

Analysis and Validation of Autophagy-Related Gene Biomarkers and Immune Cell Infiltration Characteristic in Bronchopulmonary Dysplasia by Integrating Bioinformatics and Machine Learning

Shuzhe Xiao^{1,*}, Yue Ding^{1,*}, Chen Du^{1,*}, Yiting Lv¹, Shumei Yang^{1,2}, Qi Zheng³, Zhiqiu Wang³, Qiaoli Zheng³, Meifang Huang¹, Qingyan Xiao¹, Zhuxiao Ren², Guangliang Bi¹, Jie Yang^{1,2}

¹Department of Neonatology, Nanfang Hospital, Southern Medical University, Guangzhou, 510515, People's Republic of China; ²Department of Neonatology, Guangdong Women and Children Hospital, Guangzhou, 511442, People's Republic of China; ³The First Clinical Medical College of Southern Medical University, Guangzhou, 510515, People's Republic of China

*These authors contributed equally to this work

Correspondence: Jie Yang; Guangliang Bi, Email Jieyang0830@126.com; bader515@163.com

Background: Autophagy and immunity play important regulatory roles in lung developmental disorders. However, there is currently a lack of bioinformatics analysis on autophagy-related genes (ARGs) and immune infiltration in bronchopulmonary dysplasia (BPD). We aim to screen and validate the signature genes of BPD by bioinformatics and in vivo experiment.

Methods: GSE8586 was obtained from the Gene Expression Omnibus (GEO) database. The differentially expressed genes (DEGs) were identified using the R program. Using cell-type identification with CIBERSORT to analyze the inflammatory and immune status of BPD. Subsequently, the hub genes were identified by Lasso and Cytoscape with three machine-learning algorithms (MCC, Degree and MCODE). In addition, hub genes were validated with ROC, single-cell sequence and IHC in hyperoxia rats. Finally, we searched the drug targets of these hub genes, and established a nomogram model for predicting the risk of BPD.

Results: There were 73 the differentially expressed and autophagy-related genes (DE-ARGs) by overlapping the DEGs in GSE8586 and ARGs. Five hub genes, BRX1, JUN, PES1, NR4A1 and RRP9, were lowly expressed in the BPD group and had high diagnostic value in the diagnostic model. All hub genes are mainly located in B cell, epithelial cell, fibroblast, endothelial cell, smooth muscle cell and pneumocyte in lung single-cell sequencing. Moreover, immune infiltration analysis showed immune cells were higher in the BPD group and were closely associated with hub genes. We also predict the drug targets of the genes. Finally, the IHC result in rats showed that expression of PES1, BRX1, RRP9, JUN, NR4A1 was lower in the hyperoxia group compared to the normoxia group.

Conclusion: BRX1, JUN, PES1, NR4A1, RRP9, may be promising therapeutic targets for BPD. Our findings provided researchers and clinicians with more evidence regarding immunotherapeutic strategies for BPD treatment.

Keywords: BPD, autophagy, immune cell infiltration, single cell sequencing, biomarkers

Introduction

BPD, a potentially fatal illness, is a significant public health concern because of its association with an aberrant immune response to infections and severe dyspnea.¹ Despite major efforts to minimize injurious but often life-saving postnatal interventions (such as oxygen, mechanical ventilation and corticosteroids), BPD remains the most common frequent complication of extreme preterm birth.^{2,3} Nearly one-third to one-half of all infants born before 28th week of gestation develop BPD.⁴ However, the precise mechanisms of BPD remain unclear.

An increasing body of researches suggests that inflammatory mediators, including cytokines and growth factors, play a role in BPD pathogenesis, and their activity can be upregulated by inflammatory processes.^{5,6} Autophagy has been

implicated in the regulation of immune cell infiltration and is associated with lung injury.⁷ According to an investigation, BPD is positively correlated with the occurrence of intrauterine infection.⁸ Besides, the umbilical cord tissue not only reflects but also can influence the intrauterine environment. Moreover, umbilical cord blood contains not only hematopoietic stem cells but also NK cells, lymphocytes, monocytes, mesenchymal stem cells, and other kinds of cells.⁹ The wide origins of these many types of cells make single-cell sequencing possible, allowing us to identify more extensively the links between genes and immune cells, as well as understand the probable mechanisms of action of BPD-related genes and immune cells. Therefore, we could evaluate the intrauterine infection by detecting the umbilical cord tissue to identify possible factors in the pathogenesis of BPD.

Microarray expression profiling provides a comprehensive portrait of the transcriptional world, enabling us to view the organism as a “system” that is more than the sum of its parts.¹⁰ Machine-learning algorithms have shown great promise in assessing transcriptome data and locating genes that are of biological significance.¹¹ Machine learning and bioinformatics is an innovative research method. This approach helps improve the precision of our research on the relationship between gene targets and diseases, and enables us to more efficiently discover additional gene targets that we have not previously identified. However, there are few investigations of biomarkers based on machine learning about BPD. What remains unclear is the extent to which biologic markers obtained from the umbilical cord might predict the occurrence of BPD and have an influence on preterm infants. This study provides new insights into the potential therapeutic interventions of BPD.

Method

Source of Microarray Data

In this study, we downloaded a microarray dataset (GSE8586) from the Gene Expression Omnibus (GEO) database¹² (<https://www.ncbi.nlm.nih.gov/geo/>). The GSE8586 dataset was applied to identify DEGs, which contained umbilical cord tissue samples from infants born before 28th week of gestation, including 20 infants with BPD and 34 neonates with unaffected controls, respectively. BPD is defined as the requirement for supplemental oxygen at 36 weeks postmenstrual age. To ensure the reliability and accuracy of our subsequent analyses, we used the Affy package of R to perform normalization and background correction. The “sva” package was used to perform batch normalization. Besides, we downloaded 4409 autophagy-related genes (ARGs) in GeneCards. A schematic diagram of our study is depicted in [Figure 1](#).

Sample Detection and Differential Expression Genes Analysis

The DEGs between normal and BPD groups in the GSE8586 dataset were identified using the limma package. $|\log_2 \text{FC}| > 0.5$ and $p \text{ value} < 0.05$ were regarded as significant. Besides, volcano plot was produced to highlight the differential expression of DEGs. A heatmap was generated based on the screened DEGs.

Protein–Protein Interaction and Module Analysis

Protein–protein interactions (PPIs) were analyzed using STRING for interacting Genes/Proteins dataset (<http://cn.string-db.org/version10.0>), and an interaction with a combined score > 0.4 was considered statistically significant.¹³ We used the Cytoscape, an open-source bioinformatics software platform, to visualizing molecular interaction networks.¹⁴ Besides, the most significant modules in the PPI networks were identified using MCODE, MCC and Degree. The criteria for selection were as follows: MCODE scores > 5 , degree cut-off = 2, node score cut-off = 0.2, Max depth = 100 and $\kappa\text{-score} = 2$.

Least Absolute Shrinkage and Selection Operator (LASSO)

To reduce data dimensionality and identify genetic biomarkers for BPD, we used LASSO with R (4.2.1) version and glmnet [4.1.7] to analyze the cleaned data and obtain variable coefficient values, logarithm of lambda values, likelihood values or classification error rates.¹⁵ This allows variable selection and complexity adjustment by controlling the lambda

