


RESEARCH

Open Access



# Combined analysis of single-cell and bulk transcriptome sequencing data identifies critical glycolysis genes in idiopathic pulmonary arterial hypertension

Xuan Gao<sup>1,2†</sup>, Youli Fan<sup>1,2†</sup>, Guijia Wang<sup>1†</sup>, Jiangjiang Xu<sup>1</sup>, Runwei Deng<sup>1,2</sup>, Jiangwei Song<sup>1,2</sup>, Binfeng Sun<sup>1</sup>, Yongbing Wang<sup>1,2</sup>, Zixuan Wu<sup>1,2</sup>, Ruyi Jia<sup>1,2</sup>, Jing Huang<sup>1,2</sup>, Huiyu He<sup>1,2</sup>, Lei Gao<sup>1</sup>, Yihao Zhang<sup>1</sup>, Na Sun<sup>1,2\*</sup> and Bingxiang Wu<sup>1,2\*</sup> 

## Abstract

**Background** Abnormal glycolytic metabolism plays a significant role in pulmonary vascular remodeling in idiopathic pulmonary arterial hypertension (IPAH), yet the specific mechanisms remain unclear. The primary objective of this study is to investigate the key regulatory mechanisms of glycolysis in IPAH.

**Methods** Bulk and single-cell sequencing data obtained from IPAH patient tissue samples were downloaded from the GEO database. scMetabolism and AUCcell analyses of the IPAH single-cell sequencing data were carried out to quantify the glycolytic metabolic activity and identify the main cell types regulating glycolysis, respectively. The ssGSEA method was used to assess the glycolytic activity in each bulk sample within the bulk sequencing data. Differential analysis, weighted gene co-expression network analysis (WGCNA), and protein-protein interaction (PPI) network analysis were conducted to identify key genes associated with glycolysis in IPAH samples. Single-cell sequencing and a monocrotaline (MCT)-induced model of PH in rats were utilized to validate the expression of these key genes.

**Results** Single-cell sequencing data indicated that IPAH patients displayed increased glycolytic activity, which was primarily regulated by fibroblasts. Similarly, bulk transcriptomic data revealed a significant increase in glycolytic activity in IPAH patients. Differential analysis, WGCNA, PPI network analysis, and integrated single-cell analysis further identified insulin-like growth factor-1 (IGF1), lysyl-tRNA synthetase (KARS), caspase-3 (CASP3), and cyclin-dependent kinase inhibitor 2 A (CDKN2A) as key genes associated with fibroblast-mediated glycolysis in IPAH patients. Differential expression of IGF1, KARS, CASP3, and CDKN2A was also observed in our in vivo model of PH.

<sup>†</sup>Xuan Gao, Youli Fan and Guijia Wang are first authors.

\*Correspondence:

Na Sun  
shirly\_cool@163.com  
Bingxiang Wu  
wubingxiang1964@163.com

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

**Conclusion** Our study identifies IGF1, KARS, CASP3, and CDKN2A as key regulatory genes in glycolysis in IPAH, which provides the basis for the development of targeted therapies.

**Keywords** IPAH, Glycolysis, Single-cell sequencing

## Introduction

Idiopathic pulmonary arterial hypertension (IPAH) refers to a group of cardiopulmonary circulatory disorders of unknown etiology characterized by progressive pulmonary vascular remodeling and elevated pulmonary vascular resistance [1]. The underlying mechanisms of IPAH are complex, and involve key pathological changes such as vasoconstriction, inflammation, and proliferative vascular remodeling [2]. Although the introduction of PAH-specific therapies has led to improvements in the median survival rates of IPAH patients, long-term survival outcomes remain unsatisfactory [3]. Current treatments have focused on targeting pulmonary endothelial dysfunction to regulate vasodilatory effects [4]. However, there remains a critical need for the development of new drugs that specifically target the proliferative remodeling of the pulmonary vasculature.

The metabolism of PAH patients has been shown to depend on glycolysis, even in the presence of sufficient oxygen levels, a phenomenon known as the Warburg effect [5]. The Warburg effect promotes lactate accumulation and the formation of an acidic microenvironment, activating transcription factors such as hypoxia inducible factor-1 $\alpha$  (HIF-1 $\alpha$ ), which in turn regulate phenotypic switching and proliferation of vascular smooth muscle cells [6]. In cardiovascular diseases, glycolysis not only provides energy to cells but also regulates vascular remodeling and smooth muscle cell proliferation through the generation of intermediate metabolites [7]. This preference for glycolysis over aerobic respiration has been found to be significantly increased in the remodeled pulmonary vasculature and right ventricle of rats with pulmonary hypertension (PH), as evidenced by a heightened uptake of  $^{18}\text{F}$ -fluorodeoxyglucose in positron emission tomography studies [8]. An imbalance in metabolic processes is believed to be a key factor associated with the abnormal proliferation and increased anti-apoptotic behavior of pulmonary artery smooth muscle cells (PASMCs) in patients with PAH [9]. Overactivation of phosphofructo-2-kinase/fructose-2,6-bisphosphatase 3 (PFKFB3) in PAH has been shown to drive collagen synthesis and the proliferation of PASMCs through increased glycolytic activity [10]. Cannabidiol has been shown to lower pulmonary artery pressure by targeting this dysregulated glycolysis in hypoxia-induced PH mice, specifically by reducing PFKFB3 mRNA expression levels [11]. In addition, Ras association domain

family protein-1 isoform A (RASSF1A) has been shown to interact with HIF-1 $\alpha$ , creating a feedforward loop that prevents its prolyl-hydroxylation and subsequent degradation, thereby amplifying the Warburg effect and enhancing hypoxic signaling in PH [11]. Inhibitors of HIF-1 $\alpha$  have been effective in reducing the Warburg effect through decreased pyruvate dehydrogenase kinase 1 (PDK1) activity, and thus, provide a promising strategy to treat PAH [12]. However, the precise role of glycolysis in the development of PAH remains poorly understood.

In recent years, single-cell and bulk sequencing technologies have emerged as essential tools for studying the molecular mechanisms of pulmonary vascular disease. Single-cell sequencing identifies cellular heterogeneity at high resolution, thereby enabling differentiation between different cell subtypes within complex populations, which is critical for the systematic and comprehensive exploration of disease mechanisms at the cellular level [13]. Successful application of single-cell sequencing has been reported in studies examining the role of aberrant glycolytic processes in the development of diseases such as tumors [14]. RNA sequencing technology has been widely used in the study of pulmonary vascular diseases, and its validity and reliability in the search for novel markers for the diagnosis and treatment of IPAH have been demonstrated [15]. To date, there have been no studies examining potential key glycolysis molecules associated with the pathogenicity of IPAH.

Here, we used transcriptomics combined with single-cell analysis to examine IGF1, KARS, CASP3, and CDKN2A as potential regulators of glycolysis in IPAH, which provides the basis for the development of targeted therapies.

## Materials and methods

### Processing and analysis of single-cell RNA sequencing data

Single-cell transcriptomic data from the pulmonary artery tissue of three control and three IPAH patients were obtained from the GEO dataset (GSE228644 [16]) (<http://www.ncbi.nlm.nih.gov/geo>). Cells containing more than 200 and less than 2500 genes with no more than 15% mitochondrial genes were screened using the Seurat R package (Version 4.3.1) [17]. After filtering out low-quality cells, approximately 21,000 high-quality cells were screened. The principal component was set to 10 to differentiate between the different cell subpopulations. UMAP plots were used to visualize

the different cell subpopulations. Data were normalized using the Normalize Data function, and genes with higher coefficients of variation between cells were extracted using the Find Variable Features function. The “FindAllMarkers” function was performed to identify marker genes of each cell cluster with adjusted  $P < 0.05$  and  $|\log_2(\text{fold change})| > 0.25$  set as thresholds. Finally, cell clusters were annotated using the CellMarker database [18].

#### Pathway enrichment analysis and glycolysis scores

The “scMetabolism” package (Version 5.0.1) was used to determine the metabolic pathway activation scores for the single-cell sequencing data [19]. The glycolysis score of each cell was assessed using the “AUCell” package (Version 1.22.0) [20]. Then, the glycolysis scores of different cell clusters were visualized using UMAP and violin plots.

