



OPEN Identifying pyroptosis-related prognostic genes in the co-occurrence of lung adenocarcinoma and COPD via bioinformatics analysis

Chaofan Cao¹, Zhaoshuang Zhong², Bo Wu³, Yang Yang¹, Lingfei Kong⁴, Shuyue Xia^{2✉} & Guixian Xiao^{1✉}

Studies have indicated a complex association between chronic obstructive pulmonary disease (COPD) and lung adenocarcinoma (LUAD). However, the underlying mechanisms of their coexistence are still not fully understood. Thus, this study evaluated the possible mechanisms and biomarkers of COPD and LUAD by analyzing public RNA sequencing databases via bioinformatics analysis. This study obtained the LUAD datasets (TCGA-LUAD, GSE118370, and GSE30219) and the COPD dataset (GSE11784 and GSE39874) from TCGA and GEO databases, respectively. The differentially expressed genes (DEGs) were analyzed using the DESeq2 and limma packages. These DEGs were then intersected with pyroptosis-related genes (PRGs) to produce PRDEGs, which were examined via GO analysis and KEGG enrichment analyses. Simultaneously, a prognostic model was developed using PRDEGs by the TCGA-LUAD dataset to generate diagnostic PRDEGs (DPRDEGs). The STING database was employed to develop a protein-protein interaction (PPI) network for DPRDEGs. Transcription factors-associated with DPRDEGs were also identified in the ChIPBase and hTFtarget databases. The comparative toxicogenomics database (CTD) was employed to detect possible drugs or small molecules that interacted with DPRDEGs, and results were illustrated using Cytoscape. Moreover, this study developed a prognostic model using multivariate analysis and simultaneously conducted a prognostic analysis. The results were further validated by immunohistochemistry (IHC), western blotting (WB), and qPCR of clinical specimens. A total of 273 DEGs were identified, and 12 PRDEGs were detected after intersecting with PRGs. Inflammation and infectious diseases were the primary enriched regions for these PRDEGs, as indicated by GO and KEGG enrichment analyses. The study identified six DPRDEGs (*BNIP3*, *FTO*, *NEK7*, *POLR2H*, *S100A12*, and *TLR4*) via prognosis modeling of PRDEGs. The expression of these DPRDEGs in COPD and LUAD was verified through IHC, WB, and qPCR examinations. Based on multifactorial prognosis modeling, among six, *FTO*, *POLR2H*, *S100A12*, and *TLR4* revealed enhanced prognostic predictive effects. This study demonstrated that COPD and LUAD have common pathogenic mechanisms. The identified DPRDEGs and predictive models offer new perspectives for understanding and addressing COPD and LUAD.

Keywords COPD, Lung adenocarcinoma, Co-occurrence, Biomarkers, Early diagnosis, Pyroptosis, Bioinformatics analysis

Abbreviations

COPD Chronic obstructive pulmonary disease

¹Department of Respiratory Medicine, The Second Affiliated Hospital of Shenyang Medical College, No. 64, Qishan West Road, Shenyang 110035, Liaoning, China. ²Department of Respiratory Medicine, Central Hospital Affiliated to Shenyang Medical College, No. 5, Nanqi West Road, Shenyang 110024, Liaoning, China. ³Department of Respiratory Medicine, The Shengjing Hospital of China Medical University, No. 36, Sanhao Street, Shenyang 110000, Liaoning, China. ⁴Department of Respiratory Medicine, The First Affiliated Hospital of China Medical University, No. 155, Nanjing North Road, Shenyang 110000, Liaoning, China. ✉email: syx262@126.com; 18002453148@163.com

LUAD	Lung adenocarcinoma
DEGs	Differentially expressed genes
PRGs	Pyroptosis-related genes
PRDEGs	Pyroptosis-related differentially expressed genes
PPI	Protein-protein interaction
CTD	Comparative toxicogenomics database
IHC	Immunohistochemistry
WB	Western blotting
PCV13	13-Valent pneumococcal conjugate vaccine
PPSV23	23-Valent pneumococcal polysaccharide vaccine
LC	Lung cancer
GSDM	Gasdermin
TCGA	The cancer genome atlas
GEO	Gene expression omnibus
DPRDEGs	Diagnostic pyroptosis-related differentially expressed genes
GO	Gene ontology
BP	Biological processes
MF	Molecular functions
CC	Cellular components
KEGG	Kyoto encyclopedia of genes and genomes
SVM	Support vector machine
LASSO	Least absolute shrinkage and selection operator
DCA	Decision curve analysis
TFs	Transcriptional factors
PFA	Paraformaldehyde
ROC	Receiver operating characteristic
AUC	Area under the curve
IBD	Inflammatory bowel disease

Based on the Global Initiative for Chronic Obstructive Lung Disease, COPD is defined by prolonged respiratory manifestations and airflow obstruction resulting from disruptions in the airways or alveoli¹. It is globally recognized as the third leading cause of mortality, affecting approximately 12% of the population and contributing to 3 million deaths annually². The etiology of COPD is complex, with smoking being the primary environmental risk factor, while genetic factors and abnormal lung development could also contribute to its occurrence. Currently, COPD remains an incurable chronic disease without specific treatments, highlighting the critical importance of preventive interventions over pharmacological therapies. Smoking cessation is regarded as the most critical preventive measure, while vaccination is an alternative preventive approach. All patients aged ≥ 65 years are recommended to be administered either the 13-valent pneumococcal conjugate vaccine (PCV13) or the 23-valent pneumococcal polysaccharide vaccine (PPSV23). Younger COPD patients with major comorbidities, such as chronic heart disease, are also advised to initiate treatment with the PPSV23. Available pharmacological treatments at present involve inhaled long-acting bronchodilators, corticosteroids, etc., based on quitting smoking³.

Lung cancer (LC) is the most lethal tumor disease, with COPD being the most common comorbidity. Approximately 40% to 70% of lung cancer patients exhibit varying degrees of airflow obstruction, and these patients typically have poor clinical treatment outcomes^{4,5}. There is increasing evidence of a complex interaction between COPD and lung cancer. The widespread and sustained inflammatory response in the lung tissue of COPD patients increases the risk of lung cancer development⁶. Concurrent COPD could promote primary tumor metastasis in the lungs^{7,8}.

Pyroptosis is a novel form of programmed cell death mediated by Gasdermin (GSDM) proteins, which is characterized by the release of inflammatory factors⁹. It plays dual roles in homeostasis: moderate activation clears pathogens via inflammation, while excessive activation exacerbates tissue damage. It has an important role in both COPD and lung cancer¹⁰. Cigarette smoke-induced endoplasmic reticulum stress activates pyroptosis via the ROS/NLRP3/caspase-1 signaling pathway, contributing to the progression of COPD. Particularly, ROS activates the NLRP3 inflammasome, which in turn triggers caspase-1 stimulation and the release of inflammatory mediators like IL-1 β . This process induces pyroptosis and contributes to COPD progression¹¹. Pyroptosis can induce an inflammatory microenvironment that may increase the proliferation of tumor cells¹². It could function as a “bridge” between lung cancer and COPD; however, its mechanisms have yet to be fully understood.