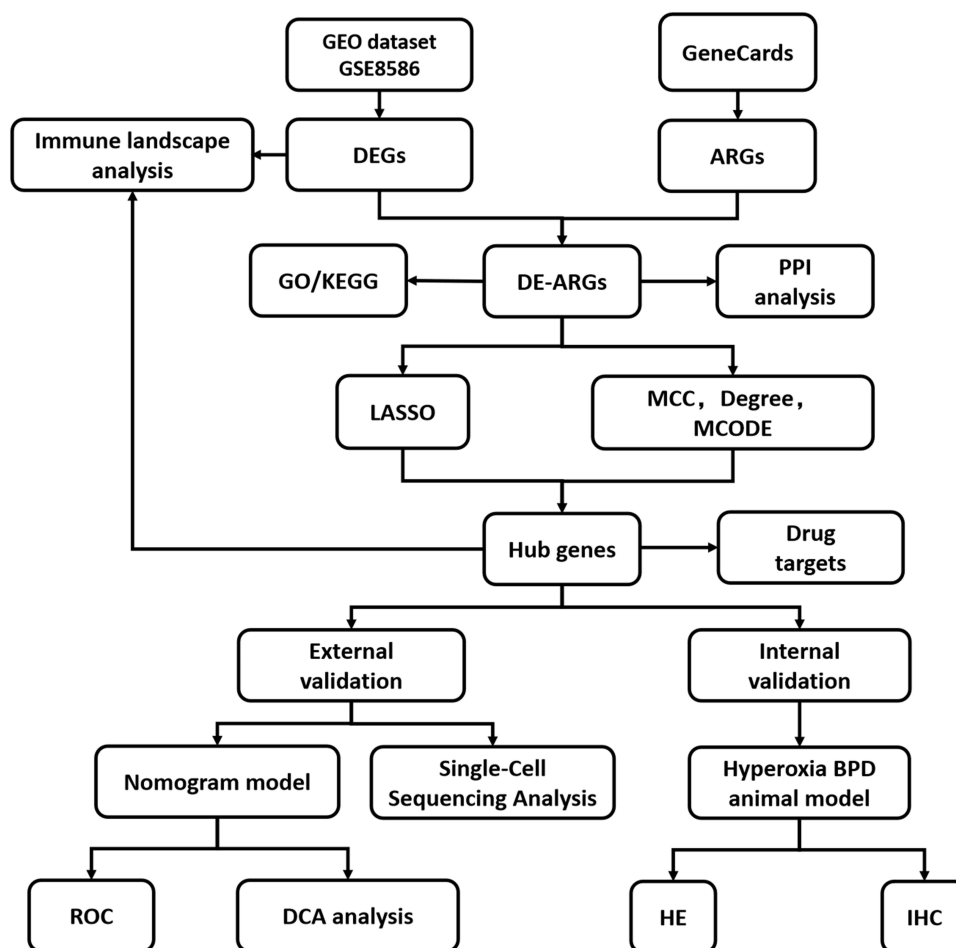


Figure 1 The overall research design. Screening DE-ARGs through GSE8586 and autophagy-related genes firstly. Using cell-type identification with CIBERSORT to analyze the immune status of BPD. The hub genes were identified by the use of Lasso and Cytoscape with four machine-learning algorithms (MCC, Degree, MCODE and LASSO). To validate hub genes, we use ROC, single cell sequence and IHC in hyperoxia rats. At last, we searched the drug targets of these hub genes, and established a nomogram model for predicting the risk of BPD.

parameter. Finally, optimum gene biomarkers for BPD were discovered using biomarkers produced by Lasso algorithm and 3 distinct algorithms in Cytoscape.

Annotation and Enrichment Analysis

Functional enrichment was applied to the data to verify the likely functions of potential targets. Gene ontology (GO) contained biological pathways (BP), cellular components (CC) and molecular functions (MF). Kyoto Encyclopedia of Genomes (KEGG) pathway enrichment was used to investigate the activities of genes and the high-level genomic information to connect to those functions.¹⁶ Both GO and KEGG analyses were carried out by cluster Profiler [4.4.4], GPlot [1.0.2], ggplot2 [3.3.6], igraph [1.4.1] and ggraph [2.1.0].

Immune Infiltration Analysis

CIBERSORTx, an analytical tool from the Alizadeh Lab and Newman Lab, was utilized to impute gene expression profiles and estimate the abundances of 22 different member immune cell types in a mixed cell population,¹⁷ using gene expression data in BPD samples from the GSE8586 dataset. All immune cell-type fractions were summed up to one for each sample. The Spearman association between unique diagnostic markers and immune invading cells was analyzed using the “ggstatsplot” and “ggplot2” packages to illustrate the results.

Predicting Drug Targets for Hub Genes

We accessed the DSigDB drug database on Enrichr (<https://maayanlab.cloud/Enrichr/>) and inputted the five hub genes we studied. Using $P < 0.05$ as a condition, we screened for drug targets to predict the drug targets of the genes.

Receiver Operating Characteristic (ROC) Curve

The ROC curve was analyzed using R (4.2.1) version, pROC [1.18.0] package, and the results were visualized using ggplot [3.3.6] package.¹⁸ The Area Under Curve (AUC) of the ROC curve is used to evaluate the accuracy of the diagnostic model for autophagy-related genes in BPD. The AUC value ranges from 0.5 to 1, with values closer to 1 indicating better diagnostic performance of the variable in predicting the outcome. We obtained the final AUC results with the function of pROC and evaluation of the AUC.

Single-Cell Sequencing Analysis

The single-cell sequencing data about the hub genes were used from Single Cell Portal. We used single-cell sequencing data from human lung tissues to explore the expression of hub genes.

Establishment of Nomogram

A diagnostic nomogram based on the selected candidate DE-ARGs was established using the “rms” package in the R software. The consistency between our predicted values and reality was assessed using a calibration curve. ROC curve analysis and decision curve analysis (DCA) were carried out to determine whether or not the decisions made based on the model were favorable to the patients.

Hematoxylin and Eosin (HE) Staining

After the left lung tissues were fixed with 4% paraformaldehyde, dehydrated with gradient ethanol, 75%, 85%, 90%, 95% alcohol and anhydrous ethanol I, II for 2 h, respectively. Xylene I, II, III for 40 min, respectively. 65°C molten paraffin I for 0.5 h, 65°C molten paraffin II for 1 h, 65°C molten paraffin III for 2 h and 45 min, and paraffin embedded, slides with 5µm paraffin sections were dewaxed and hydrated. Then, the sections incubated with hematoxylin for 5 min, stained with eosin for 2 min. Finally, using the microscope to observe the morphology of lung tissues (×200 magnification).

Immunohistochemical (IHC)

The methodology employed utilizes rats left lung tissue, initially fixed and embedded, subsequently sectioned and subjected to a process of fixing and antigen repair. The sections are then immunostained and sealed for observation. The expression of hub genes were evaluated according to the area of positive cells. The primary antibodies used for IHC staining overnight were BRX1 (Cat No. 17295-1-AP; Proteintech; 1:100), JUN (Cat No. 24909-1-AP; Proteintech; 1:100), PES1 (Cat No. 13553-1-AP; Proteintech; 1:200), NR4A1 (Cat No. 25851-1-AP; Proteintech; 1:200), RRP9 (Cat No. 13553-1-AP; Proteintech; 1:100). This study was approved by the Institutional Animal Experiment Committee of Southern Medical University Nanfang Hospital.

Results

Screening of DEGs and Immune Landscape Analysis in BPD

After normalizing each sample in GSE8586 dataset (Figure 2A), the expression patterns of the DEGs were visualized in a clustering heatmap (Figure 2B). Then, we found 647 DEGs between BPD and control groups based on the criterion with $|\log_2 FC| > 0.5$ and $p \text{ value} < 0.05$ (Figure 2C and D). In addition, we used CIBERSORT to estimate the immune infiltration between BPD and control group. The results showed that T cells follicular helper (Tfh cells), Monocytes, plasma cells, B cells memory, Eosinophils, NK cells resting, T cells CD8 and Macrophages M1 were the main immune cells that infiltrated the plaque (Figure 2E). The abundance of B cells memory, T cells gamma delta, Eosinophils ($p \text{ value} < 0.05$), Macrophages M1 and Dendritic cells resting ($p \text{ value} < 0.01$) were significantly increased in patients with BPD, while Tfh cells, NK cells resting and Monocytes ($p \text{ value} < 0.05$) were significantly decreased in preterm infants with

