Original Research *Highlight article*

Weighted gene co-expression network analysis reveals the hub genes associated with pulmonary hypertension

Shengyan Wang[1,2], Dejun Sun[2], Chuanchuan Liu[3], Yong Guo[1], Jie Ma[4], Ri-li Ge[1] and Sen Cui[4]

[1]Research Center for High Altitude Medicine, Key Laboratory of High Altitude Medicine (Ministry of Education), Key Laboratory of Application and Foundation for High Altitude Medicine Research in Qinghai Province (Qinghai-Utah Joint Research Key Lab for High Altitude Medicine), Qinghai University, Xining 810000, China; [2]Department of Clinical Medicine, Qinghai Institute of Health Sciences, Xining 810000, China; [3]Key Laboratory of Echinococcosis, Qinghai University Affiliated Hospital, Xining 810000, China; [4]Department of Hematology, Qinghai Institute of Hematology, Qinghai University Affiliated Hospital, Xining 810000, China
Corresponding author: Sen Cui. Email: cuisen4366@yeah.net

Impact Statement

Pulmonary hypertension (PH) is a prevalent cardiovascular disease that affects over 100 million people globally. Early diagnosis and timely treatment of PH are critical. Current treatment mainly relieves symptoms but cannot reverse pulmonary vascular reconstruction. PH has a complex and little understood pathogenesis, and research is needed to identify genetic biomarkers that can be used for effective targeted therapy. To achieve this, we used data from the Gene Expression Omnibus database to perform weighted gene co-expression network analysis and identified eight PH hub genes (*DDB1*, *EFTUD2*, *mTOR*, *PSMD2*, *RBM8A*, *SMARCA4*, *TARDBP*, and *UBXN7*) with diagnostic and predictive value, which were mainly correlated with macrophages. *TARDBP* had the best diagnostic potential. We validated these findings using animal experiments. This study illustrates the pathogenesis of PH and provides new insights that may be utilized to discover possible biomarker genes and treatment targets.

Abstract

Pulmonary hypertension (PH) is a cardiopulmonary vascular disease that acutely endangers human health and can be fatal. It progresses rapidly and has a high mortality rate. Its pathophysiology is complicated and still not completely elucidated; therefore, achieving treatment breakthroughs are difficult. In this study, data from 58 normal controls and 135 patients with PH were extracted from the GSE24988, GSE113439, and GSE117261 datasets in the Gene Expression Omnibus (GEO) database and screened for differentially expressed genes (DEGs). In addition, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were performed. Weighted gene co-expression network analysis (WGCNA) was used to identify the key modules and hub genes associated with PH. Eight PH-associated hub genes were identified. Furthermore, correlation analysis between immune cell infiltration and hub genes was performed, and the receiver operating characteristic (ROC) curves showed that *TARDBP* had the best diagnostic efficacy. Moreover, a rat hypoxic pulmonary hypertension (HPH) model was generated, and the expression of hub genes in the lungs and pulmonary arteries of HPH rats was verified using western blotting assays. Our results showed that *mTOR*, *PSMD2*, *RBM8A*, *SMARCA4*, *TARDBP*, and *UBXN7* were highly expressed in the lungs. In addition, *EFTUD2*, *mTOR*, *RBM8A*, *SMARCA4*, *TARDBP*, and *UBXN7* were significantly upregulated, whereas *DDB1* was significantly downregulated in the pulmonary arteries of HPH rats compared with those of controls. In conclusion, we identified PH hub genes with diagnostic and predictive value by performing WGCNA on data from the GEO database. Furthermore, we provided novel insights of PH that might be utilized to evaluate potential biomarker genes and therapeutic targets.

Keywords: Pulmonary hypertension, weighted gene co-expression network analysis, hub genes, integrated bioinformatics analysis

***Experimental Biology and Medicine* 2023; 248: 217–231. DOI: 10.1177/15353702221147557**

Introduction

Pulmonary hypertension (PH) is a common cardiovascular illness associated with high pulmonary vascular pressure, pulmonary arteriole contraction, and pulmonary vascular remodeling.[1] Aggravation of the disease can cause right heart failure or even death. There are approximately 100 million patients with PH globally, and the disease has an incidence rate of 0.015.[2] and a survival rate of 61.2% over 5 years.[3] Due to continuous improvements in diagnosis and follow-up strategies, patients can now benefit from treatment measures. However, even with regular treatment, patient mortality is still only 38–63%.[4] Early diagnosis and timely treatment may prevent or reverse the progression of the disease. Considering that more than 140 million people live in altitudes higher than 2500 m above sea level around

the world today,[5] the complications caused by exposure in plateau areas constitute a huge public health burden. When the human body is exposed to hypobaric and hypoxic environments for a long time, the functions of various organs become disordered, and a series of pathophysiological changes occur in the cardiovascular system, including pulmonary vasoconstriction, reconstruction, and increased vascular resistance, causing functional and organic changes in pulmonary arterioles and resulting in hypoxic pulmonary hypertension (HPH).[6] PH has a complex pathogenesis. Pulmonary vascular remodeling is the main factor and is associated with endothelial cell dysfunction, inflammation, immune response, and germline mutation.[7,8] The pathogenesis of PH is currently unclear, therefore, it is imperative to identify the key PH target genes, explore their function and molecular mechanism, and develop relevant targeted drugs.

Gene Expression Omnibus (GEO) is the largest and most comprehensive resource for gene expression data. Weighted gene co-expression network analysis (WGCNA) can divide the gene co-expression network into numerous highly correlated feature modules, it links modules with certain clinical characteristics to find genes with important functions and helps identify potential mechanisms underlying specific biological processes (BPs).[9] Differentially expressed genes (DEGs) are often analyzed when studying diseases. DEG analysis uses genomic data to identify genes associated with diseases. This study combined WGCNA with standardized analysis of DEGs to identify the most relevant PH modules. Hub genes with diagnostic and predictive value were also identified, providing new insights for identifying candidate biomarker genes and therapeutic targets.

Three PH datasets were selected for analysis from the GEO database and DEGs were identified in patients with PH. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were implemented on the DEGs. WGCNA was used to screen for key modules linked to the disease phenotype, and protein–protein interaction (PPI) networks were utilized to find hub genes. The correlation between hub genes and immune cell infiltration was investigated. Finally, a rat HPH model was constructed, and the accuracy of the bioinformatic analysis was verified using western blotting analysis.

Materials and methods

Data analysis

The GSE24988, GSE113439, and GSE117261 datasets were selected from the GEO database (https://www.ncbi.nlm.nih.gov/gds/). Human gene expression profiles were obtained from data in the three datasets. The GSE24988 dataset has 17 patients with severe PH, 45 patients with moderate PH, and 22 normal controls (NCs); the GSE113439 dataset has 15 patients with PH and 11 NCs; and the GSE117261 dataset has 58 patients with PH and 25 NCs. We used the "limma" package based on R software (version 4.1.1) to identify genes that showed differential expression between PH and NCs in the three datasets ($P < 0.05$ and $|\log2 \text{ fold change (FC)}| > 0$). DEGs were visualized using volcano plots. Intersections between DEGs in GSE24988, GSE113439, and GSE117261 datasets were determined using Venn diagrams (https://bioinformatics.psb.ugent.be/webtools/Venn/). Genes that were upregulated or downregulated in at least two datasets were chosen for further investigation.

Functional enrichment analysis of DEGs

We used the R software package "clusterProfiler" to analyze the functional enrichment of DEGs, and genes that were differentially expressed in PH were functionally annotated using enriched GO terms associated with BPs.[10] KEGG analysis identified important pathways of gene enrichment,[11] and networks of molecular interactions and relationships were constructed.

Screening key modules associated with disease phenotypes using WGCNA

The R package "WGCNA" was used to divide the gene network into different modules based on expression similarities to identify hub genes. Using intersection genes identified in the GSE117261 dataset, samples were clustered to identify outliers using an appropriate soft threshold ($\beta = 4$) that was chosen to ensure the best fit to a scale-free network. Correlations between input genes were calculated, a similarity matrix was constructed, and average link hierarchical clustering was performed based on the obtained dissimilarity matrix. Module genes associated with PH were then identified and their correlation with clinical features was measured. Functional modules in the co-expression network were further investigated by correlating module features with clinical features. For following investigation, modules that were closely related to clinical characteristics were chosen. Heatmaps were created using the R software package "ggplot2." KEGG analysis of key modules (MEblue module and MEbrown module) was performed using the "clusterProfiler" package.

PPI network construction and hub gene identification

The string database (https://string-db.org/) was used to analyze the relationship between gene interactions.[12] Key modules were combined to build a PPI network between DEGs, and the interaction between protein functions was analyzed. The Cytohubba plugin in the Cytoscape software (version 3.8.2) was employed to find hub genes. Five algorithms (betweenness, closeness, degree, radiance, and stress) were employed to locate the top 15 hub genes and to analyze the intersections between the hub genes. Eight hub genes in all were identified. Correlations between the eight hub genes were analyzed using data from 58 patients with PH in the GSE117261 dataset.

Correlation between hub genes and immune cell infiltration

The R software package "cibersoft" was used to evaluate the degree of immune cell infiltration in 58 PH tissues in the GSE117261 dataset. Hub genes were investigated for associations with immune cell infiltration, and significant correlations ($P < 0.05$) were plotted using the "ggplot2" package.

Correlation analysis and diagnostic efficacy of hub genes

The correlation analysis between hub genes and all genes was determined using data from 58 PH tissues in the GSE117261 dataset. Single-gene gene set enrichment analysis (GSEA) was implemented based on the Reactome pathway using the "clusterProfiler" package. Receiver operating characteristic (ROC) curves showed diagnostic efficacies of eight hub genes in the three datasets.