This study aims to identify pyroptosis-related genes shared between LUAD and COPD by downloading datasets from the Cancer Genome Atlas (TCGA) and the Gene Expression Omnibus (GEO) databases for LUAD and COPD, respectively. The datasets were analyzed to identify DEGs, which were then intersected with PRGs to obtain PRDEGs. As a result, *BNIP3*, *FTO*, *NEK7*, *POLR2H*, *S100A12*, and *TLR4* were identified as six diagnostic PRDEGs (DPRDEGs) via prognosis modeling of PRDEGs. The expression of these six DPRDEGs in COPD and LUAD was verified using IHC, WB, and qPCR techniques.

Their prognostic effects were compared through multifactorial prognostic modeling. Therefore, this study aimed to clarify the comorbid mechanisms between COPD and LUAD from the perspective of pyroptosis, thereby offering novel insights into their association and establishing the theoretical basis for their clinical diagnosis and treatment.

Materials and methods

Data acquisition

The TCGA portal (<https://portal.gdc.cancer.gov/>) was the main source to obtain the LUAD dataset (TCGA-LUAD). This dataset consists of 59 similar adjacent normal (Group: Normal) and 539 LUAD (Group: LUAD) samples. The clinical data associated with these samples originated from the UCSC Xena database (<http://genome.ucsc.edu>).

Further, LUAD patient datasets (GSE118370 and GSE30219) were obtained from the GEO database (<http://www.ncbi.nlm.nih.gov/geo>). The GSE118370 dataset comprised 12 samples (6 tumor tissues from LUAD and 6 normal lung tissues). The dataset GSE30219 consisted of 307 samples (293 tumor tissues from LUAD and 14 normal lung tissues).

Moreover, datasets of COPD patients (GSE11784 and GSE38974) were obtained from the GEO database. The GSE11784 dataset contains 171 samples (72 samples from healthy smokers and 22 samples from smokers with COPD). The GSE38974 dataset comprised 59 samples (23 lung tumor samples from COPD patients and 9 non-tumor lung tissue samples). Details on each dataset have been uploaded to (Supplementary Table 1).

Data processing

The LUAD datasets (GSE118370 and GSE30219) and the COPD datasets (GSE118370 and GSE30219) were initially merged and processed individually. The R package sva (Version 3.50.0)¹³ was employed to eliminate batch effects, and the resulting data were standardized using the limma package (Version 3.58.1)¹⁴. Therefore, both LUAD and COPD datasets were acquired. Further, the limma package was used to standardize the TCGA-LUAD dataset.

Identification of differentially expressed genes

The DESeq2 package (Version 1.42.0)¹⁵ was used to conduct differential analysis on the TCGA-LUAD dataset and the limma package on the LUAD dataset and COPD dataset to identify DEGs between distinct groups (LUAD/Normal, COPD/Normal). To ensure the broad capture of potential genes, DEGs were identified by applying inclusive thresholds, with $|\log FC| \geq 0$ and $p\text{-adj} \leq 0.05$ ¹⁶. However, PRDEGs were obtained by the intersection of DEGs and PRGs.

Functional and pathway enrichment analyses

Gene ontology (GO)¹⁷ analysis is a widely used technique for large-scale functional enrichment studies. It involves analyzing biological processes (BP), molecular functions (MF), and cellular components (CC). The Kyoto Encyclopedia of Genes and Genomes (KEGG)¹⁸ is a well-known database that stores enriched genomic data, biological mechanisms, diseases, and drugs. The R package clusterProfiler (Version 4.10.0)¹⁹ was used to conduct GO and KEGG annotation analysis on PRDEGs. The thresholds for entry selection were $p\text{-adj} \leq 0.05$ and FDR ($q \leq 0.05$ value, with the method of Benjamini-Hochberg (BH) to adjust $p\text{-adj}$ values.

Construction of PRDEGs prognostic model

First, a support vector machine (SVM)²⁰ model was constructed using the expression matrix and grouping data from the TCGA-LUAD dataset. PRDEGs were selected based on achieving the highest accuracy and lowest error rate as determined by the SVM algorithm. Subsequently, a model was developed using the randomForest package (Version 4.7-1.1)²¹.

$$I(X = x_i) = -\log_2 p(x_i) \quad (1)$$

To select PRDEGs and develop the logistic diagnostic model, an analysis was performed on PRDEGs using a significance threshold of $p \leq 0.05$. Their molecular expression within the model was visualized using a forest plot. Accordingly, the glmnet package (Version 4.1-8)²² was employed to conduct a Least Absolute Shrinkage and Selection Operator (LASSO) regression analysis on PRDEGs, yielding the Logistic-LASSO regression model. Risk scores were computed according to the following equation:

$$\text{riskScore} = \sum_i \text{Coefficient}(\text{hub gene}_i) * \text{mRNA Expression}(\text{hub gene}_i) \quad (2)$$

To illustrate the interconnections of DPRDEGs within the PRDEGs prognostic model, a nomogram was generated using the rms package (Version 6.7-1) based on the logistic LASSO regression analysis results²³. Calibration curves were plotted to assess the accuracy and discriminative ability of the PRDEGs prognostic model. Finally, Decision Curve Analysis (DCA) plots were generated using the ggDCA package (Version 1.1)²⁴ according to the risk scores to further evaluate the model's performance.

Protein-protein interaction network

A protein-protein interaction (PPI) network associated with DPRDEGs was constructed using the STRING database²⁵ (<https://string-db.org/>), with an interaction threshold set to 0.150 (low confidence), and the network was visualized using Cytoscape²⁶ (Version 3.9.1).

Construction of mRNA-TF network and mRNA-drugs network

To further explore the regulatory mechanisms of DPRDEGs, common transcription factors (TFs) interacting with the DPRDEGs were identified using the CHIPBase²⁷ (<https://rna.sysu.edu.cn/chipbase/>) and hTFtarget²⁸ (<http://bioinfo.life.hust.edu.cn/hTFtarget/>) databases, followed by the construction of an mRNA-TF interaction

network. Additionally, the comparative toxicogenomics database (CTD)²⁹ (<http://ctdbase.org/>) was employed to predict potential drugs or small molecules interacting with DPRDEGs and to construct an mRNA-Drug interaction network. The mRNA-TF and mRNA-drugs interaction networks were then visualized using Cytoscape.

Construction of multifactorial DPRDEGs prognostic model

Statistically significant DPRDEGs ($p \leq 0.05$) were initially selected using univariate Cox regression analysis. Subsequently, a multifactorial DPRDEGs prognostic model was established by performing multivariate Cox regression analysis on the selected genes. The rms package was employed to construct a nomogram for evaluating the survival prediction capability of the Cox prognostic model.

Validation and assessment of multifactorial DPRDEGs prognostic model

Model accuracy was assessed by comparing predicted survival rates with actual outcomes through calibration curves. Additionally, the clinical predictive value of the model was examined via DCA using the ggDCA package.

Human samples

Lung tissue specimens were prospectively obtained from four distinct cohorts: (1) patients with co-occurring LUAD and COPD ($n=5$), (2) LUAD patients without documented COPD ($n=5$), paired with histologically normal adjacent tissues located at least 5 cm from tumor margins ($n=5$), and (3) COPD patients without malignancy ($n=5$). All specimens were either surgically resected or biopsy-derived from the Department of Respiratory Medicine, Second Affiliated Hospital of Shenyang Medical College.

