Integrated bioinformatic analysis reveals the underlying molecular mechanism of and potential drugs for pulmonary arterial hypertension

Haoru Dong[2,*], Xiuchun Li[1,*], Mengsi Cai[1,*], Chi Zhang[2], Weiqi Mao[2], Ying Wang[2], Qian Xu[2], Mayun Chen[1], Liangxing Wang[1,&], Xiaoying Huang[1]

[1]Division of Pulmonary Medicine, The First Affiliated Hospital of Wenzhou Medical University, Key Laboratory of Heart and Lung, Wenzhou 325000, Zhejiang, P.R. China
[2]The First Clinical Medical College, Wenzhou Medical University, Wenzhou 325000, Zhejiang, P.R. China
*Equal contribution

Correspondence to: Xiaoying Huang, Liangxing Wang; **email:** zjwzhxy@126.com, https://orcid.org/0000-0003-4854-0922; wzyxywlx@163.com, https://orcid.org/0000-0001-7214-2786

ABSTRACT

Pulmonary arterial hypertension (PAH) is a devastating cardiovascular disease without a clear mechanism or drugs for treatment. Therefore, it is crucial to reveal the underlying molecular mechanism and identify potential drugs for PAH. In this study, we first integrated three human lung tissue datasets (GSE113439, GSE53408, GSE117261) from GEO. A total of 151 differentially expressed genes (DEGs) were screened, followed by KEGG and GO enrichment analyses and PPI network construction. Five hub genes (CSF3R, NT5E, ANGPT2, FGF7, and CXCL9) were identified by Cytoscape (Cytohubba). GSEA and GSVA were performed for each hub gene to uncover the potential mechanism. Moreover, to repurpose known and therapeutic drugs, the CMap database was retrieved, and nine candidate compounds (lypressin, ruxolitinib, triclabendazole, L-BSO, tiaprofenic acid, AT-9283, QL-X-138, huperzine-a, and L-741742) with a high level of confidence were obtained. Then ruxolitinib was selected to perform molecular docking simulations with ANGPT2, FGF7, NT5E, CSF3R, JAK1, JAK2, JAK3, TYK2. A certain concentration of ruxolitinib could inhibit the proliferation and migration of rat pulmonary artery smooth muscle cells (rPASMCs) *in vitro*. Together, these analyses principally identified CSF3R, NT5E, ANGPT2, FGF7 and CXCL9 as candidate biomarkers of PAH, and ruxolitinib might exert promising therapeutic action for PAH.

INTRODUCTION

Pulmonary arterial hypertension (PAH) is a serious cardiovascular disease leading to right heart failure and eventually death [1]. The main pathophysiology of PAH is incrassation of the medial and intimal layers of the pulmonary arterial wall, which might result in increased pulmonary vascular resistance and haemodynamic derangements [2]. Although many genes and related biological processes have been reported to be involved in the development of PAH in recent years [3], the underlying molecular mechanism of PAH remains unclear. In terms of PAH therapy, the curative effects of specific drugs related to the prostacyclin pathway [4], endothelin pathway [5] or nitric oxide pathway [6] are not satisfactory following the support of clinical trials consisting of all PAH subtypes. Therefore, investigation of the underlying mechanisms of PAH and a search for auxiliary potential drugs for its treatment are still desperately needed.

Recently, there has been a growing trend for PAH research to utilize high-throughput technologies to explore novel diagnostic or prognostic biomarkers and therapeutic targets. For example, Wang et al. reported that YWHAB was a diagnostic biomarker for idiopathic pulmonary arterial hypertension [3], and Zhu et al. uncovered that miR-140-5p and TNF-α might be therapeutic targets for PAH [7]. Recently, many bioinformatics tools were developed and FocusHeuristics is a competitive approach to explore disease-associated genes [8]. On the other hand, "Cytohubba", a plug-in of Cytoscape, is a useful and user-friendly tool to obtain hub genes in biological network and has been widely applied [9]. However, the limited sample sizes make these studies inconsistent and can lead to unconvincing conclusions about biological functions. Therefore, it is much more critical to elucidate the underlying molecular mechanism of PAH, which will in turn provide possible routes on which potential treatments can be designed.

Regarding PAH treatment, to develop a novel drug requires substantial costs and a lengthy process of ensuring its safety and tolerance in the human body [10, 11]. By contrast, repurposing a non-novel drug with precise and new mode of action is relatively more cost-efficient and time-saving. Recently, an increasing number of free databases and online tools have been developed, which could help our repositioning of known drugs. For example, the Connectivity map (CMap; https://clue.io/) database serves as a potentially useful tool for drug screening that can predict molecular targeted agents based on DEGs [12]. Combined with high-throughput data, it facilitates the repurposing of known drugs which has passed the toxicology and dosage analyses. Recently, a study reanalysed CMap whole-genome transcriptome data by combining 26 similarity scores with 6 different heuristics. It provided an insight to find a known drug by comparing the effects on the transcriptomes [13]. Then, molecular docking simulation and *in vitro* experiments can be performed to verify the prediction results.

In the present study, we integrated 3 datasets containing 132 human lung tissue samples to obtain DEGs. Then, Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) enrichment analyses were performed based on the DEGs. After obtaining the hub genes from the protein-protein interaction (PPI) network, Gene set enrichment analysis (GSEA) and gene set variation analysis (GSVA) were utilized to further reveal the potential biological functions of the hub genes [14]. Then, the CMap database was employed to search for potential compounds for PAH treatment. Molecular docking simulation between candidate compound and the proteins encoded by hub genes was conducted to validate their prospective

application in PAH therapy. Ultimately, *in vitro* experiments verified the effects of ruxolitinib in PAH.

RESULTS

Identification of DEGs related to PAH

The study design is illustrated in Figure 1. We downloaded the raw data of GSE113439, GSE53408, and GSE117261 from the Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/gds/) database [15]. Supplementary Table 1 shows the general characteristics of the three datasets. Next, they were inputted into R by "affy" package and the missing data were filled by "impute" package [16]. Then, "sva" package was used to remove batch effects [17]. The Q-Q plot shows the effect of batch removal (Supplementary Figure 1). The heterogeneity of the datasets before and after batch effect removal was examined by "pca" package (Figures 2A, 2B). To annotate the data, we used R software to match each probeset to the corresponding gene symbol according to "hugene10sttranscriptcluster.db" package. If multiple probesets mapped to a same gene symbol, the maximum expression value was selected. In total, 18,820 genes were obtained.

To identify the DEGs, the 132 samples were divided into 47 control samples (CON group) and 85 pulmonary arterial hypertension samples (PAH group). "Limma" package was utilized to conduct the differential gene expression analysis, and a total of 151 significant DEGs (103 upregulated and 48 downregulated) were identified (adjusted P value < 0.05, FC > 1.5; Supplementary Table 2) [18]. All DEGs were used to perform hierarchical clustering analysis, and the heatmap showed evidently different expression between the two groups (Figure 2C).

GO and KEGG pathway enrichment analyses of DEGs

We carried out GO enrichment analysis on DEGs with the help of "clusterprofiler" package to explore their biological features [19]. Biological process (BP) terms showed that the DEGs were enriched in "regulation of inflammatory response", "peptidyl-tyrosine phosphorylation" and "peptidyl-tyrosine modification" (Figure 3A), suggesting that the inflammatory response might play an important role in the development of PAH. In terms of cellular component (CC), the terms "extracellular matrix", "membrane raft" and "membrane microdomain" were significantly enriched (Figure 3B). Therefore, we hypothesized that the DEGs mainly played roles in the extracellular matrix. The major enriched molecular function (MF) terms of the DEGs

were "receptor ligand activity", "integrin binding" and "sulfur compound binding" (Figure 3C).

Moreover, to explore the more profound function of DEGs, we conducted KEGG pathway enrichment analysis with "clusterprofiler" package [19]. Based on the results, the following pathways were significantly enriched in DEGs: "PI3K-AKT signalling pathway", "focal adhesion" and "hypertrophic cardiomyopathy" (HCM) (Figure 3D).

PPI network analysis and hub gene recognition

To find the hub genes, a PPI network was constructed with Search Tool for the Retrieval of Interacting Genes (STRING; https://string-db.org/) [20], and a total of 128 nodes and 858 edges were in the PPI network (Figure 4A). Next, we took advantage of Cytoscape's plug-in "Cytohubba" to explore the PPI network. To have a more credible result, all 12 algorithms were utilized to calculate the degree of connectivity of DEGs. We selected top 40 genes in each algorithm (As many genes scored equally in the MNC, Stress, BottleNeck, and EcCentricity algorithms, the number of obtained candidate genes was over 40 (Supplementary Table 3)). Then we intersected the results of 12 algorithms. Finally, five hub genes (CSF3R, NT5E, ANGPT2, FGF7, and CXCL9) were obtained (Figure 4B and Supplementary Table 4). The five hub genes are labelled in the volcano plot to show their expression levels (Figure 4C).

GSEA and GSVA to reveal the potential functions of hub genes

To further reveal the potential function of the five hub genes, GSEA was performed based on each single hub gene (Figures 5A, 5C, 5E, 5G, 5I). The NT5E, CSF3R and ANGPT2 groups were enriched in the spliceosome pathway. Meanwhile, the cell cycle was enriched in FGF7, ANGPT2 and CSF3R groups, whereas focal adhesion was enriched in the FGF7, ANGPT2 and NT5E groups. The other enriched items were specific to single groups. For example, the chemokine signalling pathway, purine metabolism, cytokine-cytokine receptor interaction and the MAPK signalling pathway were
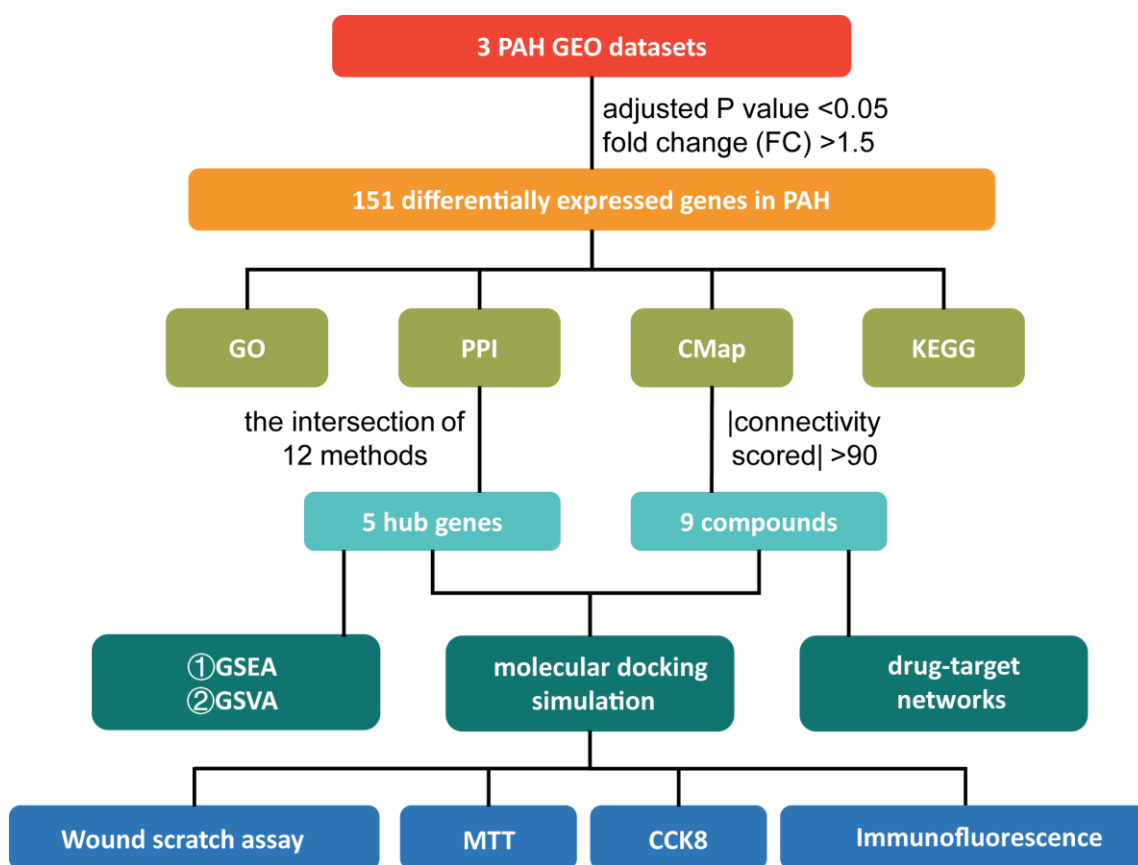
Figure 1. The study design. PAH: pulmonary arterial hypertension; GEO: Gene Expression Omnibus; GO: Gene Ontology; PPI: protein-protein interaction; CMap: Connectivity Map; KEGG: Kyoto Encyclopedia of Genes and Genomes; GSEA: Gene Set Enrichment Analysis; GSVA: Gene Set Variation Analysis; MTT: 3-[4,5-Dimethylthiazol-2-yl]-2,5-diphenyltetrazolium Bromide; CCK-8: Cell Counting Kit-8.
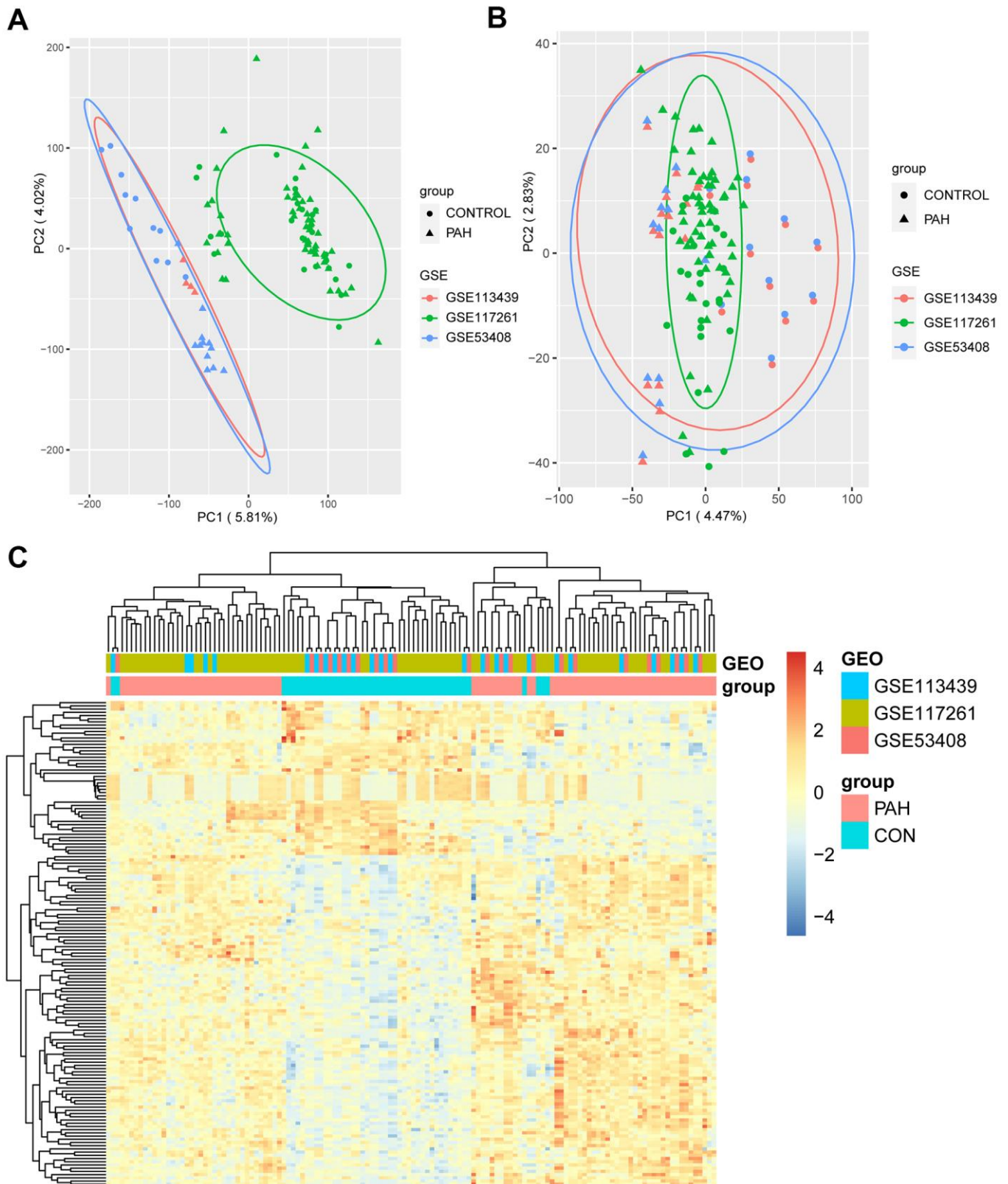
Figure 2. The result of DEGs identification. (**A**) Principal component analysis (PCA) before batch effects removement. (**B**) PCA after batch effects removement. (**C**) Heatmap of 151 DEGs screened by "Limma" package. In the differentiating gene sets (GSE113439, GSE117261 and GSE53408), samples were sorted by columns, and genes were sorted by rows. Cyan squares represented the control group, and red squares represented the PAH group. DEGs: differentially expression genes.