Construction of a rat model of HPH

The Animal Protection and Ethics Committee of the Affiliated Hospital of Qinghai University authorized this study (approval number SL-2022-059). Healthy male Sprague-Dawley (SD; mean body weight: 160 ± 20 g) rats were used in the experiment and were provided by Beijing Weitong Lihua Laboratory Animal Technology Co., Ltd (Beijing, China). A total of 16 rats were separated into two groups (HPH and control) at random. The HPH rat model was established as previously described,[5] and rats in the HPH group were bred continuously for 4 weeks in a hypobaric and hypoxic chamber (DYC-300, Guizhou Feng Lei Oxygen Chamber Co., Ltd., Guizhou, China) equivalent to 4500 m above sea level. Right heart catheterization was performed to measure the mean pulmonary artery pressure (mPAP), and the rats were weighed and anesthetized with 1% pentobarbital sodium at 40 mg/kg body weight. Then, a 16-channel physiological recorder (MP150, Biopac Systems, Inc., USA), and an irrigated PE-10 catheter (0.28 mm, BPE-T10, Solomon, USA) with 0.1% heparin were connected. The rats were fixed, the external jugular vein was exposed and distally ligated, and the proximal end was clamped with a hemostatic clip. The PE-10 catheter was placed at the same level as the heart, a small opening was cut, and the catheter was gently placed into the heart chamber, the PE-10 catheter was rotated to make it enter the pulmonary artery, and observed whether it was a PAP waveform. The PE-10 catheter was quickly secured, and the waveform was observed and recorded using the 16-channel physiological recorder. Next, the rats were euthanized through cervical dislocation, their hearts were extracted and weighed, the right ventricular (RV) free wall and the left ventricle with interventricular septum (LV + S) was separated and weighed, and RV/(LV + S) was calculated to obtain the right ventricular hypertrophy index (RVHI). The right ventricular mass index (RVMI) was calculated as follows: right ventricular weight/body weight (RV/BW). Rat pulmonary arteries and lung tissues were extracted and immediately stored in liquid nitrogen.

Western blotting

Total protein was extracted from pulmonary arteries and lung tissues. Protein concentration was measured using the bicinchoninic acid (BCA) method. Beta III Tubulin (Cat. No. ab53623), GAPDH (Cat No. ab8245), TARDBP (Cat. No. ab133547), SMARCA4 (Cat. No. ab110641), PSMD2 (Cat. No. ab26078), mTOR (Cat. No. ab32028), EFTUD2 (Cat. No. ab188327), DDB1 (Cat. No. ab97522), and Y14 (Cat. No. ab229573) antibodies (diluted 1: 1000) were supplied by Abcam (Cambridge, MA, USA). UBXN7 antibody (Cat. No. NBP2-47589; diluted 1: 1000) was obtained from Novus Biologicals (Littleton, CO, USA). Image J software was used to calculate relative protein expression.

Statistical analysis

For statistical analysis, SPSS 26.0 was utilized. The data are presented as mean ± standard deviation, and $P < 0.05$ was considered statistically significant.

Results

Identification of DEGs

PH was taken as the case group in the differential expression analysis. We identified 2741 upregulated and 3093 downregulated genes in the GSE24988 dataset (Figure 1(A)), and 5207 upregulated and 8040 downregulated genes in the GSE113439 dataset (Figure 1(B)), and 250 upregulated and 426 downregulated genes in the GSE117261 dataset (Figure 1(C)). Volcano plots were generated for all genes. A total of 456 genes that were simultaneously upregulated (Figure 1(D)) and 1073 genes that were simultaneously downregulated (Figure 1(E)) in at least two datasets were selected for subsequent analysis.

DEGs functional enrichment analysis

GO functional annotation results showed that DEGs were primarily associated with RNA splicing, pattern specification, regionalization, anterior/posterior pattern specification, transesterification reactions, and localization of protein-containing complexes (Figure 2(A)). KEGG pathway enrichment analysis demonstrated that the DEGs were primarily involved in human immunodeficiency virus 1 infection, herpes simplex virus 1 infection, maturity onset diabetes of the young, ubiquitin-mediated proteolysis, and cAMP signaling pathway (Figure 2(B)). This shows that PH is associated with complex BPs and is regulated via multiple pathways.

Screening key modules associated with disease phenotypes using WGCNA

WGCNA was used to identify gene modules that were substantially linked with PH. Hierarchical clustering was performed to detect outliers and determine optimal weighting coefficient ($\beta = 4$, $R^2 = 0.9$; Figure 3(A) to (C)). The cluster dendrogram of the modules is presented in Figure 3(D). Gene expression similarity was evaluated based on topological overlap matrix (TOM), and the co-expression module was constructed. Correlations between key modules and clinical features were measured by determining the correlation between clinical features and each color module. MEblue module ($r = -0.38$, $P = 3E\text{-}04$) and MEbrown module ($r = 0.29$, $P = 0.007$) had the highest correlation with PH. Therefore, the genes in these two modules were chosen for further investigation (Figure 4(A)). However, there was no correlation between sex and either module, indicating that sex has little influence on PH. A heatmap was used to visualize similarities between gene expression; the higher the similarity, the

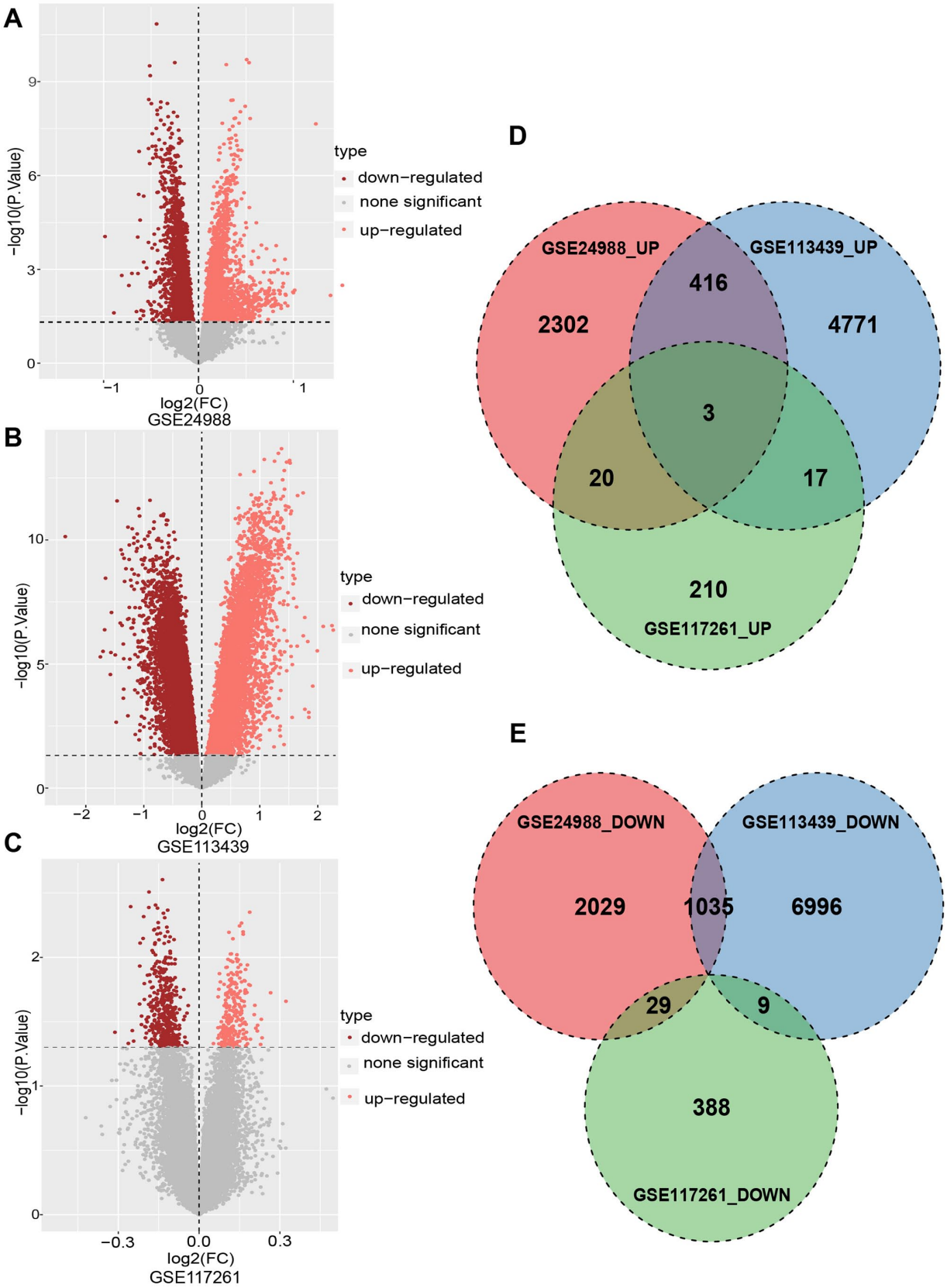

Figure 1. Identification of genes that are differentially expressed between pulmonary hypertension (PH) and normal control (NC) samples. Volcano plots of differentially expressed genes (DEGs) in (A) GSE24988, (B) GSE113439, and (C) GSE117261 datasets in the gene expression omnibus database. Dark red color indicates elevated genes, light red indicates downregulated genes, and black indicates genes whose expression was not significant. (D) The intersection of upregulated genes is depicted by a Venn diagram in the differential expression analysis. (E) The intersection of downregulated genes is depicted by a Venn diagram in the differential expression analysis.