The diagnosis of COPD was confirmed according to the 2024 Global Initiative for Chronic Obstructive Lung Disease (GOLD) guidelines (post-bronchodilator $FEV_1/FVC < 0.70$), excluding individuals with a history of malignancy or chronic respiratory comorbidities. The LUAD cohort required a pathologically confirmed diagnosis of primary lung adenocarcinoma, as defined by the 2021 WHO Classification of Tumours (5th edition), with normal lung function ($FEV_1/FVC \geq 0.70$) and tumor stages I–III, as per the AJCC 9th edition TNM staging. Adjacent normal tissues were validated by two independent pathologists to ensure the absence of tumor infiltration or atypical hyperplasia. The COPD-LUAD cohort met both diagnostic criteria.

Exclusion criteria included: concurrent pulmonary conditions (e.g., asthma, interstitial lung disease), autoimmune disorders, malignancies diagnosed within the past 5 years, use of corticosteroids or immunosuppressants within 3 months prior to enrollment, neoadjuvant therapy in LUAD, suboptimal specimens (e.g., improper fixation or autolysis), pregnancy or lactation, and severe organ dysfunction.

No statistically significant differences in age or gender were observed between the patient groups. Informed consent was obtained from all participants, and the study protocol was approved by the Ethics Committee of the Second Affiliated Hospital of Shenyang Medical College.

Immunohistochemistry

All tissue samples were placed in 4% paraformaldehyde (PFA) for fixation at 4 °C for 24 h. They were further placed in paraffin wax and sectioned (5 µm thick) for IHC analysis. The primary antibodies used for staining the sections are presented in (Supplementary Table 2). All these antibodies were used as per the provided protocols.

RNA isolation and quantitative real-time PCR

The total content of RNA from human samples was isolated and transcribed into cDNA via PrimeScript™ RT Master Mix (TAKARA, Tokyo, Japan) under the precise protocol of TRIzol reagent (Invitrogen, CA, USA). PCR procedures were carried out using the TB Green Premix Ex Taq II (TAKARA, Tokyo, Japan) on the Accurate 96 Real-time fluorescence quantitative PCR System (DLAB, Beijing, China). The primers used for PCR amplification of specific genes are presented in (Supplementary Table 3). The relative transcription pattern was estimated via the $2^{-\Delta\Delta CT}$ method and β -actin was used to normalize each expression.

Western blotting

Human samples were processed to extract the total protein content. After separation on 10% SDS-PAGE, 30 µg of proteins were transferred to PVDF membranes (Millipore, Bedford, MA, USA). Based on the provided instruction, the membranes were kept for two h blocking for 24 h at 4 °C with the following primary antibodies, which are listed in (Supplementary Table 4). Subsequently, each membrane was kept at 25 °C for 1 h with peroxidase-linked anti-mouse and anti-rabbit secondary antibodies (1:5000, ZSGB-BIO, Beijing, China). The protein bands were depicted via the respective ECL reagent (Tanon, Shanghai, China), detected under the Tanon5200 gel imaging system (Tanon, Shanghai, China), and quantified using Image J software. The relative protein expression levels were estimated by comparing them to the quantity of the control protein β -actin (1:2000, ZSGB-BIO, Beijing, China).

Statistical analysis

Bioinformatics analyses were performed using R software (version 4.1.2). For normally distributed continuous variables, comparisons between two groups were conducted using an independent Student's *t*-test; non-normally distributed variables were assessed with the Mann–Whitney U test (Wilcoxon rank-sum test). Survival analysis was carried out using the survival package in R, with Kaplan–Meier curves for visualization and log-rank tests for statistical significance. Spearman's correlation analysis was used to evaluate relationships between variables, unless otherwise specified.

Immunohistochemistry, Western Blotting, and qPCR analyses were performed using GraphPad Prism 8.0 (GraphPad Software, San Diego, CA, USA) with two-tailed tests. Data were analyzed using SPSS 17.0, and results are presented as mean $\pm$ standard deviation; statistical significance was set at $P < 0.05$.

Results

Identification and analysis of PRDEGs

Figure 1 illustrates the study workflow. Differential gene expression analysis was performed to assess expression variations among groups in the TCGA-LUAD, LUAD, and COPD datasets. The TCGA-LUAD dataset was processed using the DESeq2 package, whereas the LUAD and COPD datasets were analyzed with the limma package. This analysis identified DEGs in each dataset, and volcano plots were generated to display the results (Fig. 2A). Common DEGs were determined by intersecting the TCGA-LUAD, LUAD, and COPD datasets, and these were represented in a Venn diagram (Fig. 2B). Subsequently, intersecting the common DEGs with PRGs from the three datasets yielded 12 PRDEGs, which are depicted in a Venn diagram (Fig. 2C). These 12 PRDEGs are BNIP3, CD14, FADD, FTO, GTF2H4, IL1B, NEK7, NLRP3, POLR2H, S100A12, TLR2, and TLR4. Finally, the differential expression of these 12 PRDEGs was compared among various groups within the respective datasets, and a heatmap was generated using the pheatmap package (Version 1.0.12) (Fig. 2D).

Functional and pathway enrichment analyses

To examine the association between LUAD and COPD, GO functional enrichment and KEGG pathway enrichment analyses were performed using the 12 PRDEGs. The results revealed that the 12 PRDEGs are primarily enriched in biological processes (BP), such as positive regulation of cytokine production, modulation of the inflammatory response, phagocytosis, positive regulation of T cell-mediated immunity, and lymphocyte proliferation. In contrast, these genes were enriched in cellular components (CC) including membrane rafts, RNA polymerase II holoenzyme, secretory granule membranes, the transcription factor TFIID core complex, and phagocytic cups. Additionally, they showed enrichment in molecular functions (MF), such as pattern recognition receptor activity, lipopolysaccharide binding, NAD(P)+ nucleosidase activity, interleukin-1 receptor binding, and death receptor binding. These PRDEGs were also abundant in KEGG pathways, including Salmonella infection, Coronavirus disease—COVID-19, Tuberculosis, Pertussis, and the IL-17 signaling pathway (Fig. 3A).

Moreover, the enriched GO functional groups for BP, CC, and MF (Fig. 3A–D) were illustrated via network diagrams. The findings of the KEGG pathway enrichment analysis were represented in a circular network diagram (Fig. 3E).