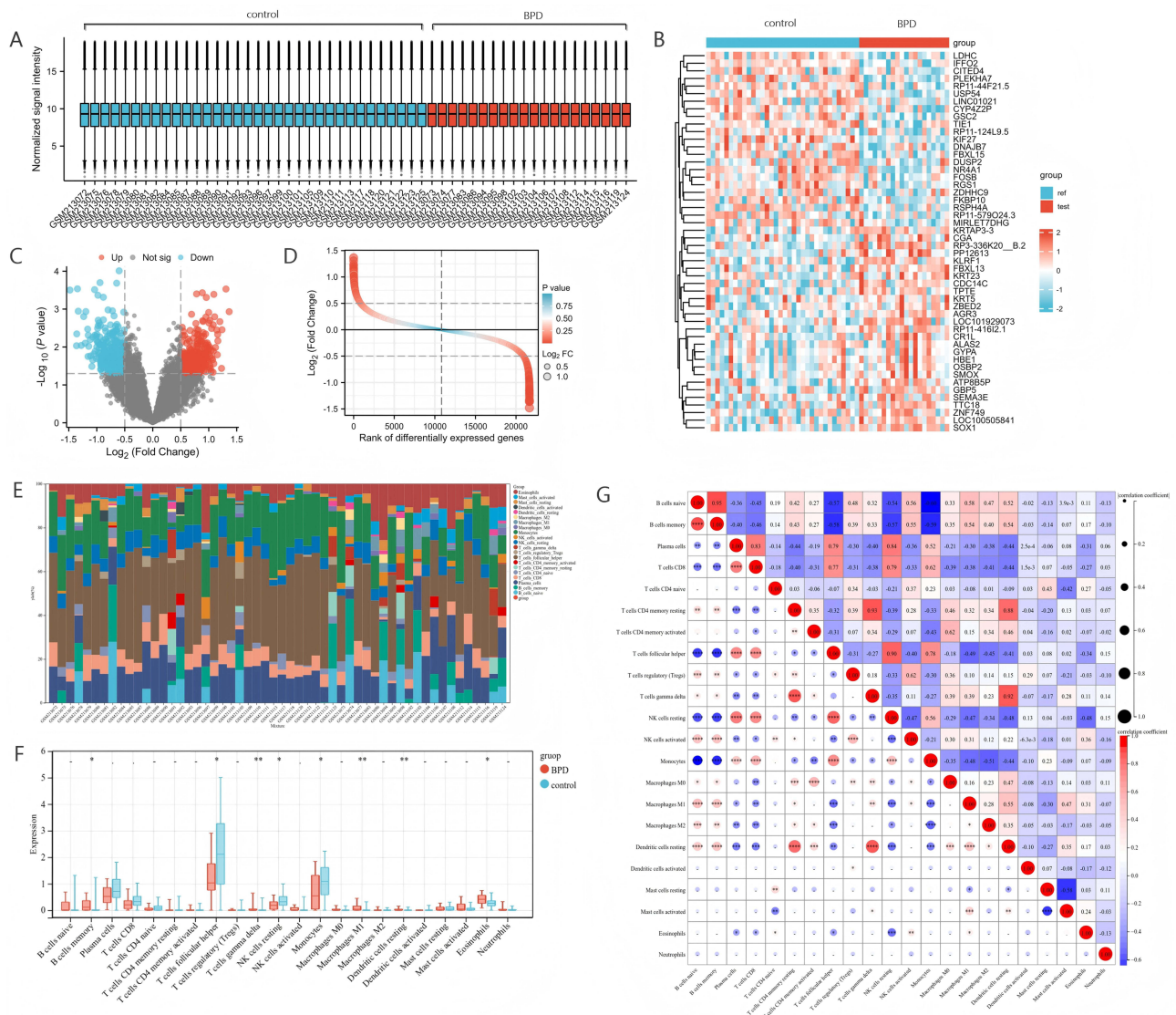


Figure 2 Selection of DEGs and immune infiltration analysis from dataset GSE8586. **(A)** Normalization of each sample data. **(B)** Heatmap displaying the expression levels of various genes. **(C)** Selection of DEGs illustrated with a volcano plot. **(D)** Differential analysis chart presenting the DEGs. **(E)** Cibersort immune infiltration analysis of all samples containing DEGs from GSE8586. **(F)** Expression profiles of 22 immune cell types in BPD and control groups. **(G)** Heatmap of the correlation among immune cells. *** $P < 0.0001$, ** $P < 0.001$, * $P < 0.01$, $P < 0.05$.

BPD (Figure 2F). The correlation between immune cells in BPD was further investigated by Pearson’s correlation analysis (Figure 2G).

Functional Enrichment of 73 DE-ARGs

To further explore whether autophagy played a role in BPD, we downloaded 4409 autophagy-related genes (ARGs) from GeneCards. There were 73 the differentially expressed and autophagy-related genes (DE-ARGs) by overlapping the DEGs in GSE8586 and ARGs, including 43 upregulated genes and 30 downregulated genes (Figure 3A). To gain a deeper understanding of the biological functions of the DE-ARGs, we conducted functional analysis. The results of GO enrichment analysis revealed that these DE-ARGs were linked to response to mechanical stimulus, release of cytochrome c from mitochondria, apoptotic mitochondrial changes, connexin complex, gap junction, preribosome, R-SMAD binding, gap junction channel activity and wide pore channel activity. KEGG analysis was associated with IL-17 signaling pathway and colorectal cancer (Figure 3B–D).

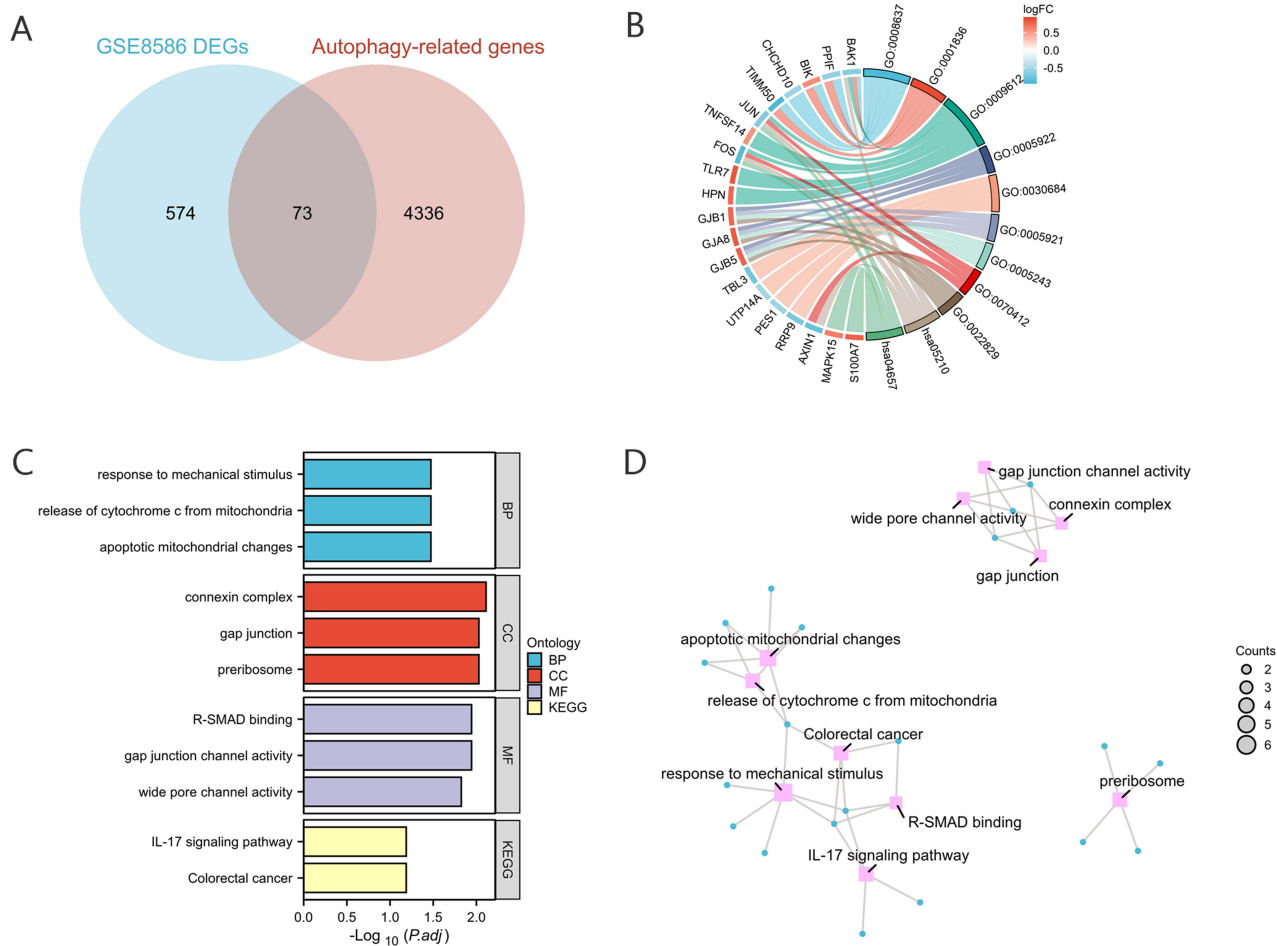


Figure 3 Enrichment analysis of differentially expressed genes and autophagy-related genes. **(A)** Intersection of DEGs from GSE8586 with autophagy-related genes (ARGs) from the Genecards dataset. **(B)** Chord diagram illustrating combined GO and KEGG analysis with LogFC. **(C)** GO and KEGG analysis of the 73 intersecting genes. **(D)** Network diagram of the GO and KEGG enrichment analysis results.