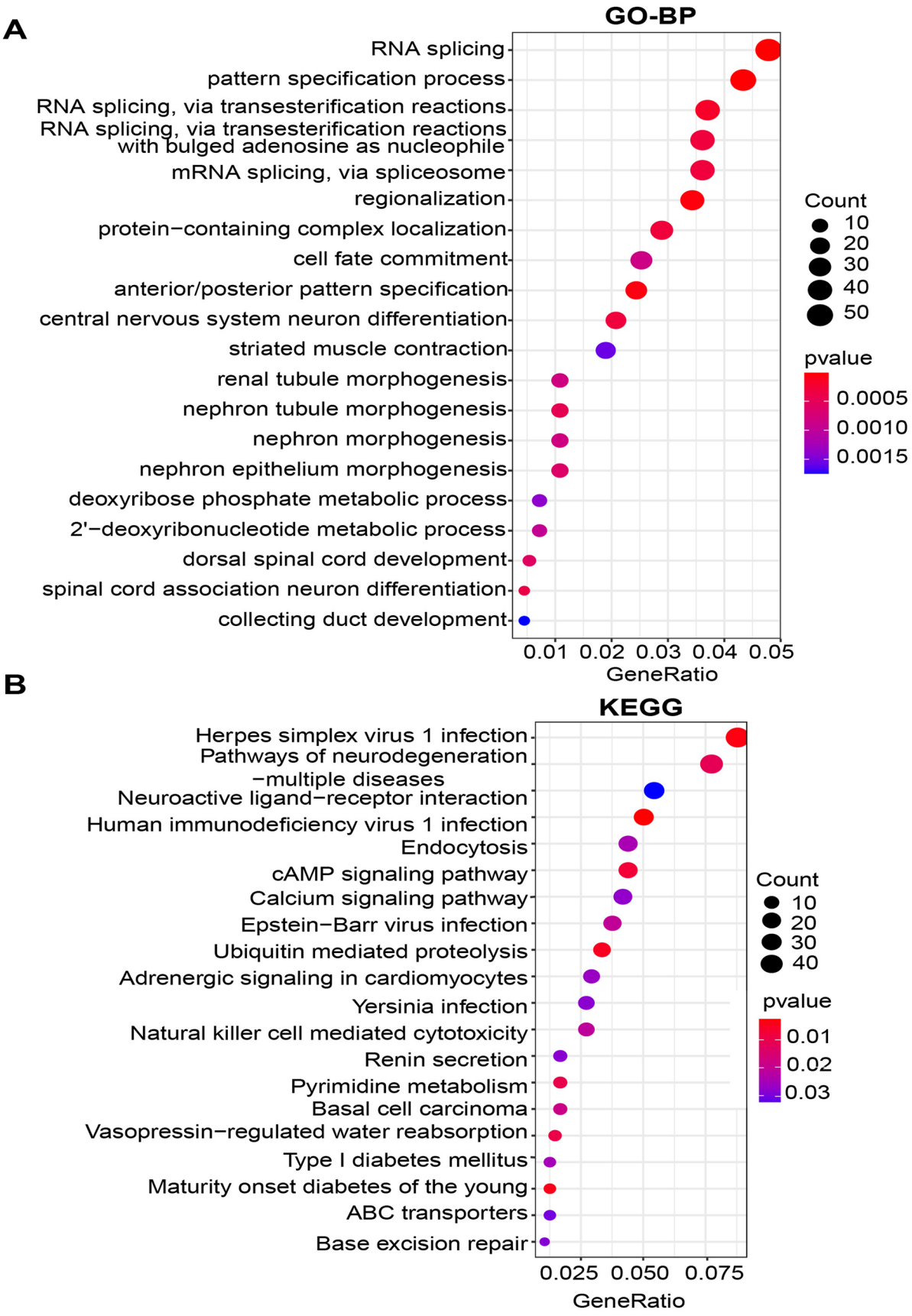

Figure 2. Enrichment analysis of differentially expressed genes (DEGs). (A) Top 20 enriched GO terms in biological process (BP). (B) Top 20 KEGG pathway enrichment analysis results.

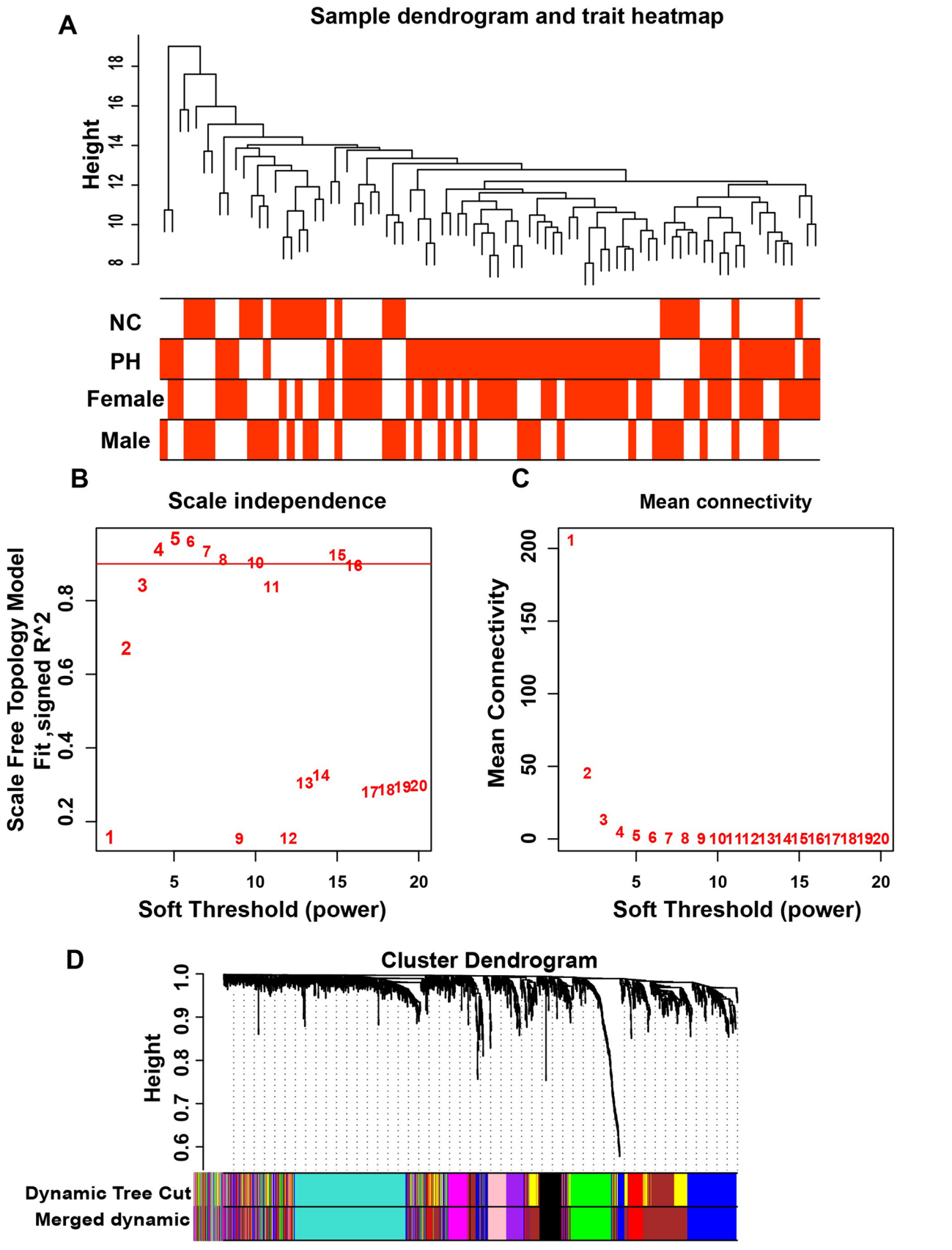

Figure 3. Sample clustering and network creation of weighted co-expressed genes. (A) Distribution of samples and clinical characteristics. (B) and (C) Analysis of scale-free topological fit index and connectivity. Soft threshold powers ($\beta = 4$, $R^2 = 0.9$) were selected to ensure the best fit to a scale-free network. (D) Cluster dendrogram. Gene expression similarity was evaluated based on topological overlap matrix (TOM). Different colors represent different co-expression modules and one branch represents one gene.