#### Analysis of bulk sequencing data

Two PAH microarray datasets (GSE113439 [21] and GSE53408 [22]) were obtained from the GEO database. Expression matrix data were provided in Supplementary Material 1. The GSE113439 dataset includes 6 IPAH patients and 11 control samples, and the GSE53408 dataset contains 12 severe PAH patients and 11 control samples. Control and IPAH samples in the datasets were selected for subsequent analysis. The raw data of the datasets were normalized using the “Affy” package [23]. The “ComBat” package was used to correct for batch effects across platforms, and principal component analysis was used to evaluate the correction. Differential expression analysis between controls and IPAH patients was performed using the “limma” package (Version 3.50.0) [24]. The linear model was constructed using the lmFit function to fit gene expression data, and the empirical Bayes correction was applied using the eBayes function to enhance statistical power. The criteria for differentially expressed genes (DEGs) included an absolute  $\log_2$  fold-change greater than 0.58 and an adjusted  $p$ -value of less than 0.05.

#### Determination of the diagnostic effectiveness of the glycolysis score

Five hundred glycolysis-related genes were identified from the GeneCards database (<https://www.genecard.s.org/>). Single-sample Gene Set Enrichment Analysis (ssGSEA) is a computational method used to evaluate the activity of gene sets in individual samples. We employed ssGSEA to calculate enrichment scores for glycolysis-related gene sets, thereby quantifying glycolytic activity in the samples. All analyses were performed using the GSVA package in R software. The

glycolysis scores, based on the expression levels of glycolysis-related genes, were compared between controls and IPAH patients. The diagnostic efficacy of the glycolysis score in predicting IPAH was analyzed using receiver operating characteristic (ROC) curves. The GSE53408 dataset was used as a validation dataset to re-validate the diagnostic efficacy of the glycolysis score.

#### Weighted co-expression network analysis (WGCNA)

The “WGCNA” package (Version 1.72-5) was used to construct co-expression networks by analyzing the expression patterns among genes [25]. Glycolysis-related genes used for grouping were selected from the GeneCard database. The optimal soft threshold was determined using the pickSoftThreshold function for network topology analysis, and the dynamic tree cut method was employed to identify modules. Modules that were highly correlated with glycolysis and IPAH were screened, and genes from the modules were selected for subsequent analysis. Finally, the correlation between module genes and the most relevant modules for glycolysis were calculated.

#### GO and KEGG enrichment analyses

The “limma” package was used to identify DEGs between controls and IPAH patients, with adjusted  $P < 0.05$  and  $|\log_2(\text{fold change})| > 0.58$  set as the cutoff value [24]. The 500 genes most closely associated with PAH were obtained from the GeneCards database (<https://www.genecards.org/>). Next, overlapping genes between modular genes, DEGs, and GeneCards genes were selected. Finally, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses of the overlapping genes were carried out using the “clusterProfiler” package (Version 4.8.2) [26].

#### Construction of the PPI network and identification of hub genes

Protein-protein interaction (PPI) networks of overlapping genes were constructed using the STRING database (<http://string-db.org/>) [27]. The top five genes with the highest connectivity in the PPI network were identified using the Maximal Clique Centrality (MCC) and Maximum Neighborhood Component (MNC) algorithms in Cytoscape software (Version 3.9.1) [28]. Intersecting genes derived from two algorithms were considered to be hub genes.

#### Analysis of hub genes

The expression levels of the hub genes in controls and IPAH patients in the GSE113439 dataset were assessed. Then, the diagnostic abilities of the hub

genes for IPAH were determined using ROC curves. Finally, the expression levels of key genes in different cell subpopulations were visualized using UMAP and bubble plots.

Establishment of an in vivo model of PH in rats

This study was conducted in accordance with the guidelines of the Declaration of Helsinki and approved by the Institutional Research Ethics Committee of the Second Affiliated Hospital of Harbin Medical University (NO. YJSDW2022-162). All animal experiments were performed in accordance with the National Institutes of Health guide for the care and use of laboratory animals. Adult male Sprague-Dawley rats had free access to food and water and were housed in the animal center of the Second Affiliated Hospital of Harbin Medical University ( $22 \pm 2^{\circ}\text{C}$ ,  $55 \pm 5\%$  relative humidity with a 12-h light/dark cycle). After 1 week of adaptive feeding, rats were randomly divided into two groups: control group (control,  $n = 6$ ) and model group (MCT,  $n = 6$ ). Rats in the model group were subcutaneously injected with 1% monocrotaline (MCT; dissolved in 20% anhydrous ethanol in saline) at a dose of 60 mg/kg. Three weeks after administration of MCT, rats were anesthetized with an intraperitoneal injection of sodium pentobarbital (50 mg/kg body weight) and their hemodynamics were measured. Then, the rats were humanely killed by cervical dislocation. A catheter was inserted through the jugular vein into the right ventricle (RV) to measure the RV systolic pressure (RVSP), then the lung tissue was removed. Left lungs were fixed in paraformaldehyde and paraffin-embedded for histological examination. Right lungs were stored at  $-80^{\circ}\text{C}$ . The data from all rats with successfully measured hemodynamics were included in the final analysis.

Real-time reverse transcription PCR

Total RNA was extracted from lung tissue samples using Trizol according to the manufacturer’s instructions (Invitrogen, California, USA). cDNA was synthesized using the Reverse Transcription Kit (Roche, Basel, Switzerland), and quantitative PCR (qPCR) was performed using SYBR Green (Roche). All data were normalized to  $\beta$ -actin mRNA expression levels. The

sequences of the gene-specific primers used in this study can be found in Table 1.

Statistical analysis

R software (Version 4.3.1) was used for all data processing and statistical analysis throughout the study. Comparison of two groups of continuous variables was carried out using Student’s t-test for normally distributed data and the Mann-Whitney U-test for non-normally distributed data. A two-tailed  $P < 0.05$  was considered statistically significant.

Results

Identification of cell types and calculation of glycolysis scores

Cells were separated into 23 clusters using the “KNN” algorithm (Fig. 1A). Next, the expression levels of marker genes within the different cell clusters were examined (Fig. 1B). Based on the expression of the cell subgroup marker genes, 13 cell types were identified (Fig. 1C). An increased proportion of fibroblasts and endothelial cells was found in IPAH patients, whereas the proportion of T cells was decreased (Fig. 1D). Glucose metabolism-related pathways were significantly activated in IPAH patients compared to controls (Fig. 2A). Among the cell subpopulations, fibroblasts were found to be the major glycolysis-regulating effector cells, followed by myeloid cells (Fig. 2B, C). Further comparison of the glycolytic activity of fibroblasts between IPAH patients and controls revealed that the glycolytic activity of IPAH fibroblasts was significantly higher than that of controls (Fig. 2D).

Diagnostic efficacy of the glycolysis score

Data from the GSE113439 dataset were normalized resulting in excellent inter-sample agreement (Fig. 3A, B). The three-dimensional PCA plot showing the distribution of samples after data processing revealed that the control and IPAH groups were significantly different (Fig. 3C). Consistent with the single-cell sequencing data, a significantly higher glycolysis score was found in the IPAH group than the control group (Fig. 3D). The glycolysis score was able to distinguish the IPAH samples from the overall samples well with an area under the curve (AUC) of 0.84 (Fig. 3E). In the validation dataset GSE53408, the glycolysis score also displayed excellent predictive efficacy (AUC = 1.00) (Fig. 3F).