Identification and Immune Landscape Analysis of 5 Hub Genes as Diagnostic Genes for BPD

Then, the protein–protein interaction network of 73 DE-ARGs was conducted, with 30 downregulated genes and 43 upregulated genes (Figure 4A). And the Cytoscape was used to visualize these more interacting genes (Figure 4B). To evaluate the diagnostic potential of autophagy-related genes between BPD and control samples, we conducted 3 distinct machine learning algorithms (MCC, Degree and MCODE) on the GSE8586 dataset (Figure 4C–G). By utilizing a Venn diagram to compare the overlapping regions of DEGs and key module genes, we were able to identify 10 genes as module genes (Figure 4H). To further minimize the potential diagnostic genes for BPD, we used the LASSO algorithm to identify TBL3, PES1, BRIX1, RRP9, JUN, DUSP1, NR4A1 (Figure 5A and B). Then, we drew a one-dimensional bar chart of 7 genes based on their lambda min, and selected 5 most influential genes, PES1, BRIX1, RRP9, JUN, and NR4A1 (Figure 5C). Finally, we used ROC to validate the 5 hub genes, each AUC of hub genes was higher than 0.67 (Figure 5D). Subsequent GO enrichment analysis showed that 5 hub genes were correlated with ribosome biogenesis, rRNA metabolic process, rRNA processing, transcription regulator complex, preribosome, preribosome large subunit precursor, MAP kinase tyrosine phosphatase activity, U3 snoRNA binding, protein tyrosine phosphatase activity, while KEGG enrichment analysis presented these genes were linked to MAPK signaling pathway, fluid shear stress and atherosclerosis and cocaine addiction (Figure 5E–G). In addition, PES1 was correlated with RRP9, while Jun was correlated with NR4A1 (Figure 5H). And the expression of 5 hub genes were significantly different between the BPD and

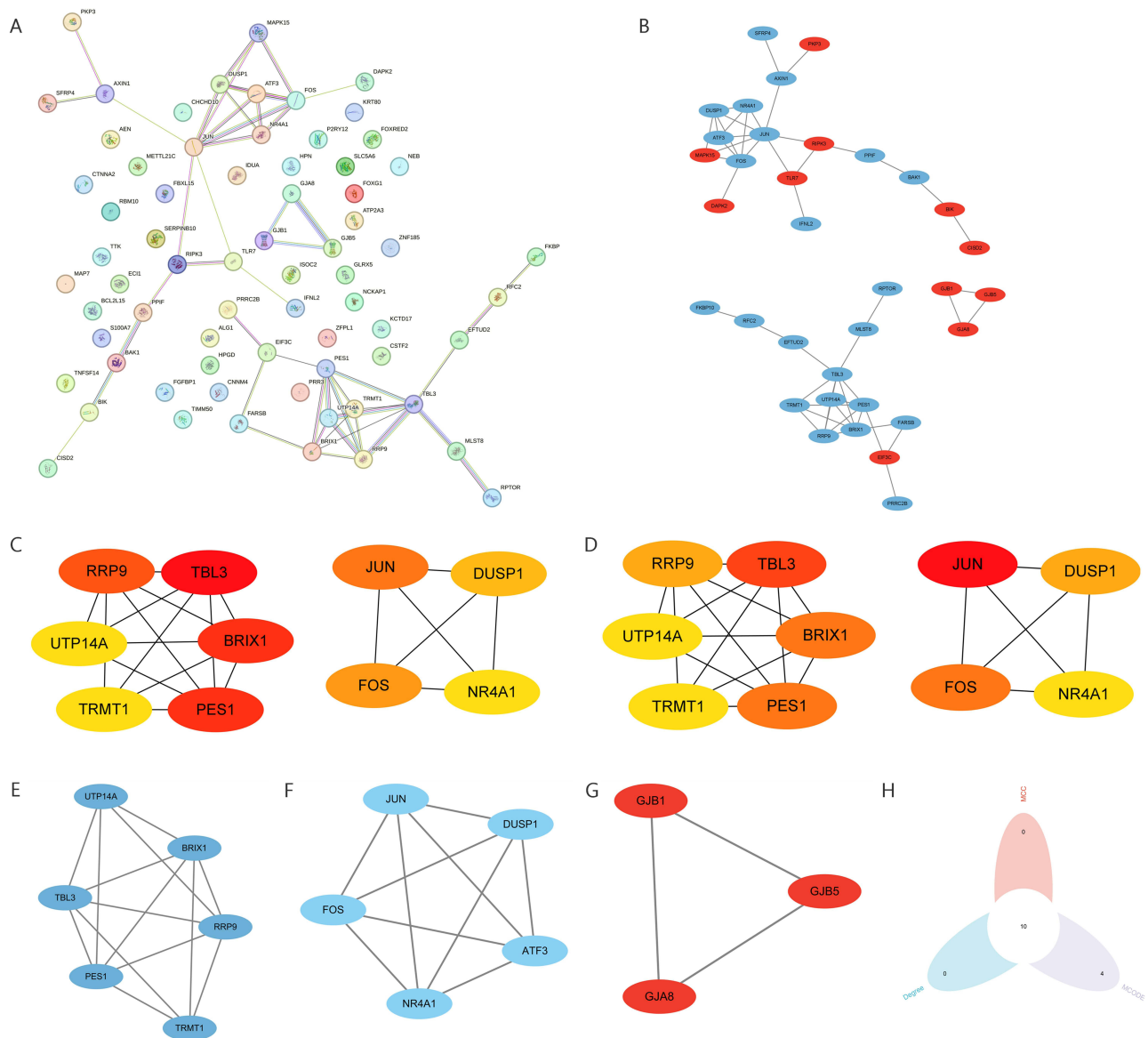


Figure 4 PPI interaction and module analysis of autophagy-related genes. **(A)** PPI network analysis of autophagy-related differentially expressed genes (DE-ARGs). **(B)** Visualization of more interacting DE-ARGs using Cytoscape software, with 23 downregulated genes (blue) and 11 upregulated genes (red). **(C and D)** Selection of the top 10 genes using the MCC (Maximal Clique Centrality) and Degree algorithms. **(E–G)** Identification of 3 clusters using the MCODE (Molecular Complex Detection) algorithm. **(H)** Intersection of the three algorithms in a Venn diagram to identify 10 common genes.

control groups (Figure 5I). Therefore, BRIX1, JUN, PES1, NR4A1, and RRP9 may be potential therapeutic targets for patients with BPD.

To study immune cell infiltration in BPD and control, the corrected expression matrix was subjected to CIBERSORT to estimate the abundances of infiltrating immune cells in a mixed-cell population.¹⁷ We analyzed immune cell infiltration in the 5 hub genes. We found that **JUN** was positively correlated with plasma cell, T cells CD8, Tfh cells, NK cells resting, and monocytes, while negative correlated with B cells, Macrophages M2 (p value < 0.01), Macrophages M1, T cells CD4, and Dendritic cells resting (p value < 0.05). **NR4A1** was positively correlated with T cells CD4, Tfh cells, monocytes (p value < 0.05). **BRIX1** was positive correlated with plasma cells, while negatively correlated with Eosinophils (p value < 0.05). **RRP9** was positive correlated with Dendritic cell, while negatively correlated with Neutrophils (p value < 0.05). However, **PES1** was not correlated with these immune cells (Figure 5J).