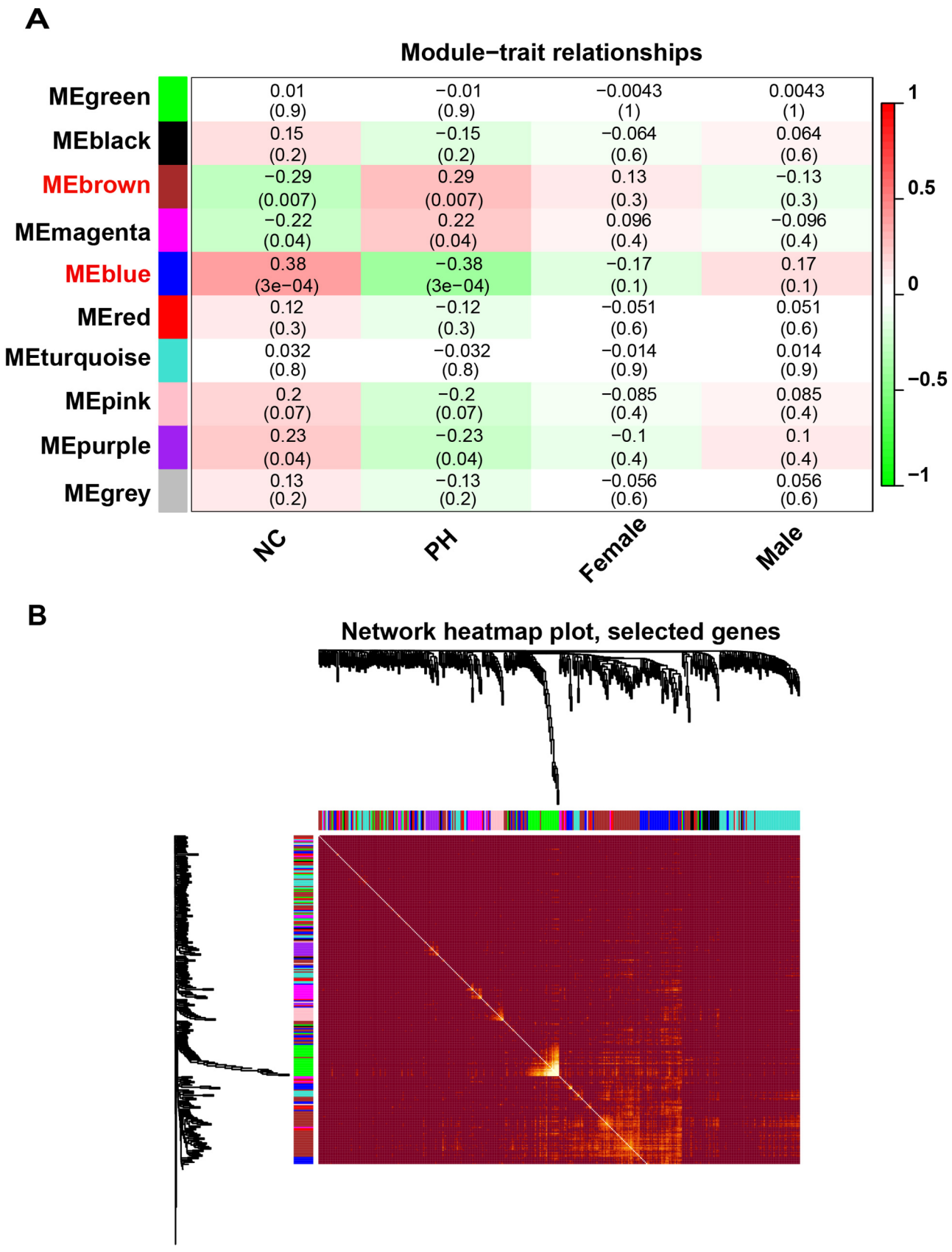

Figure 4. Key modules associated with disease phenotypes were screened. (A) Correlation between gene modules and pulmonary hypertension–associated clinical traits. Each module is represented by a different color, and the correlation coefficients and *P* values are displayed for the cells. Red shows positive correlations, blue shows negative correlations; the darker the hue, the greater the correlation. MEblue and MEbrown modules have the strongest correlations. (B) Heatmap showing similarities between gene expression; the higher the similarity, the brighter the color. Genes in the same module have a higher co-expression pattern, and different modules show significant differences.

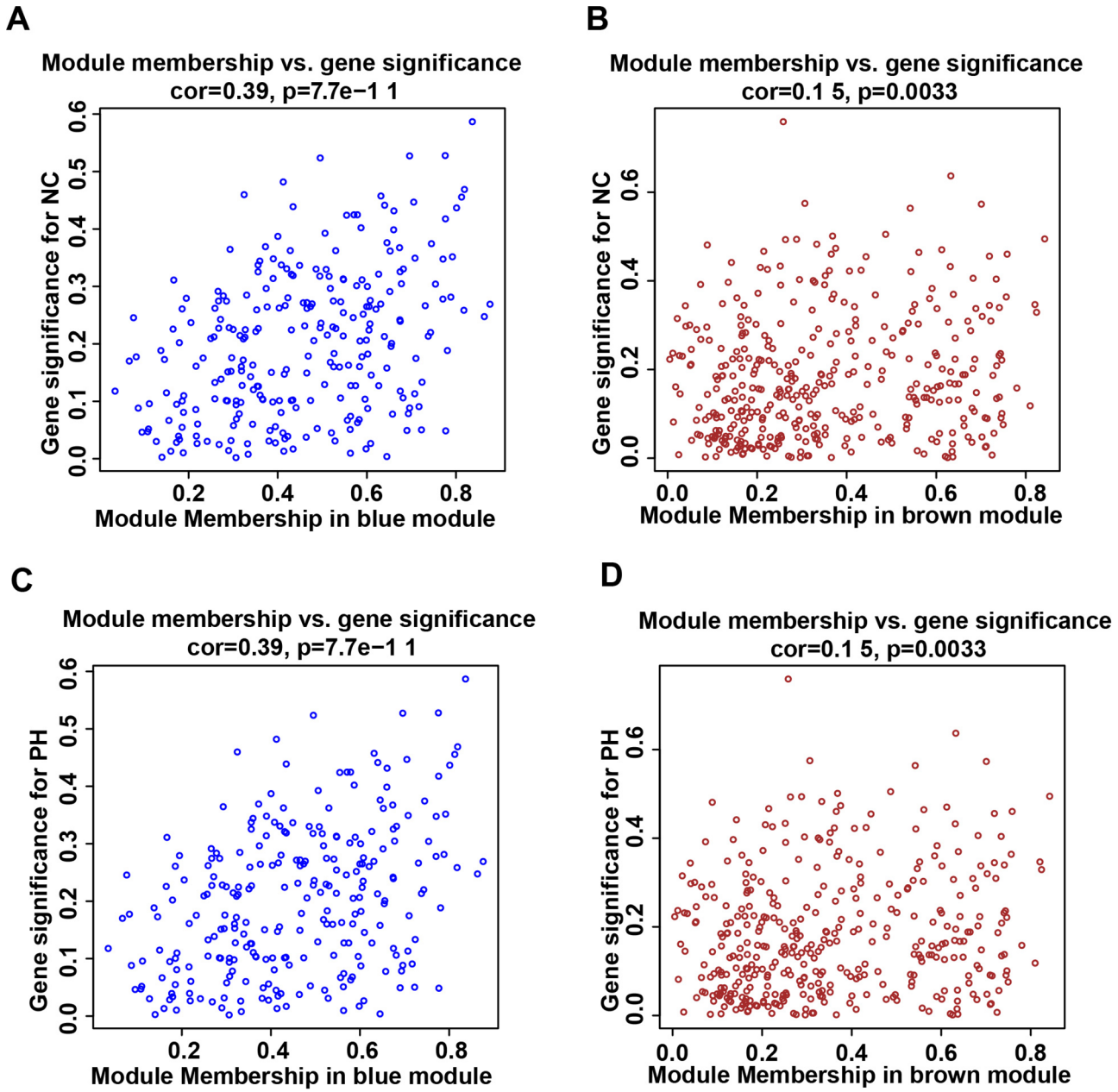

Figure 5. Correlations between key modules and clinical characteristics of NC and PH. (A) Correlation between MEblue module genes and normal control character. (B) Correlation between MEbrown module genes and normal control character. (C) Correlation between MEblue module genes and pulmonary hypertension character. (D) Correlation between MEbrown module gene and pulmonary hypertension character.

brighter the color (Figure 4(B)). Correlations between key modules and clinical characteristics of NC and PH were also demonstrated (Figure 5(A) to (D)).

Enrichment analysis of key modules

KEGG analysis was conducted using genes in MEblue and MEbrown modules, and the main pathways involved in PH were identified. Analysis of MEblue module (Figure 6(A)) showed that Fc epsilon RI signaling pathway, hepatocellular carcinoma, neurodegeneration pathways (associated with multiple diseases), and repair of nucleotide excision are involved in PH. Analysis of the MEbrown module (Figure 6(B)) showed that herpes simplex virus 1 infection, human immunodeficiency virus 1 infection, spliceosome, and nucleocytoplasmic transport are involved in PH. Merging MEblue and MEbrown modules for KEGG pathway enrichment analysis (Figure 6(C)) revealed that human immunodeficiency virus 1 infection, herpes simplex virus 1 infection, spliceosome, yersinia infection, and ubiquitin-mediated proteolysis are involved in PH.

PPI network construction and hub genes identification

A PPI network was created by combining MEblue and MEbrown modules. Cytoscape software was used to display the network. Top 15 hub genes were determined using five