enriched in a group of CXCL9. Most of the pathways were closely associated with cell proliferation.

Furthermore, the top 20 GSVA terms of the five groups are shown in Figures 5B, 5D, 5F, 5H, 5J, and the details of the 20 terms are shown in Supplementary Tables 5–9. The top GSVA term in each result was "one carbon pool by folate", "propanoate metabolism", "graft versus host disease", "valine leucine and isoleucine degradation", and "protein export". In summary, these results confirmed that these hub genes contributed to proliferation processes.

CMap analysis and molecular docking simulation

To find drugs for PAH therapy, we searched the CMap database. Nine candidate compounds (lypressin, ruxolitinib, triclabendazole, L-BSO, tiaprofenic acid, AT-9283, QL-X-138, huperzine-a, and L-741742) were

considered with high levels of confidence (Table 1). Their 2D structures were provided by PubChem (https://pubchem.ncbi.nlm.nih.gov/; Supplementary Figure 2) [21]. Because the 2D structure of QL-X-138 was not retrieved, only eight 2D structures are presented. The targets of these nine compounds provided by the CMap database were utilized to construct drug-target networks in Cytoscape software (Supplementary Figure 3).

Furthermore, molecular docking simulation was utilized to delve into the possible therapeutic mechanisms of these drugs. Five hub genes were supposed to serve as potential therapeutic targets in PAH. Among the nine compounds, lypressin ranked first, but its 3D structure was not provided on ZINC (http://zinc.docking.org/) [22]. In addition, ruxolitinib, described as a Janus-associated kinase (JAK) inhibitor in the CMap database, obtained the second highest score (97.86). Activation of
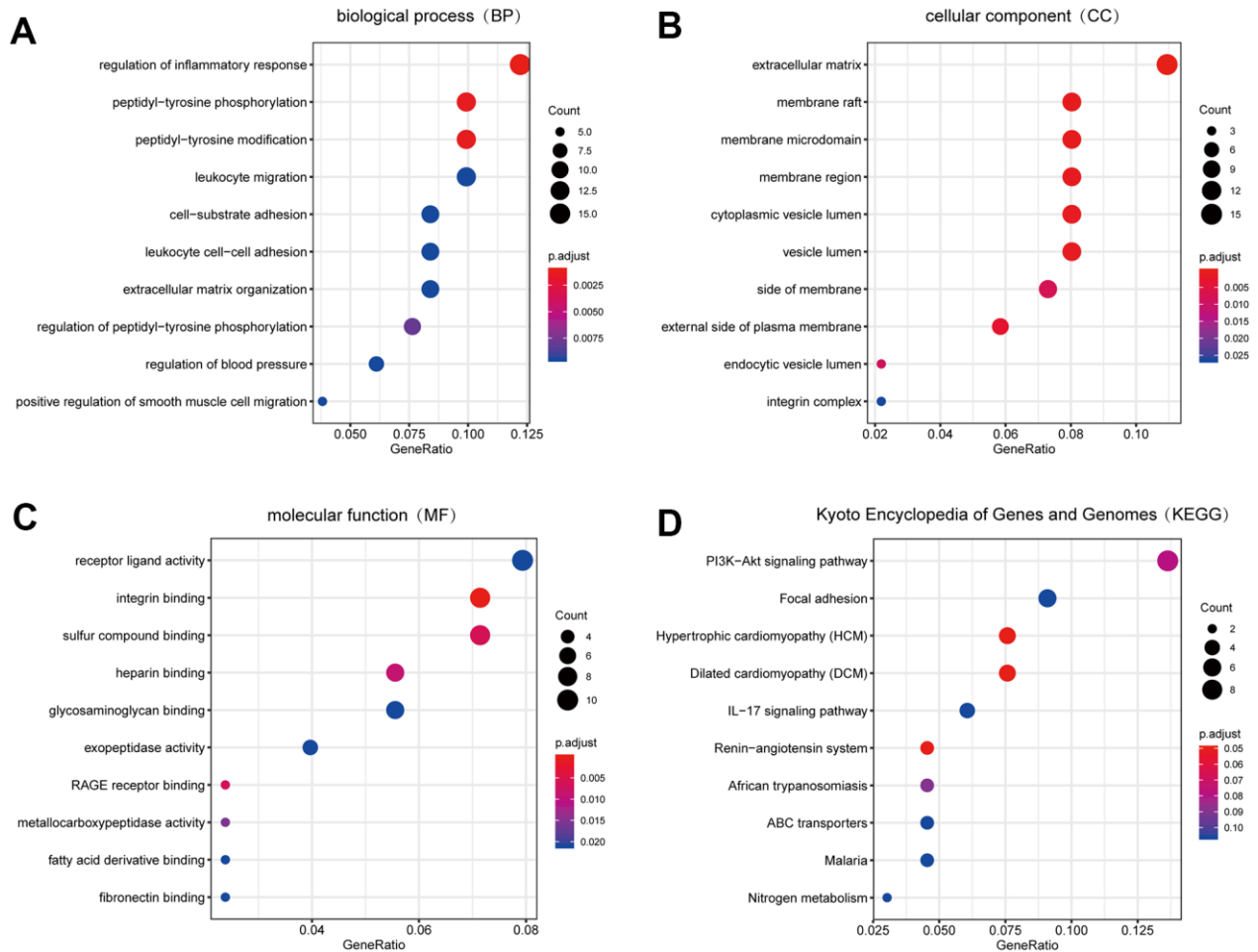
**Figure 3. The result of GO and KEGG pathway enrichment analyses.** (**A**) Biological process GO terms for DEGs. (**B**) Cellular component GO terms for DEGs. (**C**) Molecular function GO terms for DEGs. (**D**) KEGG pathways for DEGs. Top 10 sorted by GeneRatio of GO terms or KEGG pathways were shown. GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes. DEGs: differentially expression genes.
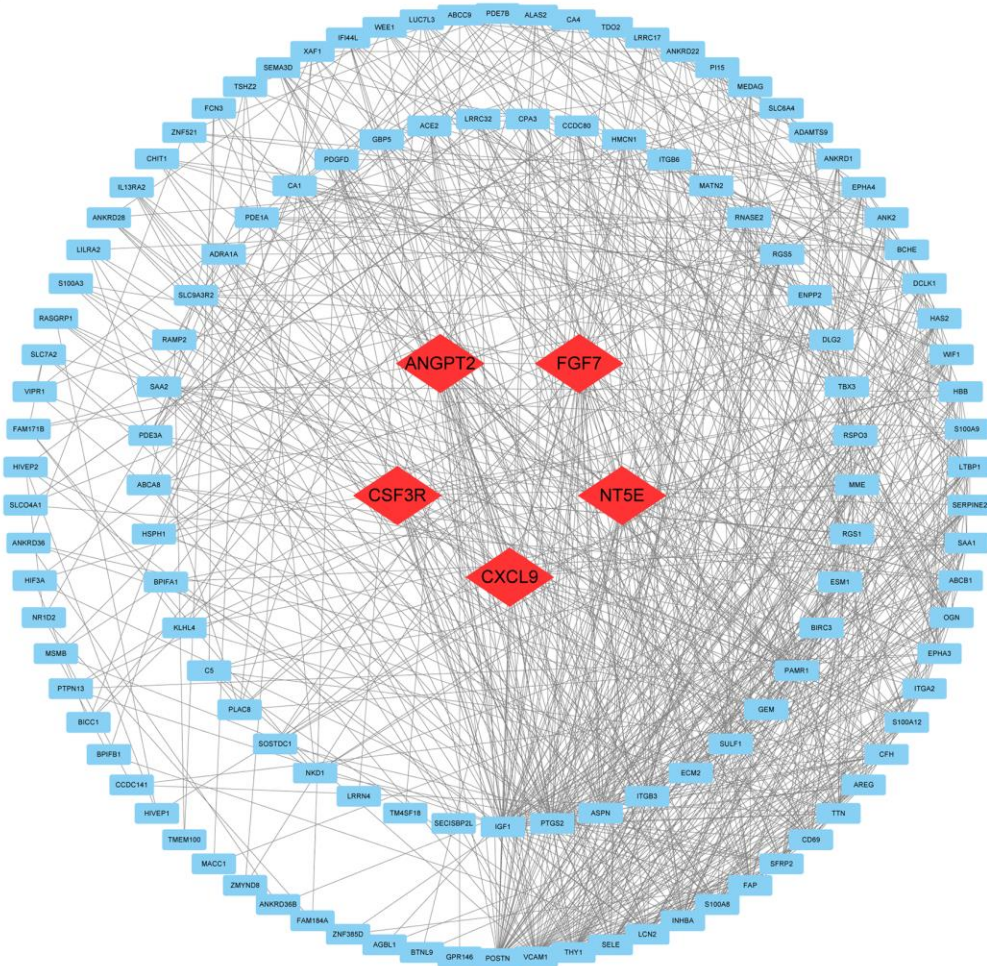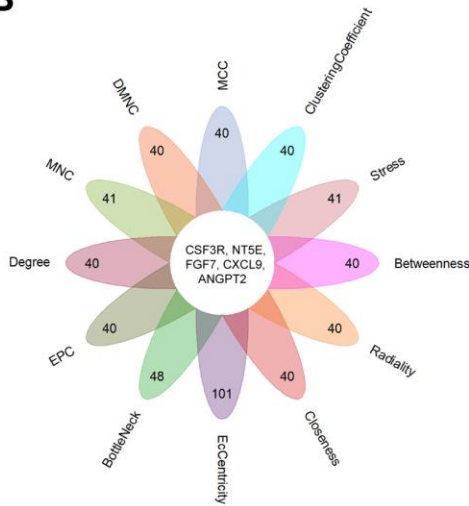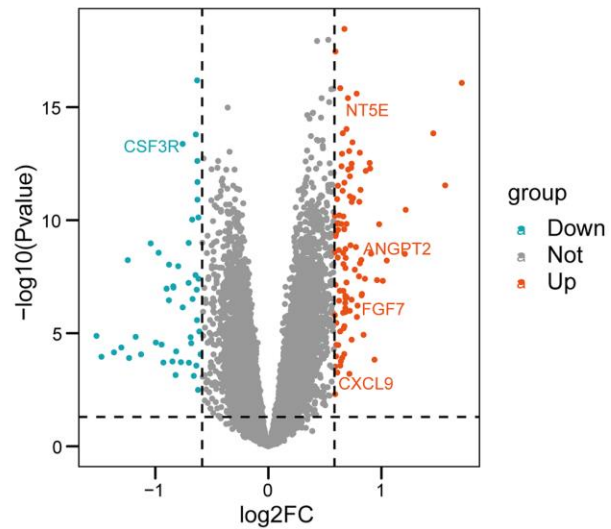
Figure 4. PPI network construction and hub genes analyses. (**A**) PPI network. Quadrangles represented proteins and lines represented interactions between proteins. hub genes were in red color. (**B**) Flower plot of results from twelve algorithms. (**C**) Volcano plot of all genes. Orang dots represented 103 up-regulated genes and cyan dots represented 48 down-regulated genes. CSF3R, NT5E, ANGPT2, FGF7, CXCL9 marked in the figure were hub genes. PPI: protein-protein interaction.
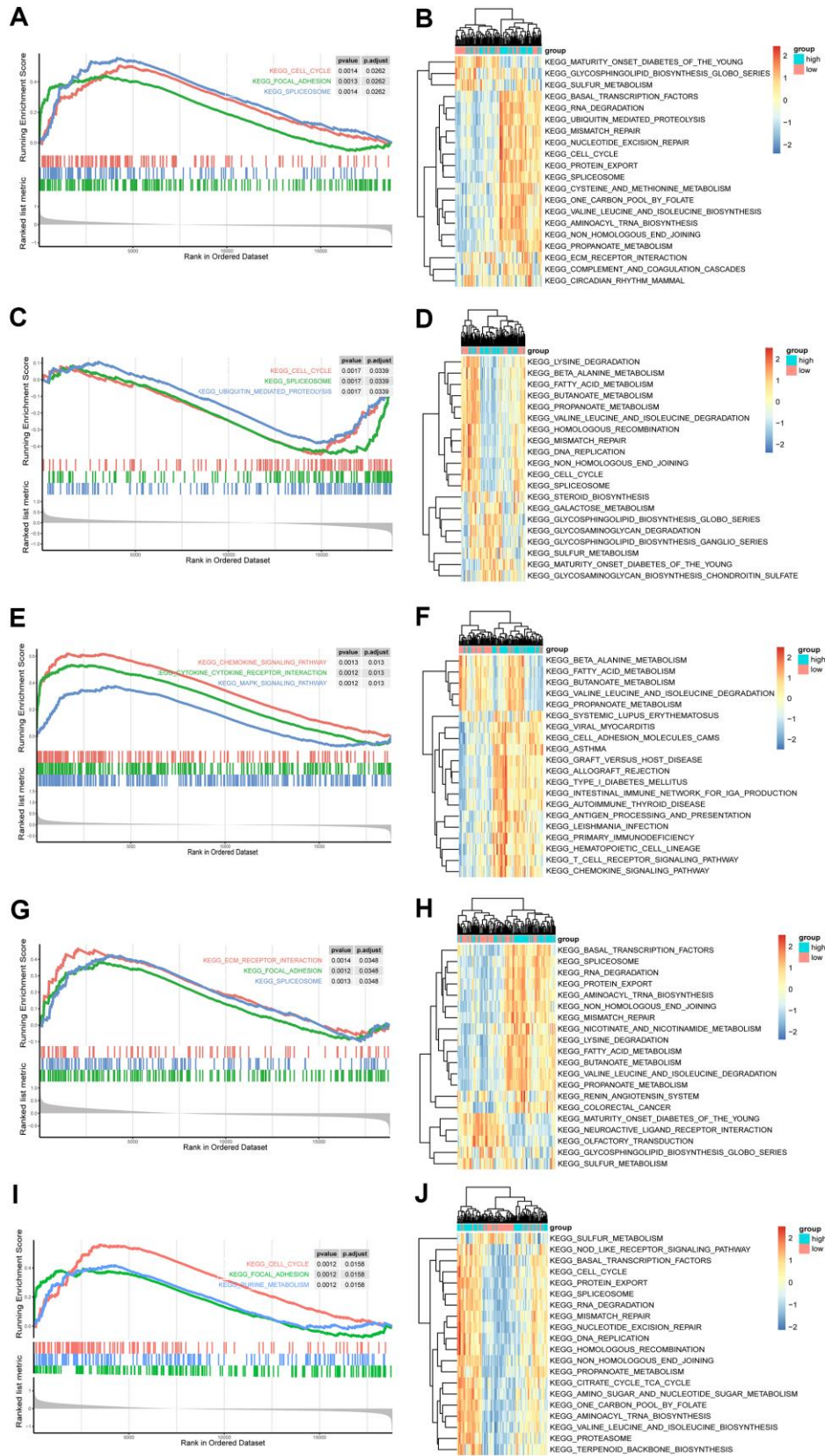
Figure 5. GSEA and GSVA of hub genes in the PAH. (**A**, **C**, **E**, **G**, **I**) Top 3 KEGG pathways in the high-expression group of single hub genes. (**A**) ANGPT2; (**C**) CSF3R; (**E**) CXCL9; (**G**) NT5E; (**I**) FGF7. (**B**, **D**, **F**, **H**, **J**) GSVA-derived clustering heatmaps of top 20 differentially expressed pathways for each hub gene. (**B**) ANGPT2; (**D**) CSF3R; (**F**) CXCL9; (**H**) NT5E; (**J**) FGF7.