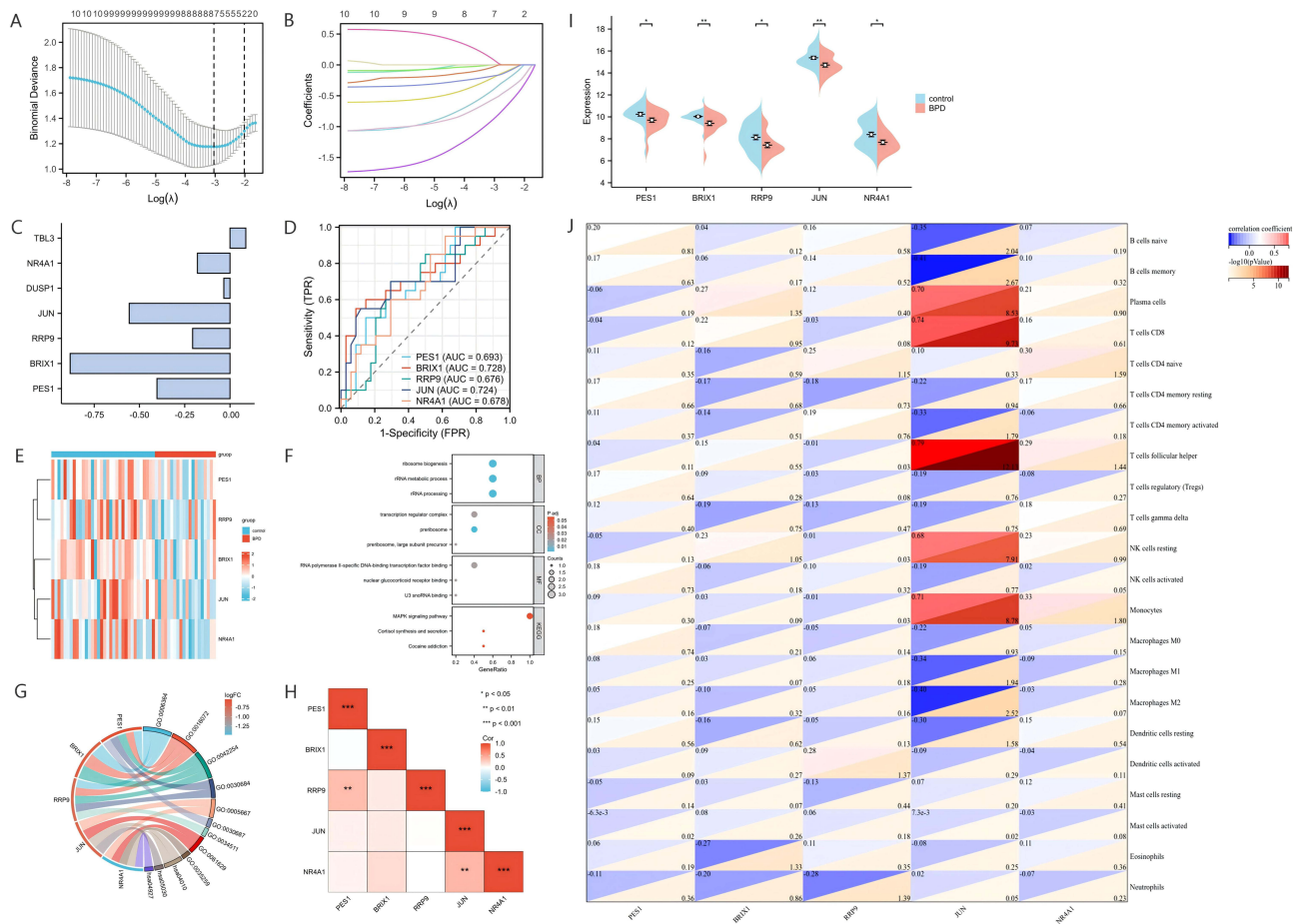


Figure 5 Screening of hub genes and analysis of their functional enrichment and immune infiltration. **(A)** Diagnostic Lasso coefficient screening. **(B)** Diagnostic Lasso variable trajectory. **(C)** One-dimensional bar chart screening of the top 5 significant impact factors. **(D)** ROC curve confirming the reliability of the 5 hub genes. **(E)** Heatmap presentation of hub gene expression. **(F)** Cluster analysis of hub genes using GO and KEGG. **(G)** Combined LogFC enrichment analysis of GO and KEGG. **(H)** Correlation among hub genes. **(I)** Expression of five hub genes between BPD and control groups; a Mann–Whitney *U*-test (Wilcoxon rank sum test) was employed for statistical analysis, with bean plots illustrating the comparative expression of the five hub genes across multiple dimensions between the two groups. **(J)** Expression of the five hub genes across 22 types of immune cells. ****P*<0.001, ***P*<0.01, **P*<0.05.

Structure and Function of Hub Genes

PES1 (Pescadillo ribosomal biogenesis factor 1) located on chromosome 22 is a protein-coding gene, which is related to rRNA processing in the nucleus and cytosol, processing of capped intron-containing pre-mRNA and processing cell cycle regulation.^{19,20} **BRX1**, located in chromosome 5 and nucleolus, is involved in ribosomal large subunit assembly.²¹ **RRP9** (Ribosomal RNA processing 9) encodes a member of the WD-repeat protein family, which participate in the processing and modification of pre-rRNA. Ectopic expression of RRP9 promotes tumor cell proliferation, colony formation, and cell migration.²² **JUN** is the putative transforming gene of avian sarcoma virus 17. It encodes a protein, which is highly similar to the viral protein, and which interacts directly with specific target DNA sequences to regulate gene expression. Together with FOSB, plays a role in activation-induced cell death of T cells by binding to the AP-1 promoter site of FASLG/CD95L, and inducing its transcription in response to activation of the TCR/CD3 signaling pathway.²³ **NR4A1** (Nuclear Receptor Subfamily 4 Group A Member 1) encodes a member of the steroid-thyroid hormone-retinoid receptor superfamily. The encoded protein acts as a nuclear transcription factor.²⁴ Translocation of the protein from the nucleus to mitochondria induces apoptosis ([Supplementary Table 1](#) and [Supplementary Figure 1](#)).

Drug Targets in Hub Genes

Using data from the DSigDB database, we identified Quinpirole HL60 DOWN, Pergolide HL60 DOWN, Atrazine CTD 00005450, Cyindrospermopsin CTD 00003136, Alprostadil HL60 DOWN, 5-Fluorouracil CTD 00005987, Genistein

CTD 00007324, Progesterone CTD 00006624, Calcitriol CTD 00005558, Estradiol CTD 00005920, Hydrogen peroxide CTD 00006118, and Trichostatin A CTD 00000660 were identified as candidate drugs for most gene interactions (Table 1). Since these iconic drugs have been detected in the core genes, they may target PES1, BRX1, RRP9, JUN and NR4A1.

Clinical Implications

To further analyze the relevance of these drugs in treating BPD and how they interact with the identified hub genes, we conducted a detailed research on these hub genes-targeted drugs' potential therapeutic effects in BPD. Clinically, Alprostadil, prostaglandin E1, can be used to dilate pulmonary blood vessels, alleviating the respiratory distress symptoms caused by BPD-related pulmonary vascular narrowing.²⁵ Estradiol may improve lung function and outcome in preterm infants.²⁶ Calcitriol is a receptor agonist of vitamin D. Vitamin D protects against hyperoxia-induced BPD in newborn rats by regulating the vitamin D-VDR signaling pathway.²⁷ Trichostatin interferes with the development of alveoli in newborn rats²⁸ (Table 2). These gene-targeting drugs provide practical and promising clinical value for the treatment of BPD diseases.

Table 1 Drug Targets in Hub Genes

Term	P-value	Genes
Quinpirole HL60 DOWN	1.78E-05	PES1, BRX1, RRP9
Pergolide HL60 DOWN	4.79E-05	PES1, BRX1, RRP9
Atrazine CTD 00005450	7.18E-05	NR4A1, JUN, PES1, BRX1, RRP9
Cylindrospermopsin CTD 00003136	1.35E-04	NR4A1, JUN, RRP9
Alprostadil HL60 DOWN	1.50E-04	PES1, BRX1, RRP9
5-Fluorouracil CTD 00005987	0.001975575	NR4A1, JUN, BRX1
Genistein CTD 00007324	0.002117354	NR4A1, JUN, RRP9
Progesterone CTD 00006624	0.007556722	NR4A1, BRX1, RRP9
Calcitriol CTD 00005558	0.008049676	NR4A1, JUN, BRX1
Estradiol CTD 00005920	0.009121609	NR4A1, JUN, BRX1, RRP9
Hydrogen peroxide CTD 00006118	0.019328008	NR4A1, JUN, BRX1
Trichostatin A CTD 00000660	0.04316521	NR4A1, JUN, BRX1

Table 2 Hub Genes-Targeted Drugs' Potential Therapeutic Effects in BPD

Drug	Function
Alprostadil	Like nitric oxide, prostaglandin E1 has the effect of dilating blood vessels. Clinically, prostaglandin E1 can be used to dilate pulmonary blood vessels, alleviating the respiratory distress symptoms caused by BPD-related pulmonary vascular narrowing. ²⁵
Estradiol	In a primate model of BPD, postnatal estradiol treatment had favorable cardiovascular impact, enhanced pulmonary function, and lowered requirements for ventilatory support in association with an up-regulation of lung NOS. Estradiol may be an efficacious postnatal therapy to improve lung function and outcome in preterm infants. ²⁶
Calcitriol	Calcitriol is a receptor agonist of vitamin D. Vitamin D protects against hyperoxia-induced BPD in newborn rats by regulating the vitamin D-VDR signaling pathway. ²⁷
Trichostatin	The application of Trichostatin alone can interfere with the development of alveoli in newborn rats. ²⁸
Genistein	When lung cells are stressed in the presence of a tyrosine kinase inhibitor (Genistein), the stress-induced redirection and elongation are inhibited. ²⁹
Progesterone	Progesterone prophylaxis is widely used to prevent preterm birth. ³⁰
Pergolide	Pirenzepine can promote the production of nerve growth factors and facilitate corneal nerve regeneration in a corneal epithelial scratch mouse model. Experimental results indicate that pirenzepine can alleviate pain and improve corneal nerve regeneration after alkaline burns. ³¹
5-Fluorouracil	5-Fluorouracil is an essential component of systemic chemotherapy for colorectal cancer (CRC) in the palliative and adjuvant settings. ³²
Hydrogen peroxide	Hydrogen peroxide can stimulate the tyrosine phosphorylation of various cytoplasmic proteins (including MAP kinases). ³³
Quinpirole	Quinpirole has neuroprotective effects on pilocarpine-induced epileptic rats. Additionally, it has been found that this treatment not only alleviates the epileptic symptoms in rats but also enhances their learning and memory abilities. ³⁴

Single-Cell Sequencing Analysis

We used the Single-Cell Portal to investigate single-cell sequencing of 5 hub genes expression in the lung, which suggested that the hub genes were highly expressed in the human lung. The most significant expression of hub genes is JUN. These data were classified for 13 subgroups: apoptotic cell, B cell, endothelial cell, epithelial cell, fibroblast, lymph node lymphatic vessel endothelial cell, lymphocyte, mast cell, mitotic cell cycle, myeloid cell, smooth muscle cell, typeI pneumocyte and typeII pneumocyte. Interestingly, all hub genes mainly located in B cell, epithelial cell, fibroblast, endothelial cell, smooth muscle cell, typeI pneumocyte and typeII pneumocyte (Figure 6).