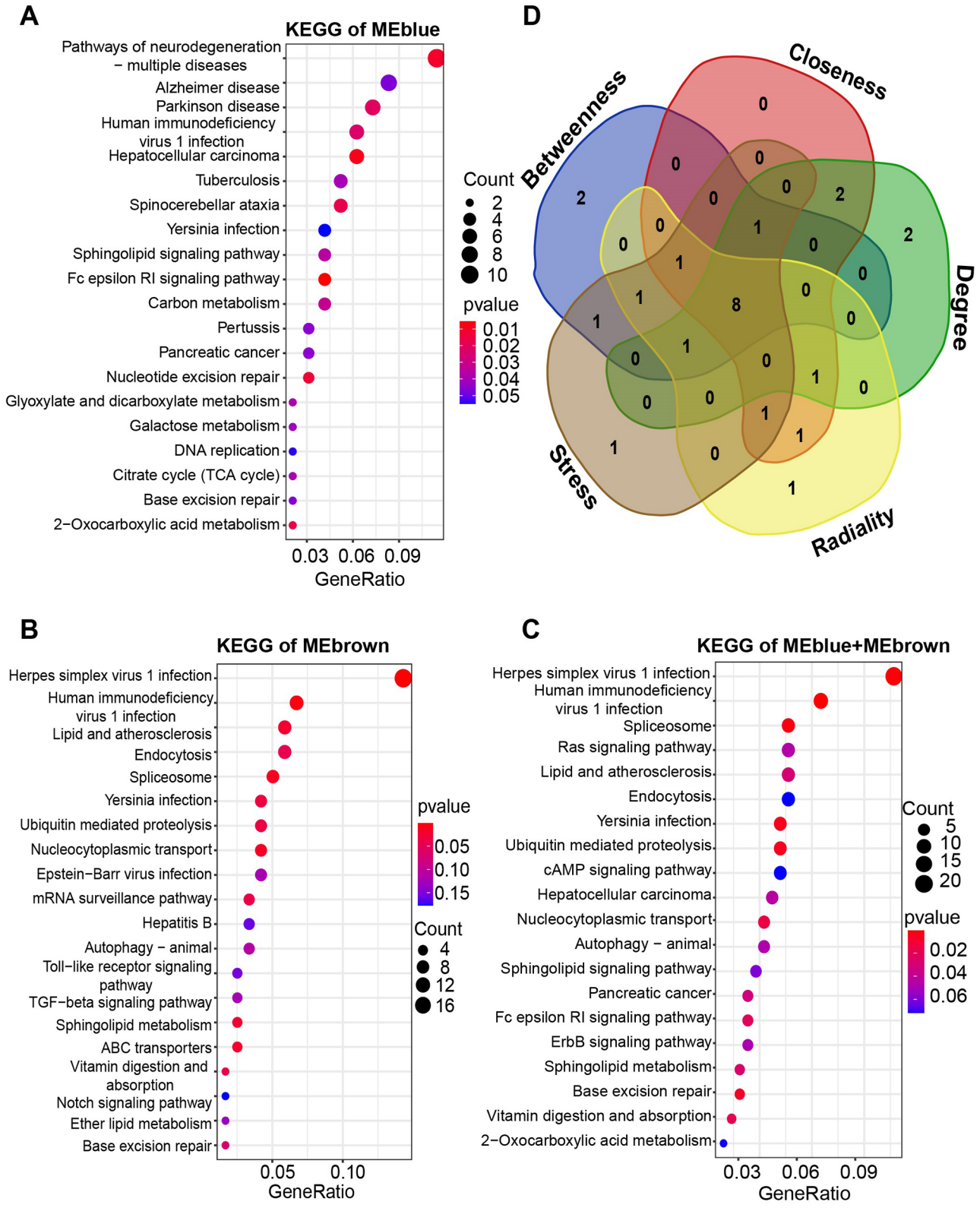

Figure 6. Analysis of the enrichment of key module genes and identification of hub genes. (A) Analysis of enriched KEGG pathways associated with MEblue module genes. (B) Analysis of enriched KEGG pathways associated with MEbrown module genes. (C) Analysis of enriched KEGG pathways in both MEblue and MEbrown modules. (D) Intersection of the hub genes screened using the five algorithms (betweenness, closeness, degree, radiance, and stress). A total of eight hub genes were obtained.

algorithms, and genes in the intersection were confirmed. The intersection of the hub genes screened using the five algorithms identified a total of eight hub genes (Figure 6(D)).

Correlation between hub genes and immune cell infiltration

Correlations among eight hub genes (*DDB1*, *EFTUD2*, *mTOR*, *PSMD2*, *RBM8A*, *SMARCA4*, *TARDBP*, and *UBXN7*) were measured using data from 58 patients with PH in the GSE117261 dataset (Figure 7(A)). The infiltration level of immune cells was evaluated using data from 58 PH tissues using GSE117261 dataset (Supplementary Tables 1 to 8). Seven hub genes (*DDB1*, *EFTUD2*, *mTOR*, *PSMD2*, *RBM8A*, *SMARCA4*, and *UBXN7*) were correlated with immune cell infiltration ($P < 0.05$). *DDB1* was positively correlated with M0 macrophages ($r = 0.417$, $P = 0.001$) and M1 macrophages ($r = 0.335$, $P = 0.013$), *EFTUD2* was positively correlated with M0 macrophages ($r = 0.324$, $P = 0.001$) and M1 macrophages ($r = 0.269$, $P = 0.041$), *mTOR* was positively correlated with M1 macrophages ($r = 0.474$, $P < 0.001$), *PSMD2* was positively correlated with M1 macrophages ($r = 0.260$, $P = 0.048$) and negatively correlated with mast cells resting ($r = –0.430$, $P < 0.001$), *RBM8A* was negatively correlated with M0 macrophages ($r = –0.354$, $P = 0.006$), *SMARCA4* was positively correlated with M0 macrophages ($r = 0.459$, $P < 0.001$) and negatively correlated with CD8T cells ($r = –0.304$, $P = 0.02$), and *UBXN7* was positively correlated with resting CD4 memory T cells ($r = 0.261$, $P < 0.05$) and negatively correlated with M0 macrophages ($r = –0.457$, $P < 0.001$). These results show that the hub genes are generally associated with macrophages.

Analysis of correlation and diagnostic efficacy of hub genes

Correlations between the eight hub genes (*DDB1*, *EFTUD2*, *mTOR*, *PSMD2*, *RBM8A*, *SMARCA4*, *TARDBP*, and *UBXN7*) and all genes were performed using data from 58 PH tissues in the GSE117261 dataset, and the heatmaps showing the expression of top 50 genes positively correlated with each hub gene were plotted (Supplementary Figures 1 to 8). Based on the results of the correlation analyses and the reactome pathway, single-gene GSEA was conducted, and the top 20 results were selected (Supplementary Figures 9 to 16). *DDB1*, *PSMD2*, *EFTUD2*, and *mTOR* were primarily associated with signal transduction, the immune system, metabolism of proteins, and metabolism. *RBM8A* was mainly involved in metabolism, cell cycle, and sensory perception. *SMARCA4* was associated with disease, signaling through receiver tyrosine kinases, and cellular responses to stress. *TARDBP* was mainly associated with gene expression (transcription), RNA polymerase II transcription, and MHC class I–mediated antigen processing and presentation. *UBXN7* mainly participated in signal transduction, generic transcription pathway, and signaling by Rho GTPases, Miro GTPases, and RHOBTB3. Furthermore, the area under the ROC curve (AUC) of the three datasets was determined (Figure 7(B) to (D)): *DDB1* (AUC = 0.651), *EFTUD2* (AUC = 0.599), *mTOR* (AUC = 0.698), *PSMD2* (AUC = 0.654), *RBM8A* (AUC = 0.686), *SMARCA4* (AUC = 0.631), *TARDBP* (AUC = 0.818), and *UBXN7* (AUC = 0.739). *TARDBP* had the best diagnostic efficacy, and we postulate that it may be a key player in the development of PH.

Construction of an HPH rat model and validation of hub genes

After the rats were bred continuously for 4 weeks in a hypobaric and hypoxic chamber, we found that mPAP (Figure 8(A)), RV/BW (Figure 8(B)), and RV/[LV + S] (Figure 8(C)) values were significantly higher in HPH rats than in control rats. Western blotting assays showed that in comparison to the lungs of controls, those of HPH rats highly expressed *mTOR*, *PSMD2*, *RBM8A*, *SMARCA4*, *TARDBP*, and *UBXN7* (Figure 8(D) and (E)). In addition, compared with the pulmonary arteries of controls, those of HPH rats exhibited significantly upregulated expression of *EFTUD2*, *mTOR*, *RBM8A*, *SMARCA4*, *TARDBP*, and *UBXN7* and significantly downregulated expression of *DDB1* (Figure 8(F) and (G)).

Discussion

PH progresses rapidly and has a high mortality. However, its pathogenesis is not fully understood. The current treatment is mainly aimed at alleviating symptoms but cannot reverse pulmonary vascular remodeling. Therefore, research on PH pathogenesis and effective targeted therapy is required.

We used three datasets in the GEO database to screen DEGs between 135 patients with PH and 58 NCs. We combined WGCNA with standardized analysis of DEGs, obtained modules that were most relevant to PH, and analyzed the correlation between key modules and clinical features. MEblue and MEbrown modules had the highest correlation with PH, and eight hub genes were identified. ROC analysis showed that *TARDBP* had the best diagnostic efficacy. Finally, a rat HPH model was constructed, and validation experiments based on western blotting analysis showed that *mTOR*, *PSMD2*, *RBM8A*, *SMARCA4*, *TARDBP*, and *UBXN7* were highly expressed in the lungs. In addition, *EFTUD2*, *mTOR*, *RBM8A*, *SMARCA4, TARDBP*, and *UBXN7* were significantly upregulated, whereas *DDB1* was significantly downregulated in the pulmonary arteries. The expression of these hub genes differed between tissues, which is consistent with the fact that the structural and functional differences between pulmonary artery and lung tissue may cause them to respond differentially to hypoxic stimulation, showing different pathophysiological processes in the formation of HPH. Therefore, the same gene functions differently in different cells. We found that seven hub genes (all except *TARDBP*) were generally correlated with M0 and M1 macrophages. Mast cells, CD8T cells, and resting memory CD4T cells were also implicated in the regulation of PH. Several investigations have found that the immune system is linked to pulmonary vascular remodeling.[13,14]

DEGs in PH regulate many pathways, including human immunodeficiency virus 1 infection, herpes simplex virus 1 infection, maturity onset diabetes in young people, ubiquitin-mediated protection, and cAMP signaling. Human immunodeficiency virus 1 infection affects the epigenetic reprogramming of host cells and adjacent immune cells through cell proliferation, differentiation, and survival and is