Table 1. The result of CMap.

Drug name	Score	Description	Target
lypressin	98.84	Vasopressin receptor agonist	AVPR1A, AVPR1B, AVPR2
ruxolitinib	97.86	JAK inhibitor	JAK1, JAK2, TYK2, JAK3
triclabendazole	94.73	Microtubule inhibitor	DNMT1
L-BSO	94.66	Glutathione transferase inhibitor	GCLM
tiaprofenic-acid	93.96	Cyclooxygenase inhibitor	PTGS2, PTGS1
AT-9283	91.00	JAK inhibitor	AURKA, AURKB, ABL1, BCR, FLT3, JAK2, JAK3, RPS6KA6, STK17A
QL-X-138	90.35	MTOR inhibitor	BTK, JAK3, MKNK2, MTOR, PRKDC
L-741742	-90.13	Dopamine receptor antagonist	DRD4, DRD3, SCN1A, SCN3A
huperzine-a	-90.46	Acetylcholinesterase inhibitor	ACHE

the JAK-STAT signalling pathway has been reported to induce the transcription of pro-angiogenesis and pro-inflammatory genes, leading to the progression of PAH [23]. Consequently, ruxolitinib was selected to dock with the proteins encoded by hub genes. We downloaded the 3D structures of NT5E (PDB: 4H2F), ANGPT2 (PDB: 4JZC), FGF7 (PDB: 1QQL), CSF3R (PDB: 2D9Q) and ruxolitinib to perform the molecular docking simulations, whereas the 3D structure of CXCL9 was not provided on RCSB PDB (https://www.rcsb.org/) [24]. In addition, JAK1 (PDB: 4GS0), JAK2 (PDB: 2B7A), JAK3 (PDB: 3ZC6) and TYK2 (PDB: 4GFO) as known targets of ruxolitinib served as positive controls. The binding energies between ruxolitinib and ANGPT2, FGF7, NT5E, CSF3R, JAK1, JAK2, JAK3, TYK2 were -6.42, -5.25, -6.30, -3.69, -3.59, -3.65, -3.90, -4.73 kcal/mol, respectively (Figure 6). It indicated that ANGPT2, FGF7 and NT5E may be the targets of ruxolitinib in treatment of PAH (The binding energies of these 3 proteins were higher than the positive controls). Other details of the results, such as the hydrogen bond, atomic distance and binding site data, are shown in Figure 6. From the above findings, ruxolitinib may have an unexpected effect in the treatment of PAH.

Ruxolitinib significantly inhibits the proliferation and migration abilities of hypoxia-induced rPASMCs

To further explore the effects of ruxolitinib in PAH, we investigated the regulation of rPASMCs migration and proliferation by ruxolitinib. Firstly, the cytotoxicity of ruxolitinib was assessed by 3-[4,5-Dimethylthiazol-2-yl]-2,5-diphenyltetrazolium Bromide (MTT) assay and evaluated by the 50% inhibitory concentration (IC50) values (6.719 µM under hypoxic and 58.26 µM under normoxic) (Figures 7A, 7B). Therefore, the concentrations of ruxolitinib was used including 0, 0.1, 0.5, 1 and 2.5 µM. The result of CCK-8 assay indicated the inhibited proliferation in PASMCs with the treatment of ruxolitinib (Figure 7C). With concentration increasing, the inhibition effect became more obvious. Coincidentally, immunofluorescence assay demonstrated that administration of ruxolitinib significantly decreased the expression of Ki67 (a marker of proliferation) in hypoxia-induced rPASMCs (Figures 7D, 7F). Moreover, scratch wound assay showed that rate of closure was increased under hypoxia exposure for 12h and 24h, nevertheless hypoxia-induced migration was significantly suppressed by ruxolitinib (Figures 7E, 7G). Collectively, these data strongly suggest that ruxolitinib could effectively inhibit cell proliferation and cell migration under hypoxia exposure *in vitro*.

DISCUSSION

Here, we identified five hub genes of PAH and potential mechanism of PAH with bioinformatic analyses. Further studies are urgently needed to verify the hub genes. Furthermore, nine candidate compounds were predicted through the CMap database, in which ruxolitinib was considered the drug with the most clinical potential given its ability to target ANGPT2, FGF7 and NT5E, especially ANGPT2 and inhibit proliferation and migration of hypoxia-induced rPASMCs. These findings will shed new light on promising therapeutic strategies to treat PAH.

PAH is a serious pulmonary vascular disease caused by multiple risk factors [17]. Although great progress has been made in the study of PAH, there is still a lack of effective methods to diagnose and treat PAH. Fortunately, with the development of high-throughput technologies, an increasing number of novel biomarkers and therapeutic targets for PAH have been uncovered. For example, Sun et al. found that Smad9, BMPR2, Eng, and IL4 were differentially expressed in PAH mice. On the other hand, lncRNAs NR-036693, NR-027783, NR-033766, and NR-001284 played an important role in PAH pathology and might serve as therapeutic targets for PAH. In the present study, we

integrated three human lung tissue datasets, obtaining 151 significant DEGs to explore the critical genes, pathways and potential drugs that are associated with PAH.