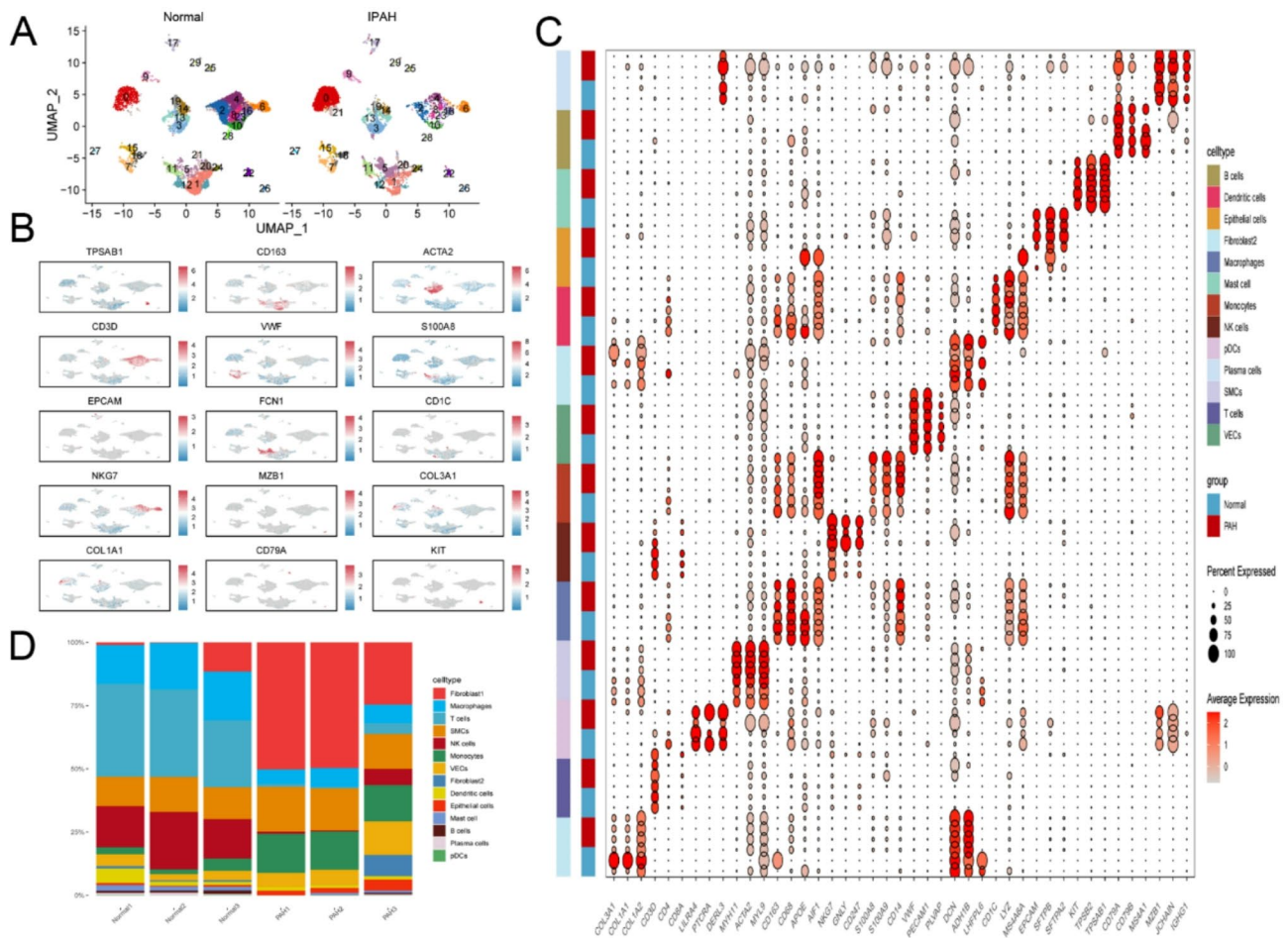
Identification of key regulatory genes for glycolysis by WGCNA

The distribution of samples without outliers is shown in the sample dendrogram in Fig. 4A. When the soft threshold of the scale-free topological model was set

Table 1 Gene primer sequences

Gene	forward primer	reverse primer
IGF1	GCCGCTTCCTTCACAGAATC	CCACCCAGTTGCTATTGCTTT
KARS	CACCAACCATACTGCTGACAAT	GTGGATGCGACCTGCTACT
CDKN2A	ATCTCCGAGAGGAAGGCGA	TTGCCATCATCATCCTGTGA
CASP3	GAGCTTGGAACGCGAAGAAA	TTGCGAGCTGACATTCCAGT
$\beta$ -actin	GAGAGGGAAATCGTGCCTGA	TGGAAGGTGGACAGTGAGGC





**Fig. 1** Transcriptomic profiles of the single-cell RNA sequencing data obtained from the lung tissue of IPAH patients. **(A)** UMAP plot showing the clustering of single cells. Different colors represent distinct cell types. **(B)** UMAP plots were used to visualize the different cell clusters. **(C)** Bubble plots showing the expression of marker genes in each cluster. **(D)** Bar plots showing the proportions of each cell subpopulation in each sample

to 8, the internal connectivity and gene similarity of the co-expression network were high (Fig. 4B). The clustering dendrogram identified a total of seven distinct modules, with genes that were close together being grouped into the same module (Fig. 4C). Among the seven modules, the blue module was significantly associated with the glycolysis score (correlation = 0.84,  $P = 2e-05$ ) (Fig. 4D). In addition, a **strong** positive correlation between the blue module and glycolysis score was observed (correlation = 0.75,  $P < 2e-100$ ) (Fig. 4E).

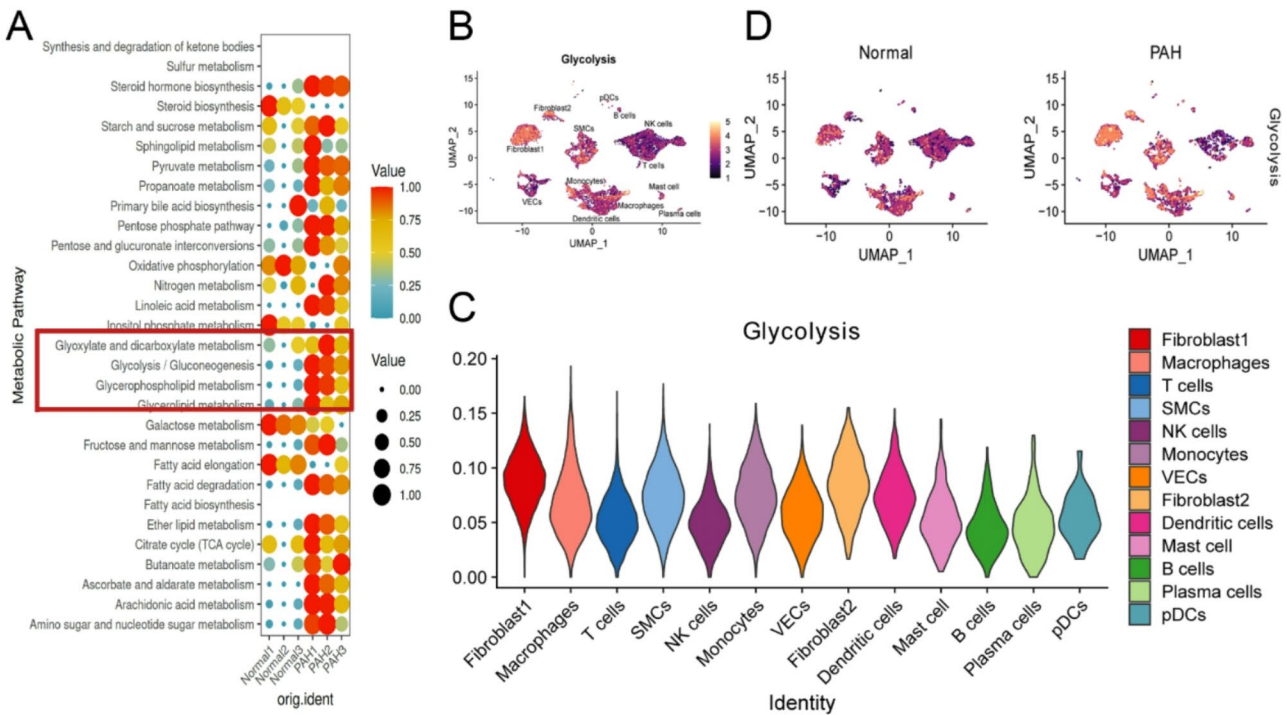
#### Functional enrichment analysis

A total of 2836 DEGs were identified in the GSE113439 dataset, including 1966 up-regulated genes and 870 down-regulated genes (Fig. 5A). There were 43 overlapping genes among the blue modules, DEGs, and genes obtained from GeneCards (Fig. 5C). These overlapping genes were most closely associated with glycolytic metabolic activity. GO functional annotation indicated that overlapping genes were associated with