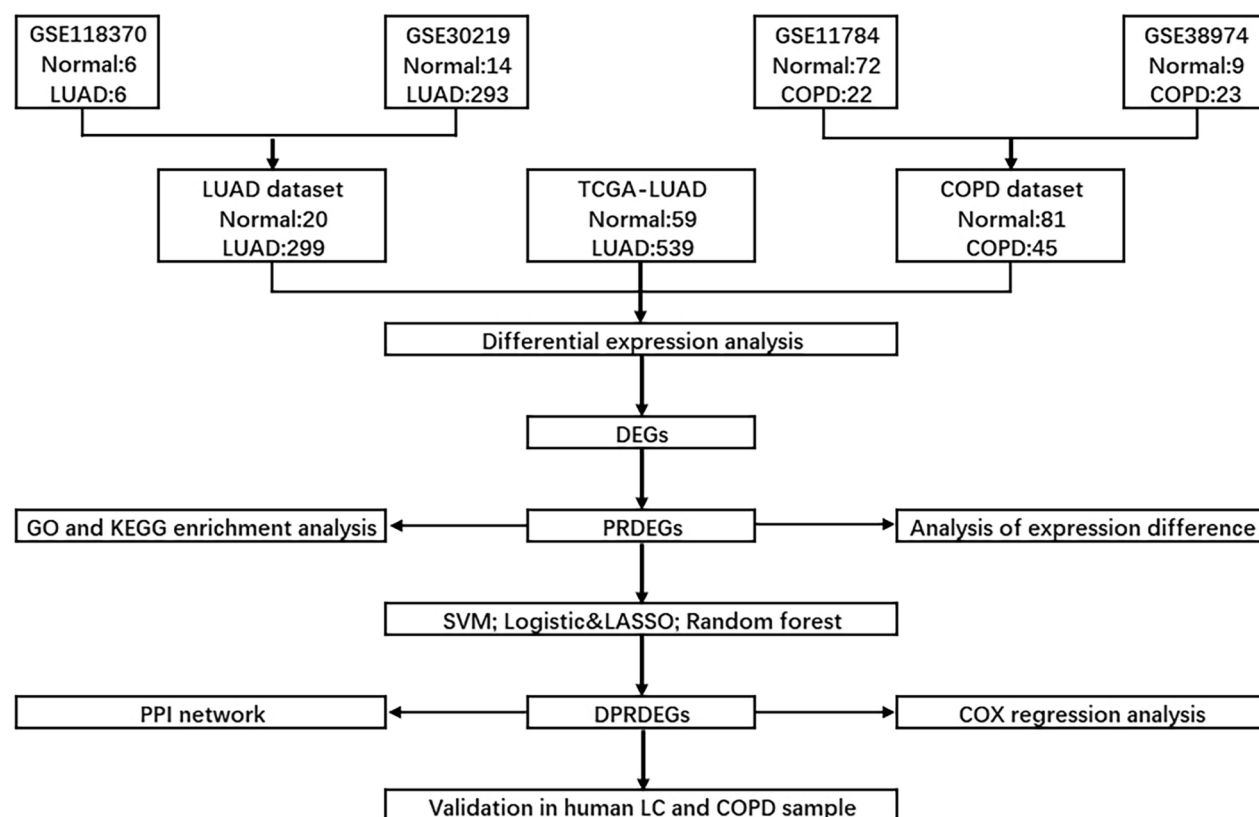


Fig. 1. Work flowchart.

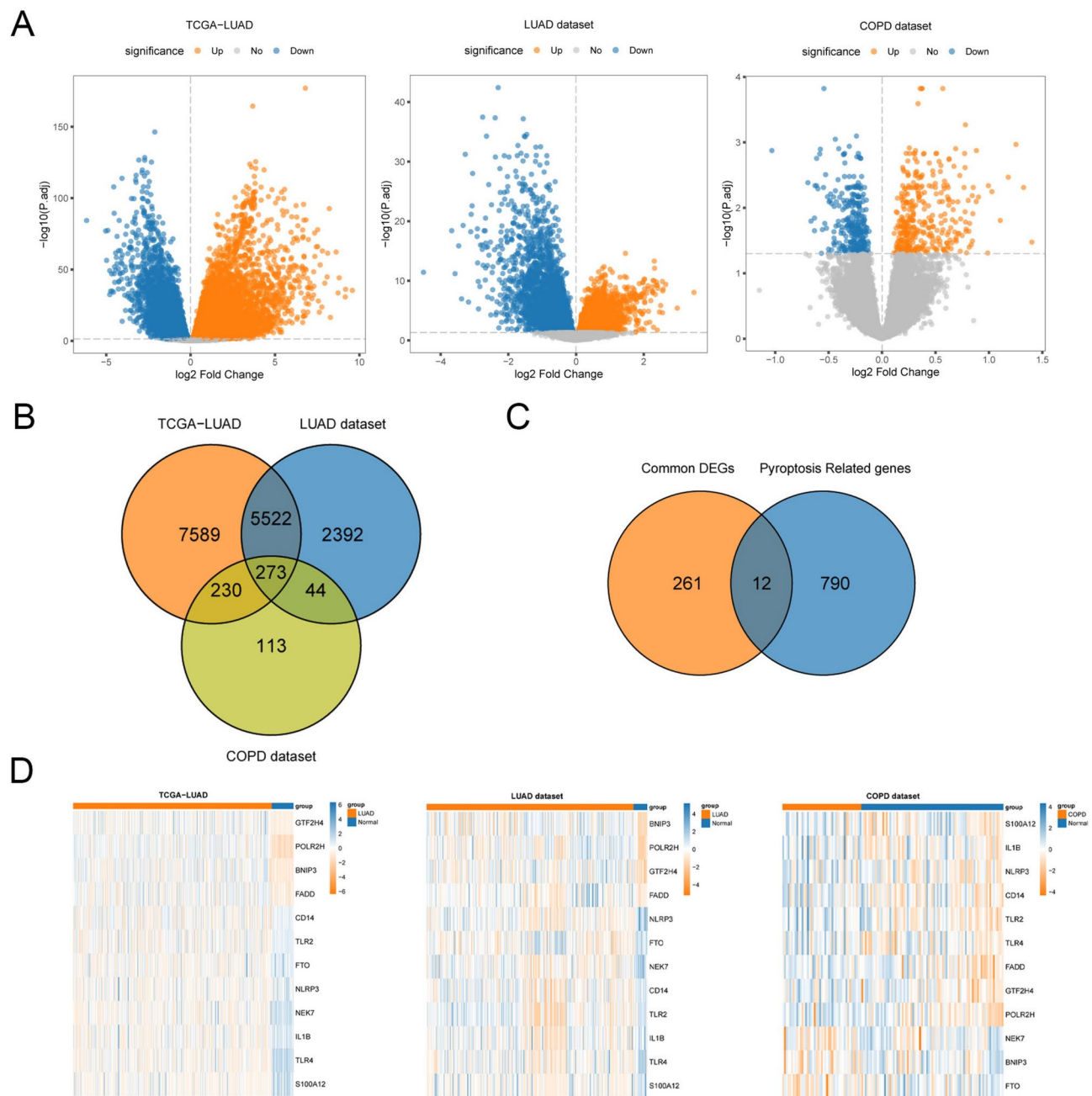


Fig. 2. Identification and analysis of PRDEGs **(A)** Volcano plots of DEGs between different groups (LUAD/Normal, COPD/Normal) in the TCGA-LUAD, LUAD, and COPD datasets. Orange dots represent upregulated genes, blue dots indicate downregulated genes, and gray dots denote non-significantly different genes. **(B)** Venn diagram of DEGs in the TCGA-LUAD, LUAD, and COPD datasets, highlighting the intersections among the three datasets. **(C)** Venn diagram showing the overlap between Common DEGs across the three datasets and PRGs, identifying a total of 12 PRDEGs. **(D)** Heatmap of PRDEGs in different groups (LUAD/Normal, COPD/Normal) within the TCGA-LUAD, LUAD, and COPD datasets, illustrating the expression differences of the 12 PRDEGs across groups. *DEGs* differentially expressed genes, *LUAD* lung adenocarcinoma, *COPD* chronic obstructive pulmonary disease, *PRGs* pyroptosis-related genes, *PRDEGs* pyroptosis related differentially expressed genes.