According to the integrated expression data, 151 DEGs were identified between the CON group and PAH group. The results of the BP term GO annotation analysis indicated that the DEGs were markedly enriched in the inflammatory response, which was consistent with the previous demonstration that inflammation was emerging as a key disease-related factor in PAH [25]. The enriched CC terms of DEGs were related to the extra-cellular matrix, which was in accordance with published literature showing that the extracellular matrix plays a central role in the pathogenesis of PAH [26]. The MF analysis of GO terms revealed that the DEGs were the most significantly enriched in receptor ligand activity, suggesting that receptor-ligand interactions are an essential component of PAH [27]. Additionally, KEGG analysis suggested that the DEGs were the most significantly enriched in the PI3K-AKT signalling pathway, which was consistent with the previous demonstration that pulmonary hypertension involves crosstalk with proliferation and apoptosis mechanisms [28, 29]. The results of the enrichment analysis exactly support the validity of the DEGs obtained in the previous analysis.
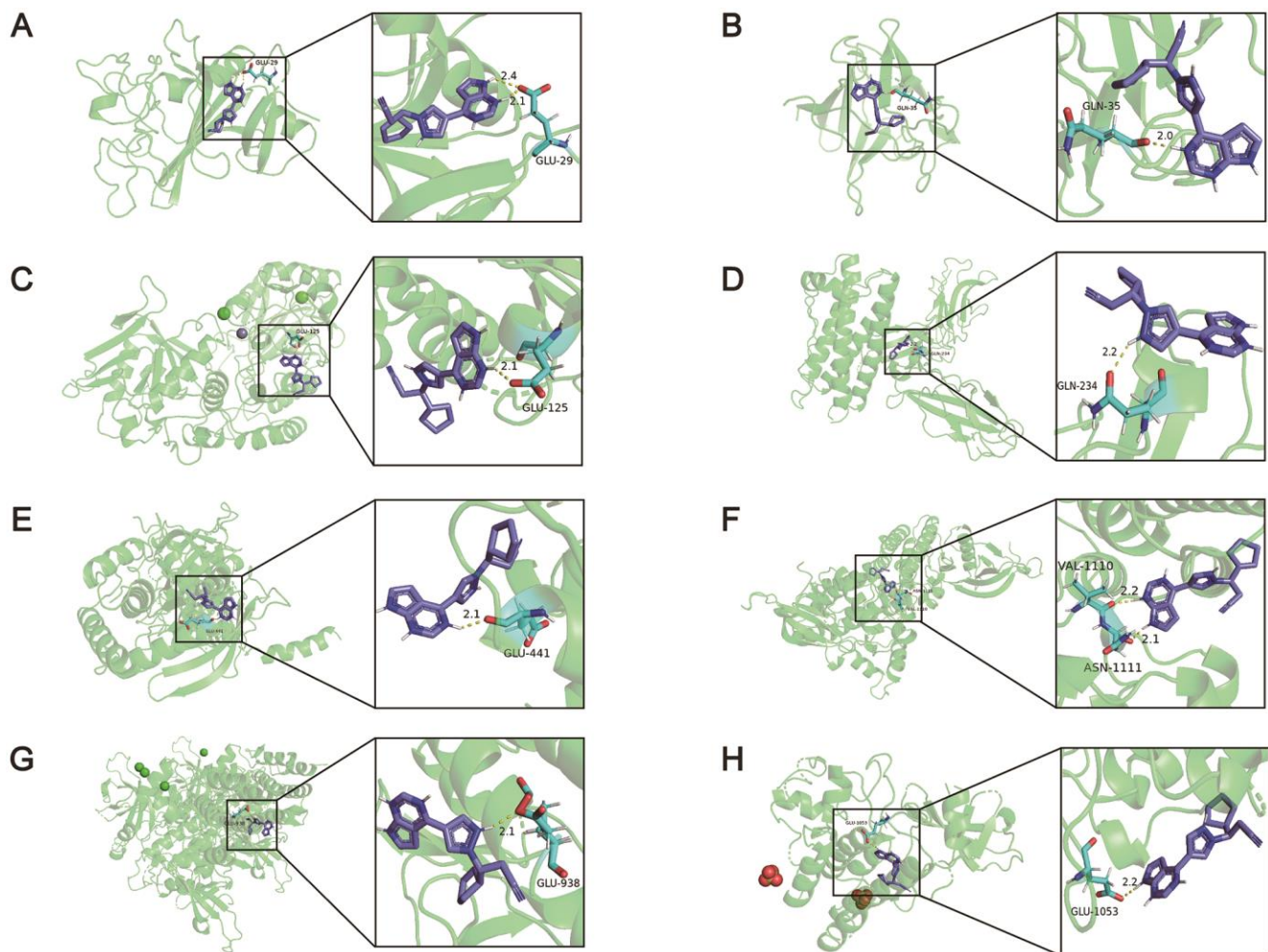
Figure 6. The results of the molecular docking simulations. (**A**) There were two hydrogen bonding between the amino acid residue of ANGPT2 (PDB: 4JZC) (GLU29) and ruxolitinib, and the distance between the atoms was 2.1Å and 2.4Å. (**B**) The amino acid residue of FGF7 (PDB: 1QQL) bound to ruxolitinib was GLN35, and the distance was 2.0Å. (**C**) The amino acid residue of NT5E (PDB: 4H2F) bound to ruxolitinib was GLU125, and the distance was 2.1Å. (**D**) The amino acid residue of CSF3R (PDB: 2D9Q) bound to ruxolitinib was GLN234, and the distance was 2.2Å. (**E**) The amino acid residue of JAK1 (PDB: 4GS0) bound to ruxolitinib was GLU441, and the distance was 2.1Å. (**F**) The amino acid residues of JAK2 (PDB: 2B7A) bound to ruxolitinib were VAL1110 and ASN1111, and the distance were 2.2Å and 2.1Å. (**G**) The amino acid residue of JAK3 (PDB: 3ZC6) bound to ruxolitinib was GLU938, and the distance was 2.1Å. (**H**) The amino acid residue of TYK2 (PDB: 4GFO) bound to ruxolitinib was GLU1053, and the distance was 2.2Å.
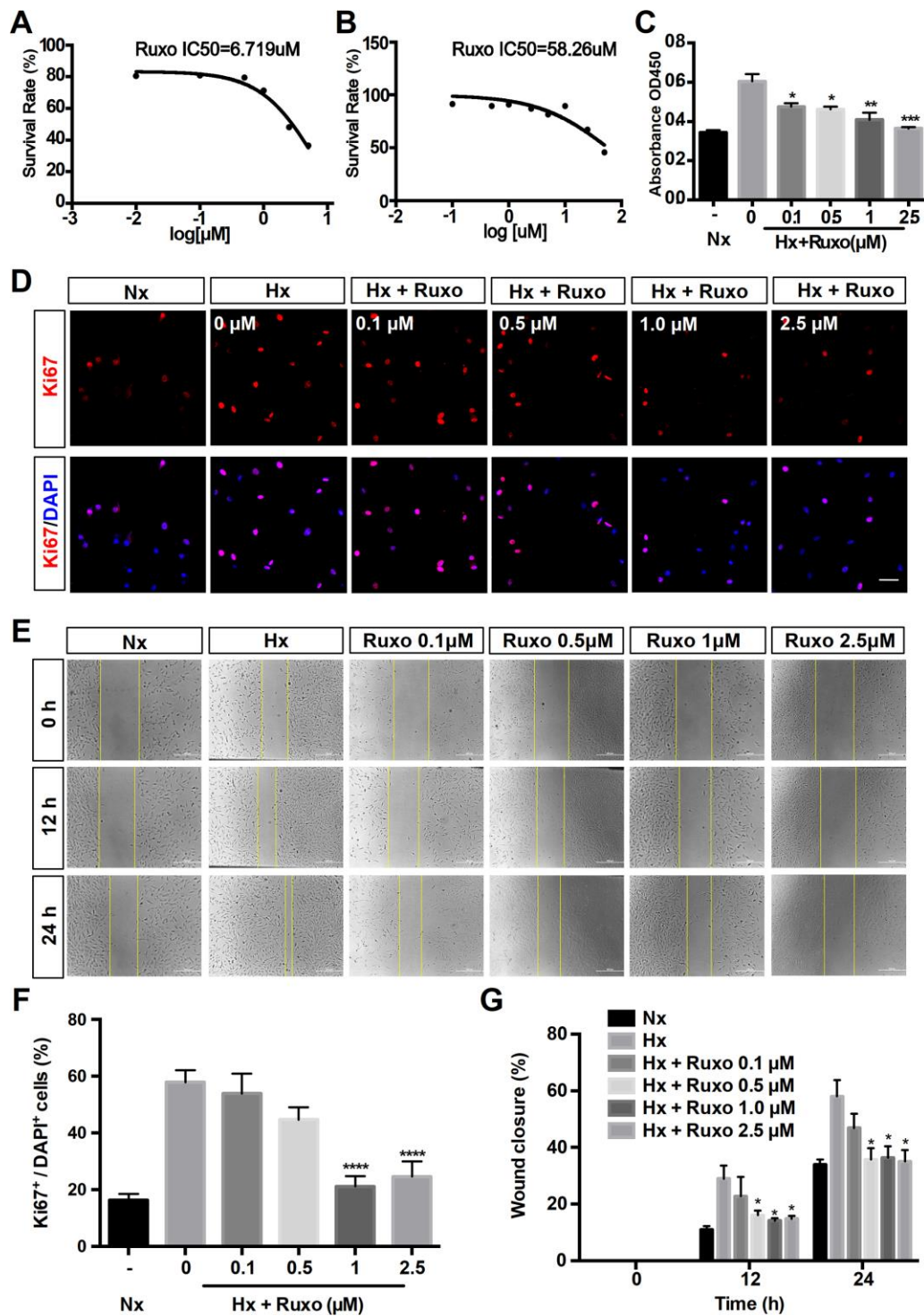
Figure 7. Ruxolitinib inhibits hypoxia-induced rPASMCs proliferation and migration. Nx = Normoxia group; Hx = Hypoxia group; Hx + Ruxo = Hypoxia plus Ruxolitinib group. (**A**, **B**) Determination of IC50 values. RPASMCs were incubated with ruxolitinib at different concentrations (0, 0.1, 0.5, 1 and 2.5 μM) for 24 h under hypoxia (**A**) and normoxia (**B**), viable cell number was determined by MTT assay. The results yielded IC50 values were 6.719 μM (**A**) and 58.26 μM (**B**), respectively. (**C**) CCK-8 assay for cell proliferation. RPASMCs were treated with indicated concentrations of ruxolitinib for 24 h, and the cell proliferation was determined by CCK-8 assay. (**D**, **F**) Immunofluorescence staining of Ki67(red) in rPASMCs of indicated groups. RPASMCs were treated with or without indicated concentrations of ruxolitinib for 24 h under hypoxic conditions, the untreated cells were treated as a normoxia control group. Cells nuclear were counterstained with DAPI (blue). (**E**, **G**) Wound scratch assay. RPASMCs were treated with ruxolitinib at a specified concentration in the presence or absence of oxygen for 12h and 24h, migration capabilities were represented by relative migration distances. Data are presented as mean ± SEM, *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001.

In the present study, five hub genes (FGF7, CXCL9, NT5E, ANGPT2 and CSF3R) were identified. Among them, FGF7 was upregulated in PAH and confirmed to be inhibited by miR-455-3p-1 via the RAS/ERK signalling pathway [30]. Belperio et al. showed that there was no significant difference in CXCL9 expression in the plasma of normal and PAH patients [31]. However, in this study, we used a much larger sample size and selected human lung tissue samples, which confirmed that CXCL9 was expressed at higher levels in the lung tissue of PAH patients than in the lung tissues of healthy controls. Moreover, ANGPT2 was upregulated in chronic obstructive pulmonary disease (COPD) patients with PAH [32]. However, the specific role of ANGPT2 in the progression of PAH is still unclear. As far as the other hub genes, the roles of NT5E and CSF3R have not been identified. Therefore, exploring the functions of these genes could provide useful information.

To further explore the role of the hub genes in the progression of PAH, GSEA and GSVA were performed based on each hub gene. GSEA is a powerful analytical method using the cross-exchange hypothesis test model to evaluate the score of the gene sets and can reveal many related biological pathways [14]. In addition, GSVA is another gene set enrichment analysis method that can estimate the variation of pathway activity over every sample in an unsupervised manner and further identify related gene sets [33]. By combining the two enrichment analytical methods, we were able to obtain more credible results. In the present research, many cell cycle-related pathways were enriched, including cell cycle, DNA replication and mismatch repair pathways, suggesting that the genes involved in these pathways may contribute to proliferation processes.

In the present study, we found nine potential small molecular compounds that can reverse the altered expression of the DEGs and improve PAH through CMap analysis. Among the nine compounds, lypressin is an antidiuretic hormone that has been found in pigs and some marsupial families [34]. Triclabendazole is an anthelminthic drug and is used in the treatment of fascioliasis [35]. Tiaprofenic acid is a non-steroidal anti-inflammatory drug that is widely used to treat pain, particularly arthritis pain [36]. As a multitarget kinase inhibitor, AT-9283 can inhibit the process of STAT3 tyrosine phosphorylation and inhibit multiple myeloma cell proliferation [37] and is widely used in cancer treatment [38, 39]. Another study showed that AT-9283 is an aurora kinase/JAK inhibitor [40]. Huperzine-a is reported as acetylcholinesterase inhibitor and neuroprotective agent [41, 42]. L-741742 is a dopamine receptor antagonist [43]. Nevertheless, previous reports have not shown whether these drugs are effective in treating PAH.

Overall, ruxolitinib obtained the second highest score (97.86) and targets JAK1, JAK2, TYK2 and JAK3 based on the CMap database. Previously, ruxolitinib was confirmed to be used to treat myelofibrosis [44]. Epidemiological studies found PAH to be a common complication of myelofibrosis [45], and ruxolitinib could effectively improve PAH in myelofibrosis patients [46]. However, a case report reported that ruxolitinib might exacerbate PAH, but its mechanism was unclear [47]. To further explore the interaction between ruxolitinib and PAH, we performed molecular docking simulations between ruxolitinib and the proteins encoded by hub genes (ANGPT2, FGF7, NT5E, and CSF3R). Among the hub genes, CXCL9 was not considered due to the lack of 3D structure. The results showed that ruxolitinib has good binding ability with ANGPT2, FGF7 and NT5E, suggesting that ruxolitinib has potential utility in PAH therapy. Therefore, in vitro experiments were conducted. The obtained results displayed the significantly suppressive effect of ruxolitinib on rPASMCs proliferation and migration. Interestingly, as an inhibitor of JAK, ruxolitinib significantly reduced cytokine-mediated lung adenocarcinoma proliferation by inhibiting the JAK/STAT signal pathway [48]. Similarly, in the basic study of JAK2-V617F-positive leukemia cells, ruxolitinib may inhibit cell proliferation through dephosphorylation of the JAK2 substrate STAT5 and further regulation of the mTORC1/S6K/4EBP1 signal pathway [49]. This study was highly correlated with our results of enrichment analysis (PI3K-AKT signalling pathway). Recently, a study reported ruxolitinib has a therapeutic effect on PAH through blocking Jak2-Stat3 signalling pathway [50]. The view in this article supports our result that ruxolitinib might exert promising therapeutic action for PAH. However, relevant molecular mechanisms that the article has studied are different than the ones we explored. In present study, the proteins encoded by 3 hub genes (ANGPT2, FGF7, NT5E) were considered as the targets. Meanwhile, ruxolitinib acted by suppressing the proliferation via PI3K-AKT signalling pathway. These demonstrated that ruxolitinib may have particular value in the treatment of PAH. However, the effects and mechanisms of ruxolitinib should still be validated by further experimental evidence and long-term clinical trials.

MATERIALS AND METHODS

Identification of DEGs

The raw data of three eligible microarray datasets (GSE113439, GSE53408, GSE117261) based on platform GPL6244 (Affymetrix Human Gene 1.0 ST Array) were downloaded from GEO database. Background correction, normalization and expression calculation of the raw data were carried out with "affy"

package using the Robust Multi-array Average (RMA) method [51]. Then the "Knn" method in "impute" package was used to fill the missing data. To remove the batch effects, "sva" package in R software was applied. We matched the probesets with corresponding gene symbols with "hugene10sttranscriptcluster.db" package in R software. To select the DEGs, "Limma" package was utilized with cut-offs of adjusted P value < 0.05 and fold change (FC) > 1.5.

GO and KEGG pathway enrichment analyses

GO enrichment analysis is a commonly used bioinformatic approach for searching comprehensive information of large-scale gene data, including BP, CC and MF. KEGG pathway enrichment analysis is widely used to understand biological mechanisms and functions. GO and KEGG enrichment analyses were performed with "clusterprofiler" package. The top 10 genes sorted by the GeneRatio of GO terms and KEGG pathways were visualized by "GOplot" package [47].

PPI network construction and hub gene identification

A PPI network of DEGs was constructed by the STRING database with a combined score > 0.7 as the cut-off point to assess the direct and indirect associations of the DEGs. "Cytohubba" (a plug-in) was utilized to identify the hub genes in Cytoscape (version 3.7.1). Then, 12 algorithms (MCC, DMNC, MNC, Degree, EPC, BottleNeck, Eccentricity, Closeness, Radiality, Betweenness, Stress and Clustering coefficient) in "Cytohubba" were used to calculate the weight of each gene in total.

GSEA and GSVA

The "clusterprofile" package was utilized to perform GSEA of the hub genes with integrated gene expression data. PAH samples were divided into two groups (high expression and low expression) based on the median expression of each hub gene. Then differential expression analysis was performed with "Limma" package. The log2 fold change is used as the rank list for GSEA analysis [52]. Finally, the top 3 GSEA terms was visualized by "enrichplot" package. Additionally, "GSVA" package was used to further explore the significant signalling pathways associated with the hub genes [33]. GSVA requires two files, a gene set and a gene expression matrix. "GSVAdata" package was used to read a file of gene set in GMT format. Then, the gene set and the gene expression matrix were integrated by "GSVA" package. Next, a matrix of KEGG pathway was obtained. Furthermore, "Limma" was used to find significant KEGG pathway. "C2.cp.kegg.v7.0.symbols.gmt" served as the reference gene set for both GSEA and GSVA, which was downloaded from MSigDB

(http://software.broadinstitute.org/gsea/msigdb/index.jsp) [53]. The top 20 GSVA terms were visualized by "pheatmap" package. P value < 0.05 was considered statistically significant.

