Khan Saddozai
Project administration: Xin-Ying Ji, Umair Ali Khan Saddozai
Software: Zhen-Dong Lu, Umair Ali Khan Saddozai
Supervision: Xin-Ying Ji
Validation: Zhen-Dong Lu, Umair Ali Khan Saddozai
Writing – original draft: Zhen-Dong Lu, Umair Ali Khan Saddozai
Writing – review & editing: Zhi-Liang Jiang, Wahab Hussain, Xin-Ying Ji

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data

All data supporting the findings of this study are included within the article and its supplementary materials.

References

1. Wu W, Chen A, Lin S, *et al*. The identification and verification of hub genes associated with pulmonary arterial hypertension using weighted gene co-expression network analysis. *BMC Pulm Med*. 2022;22(1):474.

   doi: 10.1186/s12890-022-02275-6

2. Zhang JR, Ouyang X, Hou C, *et al*. Natural ingredients from Chinese materia medica for pulmonary hypertension. *Chin J Nat Med*. 2021;19(11):801-814.

   doi: 10.1016/s1875-5364(21)60092-4

3. Bisserier M, Pradhan N, Hadri L. Current and emerging therapeutic approaches to pulmonary hypertension. *Rev Cardiovasc Med*. 2020;21(2):163-179.

   doi: 10.31083/j.rcm.2020.02.597

4. Edwards AL, Gunningham SP, Clare GC, *et al.* Professional killer cell deficiencies and decreased survival in pulmonary arterial hypertension. *Respirology*. 2013;18(8):1271-1277.

   doi: 10.1111/resp.12152

5. van Uden D, Koudstaal T, van Hulst JAC, *et al*. Peripheral blood T cells of patients with IPAH have a reduced cytokine-producing capacity. *Int J Mol Sci*. 2022;23(12):6508.

   doi: 10.3390/ijms23126508

6. Meng X, Yang J, Dong M, *et al*. Regulatory T cells in cardiovascular diseases. *Nat Rev Cardiol*. 2016;13(3):167-179.

   doi: 10.1038/nrcardio.2015.169

7. Yan T, Zhu S, Zhu M, Wang C, Guo C. Integrative identification of hub genes associated with immune cells in atrial fibrillation using weighted gene correlation network analysis. *Front Cardiovasc Med*. 2020;7:631775.

   doi: 10.3389/fcvm.2020.631775

8. Irizarry RA, Hobbs B, Collin F, *et al*. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics*. 2003;4(2):249-264.

   doi: 10.1093/biostatistics/4.2.249

9. Langfelder P, Horvath S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics*. 2008;9:559.

   doi: 10.1186/1471-2105-9-559

10. Huang W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 2009;4(1):44-57.

    doi: 10.1038/nprot.2008.211

11. Wu J, Mao X, Cai T, Luo J, Wei L. KOBAS server: A web-based platform for automated annotation and pathway identification. *Nucleic Acids Res*. 2006;34(Web Server issue):W720-W724.

    doi: 10.1093/nar/gkl167

12. Szklarczyk D, Franceschini A, Wyder S, *et al*. STRING v10: Protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res*. 2015;43(Database issue):D447-D452.

    doi: 10.1093/nar/gku1003

13. Shannon P, Markiel A, Ozier O, *et al*. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13(11):2498-2504.

    doi: 10.1101/gr.1239303

14. Zhou G, Soufan O, Ewald J, Hancock REW, Basu N, Xia J. NetworkAnalyst 3.0: A visual analytics platform for comprehensive gene expression profiling and meta-analysis. *Nucleic Acids Res*. 2019;47(W1):W234-W241.

    doi: 10.1093/nar/gkz240

15. Newman AM, Liu CL, Green MR, *et al*. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*. 2015;12(5):453-457.

    doi: 10.1038/nmeth.3337

16. Hu K. Become competent within one day in generating boxplots and violin plots for a novice without prior R experience. *Methods Protoc*. 2020;3(4):64.

    doi: 10.3390/mps3040064

17. Davis AP, Grondin CJ, Johnson RJ, *et al*. Comparative toxicogenomics database (CTD): Update 2021. *Nucleic Acids Res*. 2021;49(D1):D1138-D1143.