Establishment of Nomogram

We developed a diagnostic nomogram based on these five biomarkers (Figure 7A). ROC curves revealed that the model distinguished between BPD and unaffected control samples (Figure 7B). The predictive accuracy of the nomogram was highlighted using the calibration curves (Figure 7C). The line in the DCA curve remained above the gray and black lines from 0 to 1, suggesting that decisions based on the nomogram model may favor patients with BPD (Figure 7D).

Validation of Hub Genes in Animal Experiment

In order to show the expression levels of ARGs in rats lung, animals' studies were divided into two groups, one exposure of male and female newborn SD rats background to hyperoxia (85% O₂) for 14 days, while other group exposure to normoxia (21% O₂), as shown in Figure 8A. Alveolarization was greater in 85% O₂ rats than in 21% O₂ rats (Figure 8B). The IHC results showed that expression of PES1, BRX1, RRP9, JUN, NR4A1 was lower in the hyperoxia group compared to normoxia group (Figure 8C and D). There were 3 animals analyzed in each group, P < 0.05.

Discussion

Immune Disorder of BPD Diseases

BPD remains the most common complication associated with prematurity and is increasing in prevalence, most likely due to the increased survival of preterm infants.³⁵ Inflammatory activity in the lungs leads to the dysregulation of pro-inflammatory cytokines and growth factors, including the migration of inflammatory cells to the lungs.³⁶ In our results,

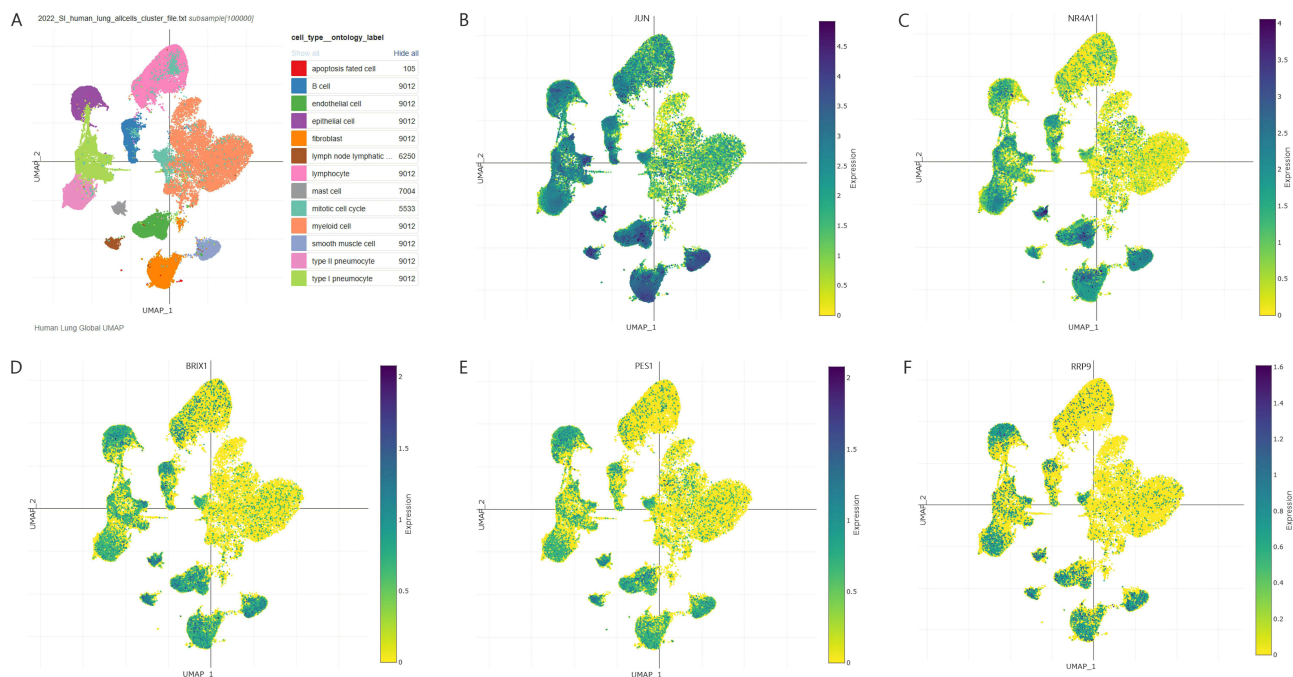


Figure 6 Expression of hub genes in single-cell sequencing dataset. (A) Using single-cell sequencing data from human lung tissue via the Single Cell Portal, cells are categorized into 13 subgroups. (B–F) In the single-cell sequencing data, the expression of hub genes ranges from strong to weak: JUN, NR4A1, BRX1, PES1, RRP9.

we found higher abundance of Macrophages M1, B cells memory, T cells gamma delta, and Eosinophils, while lower abundance of Tfh cells, NK cells resting and Monocytes in BPD.

Autophagy-Dependent Mechanism Influence BPD

Autophagy is considered a fundamental regulator of the immune response and may also influence the pathogenicity and inflammatory responses in diseases.³⁷ The progression of BPD is primarily driven by the apoptosis of alveolar epithelial cells, and autophagy plays a role in apoptosis. Our research showed that most autophagy-related genes were correlated with immune infiltration in BPD samples. Our findings were accordance with other researches. For example, an experiment demonstrated that autophagy-promoting peptide promoted autophagy and prevent apoptosis caused by exposure to hyperoxia-induced BPD model.³⁸ Another study represented that AMP-activated protein kinase-induced macrophage M2 with autophagy-dependent mechanism and played a protective role in a murine model of BPD.³⁹ A report showed that autophagy acted as an important regulator of lung development and morphogenesis, contributing to the BPD phenotype when impaired.⁴⁰ These results support the general understanding that inflammatory factors with autophagy-dependent mechanism influence the progress of BPD.

Hub Genes Involve in the Progress of Autophagy

Between the BPD and control groups, our research revealed a total 73 DE-ARGs, with 43 genes being upregulated and 30 genes being downregulated. On the basis of these DE-ARGs, we incorporated Lasso and three machine-learning algorithms to screen and identify BPD markers. We identified 5 hub genes (BRX1, JUN, PES1, NR4A1 and RRP9) that could play a pivotal role in the development of BPD, which were correlated with inflammation and autophagy. For example, the lung macrophages expansion correlated with IL4 production and was affected by the deficiency of **NR4A1**.⁴¹ Translocation of the NR4A1 protein from the nucleus to mitochondria induces apoptosis. NR4A1 can regulate the inflammatory response in macrophages by regulating metabolic adaptations during inflammation, including repressing the transcription of genes involved in the citric acid cycle (TCA).^{42,43} NR4A1 also inhibits NF-kappa-B signaling by binding to low-affinity NF-kappa-B binding sites, such as at the IL2 promoter.⁴⁴ Take another example, **PES1** is related to rRNA processing in the nucleus and cytosol and processing cell cycle regulation. There is a study revealed that overexpression of PES1 may induce TG dysregulation by inhibiting autophagy.⁴⁵ Our results demonstrated that the expression of NR4A1 and PES1 were lower in patients with BPD compared to the control group. These results support the general understanding that BPD worsen with immune dysfunction, inhibition of autophagy and increased apoptosis.

The Relationship Between Hub Genes and Single-Cell Sequencing

Based on single-cell sequencing results, we found all hub genes mainly located in B cell, epithelial cell, fibroblast, endothelial cell, smooth muscle cell, typeI pneumocyte and typeII pneumocyte. On the one hand, we speculate that these genes promote the expression of lung fibroblasts, leading to the formation of alveolar epithelial-like tissue fibrosis in the lung, which contributes to BPD. **JUN** promotes pulmonary fibrosis by regulating CD36 to initiate fibrosis, increasing the number of reticular fibroblasts, and reducing the number of adipofibroblasts.⁴⁶ **PES1** induces cell proliferation by altering pre-rRNA.⁴⁷ We speculate that this gene promotes the proliferation of fibroblasts, leading to the production of excess fibrous tissue, which in turn promotes pulmonary fibrosis. Research related to **BRX1** and **RRP9** is still limited. **BRX1** promotes ribosome synthesis through the selective translation of GLUT1.⁴⁸ We believe that excessive ribosome synthesis will lead to increased cellular physiological activity, resulting in enhanced gene transcription and protein synthesis. This, in turn, may cause immune cells or fibroblasts to excessively secrete proteins or proliferate, leading to pulmonary fibrosis. The protein-protein interaction network of **RRP9** in the SSU processome, where a reduced expression of RRP9 leads to unstable interactions between U3/pre-rRNA base pairs or a decrease in the length of their connecting segments, suggesting may cause abnormal cell proliferation and promote the proliferation of fibrosis-related cells.⁴⁹ **NR4A1** in human myofibroblasts reduces TGFβ1-induced collagen deposition and the expression of fibrosis-related genes,⁵⁰ which slow down the progression of BPD lung fibrosis. On the other hand, in our cord blood array analysis, five hub genes were lower in BPD patients, indicating genes were protective indicators of BPD. Autophagy of AT2 cells and lamellar body development enhanced the alveolarization and reduced the incidence of BPD in premature infants.⁵¹ Therefore, we

supposed that our screening autophagy-related hub genes reduced the incidence of BPD by promoting autophagy of AT2 cells. However, the exact relationship between the relevant cells and the hub genes still need more researches to validate.