Prediction of potential drugs for PAH treatment

The CMap database is an online resource that can be used to establish links between genes, compounds and diseases based on similar and opposite gene expression profiles. In the present study, the DEGs of PAH were divided into two groups (upregulated genes and downregulated genes). Then, the DEGs were loaded into the "QUERY" page. In this study, connectivity scores > 90 or < -90 were selected. 2D structures of candidate compounds were obtained from PubChem. To further investigate the specific mechanism of the compounds, the molecular docking simulation was performed on AutoDock 4 software (version 4.2.6), which was designed to forecast how small molecules bind to a receptor of known 3D structure [54]. The 3D structures of proteins and compounds were downloaded from RCSB PDB and ZINC, respectively. Then, PyMOL was utilized to remove the water molecules and separate proteins from small molecules. Next, we used AutoDock to prepare the specific coordinate file (PDBQT). Then, the file was implemented to run AutoGrid and AutoDock. We selected the molecular docking of minimal binding energy with hydrogen bonds and exported the results as a PDBQT file. The file was converted into a PDB file by OpenBabel software (version 2.4.1). Finally, PyMOL software (version 2.3) was used to view and visualize the results [55].

Chemical reagents

Ruxolitinib was purchased from Selleck (Selleck, Houston, TX, USA). MTT powder was purchased from Solarbio company (Solarbio Co., Beijing, China), dissolved by ultrasonic instrument, and finally prepared into MTT reagent (50mg/ mL, PBS). DMSO, DMEM, fetal bovine serum (FBS) were obtained from Thermo Fisher (Thermo Scientific Fisher, Wilmington, USA).

Extraction and culture of rPASMCs

Sprague-Dawley rats were anesthetized, immobilized, shaved, and disinfected. After opening the thorax of the rats, the heart and lungs were excised and the pulmonary artery was extracted. The pulmonary artery was transferred to high-pressure treated PBS for cleaning and the membrane and intima were removed carefully [56]. The pulmonary medium membrane was cut into small pieces and then cultured in DMEM containing 10% FBS (5% CO_2 in 37° C). The primary rPASMCs were passaged after being cultured for 1 week. When the

rPASMCs confluence reached 80% ~ 90%, they were subcultured to new petri dishes using trypsin–EDTA. The third to the fifth passage of rPASMCs were selected for the following experiments.

MTT assay

To explore the inhibition effect of ruxolitinib on hypoxic rPASMCs, we utilized MTT cytotoxicity assays to perform the cell viability tests [57]. RPASMCs were inoculated into 96-well plates with 8000-10000 cells/well and allowed to attach for 6-8 h. RPASMCs were treated with ruxolitinib of different concentrations (from 5 mM to 0.1 nM) under hypoxia conditions to obtain the IC50 values. After 24 h of modelling, MTT reagent (5 mg/ml, 25 µl) was added into each well and incubated for another 4 h. We discarded all the medium from the wells, next, added DMSO (150 µl) into each well to dissolve formazan crystals. Then the absorbance at 490 nm was detected by a microplate reader and the IC50 value was calculated by linear regression analysis.

Examination of cell proliferative ability

Cell proliferation assay was measured by Cell Counting Kit-8 (CCK-8) assay (Dojindo, Kumamoto, Japan). We seeded the rPASMCs into 96-well plates (8000-10000 cells/well) and preincubated until the cells attached. Then, rPASMCs were treated with different concentrations of ruxolitinib (0, 0.1, 0.5, 1 and 2.5 mM) and incubated for 24 h (5% O_2, 5% CO_2 in 37° C). Whereafter, we added 10 ul of CCK-8 assay solution to each well and the incubation continued for 4 h. The capacity of cell proliferation was determined by the absorbance measured at 450 nm through a microplate reader.

Examination of cell migratory ability

Wound scratch assay was performed to determine the ability of cell migratory. RPASMCs were plated onto 12-well plates at a density of $6\sim8\times10^3$ cells/well and cultured up to 90% cell density. Then the streak wounds were scratched with a sterile 20 µl pipette tip. The cell culture medium was changed and treated with different concentrations of ruxolitinib (0, 0.1, 0.5, 1 and 2.5 mM). We photographed the wounds in the same view at 0, 12 and 24 h time points. The relative distance of cell migration was exhibited and measured by ImageJ software (National Institutes of Health, Bethesda, MD).

Immunofluorescence

To measure the effect of ruxolitinib on the cellular proliferation phenotype, we performed cell immunofluorescence analyses to detect the expression of Ki67 in rPASMCs. RPASMCs were planted on a glass cover slip, fixed with 4% paraformaldehyde and washed with PBS. Respectively, we used 0.1% Triton X-100 and 5% BSA to permeabilize cell membranes and to block cells at room temperature. After the blocking solution was washed off, Ki67 (1:500, AF0198; Affinity) antibodies were used as primary antibody and slides were incubated overnight at 4° C. The next day, cells were incubated with the donkey anti-rabbit (1:1000; Alexa Fluor 594) for 1 h after being washed 3 times. Nuclei were counterstained with DAPI stain for 5 min. Images were observed with a fluorescence microscope (Olympus, Tokyo, Japan).

Statistical analysis

Statistical analyses were performed with GraphPad Prism 6.0 (GraphPad Software, CA, USA). All the results presented were represented from at least 3 independent experiments. All the data were expressed as the mean standard error of mean (SEM). Comparisons between two groups were analysed by unpaired two-tailed Student's t-test, and multiple comparisons were analysed by one-way analysis of variance (ANOVA). P values of <0.05 were considered statistically significant.

AUTHOR CONTRIBUTIONS

H.D. and X.L. conceived the idea and wrote the first draft of the manuscript. H.D. executed the bioinformatics analysis. M.C., C.Z. and Q.X. performed the experiments *in vitro*. M.C., W.M, Y.W. and C.Z. wrote parts of the manuscript and prepared the tables and figures. X.H., L.W. and M.C. revised the manuscript. All authors have read and agreed to the published version of the manuscript.

ACKNOWLEDGMENTS

CONFLICTS OF INTEREST

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

FUNDING

Editorial note

&This corresponding author has a verified history of publications using a personal email address for correspondence.

REFERENCES

1. Han HD, Byeon Y, Kang TH, Jung ID, Lee JW, Shin BC, Lee YJ, Sood AK, Park YM. Toll-like receptor 3-induced immune response by poly(d,l-lactide-co-glycolide) nanoparticles for dendritic cell-based cancer immunotherapy. Int J Nanomedicine. 2016; 11:5729–42.
https://doi.org/10.2147/IJN.S109001 PMID:27843314

2. Pietra GG, Capron F, Stewart S, Leone O, Humbert M, Robbins IM, Reid LM, Tuder RM. Pathologic assessment of vasculopathies in pulmonary hypertension. J Am Coll Cardiol. 2004 (Suppl S); 43:25S–32S.
https://doi.org/10.1016/j.jacc.2004.02.033 PMID:15194175

3. Wang T, Zheng X, Li R, Liu X, Wu J, Zhong X, Zhang W, Liu Y, He X, Liu W, Wang H, Zeng H. Integrated bioinformatic analysis reveals YWHAB as a novel diagnostic biomarker for idiopathic pulmonary arterial hypertension. J Cell Physiol. 2019; 234:6449–62.
https://doi.org/10.1002/jcp.27381 PMID:30317584

4. Galiè N, Manes A, Branzi A. Prostanoids for pulmonary arterial hypertension. Am J Respir Med. 2003; 2:123–37.
https://doi.org/10.1007/BF03256644 PMID:14720012

5. Galié N, Manes A, Branzi A. The endothelin system in pulmonary arterial hypertension. Cardiovasc Res. 2004; 61:227–37.
https://doi.org/10.1016/j.cardiores.2003.11.026 PMID:14736539

6. Archer SL, Michelakis ED. Phosphodiesterase type 5 inhibitors for pulmonary arterial hypertension. N Engl J Med. 2009; 361:1864–71.
https://doi.org/10.1056/NEJMct0904473 PMID:19890129

7. Zhu TT, Zhang WF, Yin YL, Liu YH, Song P, Xu J, Zhang MX, Li P. MicroRNA-140-5p targeting tumor necrosis factor-α prevents pulmonary arterial hypertension. J Cell Physiol. 2019; 234:9535–50.
https://doi.org/10.1002/jcp.27642 PMID:30367500

8. Ernst M, Du Y, Warsow G, Hamed M, Endlich N, Endlich K, Murua Escobar H, Sklarz LM, Sender S, Junghanß C, Möller S, Fuellen G, Struckmann S. FocusHeuristics - expression-data-driven network optimization and disease gene prediction. Sci Rep. 2017; 7:42638.
https://doi.org/10.1038/srep42638 PMID:28205611

9. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. cytoHubba: identifying hub objects and sub-networks from complex interactome. BMC Syst Biol. 2014 (Suppl 4); 8:S11.
https://doi.org/10.1186/1752-0509-8-S4-S11 PMID:25521941

10. Grinnan D, Trankle C, Andruska A, Bloom B, Spiekerkoetter E. Drug repositioning in pulmonary arterial hypertension: challenges and opportunities. Pulm Circ. 2019; 9:2045894019832226.
https://doi.org/10.1177/2045894019832226 PMID:30729869

11. Wilkins MR. Developing treatments for pulmonary arterial hypertension. Pulm Circ. 2013; 3:156–59.
https://doi.org/10.4103/2045-8932.109961 PMID:23662193

12. Subramanian A, Narayan R, Corsello SM, Peck DD, Natoli TE, Lu X, Gould J, Davis JF, Tubelli AA, Asiedu JK, Lahr DL, Hirschman JE, Liu Z, et al. A next generation connectivity map: L1000 platform and the first 1,000,000 profiles. Cell. 2017; 171:1437–52.e17.
https://doi.org/10.1016/j.cell.2017.10.049 PMID:29195078

13. Struckmann S, Ernst M, Fischer S, Mah N, Fuellen G, Möller S. Scoring functions for drug-effect similarity. Brief Bioinform. 2020. [Epub ahead of print].
https://doi.org/10.1093/bib/bbaa072 PMID:32484516

14. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA. 2005; 102:15545–50.
https://doi.org/10.1073/pnas.0506580102 PMID:16199517

15. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Holko M, Yefanov A, Lee H, Zhang N, et al. NCBI GEO: archive for functional genomics data sets— update. Nucleic Acids Res. 2013; 41:D991–95.
https://doi.org/10.1093/nar/gks1193 PMID:23193258

16. Lee JY, Styczynski MP. NS-kNN: a modified k-nearest neighbors approach for imputing metabolomics data. Metabolomics. 2018; 14:153.
https://doi.org/10.1007/s11306-018-1451-8 PMID:30830437

17. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics. 2012; 28:882–83. https://doi.org/10.1093/bioinformatics/bts034 PMID:22257669

18. 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:e47. https://doi.org/10.1093/nar/gkv007 PMID:25605792

19. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012; 16:284–87. https://doi.org/10.1089/omi.2011.0118 PMID:22455463

20. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, Jensen LJ, Mering CV. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. 2019; 47:D607–13. https://doi.org/10.1093/nar/gky1131 PMID:30476243

21. Kim S, Chen J, Cheng T, Gindulyte A, He J, He S, Li Q, Shoemaker BA, Thiessen PA, Yu B, Zaslavsky L, Zhang J, Bolton EE. PubChem 2019 update: improved access to chemical data. Nucleic Acids Res. 2019; 47:D1102–09. https://doi.org/10.1093/nar/gky1033 PMID:30371825

22. Sterling T, Irwin JJ. ZINC 15—ligand discovery for everyone. J Chem Inf Model. 2015; 55:2324–37. https://doi.org/10.1021/acs.jcim.5b00559 PMID:26479676

23. Luo M, Liu Z, Zhang X, Han C, Samandi LZ, Dong C, Sumer BD, Lea J, Fu YX, Gao J. Synergistic STING activation by PC7A nanovaccine and ionizing radiation improves cancer immunotherapy. J Control Release. 2019; 300:154–60. https://doi.org/10.1016/j.jconrel.2019.02.036 PMID:30844475

24. Goodsell DS, Zardecki C, Di Costanzo L, Duarte JM, Hudson BP, Persikova I, Segura J, Shao C, Voigt M, Westbrook JD, Young JY, Burley SK. RCSB protein data bank: enabling biomedical research and drug discovery. Protein Sci. 2020; 29:52–65. https://doi.org/10.1002/pro.3730 PMID:31531901

25. Soon E, Crosby A, Southwood M, Yang P, Tajsic T, Toshner M, Appleby S, Shanahan CM, Bloch KD, Pepke-Zaba J, Upton P, Morrell NW. Bone morphogenetic protein receptor type II deficiency and increased inflammatory cytokine production. A gateway to pulmonary arterial hypertension. Am J Respir Crit Care Med. 2015; 192:859–72. https://doi.org/10.1164/rccm.201408-1509OC PMID:26073741

26. Thenappan T, Chan SY, Weir EK. Role of extracellular matrix in the pathogenesis of pulmonary arterial hypertension. Am J Physiol Heart Circ Physiol. 2018; 315:H1322–31. https://doi.org/10.1152/ajpheart.00136.2018 PMID:30141981

27. Mamazhakypov A, Viswanathan G, Lawrie A, Schermuly RT, Rajagopal S. The role of chemokines and chemokine receptors in pulmonary arterial hypertension. Br J Pharmacol. 2021; 178:72–89. https://doi.org/10.1111/bph.14826 PMID:31399998