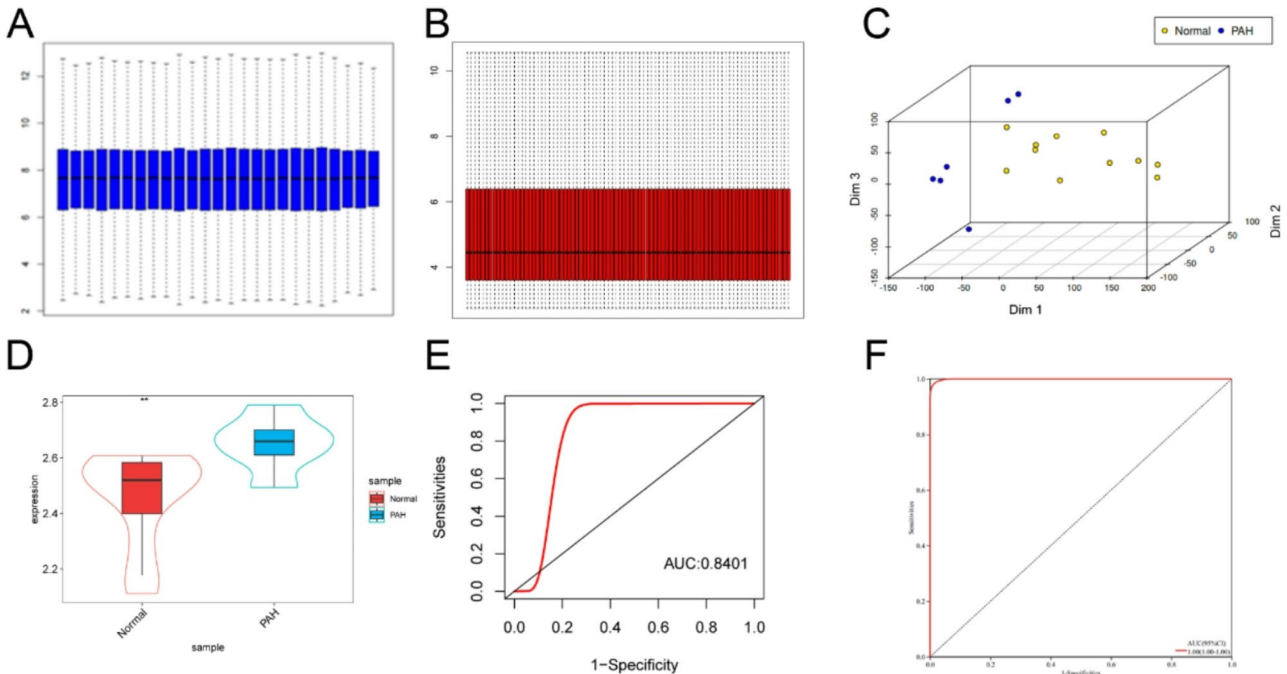
tissue migration, response to hypoxia, regulation of smooth muscle cell proliferation, membrane rafts, and phosphatidylinositol-3-kinase (PI3K) activity (Fig. 5B). KEGG annotation showed that the overlapping genes were associated with proliferation-related pathways including EGFR tyrosine kinase inhibitor resistance, the JAK-STAT signaling pathway, and PI3K-Akt signaling pathway (Fig. 5D).

#### Identification and analysis of hub genes

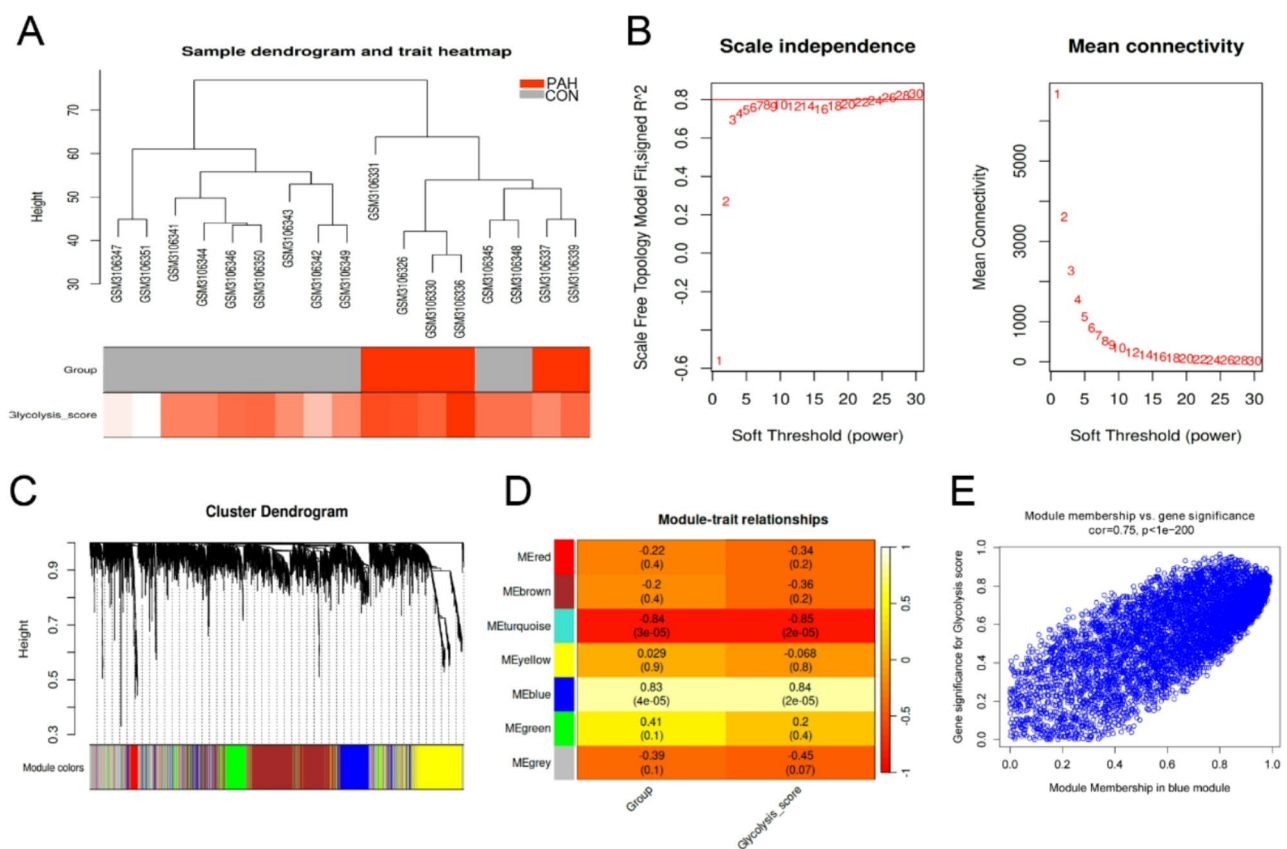
The PPI network was visualized using Cytoscape software. Genes with interaction scores greater than 0.4 were retained in the network (Fig. 6A). There were 32 nodes in the PPI network. The MCC and MNC algorithms calculated the top five most important genes in the network by evaluating the topology of the PPI network (Fig. 6B). Four common genes (insulin-like growth factor-1 (IGF1), lysyl-tRNA synthetase (KARS), caspase-3 (CASP3), and cyclin-dependent kinase inhibitor 2 A (CDKN2A) were identified as



**Fig. 2** Pathway enrichment and glycolysis scores of control and IPAH groups. **(A)** Heatmap showing the enrichment of metabolic pathways in control and IPAH groups. **(B)** UMAP plot showing the levels of glycolysis in different cell subpopulations. **(C)** Violin diagram showing a significant increase in glycolysis in fibroblasts. **(D)** Levels of glycolysis in different cell subpopulations in control and IPAH groups