Signaling Pathways Mediate BPD

Our results suggested that these genes were correlated with ribosome processing. The ribosome is the primary site of protein translation, responsible for synthesizing nearly half of the cellular proteins. Autophagy is an important pathway for the regulation of ribosomal homeostasis.⁵² Our findings also showed that in the umbilical cord tissue homogenates, the activation of MAPK and IL 17 signaling pathway was higher in neonates with BPD compared to healthy controls. These findings corroborate a previous study showing that lncRNA H19 implicated in the progression of BPD via enhanced MAPK signaling pathway.⁵³ Another study showed that in immature rats chronic airway exposure to LPS induced pulmonary inflammation, increased IL 17a expression and hypoalveolarization.⁵⁴ These results, to some extent, provide additional information regarding the pathophysiology of BPD.

Limitation of Our Research

However, our research still has some limitations. First, we discovered one of the largest BPD datasets for analysis, but the number of included samples was still small. Therefore, a large cohort is required to corroborate our results. There were not enough in experiments to verify the level of expression of the genes identified, thus reducing the precision of our study. Consequently, further validations are required to make the findings more convincing.

Conclusions

In summary, our study results suggest that autophagy-related genes may serve as biological markers for BPD, potentially regulating immune infiltration, providing new insights and approaches for the early diagnosis and potential clinical treatment of BPD.

Data Sharing Statement

The data used to support the findings of this study are available from the corresponding author upon request.

Ethical Statement

The study involving human data was approved by the Medical Ethic Committee of Nanfang Hospital of Southern Medical University, the application number is NFEC -202210-K23. Besides, our research was conducted according to the International Ethic Guidelines and the National Institutes of Health Guidelines Concerning the Care and Use of Laboratory Animals. Our research was approved by the Institutional Animal Experiment Committee of Southern Medical University Nanfang Hospital, the application number is IACUC-LAC-20240304-001.

Acknowledgments

We thank the researchers of GSE8586 for sharing the data online.

Author Contributions

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

Funding

This study was supported by the National Natural Science Foundation of China (82171714), Guangdong Provincial Natural Science Foundation of China (2022A1515010427, 2020A1515110279 and 2022A1515012021), and Guangdong Province Basic and Applied Basic Salt Roasting Fund of China (2022A1515111083).

Disclosure

All the authors declare that there are no competing interests in this work.

References

- Thébaud B, Goss KN, Laughon M, et al. Bronchopulmonary dysplasia. *Nat Rev Dis Primers*. 2019;5(1):78. PMID: 31727986; PMCID: PMC6986462. doi:10.1038/s41572-019-0127-7
- Watterberg KL, Walsh MC, Li L, et al. NICHD Neonatal Research Network. Hydrocortisone to Improve Survival without Bronchopulmonary Dysplasia. *N Engl J Med*. 2022;386(12):1121–1131. PMID: 35320643; PMCID: PMC9107291. doi:10.1056/NEJMoa2114897
- Onland W, Offringa M, van Kaam A. Late (≥ 7 days) inhaled corticosteroids to reduce bronchopulmonary dysplasia in preterm infants. *Cochrane Database Syst Rev*. 2022;12(12):CD002311. PMID: 36521169; PMCID: PMC9754672. doi:10.1002/14651858.CD002311.pub5
- Abiramalatha T, Ramaswamy VV, Bandyopadhyay T, et al. Interventions to prevent bronchopulmonary dysplasia in preterm neonates: an umbrella review of systematic reviews and meta-analyses. *JAMA Pediatr*. 2022;176(5):502–516. PMID: 35226067. doi:10.1001/jamapediatrics.2021.6619
- Hirani D, Alvira CM, Danopoulos S, et al. Macrophage-derived IL-6 trans-signalling as a novel target in the pathogenesis of bronchopulmonary dysplasia. *Eur Respir J*. 2022;59(2):2002248. PMID: 34446466; PMCID: PMC8850688. doi:10.1183/13993003.02248-2020
- Windhorst AC, Heydariyan M, Schwarz M, et al. Monocyte signature as a predictor of chronic lung disease in the preterm infant. *Front Immunol*. 2023;14:1112608. Erratum in: *Front Immunol*. 2023;14:1209992. PMID: 37090732; PMCID: PMC10113536. doi:10.3389/fimmu.2023.1112608
- Deng X, Bao Z, Yang X, et al. Molecular mechanisms of cell death in bronchopulmonary dysplasia. *Apoptosis*. 2023;28(1–2):39–54. PMID: 36369365. doi:10.1007/s10495-022-01791-4
- Staud B, Gschwendtner S, Frodermann T, et al. Microbial signatures in amniotic fluid at preterm birth and association with bronchopulmonary dysplasia. *Respir Res*. 2023;24(1):248. PMID: 37845700; PMCID: PMC10577941. doi:10.1186/s12931-023-02560-w
- Roslan FF, Yu Y, Ooi GC, et al. From banked human cord blood to induced pluripotent stem cells: new opportunities and promise in induced pluripotent stem cell banking (Review). *Int J Mol Med*. 2024;54(6):114. PMID: 39422061. doi:10.3892/ijmm.2024.5438
- Fajarda O, Almeida JR, Duarte-Pereira S, et al. Methodology to identify a gene expression signature by merging microarray datasets. *Comput Biol Med*. 2023;159:106867. PMID: 37060770. doi:10.1016/j.combiomed.2023.106867
- Robles-Remacho A, Sanchez-Martin RM, Diaz-Mochon JJ. Spatial transcriptomics: emerging technologies in tissue gene expression profiling. *Anal Chem*. 2023;95(42):15450–15460. PMID: 37814884; PMCID: PMC10603609. doi:10.1021/acs.analchem.3c02029
- Clough E, Barrett T, Wilhite SE, et al. NCBI GEO: archive for gene expression and epigenomics data sets: 23-year update. *Nucleic Acids Res*. 2024;52(D1):D138–D144. PMID: 37933855; PMCID: PMC10767856. doi:10.1093/nar/gkad965
- Franceschini A, Szklarczyk D, Frankild S, et al. STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res*. 2013;41(Database issue):D808–15. PMID: 23203871; PMCID: PMC3531103. doi:10.1093/nar/gks1094
- Otasek D, Morris JH, Bouças J, et al. Cytoscape Automation: empowering workflow-based network analysis. *Genome Biol*. 2019;20(1):185. PMID: 31477170; PMCID: PMC6717989. doi:10.1186/s13059-019-1758-4
- Alhamzawi R, Ali HTM. The Bayesian adaptive lasso regression. *Math Biosci*. 2018;303:75–82. PMID: 29920251. doi:10.1016/j.mbs.2018.06.004
- Chen L, Zhang YH, Wang S, et al. Prediction and analysis of essential genes using the enrichments of gene ontology and KEGG pathways. *PLoS One*. 2017;12(9):e0184129. PMID: 28873455; PMCID: PMC5584762. doi:10.1371/journal.pone.0184129
- Rusk N. Expanded CIBERSORTx. *Nat Methods*. 2019;16(7):577. PMID: 31249418. doi:10.1038/s41592-019-0486-8
- Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. *J Thorac Oncol*. 2010;5(9):1315–1316. PMID: 20736804. doi:10.1097/JTO.0b013e3181ec173d
- Larrea E, Fernández-Rubio C, Peña-Guerrero J, et al. The BRCT domain from the homologue of the oncogene PES1 in *Leishmania major* (LmjPES) promotes malignancy and drug resistance in mammalian cells. *Int J Mol Sci*. 2022;23(21):13203. PMID: 36361992; PMCID: PMC9655562. doi:10.3390/ijms232113203
- Li YZ, Zhang C, Pei JP, et al. The functional role of Pescadillo ribosomal biogenesis factor 1 in cancer. *J Cancer*. 2022;13(1):268–277. PMID: 34976188; PMCID: PMC8692700. doi:10.7150/jca.58982
- Ge J, Huang X, Wang P, et al. Expression of biogenesis of ribosomes BRX1 is associated with malignant progression and prognosis in colorectal cancer. *Transl Cancer Res*. 2020;9(9):5595–5602. PMID: 35117923; PMCID: PMC8798812. doi:10.21037/tcr-20-2564
- Du MG, Liu F, Chang Y, et al. Neddylation modification of the U3 snoRNA-binding protein RRP9 by Smurf1 promotes tumorigenesis. *J Biol Chem*. 2021;297(5):101307. Erratum in: *J Biol Chem*. 2022;298(11):102567. PMID: 34662580; PMCID: PMC8569593. doi:10.1016/j.jbc.2021.101307
- Baumann S, Hess J, Eichhorst ST, et al. An unexpected role for FosB in activation-induced cell death of T cells. *Oncogene*. 2003;22(9):1333–1339. PMID: 12618758. doi:10.1038/sj.onc.1206126
- Gao L, Wang H, Fang F, et al. The roles of orphan nuclear receptor 4 group A1 and A2 in fibrosis. *Int Immunopharmacol*. 2024;139:112705. PMID: 39029235. doi:10.1016/j.intimp.2024.112705
- Nair J, Lakshminrusimha S. Update on PPHN: mechanisms and treatment. *Semin Perinatol*. 2014;38(2):78–91. doi:10.1053/j.semperi.2013.11.004
- McCurnin DC, Pierce RA, Willis BC, et al. Postnatal estradiol up-regulates lung nitric oxide synthases and improves lung function in bronchopulmonary dysplasia. *Am J Respir Crit Care Med*. 2009;179(6):492–500. doi:10.1164/rccm.200805-794OC
- Wang Y, Jiang L. Role of vitamin D-vitamin D receptor signaling on hyperoxia-induced bronchopulmonary dysplasia in neonatal rats. *Pediatr Pulmonol*. 2021;56(7):2335–2344. doi:10.1002/ppul.25418
- Zhu L, Li H, Tang J, et al. Hyperoxia arrests alveolar development through suppression of histone deacetylases in neonatal rats. *Pediatr Pulmonol*. 2012;47(3):264–274. doi:10.1002/ppul.21540
- Smith PG, Garcia R, Kogerman L. Mechanical strain increases protein tyrosine phosphorylation in airway smooth muscle cells. *Exp Cell Res*. 1998;239(2):353–360. doi:10.1006/excr.1997.3905
- Norman JE, Marlow N, Messow CM, et al. Does progesterone prophylaxis to prevent preterm labour improve outcome? A randomised double-blind placebo-controlled trial (OPPTIMUM). *Health Technol Assess*. 2018;22(35):1–304. doi:10.3310/hta22350