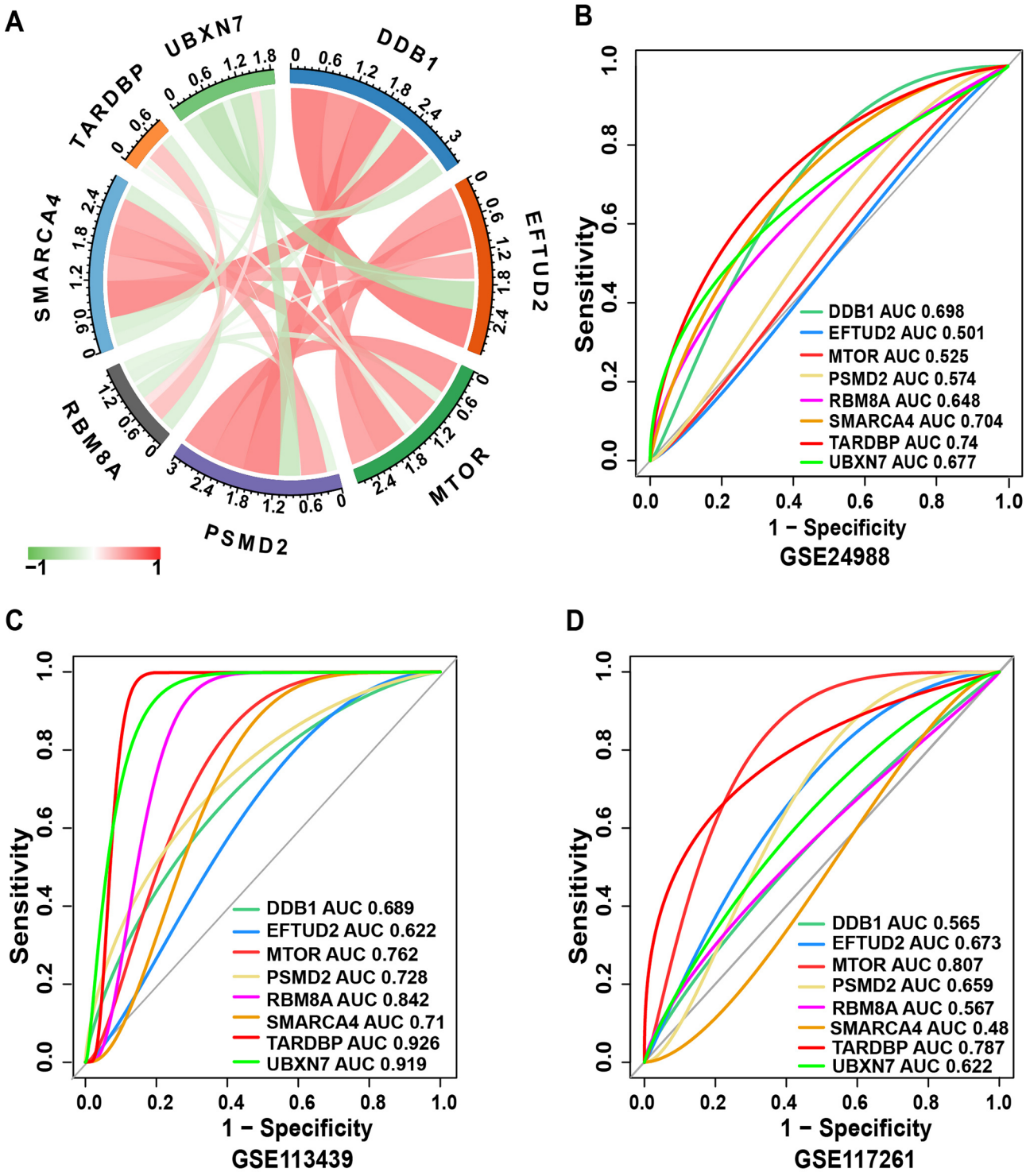

Figure 7. Correlation analysis and the diagnostic efficacies of eight hub genes. (A) Correlation analysis of hub genes. The red line shows positive correlations, whereas the green line shows negative correlations; the darker the hue, the greater the correlation. (B to D) Receiver operating characteristic (ROC) curves showing the diagnostic efficacies of eight hub genes in the three datasets: (B) GSE24988, (C) GSE113439, and (D) GSE117261.

associated with premature cell aging and abnormal immune responses.[15] Several studies have shown that inflammation and autoimmunity are correlated with PH.[16,17] Herpes simplex virus 1 infection initiates inflammation and immune responses. The increased activation of inflammasomes leads to overproduction of pro-inflammatory factors, further accelerating the remodeling of pulmonary vessels. For PH secondary to left heart disease, diabetes is known to be an independent risk factor.[18,19] It has been postulated that the regulation of ubiquitination is a factor in pulmonary vascular remodeling in PH.[20] In human pulmonary microvascular endothelial cells, E3 ubiquitin ligase regulates the function

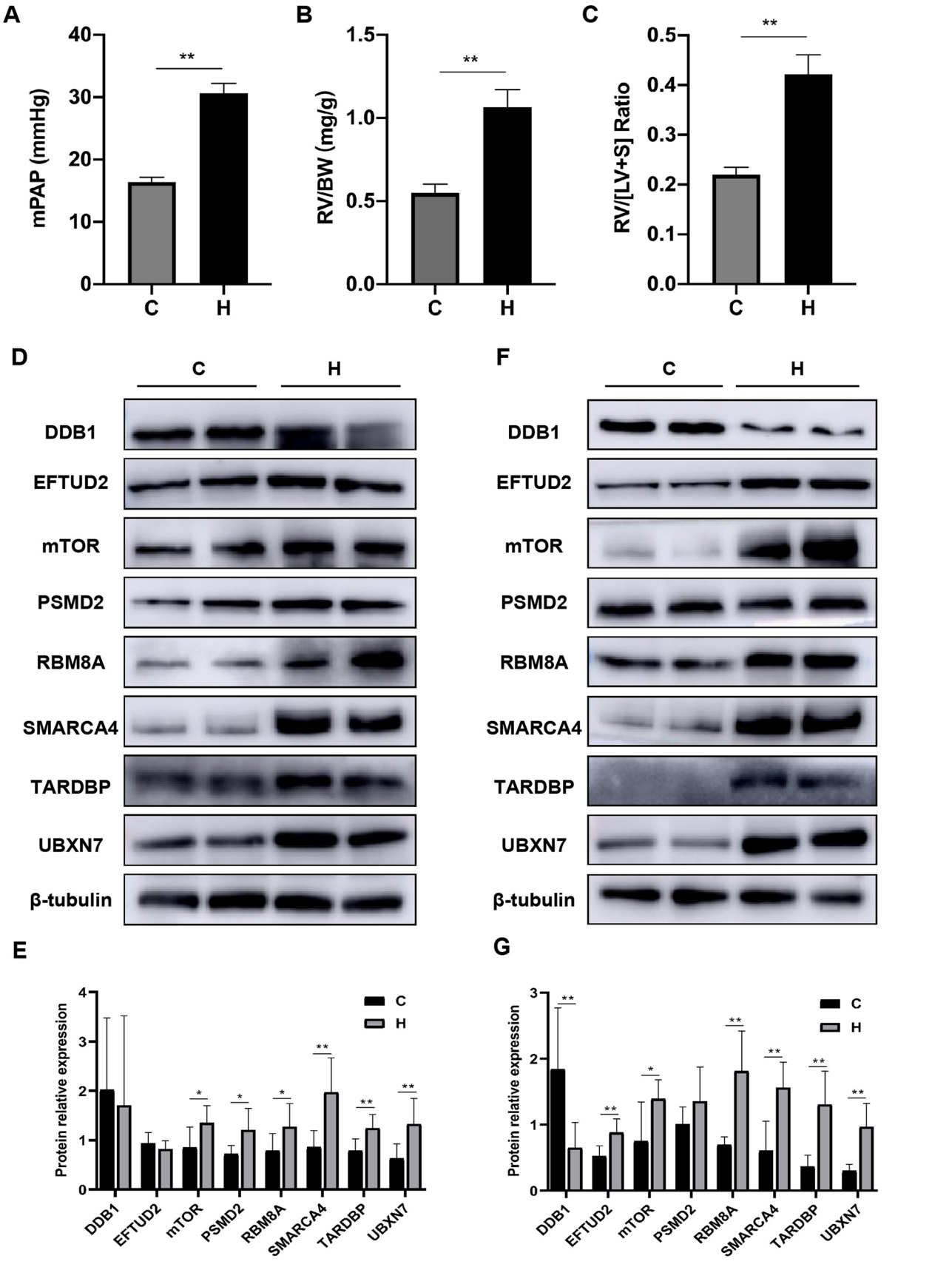

Figure 8. Rats were exposed to normoxia or 4 weeks in a hypobaric and hypoxia chamber. (A) Mean pulmonary artery pressure (mPAP). (B) Right ventricular weight/body weight (RV/BW). (C) Right ventricular (RV) free wall/the left ventricle with interventricular + septum ratio (RV/[LV + S]). (D to G) Verification of hub genes using western blotting assays. (D) and (E) show the expression of each protein in lung tissues of rats. (F) and (G) show the expression of each protein in pulmonary arteries of rats. C: control group; H: hypoxic pulmonary hypertension group. Data are presented as mean ± standard deviation.
*$P < 0.05$, **$P < 0.01$ versus control; $n = 8$.

of pulmonary artery endothelial cells.[21] Both patients with PH and HPH model mice have high levels of cAMP in their lungs. Inhibition of the cAMP pathway aggravates pulmonary vascular remodeling in HPH rats. However, infusion of exogenous cAMP into HPH rat models can prevent and reverse HPH-related pulmonary and cardiac remodeling.[22] Thus, PH pathogenesis is complex and involves multiple pathways.

The protein encoded by *DDB1* is a heterodimeric DNA damage-binding (DDB) complex involved in protein ubiquitination. It is involved in many regulatory processes, including transcription,[23] embryonic development,[24] regulation of cell proliferation,[25] and phosphorylation and post-translational modification.[26] It also mediates different types of ubiquitination reactions.[27] The role of *DDB1* in PH has not been previously studied, and more research is needed to determine how it works.