**Fig. 3** Validation of the diagnostic efficacy of the glycolysis score. **(A)** Boxplot showing the gene expression matrix before data normalization. **(B)** Boxplot showing the gene expression matrix after data normalization. **(C)** 3D PCA plot showing the separation of individual samples between control and IPAH groups. **(D)** Glycolysis scores were significantly higher in IPAH groups. **(E, F)** ROC curves of glycolysis scores in the GSE113439 and GSE53408 datasets



**Fig. 4** Identification of key regulatory genes in glycolysis by WGCNA. **(A)** Glycolysis scores of different samples. **(B)** Selection of the appropriate soft threshold for the scale-free network. The red line was set at 0.80. **(C)** Construction of the gene clustering tree based on gene expression. **(D)** Associations between glycolysis scores and different modules. **(E)** Correlation of genes in the blue module with glycolysis scores

hub genes. IGF1, KARS, and CASP3 expression levels were significantly increased in IPAHA, whereas a significant decrease in CDKN2A expression was observed (Fig. 6C). All hub genes showed excellent ability to predict IPAHA (Fig. 6D). As shown in Fig. 6E, IGF1 was expressed predominantly in fibroblasts. IGF1, KARS, and CASP3 expression levels were found to be elevated, while CDKN2A expression was decreased in IPAHA samples compared with controls (Fig. 6E, F). We successfully constructed an in vivo model of PH in rats (Fig. 7A-C). The four hub genes were validated by qPCR, and our qPCR data were found to be consistent with the RNA-seq data (Fig. 7D).

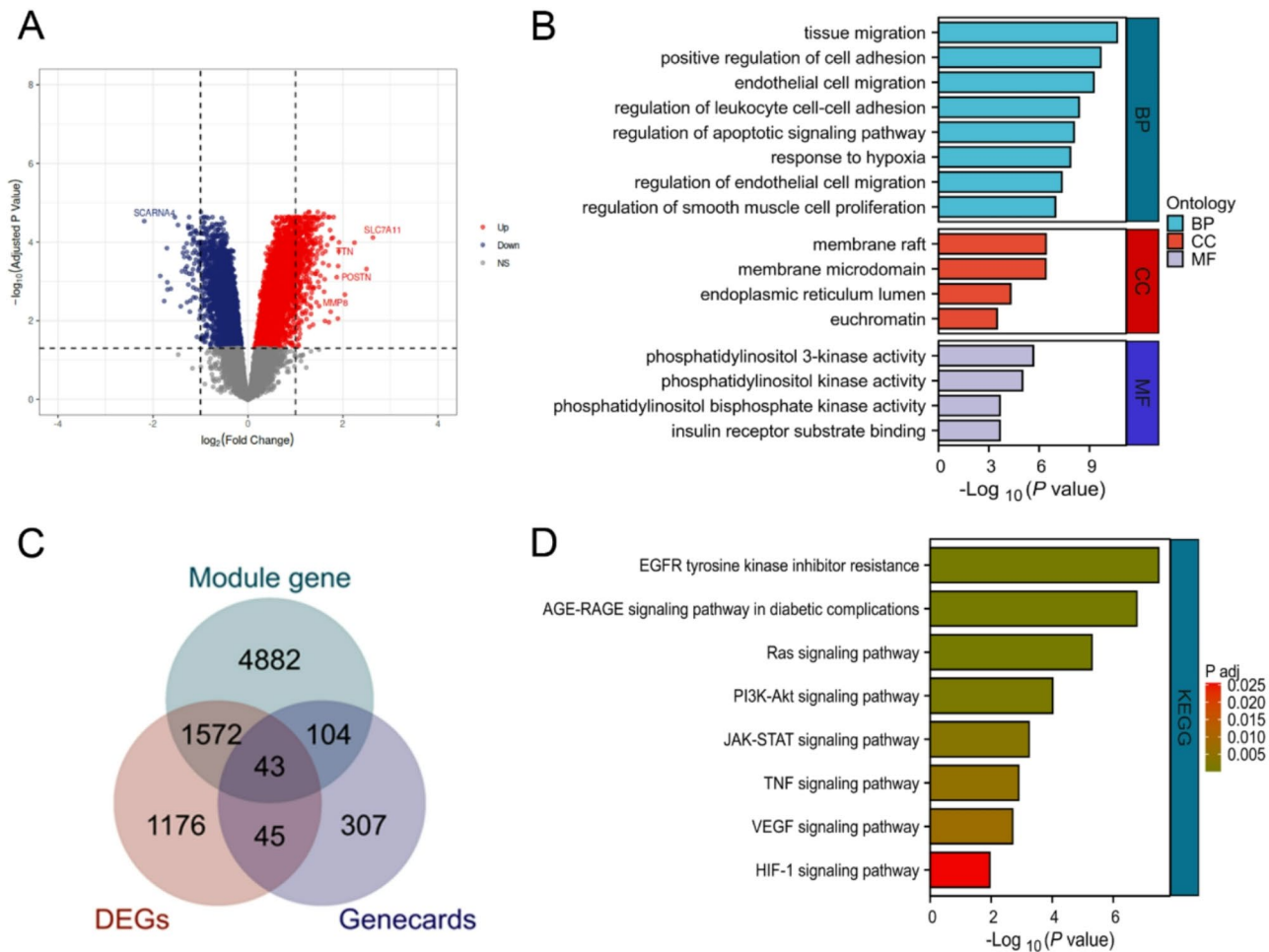
Discussion

Glycolysis plays an important role in vascular remodeling in IPAHA. Studies have shown that the glycolytic activity of PASMCs in PAH patients is significantly enhanced [29]. PDGF promotes glycolysis by activating the PI3K/AKT/mTOR/HIF-1α signaling pathway, supporting rapid cell proliferation, while lactate accumulation and inhibition of mitochondrial oxidative phosphorylation further exacerbate this process [29,

30]. Additionally, mitochondrial dysfunction promotes PASMC proliferation and vascular remodeling through HIF-1α-mediated glycolytic reprogramming [29]. The upregulation of key glycolytic enzymes, such as PFKFB3, promotes collagen synthesis and cell proliferation through ERK1/2-dependent calpain-2 phosphorylation [10]. In the present study, we combined single-cell sequencing and bulk transcriptomic analyses to identify key regulatory glycolysis-related molecules in the fibroblasts of IPAHA patients for the first time, thereby providing novel potential therapeutic targets for IPAHA.

Previous studies have not only found a metabolic shift from mitochondrial respiration to glycolysis in PASMCs obtained from lung tissues of PAH patients, but also an increase in glycolysis in animal models of PH [31, 32]. Based on single-cell RNA sequencing data, we found increased glycolytic activity in IPAHA fibroblasts compared with controls, indicating that fibroblast-mediated glycolysis is critical in mediating vascular remodeling. Aerobic glycolysis has also been shown to be present in pulmonary artery fibroblasts. In addition, the HIF-1α pathway and miR-124/PTBP1/