Differential expression analysis of PRDEGs

The expression differences of the 12 PRDEGs in various groups of the TCGA-LUAD dataset were analyzed using Wilcoxon rank sum tests (Fig. 4A). The expression levels of *BNIP3*, *FADD*, *GTF2H4*, and *POLR2H* were upregulated in the LUAD group than in the normal group. Conversely, the levels of *CD14*, *FTO*, *IL1B*, *NEK7*, *NLRP3*, *S100A12*, *TLR2*, and *TLR4* were lower in the LUAD group than in the normal group, with statistical significance ($p \leq 0.001$). The expression variations of the 12 PRDEGs in the LUAD and COPD datasets were

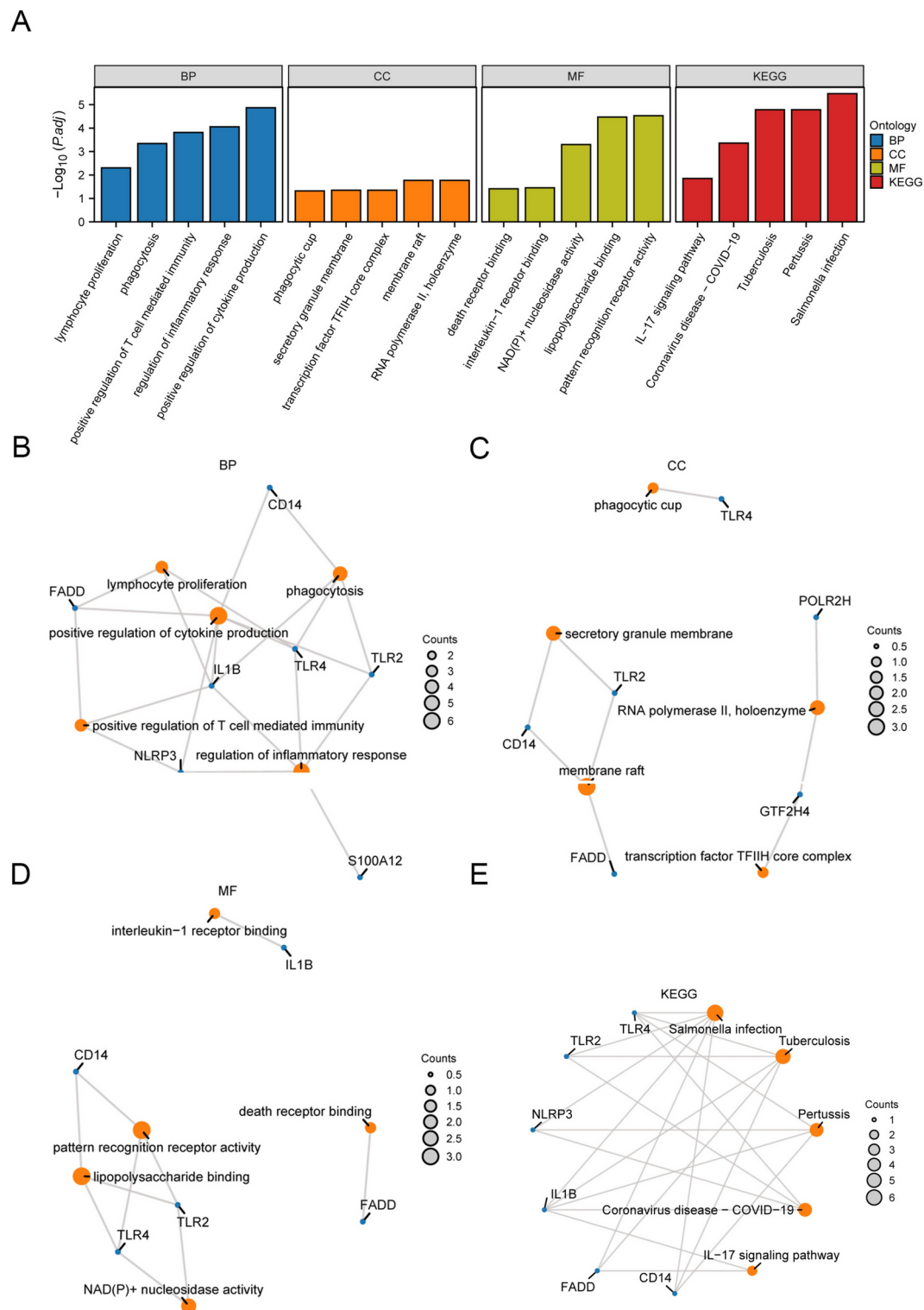


Fig. 3. Functional and pathway enrichment analyses of PRDEGs **(A)** Bar chart illustrating the results of GO functional enrichment analysis and KEGG pathway enrichment analysis for PRDEGs, including biological process (BP, blue), cellular component (CC, orange), molecular function (MF, green), and KEGG pathways (red). **(B–D)** Network diagrams of GO functional enrichment analysis for PRDEGs, depicting BP **(B)**, CC **(C)**, and MF **(D)**. Cyan nodes represent specific genes, orange circles indicate associated pathways, and the connecting lines illustrate gene-pathway relationships. **(E)** Circular network diagram of KEGG pathway enrichment analysis for PRDEGs, demonstrating the relationships between pathways and corresponding genes. Cyan nodes represent pathways, while orange nodes denote specific genes. The filtering criteria for GO and KEGG enrichment terms: $P_{adj} < 0.05$ and $FDR(q.value) < 0.05$. PRDEGs pyroptosis related differentially expressed genes, BP biological process, MF molecular function, CC cellular component.