EFTUD2 plays a role in proliferation, differentiation, and cell cycle of hepatoma cells[28] and osteoblasts.[29] It also regulates apoptosis in breast cancer cells[30] and inflammatory responses in murine colon cancer cells.[31] However, the function of *EFTUD2* in PH requires further interventions.

mTOR regulates cell proliferation and the cell cycle.[32] Activation of the mTOR pathway leads to pulmonary vascular remodeling in PH mice models.[33] Inhibition of mTOR phosphorylation reduces pulmonary artery smooth muscle cell proliferation.[34]

PSMD2, also known as 26S proteasome, is involved in the cell cycle, apoptosis, repair of DNA damage, and maintenance of protein homeostasis.[35] *PSMD2* regulates cell proliferation and participates in cell cycle progression in breast cancer[36] and hepatocellular carcinoma.[37] In a previous study, *PSMD2* was highly expressed in lung tissue extracted from a PH rat model.[38] This is consistent with our findings and suggests that *PSMD2* may participate in cell proliferation, survival, and metastasis.

RBM8A (or *Y14*) is essential for maintaining a variety of mRNA functions. *RBM8A* is abnormally expressed in many cancer cells.[39,40] It regulates the proliferation and differentiation of neural progenitor cells and regulates the apoptosis of A549 cells.[41,42] To understand the function of *RBM8A* in PH, more investigations are needed.

SMARCA4 (or *BRG1*) is associated with the pathogenic process of maintaining vascular homeostasis and stress. *BRG1* was highly expressed in human endothelial cells stimulated by hypoxic conditions and in pulmonary arteries of HPH mice. It regulates cell adhesion molecules in endothelial cells, causing endothelial dysfunction and HPH progression.[43] *BRG1* participates in the transactivation of human vascular endothelial cells under hypoxic conditions by affecting histone modification, which may be one of the reasons for the continuous vasoconstriction observed in patients with HPH. The effect of *BRG1* on HPH-derived endothelial cells has been confirmed.[43] Therefore, targeting *BRG1* is a potential new strategy for HPH treatment.

TARDBP, also called *TDP-43*, is a developmentally regulated nuclear protein.[44] It is involved in mRNA stabilization and mRNA transport and translation.[45] *TARDBP* has been studied in various neurodegenerative diseases,[46,47] including amyotrophic lateral sclerosis.[48] *TARDBP* protein lesions induce damage to the ubiquitin proteasome, resulting in intracellular ubiquitin-positive inclusions in neurons.[49] In HeLa cells, RNA interference of *TDP-43* substantially reduced cell proliferation, increased apoptosis, and dysregulated the cell cycle. *TDP-43* overexpression resulted in excessive proliferation of human embryonic kidney 293T (HEK293T) cells, impaired cellular structural integrity, and increased cell membrane permeability.[50] However, it still remains unclear whether *TDP-43* contributes to the regulation of PH.

As a co-factor protein, UBXN7 can interact with various proteins, and its expression is upregulated in lung squamous cell carcinomas.[51] UBXN7 regulates HIF-1α, glycolysis, and oxidative phosphorylation levels under hypoxia.[52] Moreover, as *HIF-1α* is highly expressed in the lungs of HPH mice,[53] our subsequent research will concentrate on investigating the regulatory relationship between UBXN7 and HIF-1α in HPH.

This study had some limitations. First, we only verified the expression of the identified hub genes in an HPH rat model; however, their expression in other types of PH and human tissues largely remain unclear. Second, this study was mostly descriptive and did not explore the mechanism underlying PH. We plan on exploring the involvement of PH hub genes in follow-up experiments.

In conclusion, this research used WGCNA to identify PH hub genes. Verification experiments showed that *DDB1*, *mTOR*, *PSMD2*, *RBM8A*, *SMARCA4*, *EFTUD2*, *TARDBP*, and *UBXN7* may be critical genes in the onset and progression of HPH. Our findings provide theoretical guidance for future PH studies and may assist in the development of novel therapies against PH.

AUTHORS' CONTRIBUTIONS

SW created the study design, evaluated the data, and drafted the manuscript's first version. DS and R-lG reviewed the proposal. YG generated the animal models and collected materials. JM collated data. CL conducted statistical analysis. SC conceptualized the research and edited the text.

ACKNOWLEDGEMENTS

The authors would like to thank Dr Bing Li and Dr Yan Liu for their assistance with data analysis.

DECLARATION OF CONFLICTING INTERESTS

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

FUNDING

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The Science and Technology Plan Natural Science Fund Team Project of Qinghai Province of China (grant number 2019-ZJ-902) and the "Thousand Talents Program" for High-end Innovation of Qinghai Province (grant number 2017-5) funded this research.

ORCID ID

Sen Cui https://orcid.org/0000-0002-1950-516X

SUPPLEMENTAL MATERIAL

Supplemental material for this article is available online.

REFERENCES

1. Schermuly RT, Ghofrani HA, Wilkins MR, Grimminger F. Mechanisms of disease: pulmonary arterial hypertension. *Nat Rev Cardiol* 2011;**8**:443–55
2. McGoon MD, Benza RL, Escribano-Subias P, Jiang X, Miller DP, Peacock AJ, Pepke-Zaba J, Pulido T, Rich S, Rosenkranz S, Suissa S, Humbert M. Pulmonary arterial hypertension: epidemiology and registries. *J Am Coll Cardiol* 2013;**62**:D51–9
3. Benza RL, Miller DP, Barst RJ, Badesch DB, Frost AE, McGoon MD. An evaluation of long-term survival from time of diagnosis in pulmonary arterial hypertension from the REVEAL Registry. *Chest* 2012;**142**:448–56
4. Leuchte HH, Ten Freyhaus H, Gall H, Halank M, Hoeper MM, Kaemmerer H, Kähler C, Riemekasten G, Ulrich S, Schwaiblmair M, Ewert R. Risk stratification strategy and assessment of disease progression in patients with pulmonary arterial hypertension: updated recommendations from the Cologne Consensus Conference 2018. Int J Cardiol 2018;272s:20–9
5. Moore LG, Niermeyer S, Zamudio S. Human adaptation to high altitude: regional and life-cycle perspectives. *Am J Phys Anthropol* 1998;**41**:25–64
6. León-Velarde F, Villafuerte FC, Richalet JP. Chronic mountain sickness and the heart. *Prog Cardiovasc Dis* 2010;**52**:540–9
7. Hoeper MM, Humbert M, Souza R, Idrees M, Kawut SM, Sliwa-Hahnle K, Jing ZC, Gibbs JS. A global view of pulmonary hypertension. *Lancet Respir Med* 2016;**4**:306–22
8. Coons JC, Pogue K, Kolodziej AR, Hirsch GA, George MP. Pulmonary arterial hypertension: a pharmacotherapeutic update. *Curr Cardiol Rep* 2019;**21**:141
9. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 2008;**9**:559
10. Gaudet P, Dessimoz C. Gene ontology: pitfalls, biases, and remedies. *Methods Mol Biol* 2017;**1446**:189–205
11. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* 2017;**45**:D353–61
12. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, Santos A, Doncheva NT, Roth A, Bork P, Jensen LJ, von Mering C. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res* 2017;**45**:D362–8
13. Goldenberg NM, Rabinovitch M, Steinberg BE. Inflammatory basis of pulmonary arterial hypertension: implications for perioperative and critical care medicine. *Anesthesiology* 2019;**131**:898–907
14. Rabinovitch M, Guignabert C, Humbert M, Nicolls MR. Inflammation and immunity in the pathogenesis of pulmonary arterial hypertension. *Circ Res* 2014;**115**:165–75
15. Paschos K, Allday MJ. Epigenetic reprogramming of host genes in viral and microbial pathogenesis. *Trends Microbiol* 2010;**18**:439–47
16. Perros F, Dorfmüller P, Montani D, Hammad H, Waelput W, Girerd B, Raymond N, Mercier O, Mussot S, Cohen-Kaminsky S, Humbert M, Lambrecht BN. Pulmonary lymphoid neogenesis in idiopathic pulmonary arterial hypertension. *Am J Respir Crit Care Med* 2012;**185**:311–21
17. Liang S, Desai AA, Black SM, Tang H. Cytokines, chemokines, and inflammation in pulmonary arterial hypertension. *Adv Exp Med Biol* 2021;**1303**:275–303
18. Aronson D, Eitan A, Dragu R, Burger AJ. Relationship between reactive pulmonary hypertension and mortality in patients with acute decompensated heart failure. *Circ Heart Fail* 2011;**4**:644–50
19. Robbins IM, Hemnes AR, Pugh ME, Brittain EL, Zhao DX, Piana RN, Fong PP, Newman JH. High prevalence of occult pulmonary venous hypertension revealed by fluid challenge in pulmonary hypertension. *Circ Heart Fail* 2014;**7**:116–22
20. Shen H, Zhang J, Wang C, Jain PP, Xiong M, Shi X, Lei Y, Chen S, Yin Q, Thistlethwaite PA, Wang J, Gong K, Yuan ZY, Yuan JX, Shyy JY. MDM2-mediated ubiquitination of angiotensin-converting enzyme 2 contributes to the development of pulmonary arterial hypertension. *Circulation* 2020;**142**:1190–204
21. Kim SY, Kim HJ, Park MK, Huh JW, Park HY, Ha SY, Shin JH, Lee YS. Mitochondrial E3 ubiquitin protein ligase 1 mediates cigarette smoke-induced endothelial cell death and dysfunction. *Am J Respir Cell Mol Biol* 2016;**54**:284–96
22. Jones C, Bisserier M, Bueno-Beti C, Bonnet G, Neves-Zaph S, Lee SY, Milara J, Dorfmüller P, Humbert M, Leopold JA, Hadri L, Hajjar RJ, Sassi Y. A novel secreted-cAMP pathway inhibits pulmonary hypertension via a feed-forward mechanism. *Cardiovasc Res* 2020;**116**:1500–13
23. Wang X, Wang HY, Hu GS, Tang WS, Weng L, Zhang Y, Guo H, Yao SS, Liu SY, Zhang GL, Han Y, Liu M, Zhang XD, Cen X, Shen HF, Xiao N, Liu CQ, Wang HR, Huang J, Liu W, Li P, Zhao TJ. DDB1 binds histone reader BRWD3 to activate the transcriptional cascade in adipogenesis and promote onset of obesity. *Cell Rep* 2021;**35**:109281
24. Cang Y, Zhang J, Nicholas SA, Bastien J, Li B, Zhou P, Goff SP. Deletion of DDB1 in mouse brain and lens leads to p53-dependent elimination of proliferating cells. *Cell* 2006;**127**:929–40
25. Zheng W, Nazish J, Wahab F, Khan R, Jiang X, Shi Q. DDB1 regulates Sertoli cell proliferation and testis cord remodeling by TGFβ pathway. *Genes* 2019;**10**:974
26. Dhiani BA, Mehellou Y. The Cul4-DDB1-WDR3/WDR6 complex binds SPAK and OSR1 kinases in a phosphorylation-dependent manner. *ChemBioChem* 2020;**21**:638–43
27. Fischer ES, Böhm K, Lydeard JR, Yang H, Stadler MB, Cavadini S, Nagel J, Serluca F, Acker V, Lingaraju GM, Tichkule RB, Schebesta M, Forrester WC, Schirle M, Hassiepen U, Ottl J, Hild M, Beckwith RE, Harper JW, Jenkins JL, Thomä NH. Structure of the DDB1-CRBN E3 ubiquitin ligase in complex with thalidomide. *Nature* 2014;**512**:49–53
28. Lv C, Li XJ, Hao LX, Zhang S, Song Z, Ji XD, Gong B. Over-activation of EFTUD2 correlates with tumor propagation and poor survival outcomes in hepatocellular carcinoma. *Clin Transl Oncol* 2022;**24**:93–103
29. Wu J, Yang Y, He Y, Li Q, Wang X, Sun C, Wang L, An Y, Luo F. EFTUD2 gene deficiency disrupts osteoblast maturation and inhibits chondrocyte differentiation via activation of the p53 signaling pathway. *Hum Genomics* 2019;**13**:63
30. Sato N, Maeda M, Sugiyama M, Ito S, Hyodo T, Masuda A, Tsunoda N, Kokuryo T, Hamaguchi M, Nagino M, Senga T. Inhibition of SNW1 association with spliceosomal proteins promotes apoptosis in breast cancer cells. *Cancer Med* 2015;**4**:268–77
31. Lv Z, Wang Z, Luo L, Chen Y, Han G, Wang R, Xiao H, Li X, Hou C, Feng J, Shen B, Wang Y, Peng H, Guo R, Li Y, Chen G. Spliceosome protein Eftud2 promotes colitis-associated tumorigenesis by modulating inflammatory response of macrophage. *Mucosal Immunol* 2019;**12**:1164–73
32. Zhu M, Miao S, Zhou W, Elnesr SS, Dong X, Zou X. MAPK, AKT/FoxO3a and mTOR pathways are involved in cadmium regulating the cell cycle, proliferation and apoptosis of chicken follicular granulosa cells. *Ecotoxicol Environ Saf* 2021;**214**:112091
33. Tang H, Wu K, Wang J, Vinjamuri S, Gu Y, Song S, Wang Z, Zhang Q, Balistrieri A, Ayon RJ, Rischard F, Vanderpool R, Chen J, Zhou G, Desai AA, Black SM, Garcia JGN, Yuan JX, Makino A. Pathogenic role of mTORC1 and mTORC2 in pulmonary hypertension. *JACC Basic Transl Sci* 2018;**3**:744–62
34. Shoji H, Yoshida Y, Sanada TJ, Naito A, Maruyama J, Zhang E, Sumi K, Sakao S, Maruyama K, Hidaka H, Tatsumi K. The isoquinolinesulfonamide compound H-1337 attenuates SU5416/hypoxia-induced pulmonary arterial hypertension in rats. *Cells* 2021;**11**:66
35. Bi W, Zhu L, Zeng Z, Jing X, Liang Y, Guo L, Shi Q, Xu A, Tao E. Investigations into the role of 26S proteasome non-ATPase regulatory subunit 13 in neuroinflammation. *Neuroimmunomodulation* 2014;**21**:331–7
36. Li Y, Huang J, Zeng B, Yang D, Sun J, Yin X, Lu M, Qiu Z, Peng W, Xiang T, Li H, Ren G. PSMD2 regulates breast cancer cell proliferation and cell cycle progression by modulating p21 and p27 proteasomal degradation. *Cancer Lett* 2018;**430**:109–22