28. Zhang Q, Cao Y, Luo Q, Wang P, Shi P, Song C, E M, Ren J, Fu B, Sun H. The transient receptor potential vanilloid-3 regulates hypoxia-mediated pulmonary artery smooth muscle cells proliferation via PI3K/AKT signaling pathway. Cell Prolif. 2018; 51:e12436. https://doi.org/10.1111/cpr.12436 PMID:29359496

29. Wu J, Yu Z, Su D. BMP4 protects rat pulmonary arterial smooth muscle cells from apoptosis by PI3K/AKT/Smad1/5/8 signaling. Int J Mol Sci. 2014; 15:13738–54. https://doi.org/10.3390/ijms150813738 PMID:25110865

30. Zhou C, Chen Y, Kang W, Lv H, Fang Z, Yan F, Li L, Zhang W, Shi J. Mir-455-3p-1 represses FGF7 expression to inhibit pulmonary arterial hypertension through inhibiting the RAS/ERK signaling pathway. J Mol Cell Cardiol. 2019; 130:23–35. https://doi.org/10.1016/j.yjmcc.2019.03.002 PMID:30858037

31. Ross DJ, Strieter RM, Fishbein MC, Ardehali A, Belperio JA. Type I immune response cytokine-chemokine cascade is associated with pulmonary arterial hypertension. J Heart Lung Transplant. 2012; 31:865–73. https://doi.org/10.1016/j.healun.2012.04.008 PMID:22658713

32. García-Lucio J, Argemi G, Tura-Ceide O, Diez M, Paul T, Bonjoch C, Coll-Bonfill N, Blanco I, Barberà JA, Musri MM, Peinado VI. Gene expression profile of angiogenic factors in pulmonary arteries in COPD: relationship with vascular remodeling. Am J Physiol Lung Cell Mol Physiol. 2016; 310:L583–92. https://doi.org/10.1152/ajplung.00261.2015 PMID:26801565

33. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics. 2013; 14:7.

https://doi.org/10.1186/1471-2105-14-7
PMID:23323831

34. Gorbulev V, Büchner H, Akhundova A, Fahrenholz F. Molecular cloning and functional characterization of V2 [8-lysine] vasopressin and oxytocin receptors from a pig kidney cell line. Eur J Biochem. 1993; 215:1–7.
https://doi.org/10.1111/j.1432-1033.1993.tb18000.x
PMID:8393786

35. Manouchehri Naeini K, Mohammad Nasiri F, Rokni MB, Kheiri S. Seroprevalence of human fascioliasis in chaharmahal and Bakhtiyari Province, Southwestern Iran. Iran J Public Health. 2016; 45:774–80.
PMID:27648421

36. Nichol FE, Samanta A, Rose CM. Synovial fluid and plasma kinetics of repeat dose sustained action tiaprofenic acid in patients with rheumatoid arthritis. Drugs. 1988 (Suppl 1); 35:46–51.
https://doi.org/10.2165/00003495-198800351-00010
PMID:3359945

37. Santo L, Hideshima T, Cirstea D, Bandi M, Nelson EA, Gorgun G, Rodig S, Vallet S, Pozzi S, Patel K, Unitt C, Squires M, Hu Y, et al. Antimyeloma activity of a multitargeted kinase inhibitor, AT9283, via potent Aurora kinase and STAT3 inhibition either alone or in combination with lenalidomide. Clin Cancer Res. 2011; 17:3259–71.
https://doi.org/10.1158/1078-0432.CCR-10-3012
PMID:21430070

38. Duong JK, Griffin MJ, Hargrave D, Vormoor J, Edwards D, Boddy AV. A population pharmacokinetic model of AT9283 in adults and children to predict the maximum tolerated dose in children with leukaemia. Br J Clin Pharmacol. 2017; 83:1713–22.
https://doi.org/10.1111/bcp.13260
PMID:28177130

39. Moreno L, Marshall LV, Pearson AD, Morland B, Elliott M, Campbell-Hewson Q, Makin G, Halford SE, Acton G, Ross P, Kazmi-Stokes S, Lock V, Rodriguez A, et al. A phase I trial of AT9283 (a selective inhibitor of aurora kinases) in children and adolescents with solid tumors: a cancer research UK study. Clin Cancer Res. 2015; 21:267–73.
https://doi.org/10.1158/1078-0432.CCR-14-1592
PMID:25370467

40. Shah M, Gallipoli P, Lyons J, Holyoake T, Jørgensen H. Effects of the novel aurora kinase/JAK inhibitor, AT9283 and imatinib on Philadelphia positive cells *in vitro*. Blood Cells Mol Dis. 2012; 48:199–201.
https://doi.org/10.1016/j.bcmd.2012.01.005
PMID:22325915

41. Damar U, Gersner R, Johnstone JT, Schachter S, Rotenberg A. Huperzine a as a neuroprotective and antiepileptic drug: a review of preclinical research. Expert Rev Neurother. 2016; 16:671–80.
https://doi.org/10.1080/14737175.2016.1175303
PMID:27086593

42. Haviv H, Wong DM, Silman I, Sussman JL. Bivalent ligands derived from Huperzine A as acetylcholinesterase inhibitors. Curr Top Med Chem. 2007; 7:375–87.
https://doi.org/10.2174/156802607779941215
PMID:17305579

43. Cao BJ, Rodgers RJ. Dopamine D4 receptor and anxiety: behavioural profiles of clozapine, L-745,870 and L-741,742 in the mouse plus-maze. Eur J Pharmacol. 1997; 335:117–25.
https://doi.org/10.1016/s0014-2999(97)01234-x
PMID:9369363

44. Haznedaroglu IC. Ruxolitinib for myelofibrosis. N Engl J Med. 2012; 366:2032.
https://doi.org/10.1056/NEJMc1203704
PMID:22621635

45. Cortelezzi A, Gritti G, Del Papa N, Pasquini MC, Calori R, Gianelli U, Cortiana M, Parati G, Onida F, Sozzi F, Vener C, Bianchi P, Deliliers GL. Pulmonary arterial hypertension in primary myelofibrosis is common and associated with an altered angiogenic status. Leukemia. 2008; 22:646–49.
https://doi.org/10.1038/sj.leu.2404943
PMID:17851555

46. Tabarroki A, Lindner DJ, Visconte V, Zhang L, Rogers HJ, Parker Y, Duong HK, Lichtin A, Kalaycio ME, Sekeres MA, Mountantonakis SE, Heresi GA, Tiu RV. Ruxolitinib leads to improvement of pulmonary hypertension in patients with myelofibrosis. Leukemia. 2014; 28:1486–93.
https://doi.org/10.1038/leu.2014.5 PMID:24406841

47. Low AT, Howard L, Harrison C, Tulloh RM. Pulmonary arterial hypertension exacerbated by ruxolitinib. Haematologica. 2015; 100:e244–45.
https://doi.org/10.3324/haematol.2014.120816
PMID:25682609

48. Mohrherr J, Haber M, Breitenecker K, Aigner P, Moritsch S, Voronin V, Eferl R, Moriggl R, Stoiber D, Győrffy B, Brcic L, László V, Döme B, et al. JAK-STAT inhibition impairs K-RAS-driven lung adenocarcinoma progression. Int J Cancer. 2019; 145:3376–88.
https://doi.org/10.1002/ijc.32624 PMID:31407334

49. Ishida S, Akiyama H, Umezawa Y, Okada K, Nogami A, Oshikawa G, Nagao T, Miura O. Mechanisms for mTORC1 activation and synergistic induction of apoptosis by ruxolitinib and BH3 mimetics or autophagy inhibitors in JAK2-V617F-expressing leukemic cells including newly established PVTL-2.

Oncotarget. 2018; 9:26834–51.
https://doi.org/10.18632/oncotarget.25515
PMID:29928488

50. Yerabolu D, Weiss A, Kojonazarov B, Boehm M, Schlueter BC, Ruppert C, Günther A, Jonigk D, Grimminger F, Ghofrani HA, Seeger W, Weissmann N, Schermuly RT. Targeting Jak-Stat signaling in experimental pulmonary hypertension. Am J Respir Cell Mol Biol. 2021; 64:100–14.
https://doi.org/10.1165/rcmb.2019-0431OC
PMID:33052714

51. Gautier L, Cope L, Bolstad BM, Irizarry RA. Affy—analysis of affymetrix GeneChip data at the probe level. Bioinformatics. 2004; 20:307–15.
https://doi.org/10.1093/bioinformatics/btg405
PMID:14960456

52. Peng G, Suo S, Chen J, Chen W, Liu C, Yu F, Wang R, Chen S, Sun N, Cui G, Song L, Tam PP, Han JD, Jing N. Spatial transcriptome for the molecular annotation of lineage fates and cell identity in mid-gastrula mouse embryo. Dev Cell. 2016; 36:681–97.
https://doi.org/10.1016/j.devcel.2016.02.020
PMID:27003939

53. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. Bioinformatics.

2011; 27:1739–40.
https://doi.org/10.1093/bioinformatics/btr260
PMID:21546393

54. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ. AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. J Comput Chem. 2009; 30:2785–91.
https://doi.org/10.1002/jcc.21256 PMID:19399780

55. Rigsby RE, Parker AB. Using the PyMOL application to reinforce visual understanding of protein structure. Biochem Mol Biol Educ. 2016; 44:433–37.
https://doi.org/10.1002/bmb.20966 PMID:27241834

56. Cao YY, Ba HX, Li Y, Tang SY, Luo ZQ, Li XH. Regulatory effects of prohibitin 1 on proliferation and apoptosis of pulmonary arterial smooth muscle cells in monocrotaline-induced PAH rats. Life Sci. 2020; 250:117548.
https://doi.org/10.1016/j.lfs.2020.117548
PMID:32173312