    doi: 10.1093/nar/gkaa891

18. Singh AV, Bhardwaj P, Laux P, *et al*. AI and ML-based risk assessment of chemicals: Predicting carcinogenic risk from chemical-induced genomic instability. *Front Toxicol*. 2024;6:1461587.

    doi: 10.3389/ftox.2024.1461587

19. Yang Z, Yan WX, Cai H, *et al*. S100A12 provokes mast cell activation: A potential amplification pathway in asthma and innate immunity. *J Allergy Clin Immunol*. 2007;119(1):106-114.

    doi: 10.1016/j.jaci.2006.08.021

20. Yan WX, Armishaw C, Goyette J, *et al*. Mast cell and monocyte recruitment by S100A12 and its hinge domain.

J Biol Chem. 2008;283(19):13035-13043.

doi: 10.1074/jbc.M710388200

21. Ryckman C, Vandal K, Rouleau P, Talbot M, Tessier PA. Proinflammatory activities of S100: Proteins S100A8, S100A9, and S100A8/A9 induce neutrophil chemotaxis and adhesion. *J Immunol*. 2003;170(6):3233-3242.

doi: 10.4049/jimmunol.170.6.3233

22. Tzouvelekis A, Herazo-Maya JD, Ryu C, *et al*. S100A12 as a marker of worse cardiac output and mortality in pulmonary hypertension. *Respirology*. 2018;23(8):771-779.

doi: 10.1111/resp.13302

23. Liu Y, Shi JZ, Jiang R, *et al*. Regulatory T cell-related gene indicators in pulmonary hypertension. *Front Pharmacol*. 2022;13:908783.

doi: 10.3389/fphar.2022.908783

24. Taz TA, Ahmed K, Paul BK, Al-Zahrani FA, Mahmud SMH, Moni MA. Identification of biomarkers and pathways for the SARS-CoV-2 infections that make complexities in pulmonary arterial hypertension patients. *Brief Bioinform*. 2021;22(2):1451-1465.

doi: 10.1093/bib/bbab026

25. Yang J, Goetz D, Li JY, *et al*. An iron delivery pathway mediated by a lipocalin. *Mol Cell*. 2002;10(5):1045-1056.

doi: 10.1016/s1097-2765(02)00710-4

26. Shields-Cutler RR, Crowley JR, Miller CD, Stapleton AE, Cui W, Henderson JP. Human metabolome-derived cofactors are required for the antibacterial activity of siderocalin in urine. *J Biol Chem*. 2016;291(50):25901-25910.

doi: 10.1074/jbc.M116.759183

27. Bao G, Clifton M, Hoette TM, *et al*. Iron traffics in circulation bound to a siderocalin (Ngal)-catechol complex. *Nat Chem Biol*. 2010;6(8):602-609.

doi: 10.1038/nchembio.402

28. Li C, Zhang Z, Xu Q, Wu T, Shi R. Potential mechanisms and serum biomarkers involved in sex differences in pulmonary arterial hypertension. *Medicine*. 2020;99(13):e19612.

doi: 10.1097/md.0000000000019612

29. Snider P, Hinton RB, Moreno-Rodriguez RA, *et al*. Periostin is required for maturation and extracellular matrix stabilization of noncardiomyocyte lineages of the heart. *Circ Res*. 2008;102(7):752-760.

doi: 10.1161/circresaha.107.159517

30. Seki M, Furukawa N, Koitabashi N, *et al*. Periostin-expressing cell-specific transforming growth factor-β inhibition in pulmonary artery prevents pulmonary arterial hypertension. *PLoS One*. 2019;14(8):e0220795.

doi: 10.1371/journal.pone.0220795

31. Kim BR, Yoon JW, Choi H, Kim D, Kang S, Kim JH. Application of periostin peptide-decorated self-assembled protein cage nanoparticles for therapeutic angiogenesis. *BMB Rep*. 2022;55(4):175-180.

doi: 10.5483/BMBRep.2022.55.4.137

32. Yoshida T, Nagaoka T, Nagata Y, *et al*. Periostin-related progression of different types of experimental pulmonary hypertension: A role for M2 macrophage and FGF-2 signalling. *Respirology*. 2022;27(7):529-538.