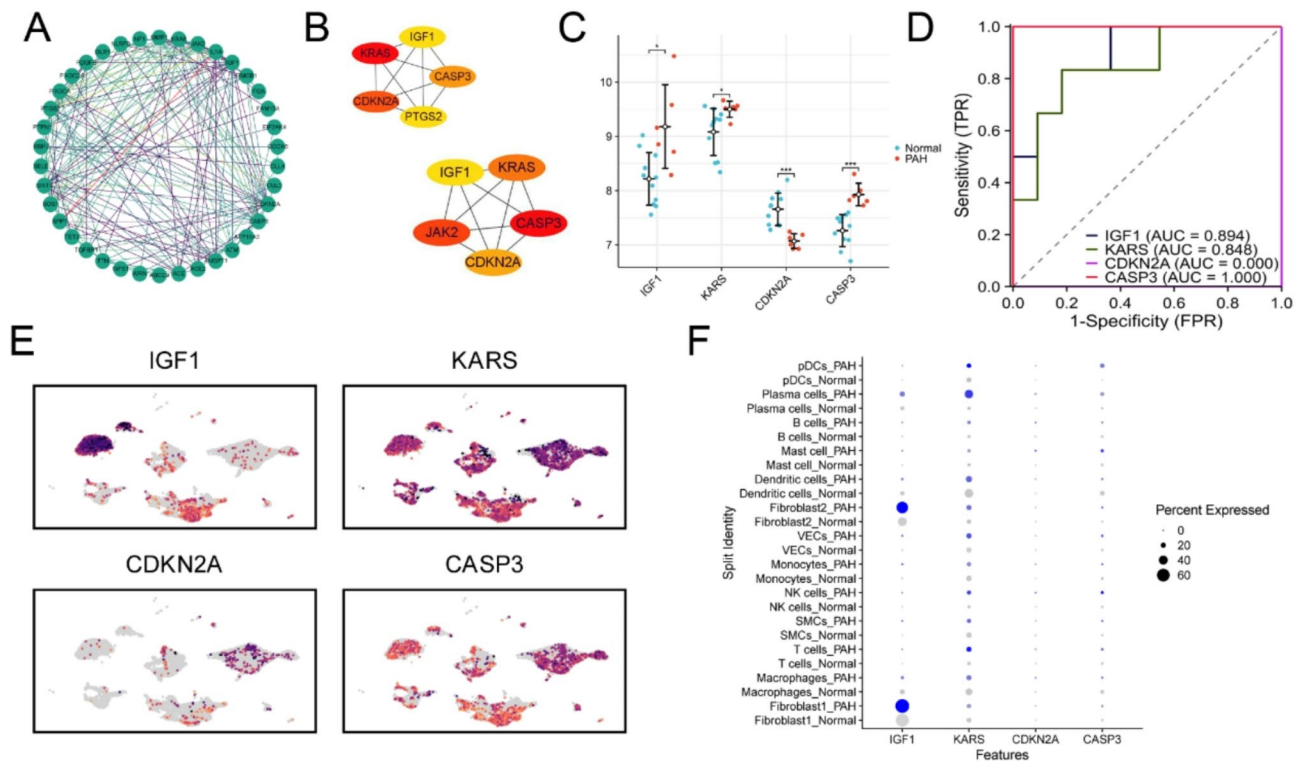
**Fig. 5** Enrichment analysis and Venn diagram. **(A)** Differential gene analysis of the GSE113439 dataset. **(B)** GO enrichment analysis of DEGs. **(C)** Overlapping genes among the module genes, DEGs and genes obtained from GeneCards. **(D)** KEGG analysis of DEGs

PKM1/2 pathway have been shown to induce the Warburg effect in fibroblasts [33]. Overexpression of miR-124 or knockdown of PTBP1 reversed glycolysis and inhibited the proliferative phenotype, thereby delaying pulmonary vascular remodeling. Thus, glycolysis may play an important role in the development of IPAH.

Transcriptomic studies of IPAH have proved to be valuable in elucidating the pathogenic mechanism of IPAH. For example, transcriptomic analysis revealed that abnormal metabolic processes, such as the dysregulation of iron metabolism, contributed to the pathogenicity of IPAH [34]. Here, we evaluated the role of glycolysis in IPAH further by analyzing transcriptomic data to determine the glycolytic metabolic activity and identify key molecules associated with IPAH. We found that IGF1, KARS and CASP3 expression levels were increased in IPAH, while CDKN2A expression levels were decreased. Our transcriptomic findings were validated by qPCR. In terms of IPAH treatment, 2-deoxy-D-glucose and metformin regulate cellular metabolism by inhibiting key glycolytic enzymes (e.g., HK2, PFKFB3), thereby suppressing PASM C proliferation and vascular remodeling, demonstrating potential therapeutic prospects [35, 36]. Based on the key molecules identified in this study (e.g., IGF1, KARS), we further propose the possibility of developing novel therapeutic strategies targeting glycolytic metabolism. In the future, research integrating single-cell metabolomics analysis will provide a more comprehensive understanding of the mechanisms underlying cellular metabolic reprogramming.

IGF1 has been shown to induce the Warburg metabolic phenotype in cells by regulating the interactions between multiple signaling pathways including activation of key glycolytic enzymes and MAPK [37]. KARS not only participates in protein synthesis but also influence glycolytic metabolism by regulating mitochondrial function. By modulating amino acid metabolism and protein synthesis, KARS could indirectly alter cellular metabolic states and influence apoptosis sensitivity [38]. Studies have demonstrated that CASP3 plays a significant role in apoptosis and metabolic





**Fig. 6** Identification and reanalysis of key glycolysis-related genes. **(A)** PPI network of 43 overlapping genes. **(B)** Different algorithms were used to identify the top five genes with the highest degree potential for IPAH. **(C)** Expression levels of key glycolysis-related genes in the GSE53408 dataset. **(D)** ROC curves of key glycolysis-related genes in the GSE53408 dataset. **(E)** Expression levels of key glycolysis-related genes in different cell subpopulations. **(F)** Expression levels of key glycolysis-related genes in different cell subpopulations of controls and IPAH patients

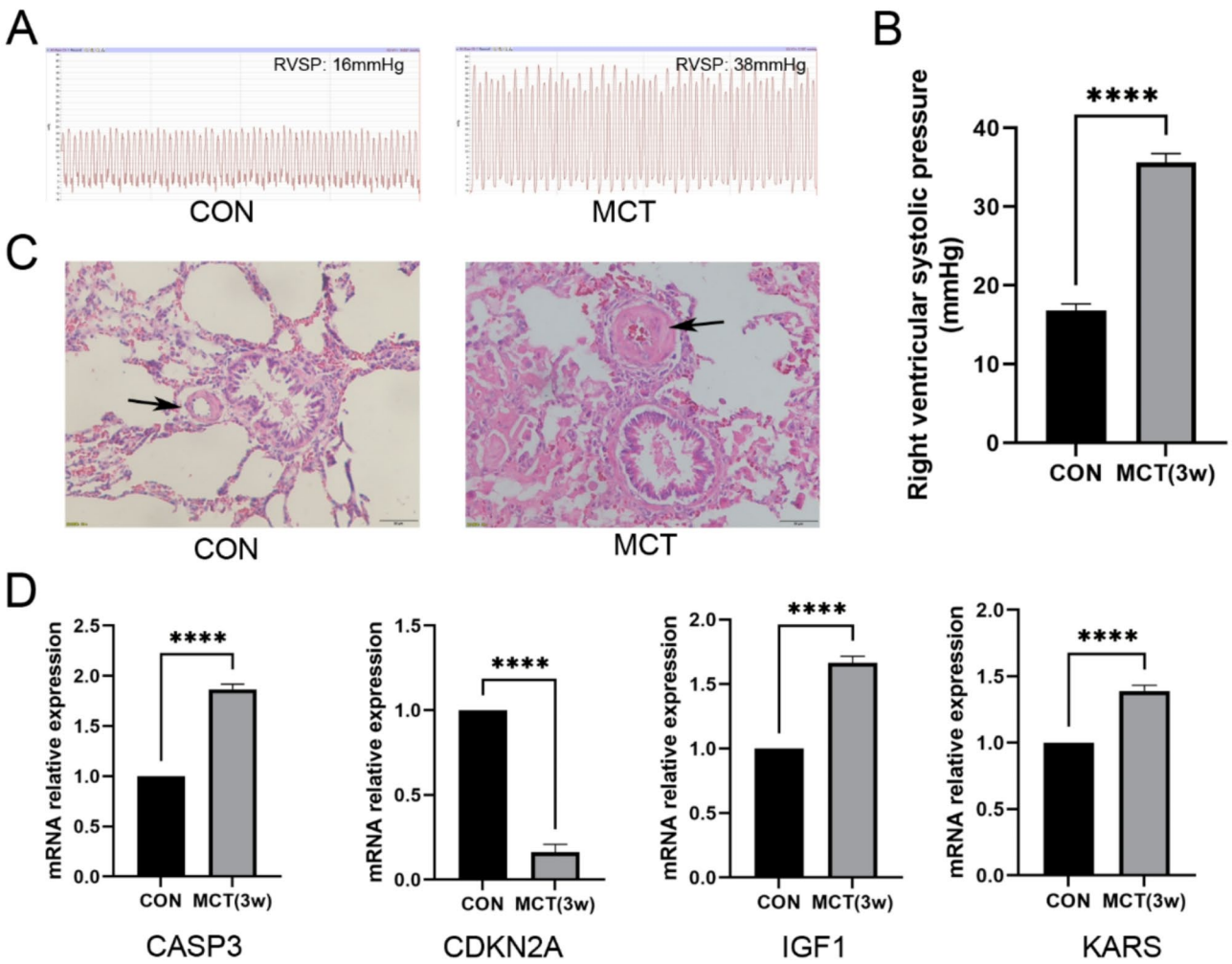
reprogramming. The activation of CASP3 may lead to mitochondrial dysfunction, thereby promoting cellular reliance on glycolysis to meet energy demands [39]. Additionally, in certain cancers, the loss or mutation of CDKN2A has been shown to enhance glycolytic pathways, thereby promoting the rapid proliferation of tumor cells [40]. Therefore, we hypothesized that IGF1, KARS, CASP3 and CDKN2A are key factors in regulating fibroblast glycolysis, which in turn promotes pulmonary vascular remodeling. Further studies are needed to understand the specific roles of IGF1, KARS, CASP3 and CDKN2A in mediating glycolysis in IPAH.