57. Cerretani D, Collodel G, Brizzi A, Fiaschi AI, Menchiari A, Moretti E, Moltoni L, Micheli L. Cytotoxic effects of cannabinoids on human HT-29 colorectal adenocarcinoma cells: different mechanisms of THC, CBD, and CB83. Int J Mol Sci. 2020; 21:5533.
https://doi.org/10.3390/ijms21155533
PMID:32752303
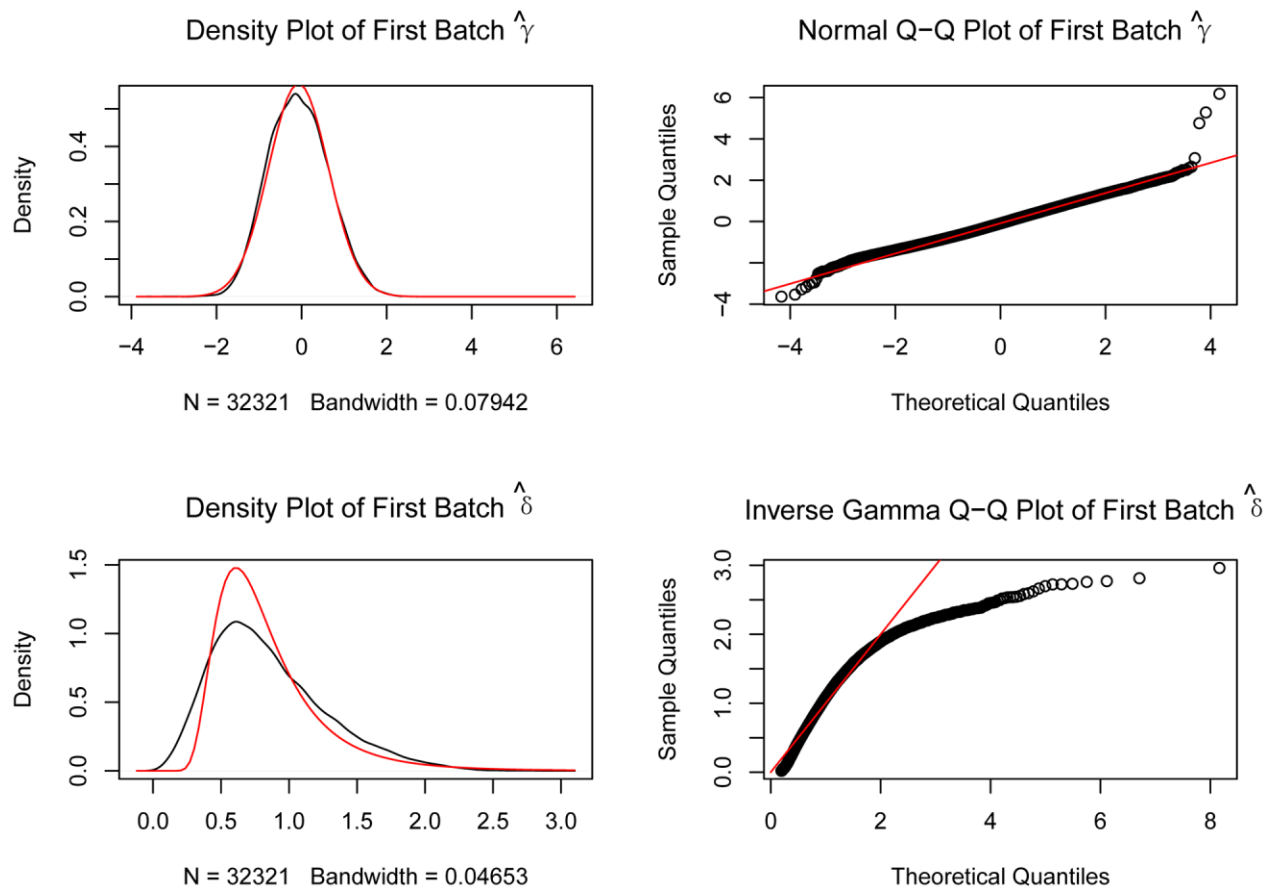
Supplementary Figure 1. The quantile-quantile (Q-Q) plot of the batch effect of GSE113439, GSE53408 and GSE117261 datasets is removed.
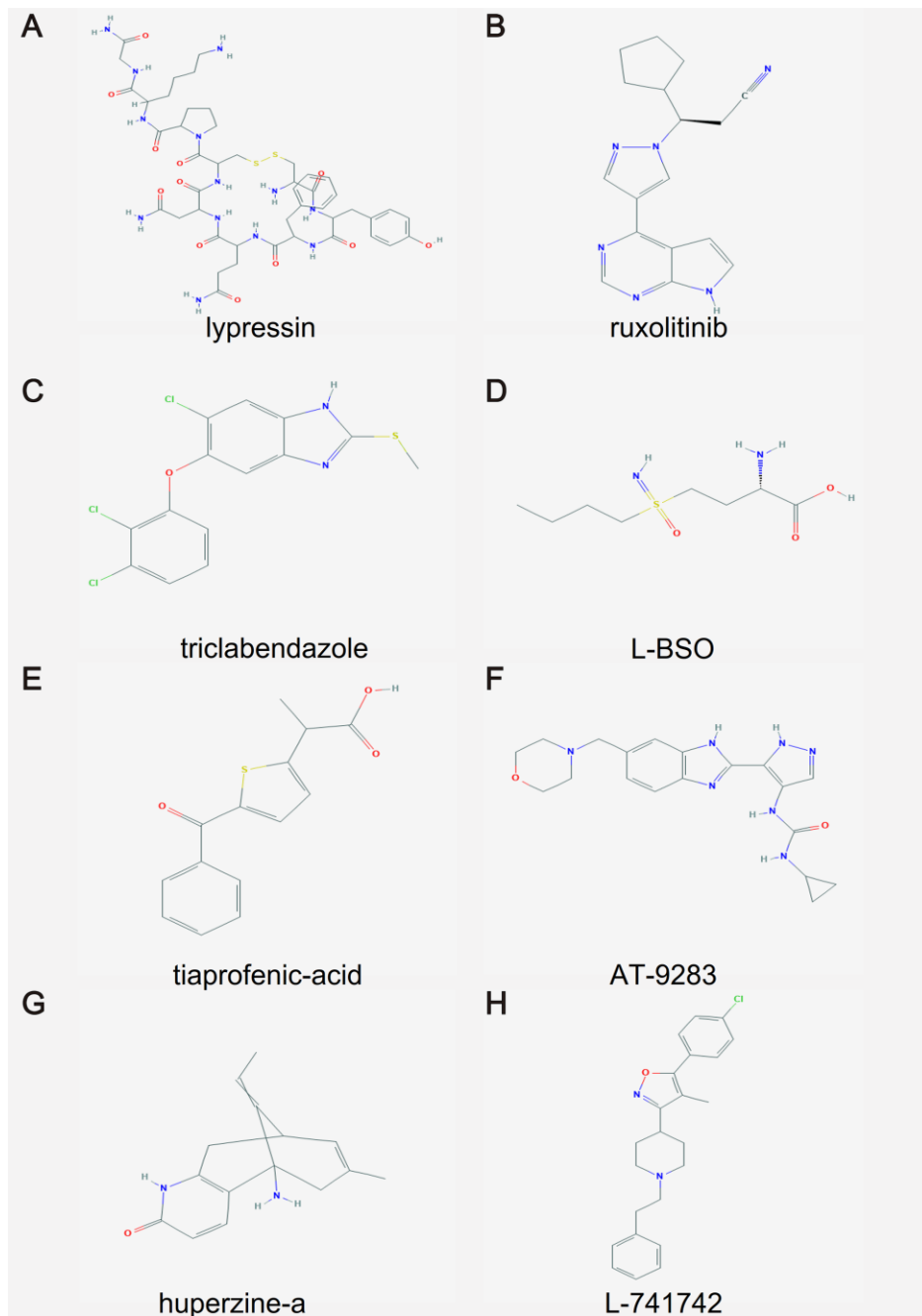
Supplementary Figure 2. 2D structure of the candidate compounds. (**A**) lypressin, (**B**) ruxolitinib, (**C**) triclabendazole, (**D**) L-BSO, (**E**) tiaprofenic acid, (**F**) AT-9283, (**G**) huperzine-a, (**H**) L-741742.
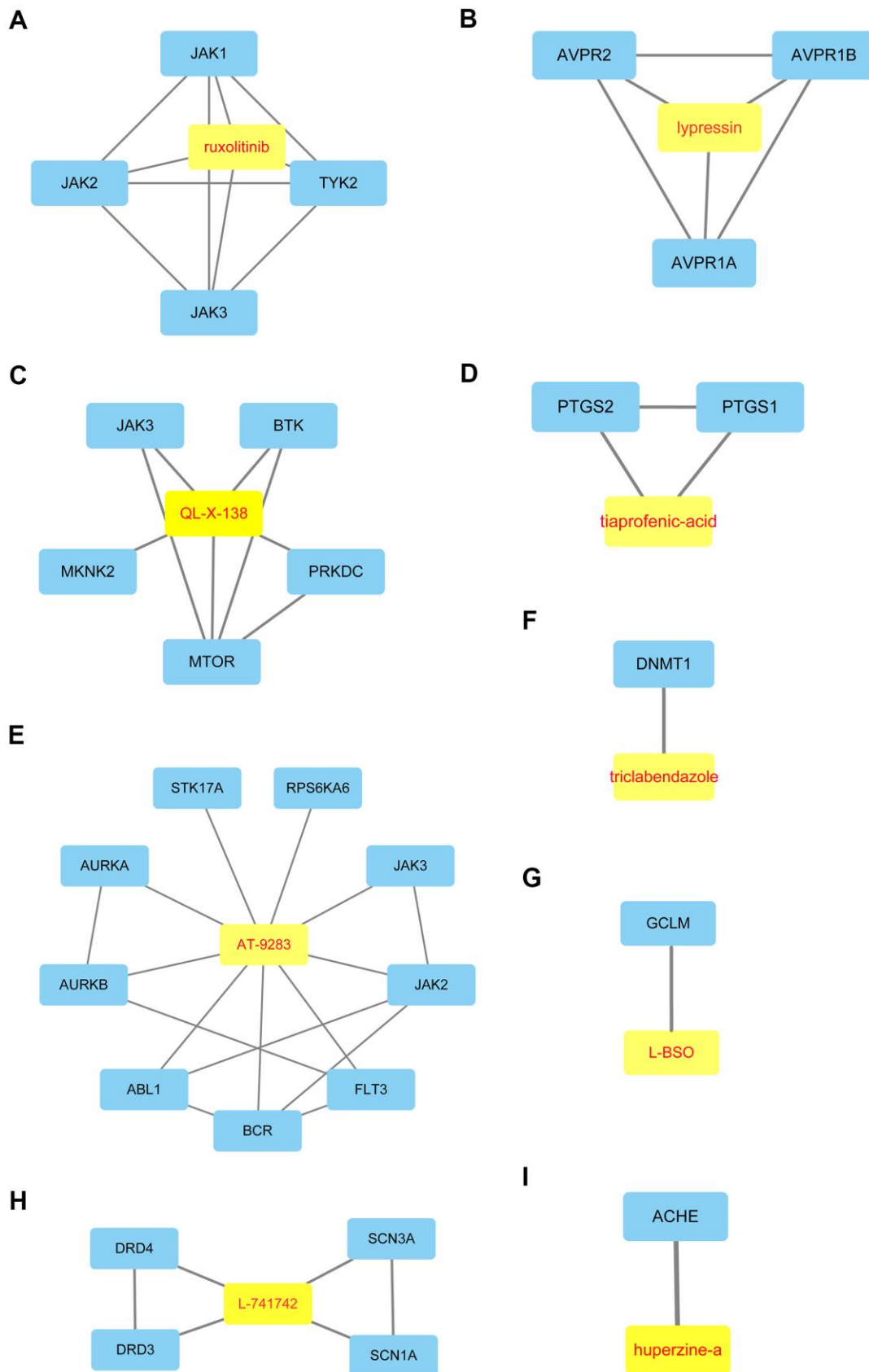
Supplementary Figure 3. The drug-target network for nine candidate drugs. Yellow represented drug and blue represented target. (**A**) ruxolitinib. (**B**) lypressin. (**C**) QL-X-138. (**D**) tiaprofenic-acid. (**E**) AT-9283. (**F**) triclabendazole. (**G**) L-BSO. (**H**) huperzine-a. (**I**) L-741742.

Supplementary Tables

Please browse Full Text version to see the data of Supplementary Table 2.

Supplementary Table 1. Characteristics of the individual dataset.

GEO ID	Platform	Tissue type	Sample size	Country	Time
GSE113439	GPL6244 (Affymetrix Human Gene 1.0 ST Array)	human lung tissue	15 vs. 11	Canada	2018
GSE53408	GPL6244 (Affymetrix Human Gene 1.0 ST Array)	human lung tissue	12 vs. 11	Canada	2013
GSE117261	GPL6244 (Affymetrix Human Gene 1.0 ST Array)	human lung tissue	58 vs. 25	United States	2018

Supplementary Table 2. All differentially expressed genes.

Supplementary Table 3. The result of PPI.

Algorithm	Gene
MCC	FGF7, POSTN, LCN2, S100A8, CSF3R, SAA1, S100A12, S100A9, ITGA2, ITGB3, CD69, MME, FAP, ABCB1, ASPN, SERPINE2, HAS2, SFRP2, INHBA, BIRC3, CFH, OGN, HBB, ECM2, SULF1, PDGFD, WIF1, RGS5, RSPO3, RNASE2, SAA2
DMNC	MME, AREG, CCDC80, SAA2, PDGFD, CSF3R, ITGA2, ANGPT2, IFI44L, S100A9, FGF7, FAP, SELE, SERPINE2, BIRC3, CXCL9, C5, SFRP2, HAS2, ANKRD22, NT5E, RGS5, SULF1, ACE2, S100A8, SAA1, ECM2, MEDAG, ADRA1A, VCAM1, THY1, CD69, S100A12, LUC7L3, RAMP2, PTPN13, LRRC17, RSPO3, LCN2, CA4
MNC	IGF1, PTGS2, POSTN, VCAM1, THY1, SELE, CXCL9, ASPN, LCN2, ITGB3, INHBA, FGF7, NT5E, ANGPT2, S100A8, SFRP2, AREG, FAP, CD69, ITGA2, S100A12, OGN, EPHA3, CSF3R, SAA1, CFH, TTN, S100A9, SERPINE2, HBB, ABCB1, DCLK1, HAS2, WIF1, LTBP1, BCHE, ANK2, SULF1, ECM2, BIRC3, EPHA4
Degree	IGF1, PTGS2, POSTN, VCAM1, THY1, SELE, CXCL9, ASPN, LCN2, ITGB3, INHBA, FGF7, NT5E, ANGPT2, S100A8, FAP, SFRP2, CD69, AREG, CFH, TTN, ITGA2, S100A12, OGN, EPHA3, CSF3R, ABCB1, SAA1, SERPINE2, S100A9, HBB, DCLK1, HAS2, WIF1, LTBP1, BCHE, ANK2, SULF1, ECM2, EPHA4
EPC	IGF1,PTGS2,POSTN,VCAM1,THY1,SELE,CXCL9,ASPN,LCN2,ITGB3,INHBA,FGF7,NT5E,ANGPT2,S100A8,FAP,SFRP2,CD69,AREG,CFH,TTN,ITGA2,S100A12,OGN,EPHA3,CSF3R,ABCB1,SAA1,SERPINE2,S100A9,DCLK,1,HAS2,WIF1,LTBP1,BCHE,SULF1,ECM2,BIRC3,ESM1,MME
BottleNeck	IGF1,PTGS2,POSTN,THY1,CXCL9,LCN2,FGF7,NT5E,ANGPT2,FAP,CD69,AREG,CFH,TTN,ITGA2,CSF3R,ABCB1,S100A9,DCLK1,WIF1,LTBP1,BCHE,EPHA4,PAMR1,GEM,RSPO3,TBX3,RGS5,DLG2,CCDC80,PDGFD,CPA3,LRRC32,ADAMTS9,MEDAG,LRRC17,TDO2,ANKRD22,ABCC9,WEE1,XAF1,LUC7L3,RAMP2,TSHZ2,LILRA2,SLC7A2,HIF3A,MSMB
Eccentricity	IGF1,PTGS2,POSTN,VCAM1,THY1,SELE,CXCL9,ASPN,LCN2,ITGB3,INHBA,FGF7,NT5E,ANGPT2,S100A8,FAP,SFRP2,CD69,AREG,CFH,TTN,ITGA2,S100A12,OGN,EPHA3,CSF3R,ABCB1,SAA1,SERPINE2,S100A9,HBB,DCLK1,HAS2,WIF1,LTBP1,BCHE,ANK2,SULF1,ECM2,EPHA4,BIRC3,ESM1,PAMR1,GEM,MME,RGS1,RSPO3,TBX3,ENPP2,RGS5,RNASE2,MATN2,DLG2,HMCN1,ITGB6,CCDC80,CA1,PDGFD,ACE2,GBP5,CPA3,LRRC32,PDE1A,SLC6A4,ADAMTS9,ANKRD1,ADRA1A,CA4,PI15,MEDAG,LRRC17,ALAS2,TDO2,ANKRD22,SEMA3D,WEE1,IFI44L,XAF1,LUC7L3,SAA2,SLC9A3R2,ABCA8,RAMP2,C5,BPIFA1,HSPH1,PLAC8,CHIT1,IL13RA2,ZNF521,ANKRD28,FCN3,S100A3,SLC7A2,RASGRP1,HIVEP2,VIPR1,PTPN13,HIF3A,MSMB,SLCO4A1,

Closeness	IGF1,PTGS2,POSTN,VCAM1,THY1,SELE,CXCL9,ASPN,LCN2,ITGB3,INHBA,FGF7,NT5E,ANGPT2,S100A8,FAP,SFRP2,CD69,AREG,CFH,TTN,ITGA2,S100A12,OGN,EPHA3,CSF3R,ABCB1,SAA1,SERPINE2,S100A9,HBB,DCLK1,HAS2,WIF1,LTBP1,BCHE,ANK2,ECM2,EPHA4,BIRC3,
Radiality	IGF1,PTGS2,POSTN,VCAM1,THY1,SELE,CXCL9,ASPN,LCN2,ITGB3,INHBA,FGF7,NT5E,ANGPT2,S100A8,FAP,SFRP2,CD69,AREG,CFH,TTN,ITGA2,OGN,EPHA3,CSF3R,ABCB1,SAA1,SERPINE2,HBB,DCLK1,HAS2,WIF1,LTBP1,BCHE,ANK2,ECM2,EPHA4,BIRC3,ESM1,TBX3,
Betweenness	IGF1,PTGS2,POSTN,VCAM1,THY1,SELE,CXCL9,ASPN,LCN2,ITGB3,INHBA,FGF7,NT5E,ANGPT2,S100A8,FAP,SFRP2,CD69,CFH,TTN,S100A12,OGN,EPHA3,CSF3R,ABCB1,SAA1,SERPINE2,DCLK1,WIF1,ANK2,EPHA4,PAMR1,GEM,TBX3,RNASE2,PDGFD,MEDAG,ABCC9,LUC7L3,LILRA2,
Stress	IGF1,PTGS2,POSTN,VCAM1,THY1,SELE,CXCL9,ASPN,LCN2,ITGB3,INHBA,FGF7,NT5E,ANGPT2,S100A8,FAP,SFRP2,CD69,CFH,TTN,S100A12,OGN,EPHA3,CSF3R,ABCB1,SAA1,SERPINE2,HBB,DCLK1,WIF1,ANK2,EPHA4,ESM1,PAMR1,GEM,TBX3,RNASE2,PDGFD,ABCC9,LUC7L3,LILRA2,
Clustering, coefficient	SELE,CXCL9,FGF7,NT5E,ANGPT2,S100A8,SFRP2,AREG,ITGA2,S100A12,CSF3R,SAA1,SERPINE2,S100A9,HAS2,SULF1,ECM2,BIRC3,MME,ENPP2,RGS5,HMCN1,CCDC80,CA1,PDGFD,ACE2,SLC6A4,ADAMTS9,CA4,SEMA3D,IFI44L,SAA2,SOSTDC1,C5,CHIT1,IL13RA2,LILRA2,PTPN13,BPIFB1,ANKRD36B,

Supplementary Table 4. The information of five hub genes.