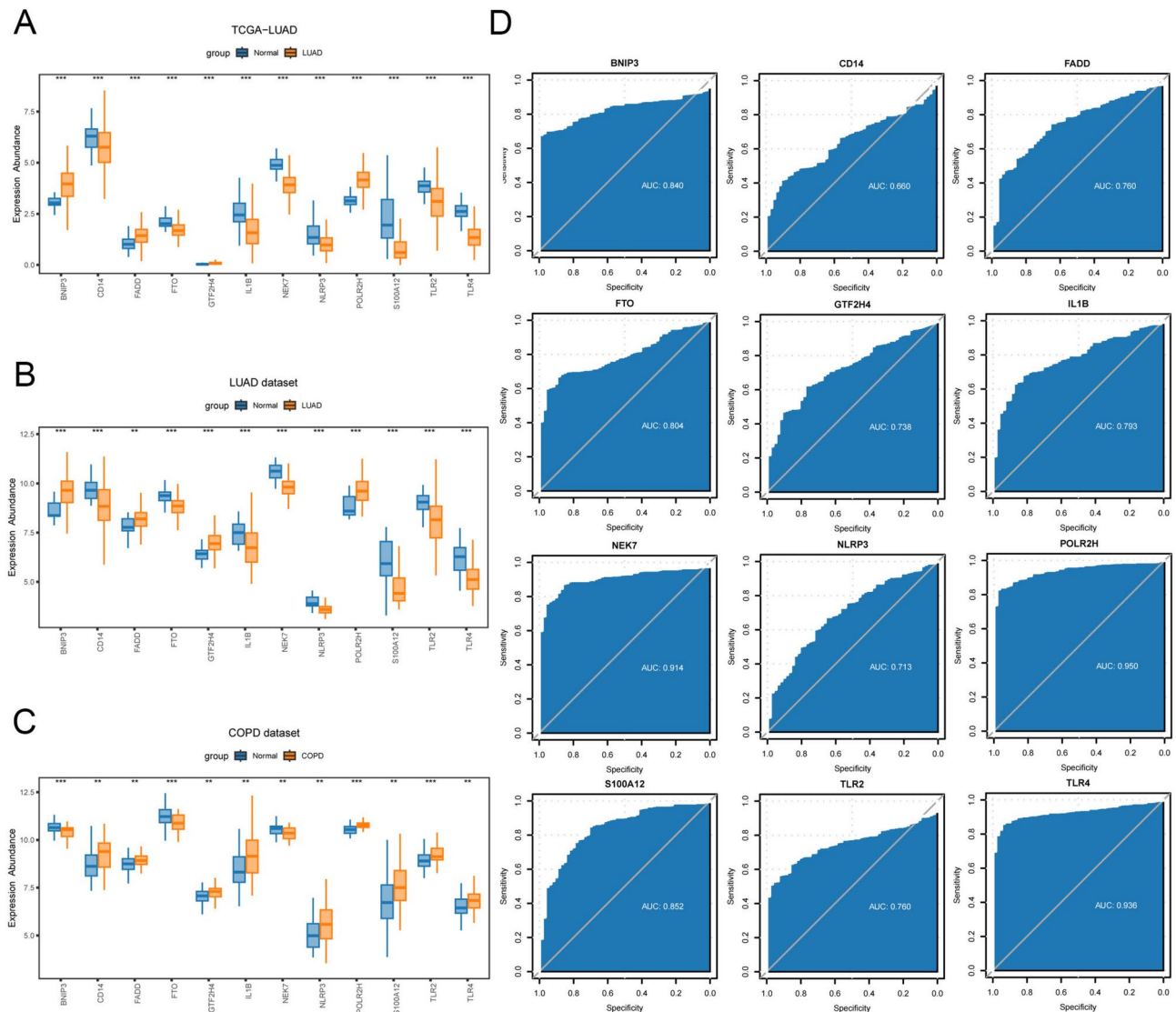


Fig. 4. Differential expression analysis of PRDEGs **(A)** Comparison of PRDEG expression levels between different groups (LUAD/Normal) in the TCGA-LUAD dataset, with all genes showing statistically significant differences. **(B)** Expression comparison of PRDEGs between LUAD and Normal groups in the LUAD dataset, with all genes exhibiting at least highly significant statistical differences. **(C)** Comparison of PRDEG expression levels between COPD and Normal groups in the COPD dataset, with all genes displaying statistically significant differences. **(D)** ROC curves of PRDEGs in the TCGA-LUAD dataset. The AUC values indicate the diagnostic accuracy of each gene in distinguishing between LUAD and Normal groups. “NS” means $P \geq 0.05$, suggesting no statistical significance. “*” $P < 0.05$, “**” $P < 0.01$, and “***” $P < 0.001$, suggesting statistical significance. PRDEGs pyroptosis related differentially expressed genes, LUAD lung adenocarcinoma, COPD chronic obstructive pulmonary disease, ROC receiver operating characteristic curve, AUC the area under the curve.

examined via the same method (Fig. 4B,C). In the LUAD dataset, the expression patterns of *BNIP3*, *FADD*, *GTF2H4*, and *POLR2H* were enhanced in the LUAD group than in the normal group. Conversely, the levels of *CD14*, *FTO*, *IL1B*, *NEK7*, *NLRP3*, *S100A12*, *TLR2*, and *TLR4* were reduced in the LUAD group than in the normal group ($p \leq 0.01$). In the COPD dataset, the levels of *CD14*, *FADD*, *GTF2H4*, *IL1B*, *NLRP3*, *POLR2H*, *S100A12*, *TLR2*, and *TLR4* were higher in the COPD group in contrast to the normal group, while the levels of *BNIP3*, *FTO*, and *NEK7* were lower in the COPD group relative to the normal group ($p \leq 0.001$). Furthermore, to identify the predictive value of the expression variances of the 12 PRDEGs in different groups of the TCGA-LUAD dataset, receiver operating characteristic (ROC) curves were plotted (Fig. 4D). Among them, *CD14* (AUC=0.660) showed relatively low accuracy in diagnosis. However, the expression of *BNIP3* (AUC=0.840), *FADD* (AUC=0.760), *FTO* (AUC=0.804), *GTF2H4* (AUC=0.738), *IL1B* (AUC=0.793), *NLRP3* (AUC=0.713), *S100A12* (AUC=0.852), *TLR2* (AUC=0.760) demonstrated moderate accuracy in diagnosis, while *NEK7* (AUC=0.914), *POLR2H* (AUC=0.950), and *TLR4* (AUC=0.936) revealed high accuracy in diagnosis.

Construction of PRDEGs prognostic model

A logistic regression analysis was performed using the expression pattern and clustering information of the 12 PRDEGs in the TCGA-LUAD dataset to assess their diagnostic significance. The results were presented as a forest plot (Fig. 5A). Subsequently, the SVM algorithm was applied to develop an SVM model, which exhibited the highest accuracy and lowest error rate when the number of genes was 12 (Fig. 5B,C). Furthermore, the random forest algorithm was used to observe the expression pattern of PRDEGs in the TCGA-LUAD dataset, identifying 10 diagnostic markers (*BNIP3*, *CD14*, *FADD*, *FTO*, *IL1B*, *NEK7*, *POLR2H*, *S100A12*, *TLR2*, and *TLR4*) (Fig. 5D,E).

The diagnostic model was then constructed via LASSO regression analysis and visualized by LASSO regression model and LASSO variable trajectory plots (Fig. 5F,G). This diagnostic model used seven PRDEGs (*BNIP3*, *FTO*, *NEK7*, *NLRP3*, *POLR2H*, *S100A12*, and *TLR4*). Moreover, six DPRDEGs were obtained by combining the PRDEGs from the Logistic-LASSO regression, SVM, and random forest models. These DPRDEGs were visualized in a Venn diagram (Fig. 5H) and were as follows: *BNIP3*, *FTO*, *NEK7*, *POLR2H*, *S100A12*, and *TLR4*. Next a prognostic model was established as per their expression pattern in the TCGA-LUAD dataset.

$$\text{RiskScore} = \text{BNIP3} * 0.641 + \text{FTO} * -0.185 + \text{NEK7} * -0.16 + \text{POLR2H} * 2.58 + \text{S100A12} * -0.722 + \text{TLR4} * -0.964 \quad (3)$$

The prognostic model was validated by plotting ROC curves based on the Risk Score from the PRDEGs prognostic model and clustering information from the TCGA-LUAD dataset (Fig. 5I). The TCGA-LUAD dataset demonstrated that the Risk Score from the PRDEGs model provided high diagnostic accuracy for LUAD (AUC=0.984). Moreover, a nomogram was developed based on the six DPRDEGs in the prognostic model (Fig. 5J). This nomogram indicated that *POLR2H* levels had a significantly greater impact on the model's performance compared to other variables. Furthermore, the predictive performance of the model was evaluated using calibration analysis and DCA, as illustrated in (Fig. 5K,L). These results suggest that the model is highly accurate in predicting the occurrence of LUAD in the TCGA-LUAD dataset.