There are some limitations associated with this study. First, the single-cell RNA-seq analysis in this study is based on 3 control and 3 IPAH samples. While this sample size provides preliminary insights into disease-related cellular heterogeneity and molecular features, it may limit statistical power and the generalizability of the findings. The small sample size might result in insufficient identification of rare cell subpopulations and the omission of certain differentially expressed genes. Future studies should expand the sample size to validate the current findings and further explore the cellular and molecular mechanisms

of IPAH. Secondly, due to the relatively limited sample size of the GSE53408 dataset, there is a potential risk of overfitting in model evaluation. Thirdly, although our RNA-seq analysis determined the role and identified key glycolysis-related molecules at the transcriptional level, as well as experimentally verified the expression levels of these genes, functional experiments were not performed to validate the role of fibroblasts and hub genes of glycolysis in IPAH. Thus, in the future, functional experiments will be carried out to further validate our findings. Finally, we used AUCcell and scMetabolism for metabolic pathway scoring. While these methods effectively assess the activity of predefined gene sets, they may not fully capture all cell type-specific metabolic features.

## Conclusion

In conclusion, we found that glycolysis levels were significantly elevated in IPAH and that fibroblasts were the predominant effector cells regulating glycolysis. In addition, we identified and validated four glycolysis-related genes, which could serve as potential biomarkers for predicting the development of IPAH. Our results clarify the precise role of glycolysis in the progression of IPAH and provide novel perspectives on



**Fig. 7** Validation of key glycolysis-related genes in an in vivo model of PH in rats. **(A)** Pulmonary artery pressure of control and MCT-treated rats. **(B)** Right ventricular systolic pressure of control and MCT-treated rats. **(C)** H&E staining of pulmonary artery tissue. **(D)** qPCR analysis of key regulatory genes in glycolysis

therapeutic strategies for IPAH. In the future, combining metabolomics analysis and personalized treatment strategies, the regulation of glycolytic metabolism is expected to provide more precise therapeutic options for IPAH patients.

**Abbreviations**

AUC	Area under the curve
CASP3	Caspase-3
CDKN2A	Cyclin-dependent kinase inhibitor 2 A
DEGs	Differentially expressed genes
HIF-1α	Hypoxia inducible factor-1α
IGF1	Insulin-like growth factor-1
IPAH	Idiopathic pulmonary arterial hypertension
KARS	Kainate-type glutamate receptors
MCC	Maximal Clique Centrality
MCT	Monocrotaline
MCT	Monocrotaline
MNC	Maximum Neighborhood Component
MSigDB	Molecular Signature Database
PASMCs	Pulmonary artery smooth muscle cells
PDK1	Pyruvate dehydrogenase kinase 1

PFKFB3	Phosphofructo-2-kinase/fructose-2,6-bisphosphatase 3
PH	Pulmonary hypertension
PPI	Protein-protein interaction
PPI	Protein-protein interaction
qPCR	Quantitative PCR
RASSF1A	Ras association domain family protein-1 isoform A
ROC	Receiver operating characteristic
RV	Right ventricle
RVSP	RV systolic pressure
WGCNA	Weighted gene co-expression network analysis

**Supplementary Information**

The online version contains supplementary material available at <https://doi.org/10.1186/s12967-025-06373-x>.

Supplementary Material 1

**Acknowledgements**

We really appreciate the Crnkovic et al.[16], Mura et al.[21] and Zhao et al. [19] offered gene expression matrix data and the public database GEO for allowing access to their data.

### Author contributions

XG, YF and GW contributed equally to this work. XG designed the study and carried out the data analyses. XG, YF, JX, GW, RD, JS, BS, YW, ZW, RJ, JH, HH, LG and YZ interpreted the results. XG, RD and YF draft the manuscript, NS and BW revised the manuscript. All authors read and approved the final manuscript.

### Funding

This work was supported by the National Natural Science Foundation of China (Grant Number: 82270049 and 82470045 to BW), the National Key Research and Development Program of China (Grant Number: 2022YFC2703901 to BW), and the Practice Innovation Foundation of Harbin Medical University (Grant Number: 0111-31021220041 to RD).

### Data availability

The supporting data the conclusions of this paper can be found in the GEO database at: <https://www.ncbi.nlm.nih.gov/geo/>.

### Declarations

#### Ethics approval and consent to participate

This study was conducted in accordance with the guidelines of the Declaration of Helsinki and approved by the Institutional Research Ethics Committee of the Second Affiliated Hospital of Harbin Medical University (NO. YJSDW2022-162).

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare no competing interests.

#### Author details

<sup>1</sup>Department of Cardiology, The Key Laboratory of Myocardial Ischemia, Ministry of Education, The Second Affiliated Hospital of Harbin Medical University, NO.246 Xuefu Road, Nangang District, Harbin 150086, China

<sup>2</sup>Key Laboratory of Myocardial Ischemia, Chinese Ministry of Education, Harbin 150086, China