31. Zhang X, Trempey K, Hunter K, et al. Pergolide is comparable to nerve growth factor in recovery after alkali trauma in mouse corneas. *Invest Ophthalmol Vis Sci.* 2023;64(8):722.
32. Vodenkova S, Buchler T, Cervena K, et al. 5-fluorouracil and other fluoropyrimidines in colorectal cancer: past, present and future. *Pharmacol Ther.* 2020;206:107447. doi:10.1016/j.pharmthera.2019.107447
33. Abe MK, Chao TS, Solway J, et al. Hydrogen peroxide stimulates mitogen-activated protein kinase in bovine tracheal myocytes: implications for human airway disease. *Am J Respir Cell Mol Biol.* 1994;11(5):577–585. doi:10.1165/ajrcmb.11.5.7946386
34. Wang H, Zhao Y, Zhang D, et al. Neuroprotective effects of quinirole on lithium chloride pilocarpine-induced epilepsy in rats and its underlying mechanisms. *Eur J Med Res.* 2024;29(1):121. doi:10.1186/s40001-024-01694-x
35. Stoll BJ, Hansen NI, Bell EF, et al. National Institute of Child Health and Human Development Neonatal Research Network. Neonatal outcomes of extremely preterm infants from the NICHD neonatal research network. *Pediatrics.* 2010;126(3):443–456. PMID: 20732945; PMCID: PMC2982806. doi:10.1542/peds.2009-2959
36. Hurskainen M, Mižiková I, Cook DP, et al. Single cell transcriptomic analysis of murine lung development on hyperoxia-induced damage. *Nat Commun.* 2021;12(1):1565. PMID: 33692365; PMCID: PMC7946947. doi:10.1038/s41467-021-21865-2
37. Levine B, Mizushima N, Virgin HW. Autophagy in immunity and inflammation. *Nature.* 2011;469(7330):323–335. PMID: 21248839; PMCID: PMC3131688. doi:10.1038/nature09782
38. Zhou Y, Zhu Y, Jin W, et al. Tat-P combined with GAPR1 releases Beclin1 to promote autophagy and improve Bronchopulmonary dysplasia model. *iScience.* 2023;26(9):107509. PMID: 37636035; PMCID: PMC10448080. doi:10.1016/j.isci.2023.107509
39. Soni S, Jiang Y, Zhang L, et al. AMPK-driven macrophage responses are autophagy dependent in experimental bronchopulmonary dysplasia. *Am J Respir Cell Mol Biol.* 2023;68(3):279–287. PMID: 36306501; PMCID: PMC9989474. doi:10.1165/rmb.2022-0282OC
40. Yeganeh B, Lee J, Ermini L, et al. Autophagy is required for lung development and morphogenesis. *J Clin Invest.* 2019;129(7):2904–2919. PMID: 31162135; PMCID: PMC6597208. doi:10.1172/JCI127307
41. Strickland AB, Chen Y, Sun D, et al. Alternatively activated lung alveolar and interstitial macrophages promote fungal growth. *iScience.* 2023;26(5):106717. PMID: 37216116; PMCID: PMC10193231. doi:10.1016/j.isci.2023.106717
42. Vyawahare A, Prakash R, Jori C, et al. Caffeic acid modified nanomicelles inhibit articular cartilage deterioration and reduce disease severity in experimental inflammatory arthritis. *ACS Nano.* 2022;16(11):18579–18591. PMID: 36222569. doi:10.1021/acsnano.2c07027
43. Wang H, Xu X, Guan X, et al. Liposomal 9-aminoacridine for treatment of ischemic stroke: from drug discovery to drug delivery. *Nano Lett.* 2020;20(3):1542–1551. PMID: 32039606. doi:10.1021/acs.nanolett.9b04018
44. Harant H, Lindley JJ. Negative cross-talk between the human orphan nuclear receptor Nur77/NAK-1/TR3 and nuclear factor-kappaB. *Nucleic Acids Res.* 2004;32(17):5280–5290. PMID: 15466594; PMCID: PMC521667. doi:10.1093/nar/gkh856
45. Zhou J, Lu Y, Lin Y, et al. Overexpression of hepatic pescadillo 1 in obesity induces lipid dysregulation by inhibiting autophagy. *Transl Res.* 2023;258:1–15. PMID: 36775058. doi:10.1016/j.trsl.2023.02.003
46. Griffin MF, Borrelli MR, Garcia JT, et al. JUN promotes hypertrophic skin scarring via CD36 in preclinical in vitro and in vivo models. *Sci Transl Med.* 2021;13(609):eabb3312. doi:10.1126/scitranslmed.abb3312
47. He Y, Xiang J, Li Y, et al. PES1 is a biomarker of head and neck squamous cell carcinoma and is associated with the tumor microenvironment. *Cancer Med.* 2023;12(11):12622–12638. doi:10.1002/cam4.5948
48. Jiang C, Sun L, Wen S, et al. BRX1 promotes ribosome synthesis and enhances glycolysis by selected translation of GLUT1 in colorectal cancer. *J Gene Med.* 2024;26(1):e3632. doi:10.1002/jgm.3632
49. Clerget G, Bourguignon-Igel V, Marmier-Gourrier N, et al. Synergistic defects in pre-rRNA processing from mutations in the U3-specific protein Rrp9 and U3 snoRNA. *Nucleic Acids Res.* 2020;48(7):3848–3868. doi:10.1093/nar/gkaa066
50. Pulakazhi Venu VK, Alston L, Iftinca M, et al. Nr4A1 modulates inflammation-associated intestinal fibrosis and dampens fibrogenic signaling in myofibroblasts. *Am J Physiol Gastrointest Liver Physiol.* 2021;321(3):G280–G297. doi:10.1152/ajpgi.00338.2019
51. Li X, Huang L, Mao M, et al. HucMSCs-derived exosomes promote lung development in premature birth via Wnt5a/ROCK1 Axis. *Stem Cell Rev Rep.* 2024. PMID: 39565502. doi:10.1007/s12015-024-10824-1
52. Yao RQ, Ren C, Xia ZF, et al. Organelle-specific autophagy in inflammatory diseases: a potential therapeutic target underlying the quality control of multiple organelles. *Autophagy.* 2021;17(2):385–401. PMID: 32048886; PMCID: PMC8007140. doi:10.1080/15548627.2020.1725377
53. Mo W, Li Y, Chang W, et al. The Role of LncRNA H19 in MAPK signaling pathway implicated in the progression of bronchopulmonary dysplasia. *Cell Transplant.* 2020;29:963689720918294. PMID: 32308025; PMCID: PMC7289048. doi:10.1177/0963689720918294
54. Cui TX, Brady AE, Zhang YJ, et al. IL-17a-producing $\gamma\delta$ T cells and NKG2D signaling mediate bacterial endotoxin-induced neonatal lung injury: implications for bronchopulmonary dysplasia. *Front Immunol.* 2023;14:1156842. PMID: 37744375; PMCID: PMC10514485. doi:10.3389/fimmu.2023.1156842