Construction of PPI network and module analysis

The PPI analysis was carried out to examine the six DPRDEGs in the PRDEGs prognostic model within the TCGA-LUAD dataset via the STRING database. These DPRDEGs include *BNIP3*, *FTO*, *NEK7*, *POLR2H*, *S100A12*, and *TLR4*. The PPI network was developed with a minimum required interaction score of 0.150 to lower confidence, resulting in a network that includes the six DPRDEGs (Fig. 6A). All genes within the PPI network were designated as hub genes for the TCGA-LUAD dataset. The findings of the functional similarity examinations were then illustrated using a boxplot (Fig. 6B) based on the six hub genes in the PPI network. Among the six hub genes, *S100A12* showed the highest functional similarity to other network components.

Next, the CHIPBase database (3.0 V) and the hTFtarget database were then employed to identify TFs that interact with the six DPRDEGs. A total of 53 TF-gene interaction data were obtained by intersecting the association retrieved from both databases with these genes. This data was then displayed via Cytoscape software (Fig. 6C).

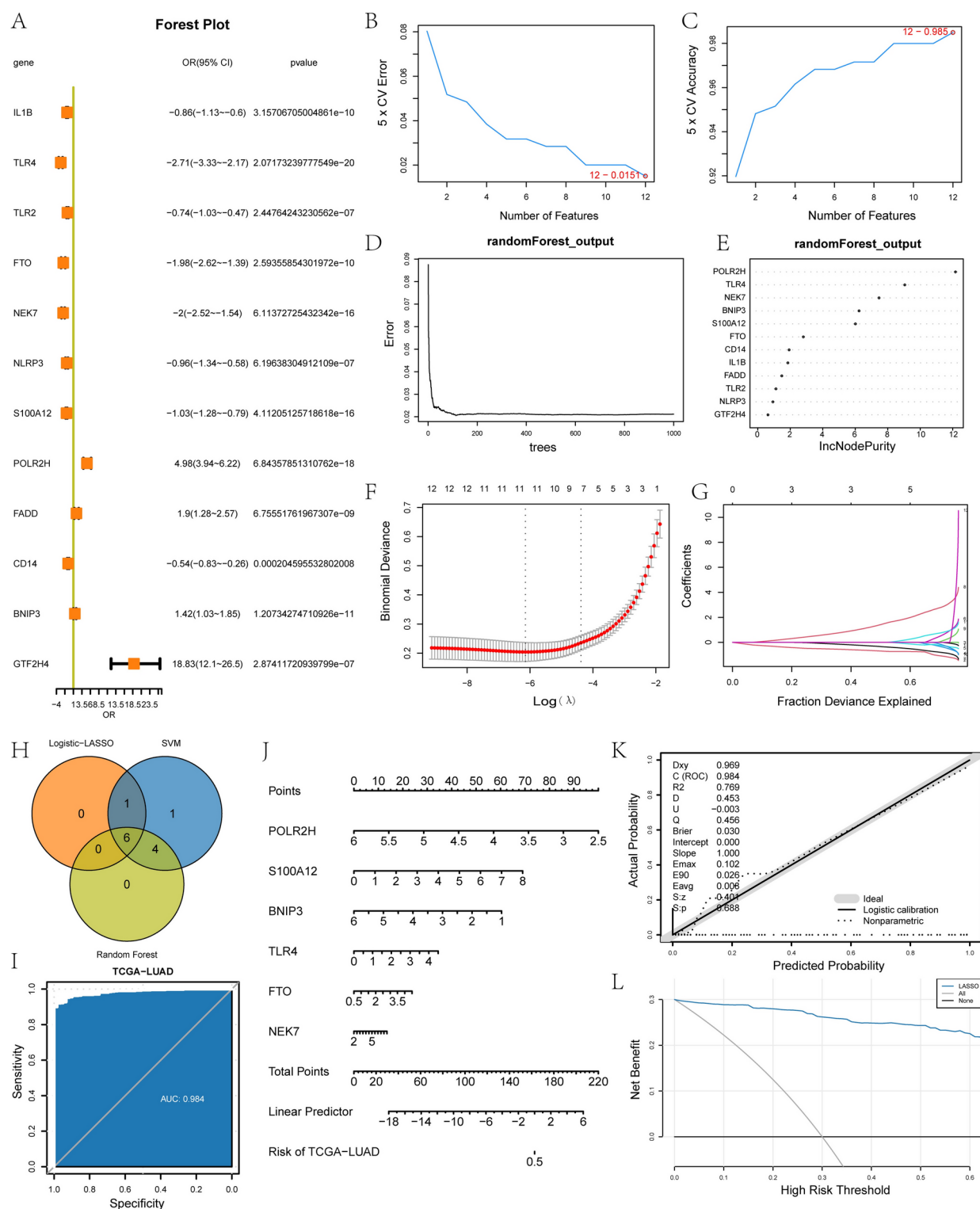
Further, the CTD database was implemented to locate possible molecular compounds or drugs that could target the six DPRDEGs. Approximately 35 effective drugs or molecules were detected that correspond to four DPRDEGs (*BNIP3*, *FTO*, *NEK7*, and *POLR2H*) via this analysis. Of these, 22 drugs or compounds target the *BNIP3* (Fig. 6D).

Construction and prognostic analysis of the multifactorial DPRDEGs prognostic model in TCGA-LUAD dataset

First, a single-factor Cox regression analysis was carried out using the expression profile of the six DPRDEGs (*BNIP3*, *FTO*, *NEK7*, *POLR2H*, *S100A12*, *TLR4*) in the TCGA-LUAD dataset to evaluate their predictive significance (excluding samples with missing clinical data). Therefore, variables that met $p \leq 0.10$ in the single-factor analysis were incorporated into the multifactorial Cox regression analysis to develop the multifactorial DPRDEGs prognostic model. The findings of this analysis were then organized and displayed in a forest plot (Fig. 7A). The formula for estimating the RiskScore in the multifactorial DPRDEGs prognostic model is as follows:

$$\text{RiskScore} = \text{FTO} * -0.164 + \text{POLR2H} * 0.032 + \text{S100A12} * 0.173 + \text{TLR4} * -0.297 \quad (4)$$

Next, nomogram analysis evaluated the prognostic capacity of the multifactorial DPRDEGs prognostic model and generated a nomogram plot (Fig. 7B). This plot showed that among the four DPRDEGs in the multifactorial prognostic model, *S100A12* and *TLR4* have increased efficiency in the model relative to other variables. The calibration analysis was performed for 1-, 3-, and 5-year prognoses based on the nomogram of the multifactorial DPRDEGs prognostic model, and calibration curves were plotted (Fig. 7C). The findings indicated that the predictive performance of the DEGs model ranked as follows: 3-year $\geq$ 5-year $\geq$ 1-year. The DCA results are presented in (Fig. 7D), and the clinical utility of the developed multifactorial Cox regression model was assessed at 1-, 3-, and 5-year time points. The analysis further demonstrated that the clinical predictive efficacy of the multifactorial DPRDEGs prognostic model was highest for 3-year outcomes, followed by 5-year outcomes, and lowest for 1-year outcomes.