Gene symbol	Full name	Log (fold change)	P value
CSF3R	colony stimulating factor 3 receptor	-0.641687	1.61E^{-14}
NT5E	5'-nucleotidase ecto	0.703553	3.98E^{-16}
ANGPT2	angiopoietin 2	1.047516	6.1E^{-9}
FGF7	fibroblast growth factor 7	0.736298	3.27E^{-7}
CXCL9	C-X-C motif chemokine ligand 9	0.609940	0.000549

Supplementary Table 5. The top 20 GSVA terms of ANGPT2.

Pathways	LogFC	P.value	FDR
KEGG_ONE_CARBON_POOL_BY_FOLATE	0.243792192	5.86E-05	0.001317033
KEGG_PROTEIN_EXPORT	0.233239809	0.000238508	0.002194269
KEGG_VALINE_LEUCINE_AND_ISOLEUCINE_BIOSYNTHESIS	0.232604737	0.000427454	0.003292363
KEGG_SPLICEOSOME	0.231820688	1.50E-05	0.000552429
KEGG_NON_HOMOLOGOUS_END_JOINING	0.210725052	0.002509695	0.012152209
KEGG_PROPANOATE_METABOLISM	0.192878556	0.002170981	0.011568672
KEGG_RNA_DEGRADATION	0.189000949	0.000137356	0.001534455
KEGG_CELL_CYCLE	0.186071751	7.77E-06	0.000476692
KEGG_MISMATCH_REPAIR	0.185866007	0.001760087	0.010120498
KEGG_NUCLEOTIDE_EXCISION_REPAIR	0.184198544	0.00014177	0.001534455
KEGG_MATURITY_ONSET_DIABETES_OF_THE_YOUNG	-0.179973823	0.001329905	0.008288554
KEGG_AMINOACYL_TRNA_BIOSYNTHESIS	0.176774474	0.004419167	0.01805783
KEGG_BASAL_TRANSCRIPTION_FACTORS	0.170206927	0.000470546	0.003463222
KEGG_GLYCOSPHINGOLIPID_BIOSYNTHESIS_GLOBO_SERIES	-0.167989923	0.000584972	0.004139802
KEGG_CYSTEINE_AND_METHIONINE_METABOLISM	0.166307349	1.32E-05	0.000552429
KEGG_SULFUR_METABOLISM	-0.165247017	0.002200563	0.011568672
KEGG_ECM_RECEPTOR_INTERACTION	0.156670637	3.40E-05	0.001042607
KEGG_COMPLEMENT_AND_COAGULATION_CASCADES	0.152255983	4.24E-05	0.001113465
KEGG_UBIQUITIN_MEDIATED_PROTEOLYSIS	0.150071579	9.25E-05	0.001511601
KEGG_CIRCADIAN_RHYTHM_MAMMAL	0.148754719	0.006024413	0.023146939

Supplementary Table 6. The top 20 GSVA terms of CSF3R.

Pathways	LogFC	P.value	FDR
KEGG_PROPANOATE_METABOLISM	-0.295856535	1.43E-06	0.000131203
KEGG_GLYCOSPHINGOLIPID_BIOSYNTHESIS_GLOBO_SERIES	0.275953058	4.08E-09	7.50E-07
KEGG_VALINE_LEUCINE_AND_ISOLEUCINE_DEGRADATION	-0.268552435	6.01E-05	0.001381529
KEGG_STEROID_BIOSYNTHESIS	0.237527029	0.000184989	0.003403804
KEGG_BUTANOATE_METABOLISM	-0.226972724	9.60E-06	0.000353348
KEGG_MISMATCH_REPAIR	-0.212629125	0.000318776	0.003910318
KEGG_GLYCOSAMINOGLYCAN_DEGRADATION	0.207225325	7.50E-06	0.000344979
KEGG_FATTY_ACID_METABOLISM	-0.187955978	0.000315136	0.003910318
KEGG_NON_HOMOLOGOUS_END_JOINING	-0.18645673	0.007698082	0.03219198
KEGG_CELL_CYCLE	-0.169366017	5.27E-05	0.001381529
KEGG_DNA_REPLICATION	-0.165048486	0.002357046	0.015489158
KEGG_SPLICEOSOME	-0.162908395	0.002795828	0.01607601
KEGG_MATURITY_ONSET_DIABETES_OF_THE_YOUNG	0.157441906	0.005180597	0.025947799
KEGG_SULFUR_METABOLISM	0.156650672	0.003757833	0.02033651
KEGG_GALACTOSE_METABOLISM	0.156212812	0.000123689	0.002528749
KEGG_LYSINE_DEGRADATION	-0.155685152	0.000437974	0.005036704
KEGG_GLYCOSPHINGOLIPID_BIOSYNTHESIS_GANGLIO_SERIES	0.151678381	0.001026703	0.008586969
KEGG_BETA_ALANINE_METABOLISM	-0.145542379	0.000570707	0.00583389
KEGG_GLYCOSAMINOGLYCAN_BIOSYNTHESIS_CHONDROITIN _SULFATE	0.144237139	0.000886173	0.007764567
KEGG_HOMOLOGOUS_RECOMBINATION	-0.141959376	0.002459603	0.015605755

Supplementary Table 7. The top 20 GSVA terms of CXCL9.

Pathways	LogFC	P.value	FDR
KEGG_GRAFT_VERSUS_HOST_DISEASE	0.282957086	1.27E-06	2.93E-05
KEGG_ALLOGRAFT_REJECTION	0.28130979	1.62E-07	7.47E-06
KEGG_PRIMARY_IMMUNODEFICIENCY	0.274200636	1.34E-08	1.24E-06
KEGG_INTESTINAL_IMMUNE_NETWORK_FOR_IGA_PRODUCTION	0.273124844	5.38E-09	9.90E-07
KEGG_TYPE_I_DIABETES_MELLITUS	0.216729898	5.40E-06	0.000110495
KEGG_AUTOIMMUNE_THYROID_DISEASE	0.205178161	1.03E-06	2.71E-05
KEGG_VALINE_LEUCINE_AND_ISOLEUCINE_DEGRADATION	-0.194472279	0.00419016	0.027535338
KEGG_PROPANOATE_METABOLISM	-0.186295189	0.003106107	0.021167542
KEGG_ASTHMA	0.178999868	0.000488451	0.005617185
KEGG_ANTIGEN_PROCESSING_AND_PRESENTATION	0.178873662	1.78E-05	0.000297079
KEGG_T_CELL_RECEPTOR_SIGNALING_PATHWAY	0.16991138	2.46E-07	9.06E-06
KEGG_BUTANOATE_METABOLISM	-0.159434902	0.002278434	0.017467998
KEGG_VIRAL_MYOCARDITIS	0.152406694	7.50E-06	0.00013791
KEGG_FATTY_ACID_METABOLISM	-0.146538754	0.005412575	0.034341853
KEGG_SYSTEMIC_LUPUS_ERYTHEMATOSUS	0.144855573	0.000747094	0.007650551
KEGG_CHEMOKINE_SIGNALING_PATHWAY	0.141466051	1.23E-07	7.47E-06
KEGG_BETA_ALANINE_METABOLISM	-0.140925451	0.000866423	0.008390627
KEGG_CELL_ADHESION_MOLECULES_CAMS	0.140619537	6.53E-07	2.00E-05
KEGG_HEMATOPOIETIC_CELL_LINEAGE	0.137185897	4.99E-05	0.000765681
KEGG_LEISHMANIA_INFECTION	0.132412529	0.000748424	0.007650551

Supplementary Table 8. The top 20 GSVA terms of NT5E.

Pathways	LogFC	P.value	FDR
KEGG_VALINE_LEUCINE_AND_ISOLEUCINE_DEGRADATION	0.268767259	5.92E-05	0.000726073
KEGG_PROPANOATE_METABOLISM	0.26782988	1.49E-05	0.000355897
KEGG_NON_HOMOLOGOUS_END_JOINING	0.241074038	0.000507512	0.003735287
KEGG_MATURITY_ONSET_DIABETES_OF_THE_YOUNG	-0.238695958	1.55E-05	0.000355897
KEGG_MISMATCH_REPAIR	0.237927072	5.06E-05	0.000664519
KEGG_PROTEIN_EXPORT	0.210218429	0.000977788	0.005758475
KEGG_GLYCOSPHINGOLIPID_BIOSYNTHESIS_GLOBO_SERIES	-0.206021643	1.99E-05	0.000366165
KEGG_FATTY_ACID_METABOLISM	0.203952086	8.62E-05	0.000933061
KEGG_NICOTINATE_AND_NICOTINAMIDE_METABOLISM	0.203586763	2.54E-08	4.68E-06
KEGG_AMINOACYL_TRNA_BIOSYNTHESIS	0.202799915	0.001030024	0.005758475
KEGG_RENIN_ANGIOTENSIN_SYSTEM	0.189854677	3.03E-06	0.000139364
KEGG_SPLICEOSOME	0.183352551	0.000725563	0.004450122
KEGG_LYSINE_DEGRADATION	0.182695153	3.14E-05	0.000480322
KEGG_BUTANOATE_METABOLISM	0.177215104	0.000656871	0.004167734
KEGG_SULFUR_METABOLISM	-0.176625112	0.00103277	0.005758475
KEGG_BASAL_TRANSCRIPTION_FACTORS	0.167026339	0.000605049	0.003976034
KEGG_COLORECTAL_CANCER	0.16453329	8.69E-08	8.00E-06
KEGG_NEUROACTIVE_LIGAND_RECEPTOR_INTERACTION	-0.164283705	8.32E-05	0.000933061
KEGG_OLFACTORY_TRANSDUCTION	-0.160357734	0.000263719	0.002426211
KEGG_RNA_DEGRADATION	0.159304048	0.001426328	0.007718949

Supplementary Table 9. The top 20 GSVA terms of FGF7.

Pathways	LogFC	P.value	FDR
KEGG_PROTEIN_EXPORT	0.387193054	1.47E-10	1.35E-08
KEGG_NON_HOMOLOGOUS_END_JOINING	0.350637284	2.03E-07	1.78E-06
KEGG_AMINOACYL_TRNA_BIOSYNTHESIS	0.341186071	9.03E-09	1.85E-07
KEGG_ONE_CARBON_POOL_BY_FOLATE	0.327332488	3.01E-08	4.61E-07
KEGG_MISMATCH_REPAIR	0.31217779	5.10E-08	5.87E-07
KEGG_SPLICEOSOME	0.292664541	2.25E-08	3.76E-07
KEGG_VALINE_LEUCINE_AND_ISOLEUCINE_BIOSYNTHESIS	0.291787649	7.32E-06	3.74E-05
KEGG_CELL_CYCLE	0.274431256	4.17E-12	7.68E-10
KEGG_PROTEASOME	0.267658925	1.61E-05	7.41E-05
KEGG_RNA_DEGRADATION	0.265276504	3.62E-08	5.12E-07
KEGG_NOD_LIKE_RECEPTOR_SIGNALING_PATHWAY	0.263666936	1.68E-09	7.74E-08
KEGG_SULFUR_METABOLISM	-0.26215574	5.76E-07	3.99E-06
KEGG_NUCLEOTIDE_EXCISION_REPAIR	0.24581461	2.03E-07	1.78E-06
KEGG_BASAL_TRANSCRIPTION_FACTORS	0.245156814	2.30E-07	1.92E-06
KEGG_CITRATE_CYCLE_TCA_CYCLE	0.236617885	3.59E-05	0.000157218
KEGG_DNA_REPLICATION	0.235961202	9.30E-06	4.60E-05
KEGG_AMINO_SUGAR_AND_NUCLEOTIDE_SUGAR_METABOLISM	0.2267432	2.36E-06	1.31E-05
KEGG_PROPANOATE_METABOLISM	0.221093362	0.000404135	0.001352016
KEGG_HOMOLOGOUS_RECOMBINATION	0.218482259	1.75E-06	1.08E-05
KEGG_TERPENOID_BACKBONE_BIOSYNTHESIS	0.211187654	0.00022772	0.00080578