Received: 4 February 2025 / Accepted: 8 March 2025

Published online: 26 March 2025

### References

- Ruopp NF, Cockrill BA. Diagnosis and treatment of pulmonary arterial hypertension: a review. *JAMA*. 2022;327(14):1379–91.
- Hassoun PM. Pulmonary arterial hypertension. *N Engl J Med*. 2021;385(25):2361–76.
- Hendriks PM, Staal DP, van de Groep LD, van den Toorn LM, Chandoesing PP, Kauling RM, Mager HJ, van den Bosch AE, Post MC, Boomars KA. The evolution of survival of pulmonary arterial hypertension over 15 years. *Pulm Circ*. 2022;12(4):e12137.
- Weatherald J, Boucly A, Peters A, Montani D, Prasad K, Psotka MA, Zannad F, Gombert-Maitland M, McLaughlin V, Simonneau G, et al. The evolving landscape of pulmonary arterial hypertension clinical trials. *Lancet*. 2022;400(10366):1884–98.
- Xu W, Comhair SAA, Chen R, Hu B, Hou Y, Zhou Y, Mavrikis LA, Janocha AJ, Li L, Zhang D, et al. Integrative proteomics and phosphoproteomics in pulmonary arterial hypertension. *Sci Rep*. 2019;9(1):18623.
- Chen Z, Liu M, Li L, Chen L. Involvement of the Warburg effect in non-tumor diseases processes. *J Cell Physiol*. 2018;233(4):2839–49.
- Chen S, Zou Y, Song C, Cao K, Cai K, Wu Y, Zhang Z, Geng D, Sun W, Ouyang N, et al. The role of glycolytic metabolic pathways in cardiovascular disease and potential therapeutic approaches. *Basic Res Cardiol*. 2023;118(1):48.
- Marsboom G, Wietholt C, Haney CR, Toth PT, Ryan JJ, Morrow E, Thenappan T, Bache-Wiig P, Piao L, Paul J, et al. Lung (18)F-fluorodeoxyglucose positron emission tomography for diagnosis and monitoring of pulmonary arterial hypertension. *Am J Respir Crit Care Med*. 2012;185(6):670–9.
- Xu W, Janocha AJ, Erzurum SC. Metabolism in pulmonary hypertension. *Annu Rev Physiol*. 2021;83:551–76.
- Kovacs L, Cao Y, Han W, Meadows L, Kovacs-Kasa A, Kondrikov D, Verin AD, Barman SA, Dong Z, Huo Y, et al. PFKFB3 in smooth muscle promotes vascular remodeling in pulmonary arterial hypertension. *Am J Respir Crit Care Med*. 2019;200(5):617–27.
- Lu X, Zhang J, Liu H, Ma W, Yu L, Tan X, Wang S, Ren F, Li X, Li X. Cannabidiol attenuates pulmonary arterial hypertension by improving vascular smooth muscle cells mitochondrial function. *Theranostics*. 2021;11(11):5267–78.
- Arai MA, Sakuraba K, Makita Y, Hara Y, Ishibashi M. Evaluation of naturally occurring HIF-1 inhibitors for pulmonary arterial hypertension. *ChemBioChem*. 2021;22(18):2799–804.
- Li X, Ma S, Wang Q, Li Y, Ji X, Liu J, Ma J, Wang Y, Zhang Z, Zhang H, et al. A new integrative analysis of histopathology and single cell RNA-seq reveals the CCL5 mediated T and NK cell interaction with vascular cells in idiopathic pulmonary arterial hypertension. *J Transl Med*. 2024;22(1):502.
- Xiao Z, Dai Z, Locasale JW. Metabolic landscape of the tumor microenvironment at single cell resolution. *Nat Commun*. 2019;10(1):3763.
- Du Y, Zhang J, Guo K, Yin Y. Identification of potential biomarkers for idiopathic pulmonary arterial hypertension using single-cell and bulk RNA sequencing analysis. *Front Genet*. 2024;15:1328234.
- Crnkovic S, Valzano F, Fliesser E, Gindlhuber J, Thekkekara Puthenparampil H, Basil M, Morley MP, Katzen J, Gschwandtner E, Klepetko W, et al. Single-cell transcriptomics reveals skewed cellular communication and phenotypic shift in pulmonary artery remodeling. *JCI Insight*. 2022;7(20).
- Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM 3rd, Hao Y, Stoeckius M, Smibert P, Satija R. Comprehensive integration of single-cell data. *Cell*. 2019;177(7):1888–902. e1821.
- Zhang X, Lan Y, Xu J, Quan F, Zhao E, Deng C, Luo T, Xu L, Liao G, Yan M, et al. CellMarker: a manually curated resource of cell markers in human and mouse. *Nucleic Acids Res*. 2019;47(D1):D721–8.
- Wu Y, Yang S, Ma J, Chen Z, Song G, Rao D, Cheng Y, Huang S, Liu Y, Jiang S, et al. Spatiotemporal immune landscape of colorectal cancer liver metastasis at single-cell level. *Cancer Discov*. 2022;12(1):134–53.
- Aibar S, Gonzalez-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselman G, Rambow F, Marine JC, Geurts P, Aerts J, et al. SCENIC: single-cell regulatory network inference and clustering. *Nat Methods*. 2017;14(11):1083–6.
- 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):1104–10.
- Zhao YD, Yun HZH, Peng J, Yin L, Chu L, Wu L, Michalek R, Liu M, Keshavjee S, Waddell T, et al. De novo synthesis of bile acids in pulmonary arterial hypertension lung. *Metabolomics*. 2014;10(6):1169–75.
- Gautier L, Cope L, Bolstad BM, Irizarry RA. affy-analysis of affymetrix genechip data at the probe level. *Bioinformatics*. 2004;20(3):307–15.
- Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43(7):e47.
- Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 2008;9:559.
- Yu G, Wang LG, Han Y, He QY. ClusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. 2012;16(5):284–7.
- Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryar F, Hachilif R, Gable AL, Fang T, Doncheva NT, Pyysalo S, et al. The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res*. 2023;51(D1):D638–46.
- Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13(11):2498–504.
- Wang J, Liu C, Huang SS, Wang HF, Cheng CY, Ma JS, Li RN, Lian TY, Li XM, Ma YJ, et al. Functions and novel regulatory mechanisms of key glycolytic enzymes in pulmonary arterial hypertension. *Eur J Pharmacol*. 2024;970:176492.
- Xiao Y, Peng H, Hong C, Chen Z, Deng X, Wang A, Yang F, Yang L, Chen C, Qin X. PDGF promotes the Warburg effect in pulmonary arterial smooth muscle cells via activation of the PI3K/AKT/mTOR/HIF-1 $\alpha$  signaling pathway. *Cell Physiol Biochem*. 2017;42(4):1603–13.
- Dai J, Zhou Q, Chen J, Rexius-Hall ML, Rehman J, Zhou G. Alpha-enolase regulates the malignant phenotype of pulmonary artery smooth muscle cells via the AMPK-Akt pathway. *Nat Commun*. 2018;9(1):3850.
- Bonnet S, Michelakis ED, Porter CJ, Andrade-Navarro MA, Thebaud B, Bonnet S, Haromy A, Harry G, Moudgil R, McMurtry MS, et al. An abnormal mitochondrial-hypoxia inducible factor-1 $\alpha$ -Kv channel pathway disrupts

- oxygen sensing and triggers pulmonary arterial hypertension in fawn hooded rats: similarities to human pulmonary arterial hypertension. *Circulation*. 2006;113(22):2630–41.
33. Taniguchi K, Sugito N, Kumazaki M, Shinohara H, Yamada N, Nakagawa Y, Ito Y, Otsuki Y, Uno B, Uchiyama K, et al. MicroRNA-124 inhibits cancer cell growth through PTB1/PKM1/PKM2 feedback cascade in colorectal cancer. *Cancer Lett*. 2015;363(1):17–27.
34. Zou HX, Qiu BQ, Lai SQ, Zhou XL, Gong CW, Wang LJ, Yuan MM, He AD, Liu JC, Huang H. Iron metabolism and idiopathic pulmonary arterial hypertension: new insights from bioinformatic analysis. *Biomed Res Int*. 2021;2021:5669412.
35. Pajak B, Siwiak E, Soltyka M, Priebe A, Zielinski R, Fokt I, Ziemniak M, Jaskiewicz A, Borowski R, Domoradzki T et al. 2-deoxy-d-glucose and its analogs: from diagnostic to therapeutic agents. *Int J Mol Sci*. 2019;21(1).
36. Hu L, Zeng Z, Xia Q, Liu Z, Feng X, Chen J, Huang M, Chen L, Fang Z, Liu Q, et al. Metformin attenuates hepatoma cell proliferation by decreasing glycolytic flux through the HIF-1 $\alpha$ /PFKFB3/PFK1 pathway. *Life Sci*. 2019;239:116966.
37. Kasprzak A. Insulin-like growth factor 1 (IGF-1) signaling in glucose metabolism in colorectal cancer. *Int J Mol Sci*. 2021;22(12).
38. Young HJ, Lee JW, Kim S. Function of membranous lysyl-tRNA synthetase and its implication for tumorigenesis. *Biochim Biophys Acta*. 2016;1864(12):1707–13.
39. Yao F, Jin Z, Zheng Z, Lv X, Ren L, Yang J, Chen D, Wang B, Yang W, Chen L, et al. HDAC11 promotes both NLRP3/caspase-1/GSDMD and caspase-3/GSDME pathways causing pyroptosis via ERG in vascular endothelial cells. *Cell Death Discov*. 2022;8(1):112.
40. Zhang D, Wang T, Zhou Y, Zhang X. Comprehensive analyses of cuproptosis-related gene CDKN2A on prognosis and immunologic therapy in human tumors. *Med (Baltim)*. 2023;102(14):e33468.

## Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.