doi: 10.1111/resp.14249

33. Lee D, Lee H, Jo HN, *et al*. Endothelial periostin regulates vascular remodeling by promoting endothelial dysfunction in pulmonary arterial hypertension. *Anim Cells Syst*. 2024;28(1):1-14.

doi: 10.1080/19768354.2023.2300437

34. Liang X, Gao J, Wang Q, Hou S, Wu C. ECRG4 represses cell proliferation and invasiveness via NFIC/OGN/NF-κB signaling pathway in bladder cancer. *Front Genet*. 2020;11:846.

doi: 10.3389/fgene.2020.00846

35. Xu T, Zhang R, Dong M, *et al*. Osteoglycin (OGN) inhibits cell proliferation and invasiveness in breast cancer via PI3K/Akt/mTOR signaling pathway. *Onco Targets Ther*. 2019;12:10639-10650.

doi: 10.2147/ott.S222967

36. Rochette A, Boufaied N, Scarlata E, *et al*. Asporin is a stromally expressed marker associated with prostate cancer progression. *Br J Cancer*. 2017;116(6):775-784.

doi: 10.1038/bjc.2017.15

37. Li H, Yang HH, Sun ZG, Tang HB, Min JK. Whole-transcriptome sequencing of knee joint cartilage from osteoarthritis patients. *Bone Joint Res*. 2019;8(7):290-303.

doi: 10.1302/2046-3758.87.Bjr-2018-0297.R1

38. Dai B, Ding L, Zhao L, Zhu H, Luo H. Contributions of immune cells and stromal cells to the pathogenesis of systemic sclerosis: Recent insights. *Front Pharmacol*. 2022;13:826839.

doi: 10.3389/fphar.2022.826839

39. Wang L, Sun J. ASPN Is a potential biomarker and associated with immune infiltration in endometriosis. *Genes (Basel)*. 2022;13(8):1352.

doi: 10.3390/genes13081352

40. Koudouna A, Gkioka AI, Gkiokas A, *et al*. Serum-soluble CD163 levels as a prognostic biomarker in patients with diffuse large B-cell lymphoma treated with chemoimmunotherapy. *Int J Mol Sci*. 2024;25(5):2862.

doi: 10.3390/ijms25052862

41. Møller HJ. Soluble CD163. *Scand J Clin Lab Invest*. 2012;72(1):1-13.

doi: 10.3109/00365513.2011.626868

42. Lekva T, Gullestad L, Broch K, Aukrust P, Andreassen AK, Ueland T. Distinct patterns of soluble leukocyte activation markers are associated with etiology and outcomes in precapillary pulmonary hypertension. *Sci Rep.* 2020;10(1):18540.

doi: 10.1038/s41598-020-75654-w

43. D'Addario CA, Lanier GM, Jacob C, *et al.* Differences in the expression of DNA methyltransferases and demethylases in leukocytes and the severity of pulmonary arterial hypertension between ethnic groups. *Physiol Rep.* 2022;10(10):e15282.

doi: 10.14814/phy2.15282

44. Meng L, Liu X, Teng X, *et al*. Osteopontin plays important roles in pulmonary arterial hypertension induced by systemic-to-pulmonary shunt. *FASEB J*. 2019;33(6):7236-7251.

doi: 10.1096/fj.201802121RR

45. Mura M, Cecchini MJ, Joseph M, Granton JT. Osteopontin lung gene expression is a marker of disease severity in pulmonary arterial hypertension. *Respirology*. 2019;24(11):1104-1110.

doi: 10.1111/resp.13557

46. Hoshikawa Y, Matsuda Y, Sakuma M, Kondo T. Potential therapeutic target for pulmonary arterial hypertension--osteopontin. *Nihon rinsho Jpn J Clin Med*. 2008;66(11):2097-2101.

47. da Silva IV, Soveral G. Aquaporins in immune cells and inflammation: New targets for drug development. *Int J Mol Sci*. 2021;22(4):1845.

doi: 10.3390/ijms22041845

48. da Silva IV, Garra S, Calamita G, Soveral G. The multifaceted role of aquaporin-9 in health and its potential as a clinical biomarker. *Biomolecules*. 2022;12(7):897.

doi: 10.3390/biom12070897