37. Tan Y, Jin Y, Wu X, Ren Z. PSMD1 and PSMD2 regulate HepG2 cell proliferation and apoptosis via modulating cellular lipid droplet metabolism. *BMC Mol Biol* 2019;**20**:24
38. Wang J, Uddin MN, Li Q, Aierken A, Li MY, Wang R, Yan QZ, Adi D, Gai MT, Wu Y. Identifying potential mitochondrial proteome signatures associated with the pathogenesis of pulmonary arterial hypertension in the rat model. *Oxid Med Cell Longev* 2022;**2022**:8401924
39. Song X, Chen B, Liang Y, Li Y, Zhang H, Han D, Wang Y, Ye F, Wang L, Zhao W, Yang Q. CircEIF3H-IGF2BP2-HuR scaffold complex promotes TNBC progression via stabilizing HSPD1/RBM8A/G3BP1 mRNA. *Cell Death Discov* 2022;**8**:261
40. Wei L, Zou C, Chen L, Lin Y, Liang L, Hu B, Mao Y, Zou D. Molecular insights and prognosis associated with RBM8A in glioblastoma. *Front Mol Biosci* 2022;**9**:876603
41. Zou D, McSweeney C, Sebastian A, Reynolds DJ, Dong F, Zhou Y, Deng D, Wang Y, Liu L, Zhu J, Zou J, Shi Y, Albert I, Mao Y. A critical role of RBM8a in proliferation and differentiation of embryonic neural progenitors. *Neural Dev* 2015;**10**:18
42. Ishigaki Y, Nakamura Y, Tatsuno T, Hashimoto M, Shimasaki T, Iwabuchi K, Tomosugi N. Depletion of RNA-binding protein RBM8A (Y14) causes cell cycle deficiency and apoptosis in human cells. *Exp Biol Med* 2013;**238**:889–97
43. Chen D, Fang F, Yang Y, Chen J, Xu G, Xu Y, Gao Y. Brahma-related gene 1 (Brg1) epigenetically regulates CAM activation during hypoxic pulmonary hypertension. *Cardiovasc Res* 2013;**100**:363–73
44. Huang C, Xia PY, Zhou H. Sustained expression of TDP-43 and FUS in motor neurons in rodent's lifetime. *Int J Biol Sci* 2010;**6**:396–406
45. Ratti A, Buratti E. Physiological functions and pathobiology of TDP-43 and FUS/TLS proteins. *J Neurochem* 2016;**138**:95–111
46. Lee S, Jeon YM, Cha SJ, Kim S, Kwon Y, Jo M, Jang YN, Lee S, Kim J, Kim SR, Lee KJ, Lee SB, Kim K, Kim HJ. PTK2/FAK regulates UPS impairment via SQSTM1/p62 phosphorylation in TARDBP/TDP-43 proteinopathies. *Autophagy* 2020;**16**:1396–412
47. van Rooij J, Mol MO, Melhem S, van der Wal P, Arp P, Paron F, Donker Kaat L, Seelaar H, Miedema SSM, Oshima T, Eggen BJL, Uitterlinden A, van Meurs J, van Kesteren RE, Smit AB, Buratti E, van Swieten JC. Somatic TARDBP variants as a cause of semantic dementia. *Brain* 2020;**143**:3827–41
48. Riva N, Gentile F, Cerri F, Gallia F, Podini P, Dina G, Falzone YM, Fazio R, Lunetta C, Calvo A, Logroscino G, Lauria G, Corbo M, Iannaccone S, Chiò A, Lazzerini A, Nobile-Orazio E, Filippi M, Quattrini A. Phosphorylated TDP-43 aggregates in peripheral motor nerves of patients with amyotrophic lateral sclerosis. *Brain* 2022;**145**:276–84
49. Bendotti C, Marino M, Cheroni C, Fontana E, Crippa V, Poletti A, De Biasi S. Dysfunction of constitutive and inducible ubiquitin-proteasome system in amyotrophic lateral sclerosis: implication for protein aggregation and immune response. *Prog Neurobiol* 2012;**97**:101–26
50. Hergesheimer R, Lanznaster D, Bourgeais J, Hérault O, Vourc'h P, Andres CR, Corcia P, Blasco H. Conditioned medium from cells overexpressing TDP-43 alters the metabolome of recipient cells. *Cells* 2020;**9**:2198
51. Wang J, Qian J, Hoeksema MD, Zou Y, Espinosa AV, Rahman SM, Zhang B, Massion PP. Integrative genomics analysis identifies candidate drivers at 3q26-29 amplicon in squamous cell carcinoma of the lung. *Clin Cancer Res* 2013;**19**:5580–90
52. Cilenti L, Di Gregorio J, Ambivero CT, Andl T, Liao R, Zervos AS. Mitochondrial MUL1 E3 ubiquitin ligase regulates hypoxia inducible factor (HIF-1α) and metabolic reprogramming by modulating the UBXN7 cofactor protein. *Sci Rep* 2020;**10**:1609
53. Deng H, Tian X, Sun H, Liu H, Lu M, Wang H. Calpain-1 mediates vascular remodelling and fibrosis via HIF-1α in hypoxia-induced pulmonary hypertension. *J Cell Mol Med* 2022;**26**:2819–30

(Received September 1, 2022, Accepted November 25, 2022)