Detection of DPRDEGs expression levels

To evaluate the expression profiles of the six DPRDEGs in clinical samples, IHC (Fig. 8A,B), qPCR (Fig. 9A), and WB (Fig. 9B) analyses were carried out. In contrast to the control group, the expression patterns of gene *FTO*, *NEK7*, *POLR2H*, and *S100A12* were considerably higher in the COPD, LUAD, and COPD with LUAD groups. The *BNIP3* expression levels were only enhanced in the COPD group, while they were lower in the LUAD and COPD with LUAD groups than in the control group. In contrast to the control group, the levels of *TLR4* were lower in the COPD, LUAD, and COPD with LUAD groups.

Discussion

Globally, COPD and LC are among the leading causes of mortality, presenting significant challenges to public health, safety, and socioeconomic development. These diseases share several common risk factors, including

Fig. 5. Construction of the prognostic model for PRDEGs. **(A)** Forest plot of the PRDEGs logistic regression model. **(B)** Number of genes with the lowest error rate identified by the SVM algorithm. **(C)** Number of genes with the highest accuracy identified by the SVM algorithm. **(D)** Model training error plot for the random forest algorithm. **(E)** Feature importance ranking of PRDEGs in the random forest model, demonstrating the contribution of each gene to the model. **(F)** Diagram of the diagnostic model based on the LASSO regression. **(G)** Variable trajectory plot of the LASSO regression model. **(H)** Venn diagram showing the overlap of PRDEGs in the logistic-LASSO regression model, SVM model, and random forest model. **(I)** ROC curve of the PRDEGs prognostic model in the TCGA-LUAD dataset. **(J)** Nomogram of the six DPRDEGs in the PRDEGs prognostic model. **(K)** Calibration curve for the PRDEGs prognostic model. **(L)** DCA plot of the PRDEGs prognostic model. *PRDEGs* pyroptosis related differentially expressed genes, *SVM* support vector machine, *LASSO* least absolute shrinkage and selection operator, *DPRDEGs* diagnosis pyroptosis related differentially expressed genes, *DCA* decision curve analysis, *ROC* receiver operating characteristic curve, *AUC* area under the curve.

genetic variations, modifications in gene expression patterns, and dysregulation of cellular signaling pathways³⁰. Identified factors include mutations in VEGFR1, upregulation of the inflammatory factor TNF- α , and abnormal activation of the IL-6/STAT3 signaling pathway^{31,32}. Interestingly, COPD has been identified as an independent risk factor for LC, with statistics indicating that the incidence of LC in patients diagnosed with COPD is 15.3%³³. The development of LC from COPD may be influenced by oxidative stress, DNA damage, and the inflammatory response. However, COPD and LC remain distinct lung diseases with different pathogenesis. Specifically, COPD is characterized by matrix degradation, impaired tissue repair, and mucosal lesions, whereas LC is primarily marked by increased tumor cell proliferation, invasion, and metastasis³⁴. Nevertheless, elucidating the pathogenic mechanisms underlying both the divergences and convergences between these two diseases will advance a holistic understanding of their comorbid progression.

This study explores the relationship between COPD and LUAD comorbidity and elucidated the mechanisms of COPD and LUAD comorbidity from the perspective of pyroptosis. This study detected six DPRDEGs and developed a multifactorial prognostic model, which demonstrated good performance in predicting the prognosis of LUAD.

In recent years, an increasing number of studies have reported a connection between COPD and LUAD. Some perspectives suggest that the continuous synthesis of ROS in the lungs of COPD patients results from inflammation. These ROS can oxidize and damage DNA, leading to an increased DNA repair rate and making epithelial cells more susceptible to malignant transformation^{35,36}. Others propose that the aberrant immune microenvironment in COPD patients disrupts the regulation of inflammatory cytokines, promoting tumor cell proliferation and allowing them to evade immune surveillance^{37,38}. Furthermore, the mechanism underlying the comorbidity of COPD and LUAD has been linked to variations in microbial communities between healthy and diseased airways, with certain bacterial metabolites, such as acetaldehyde and deoxycholic acid, acting as carcinogens³⁹. Although these perspectives focus on different aspects, they all highlight the significant role of inflammation in the comorbid mechanisms of COPD and LUAD.

Pyroptosis is closely associated with cellular inflammatory responses. In the chronic inflammatory microenvironment of COPD, persistent oxidative stress and inflammatory cytokine stimulation activate the NLRP3 inflammasome, a key sensor of cellular damage that integrates innate immunity with pyroptotic signaling. The role of pyroptosis in LUAD pathogenesis is complex and dual-faceted. On one hand, the release of inflammatory cytokines mediated by pyroptosis enhances the pro-inflammatory tumor microenvironment, thereby promoting tumor cell proliferation and metastasis^{40,41}. On the other hand, excessive activation of pyroptosis may induce tumor cell death, potentially inhibiting tumor progression under certain conditions⁴². These findings underscore pyroptosis as a critical pathway in the progression from COPD to LUAD. This has resulted in the identification of six DPRDEGs: *BNIP3*, *FTO*, *NEK7*, *POLR2H*, *TLR4*, and *S100A12*. Among them, *BNIP3* expression levels show significant correlation with cellular autophagy activity⁴³. Given that autophagy is a key mechanism for maintaining homeostasis, its upregulation could offer an effective strategy to counteract excessive pyroptosis. This observation indicates that *BNIP3* may serve as a potential therapeutic target to inhibit COPD-LUAD progression by modulating the balance between pyroptosis and autophagy within the shared inflammatory microenvironment. *FTO* acts as a demethylase for RNA m6A methylation modification, and abnormal m6A methylation modification has been observed in COPD and LUAD. Unlike other genes, *FTO* is a potential biomarker, as it plays a role in the progression of COPD and LUAD coexistence via epigenetic pathways^{44,45}. *NEK7* interacts with NLRP3 to regulate inflammasome activation, affecting inflammation in a pyroptosis-dependent manner. However, no reports have linked *NEK7* to COPD or LUAD⁴⁶. The results demonstrated that the expression levels of *NEK7* are higher in tissue samples from patients with COPD, LUAD, and COPD concomitant with LUAD, compared to normal lung tissue. This suggests that *NEK7* may serve as a potential biomarker. This study is the first to uncover the potential significance of *POLR2H* in the comorbidity of COPD and LUAD. Given its diagnostic value in the Prognostic Model, *POLR2H* may serve as a novel biomarker for the early detection of COPD and LUAD comorbidity. However, current understanding of *POLR2H* remains limited, even though it encodes a critical subunit of RNA Polymerase II. Previous reports have linked the abnormal expression of this gene to the survival rate of lung cancer. Its role in lung cancer may be associated with functions such as DNA damage repair, gene expression regulation, and cell cycle control⁴⁷. In contrast, *S100A12* is highly expressed in COPD and closely associated with inflammatory responses, though fewer reports exist regarding its role in LUAD⁴⁸. Our study reveals that *S100A12* expression is also significantly upregulated in tissue samples from LUAD and COPD concomitant with LUAD patients, suggesting that *S100A12* may facilitate

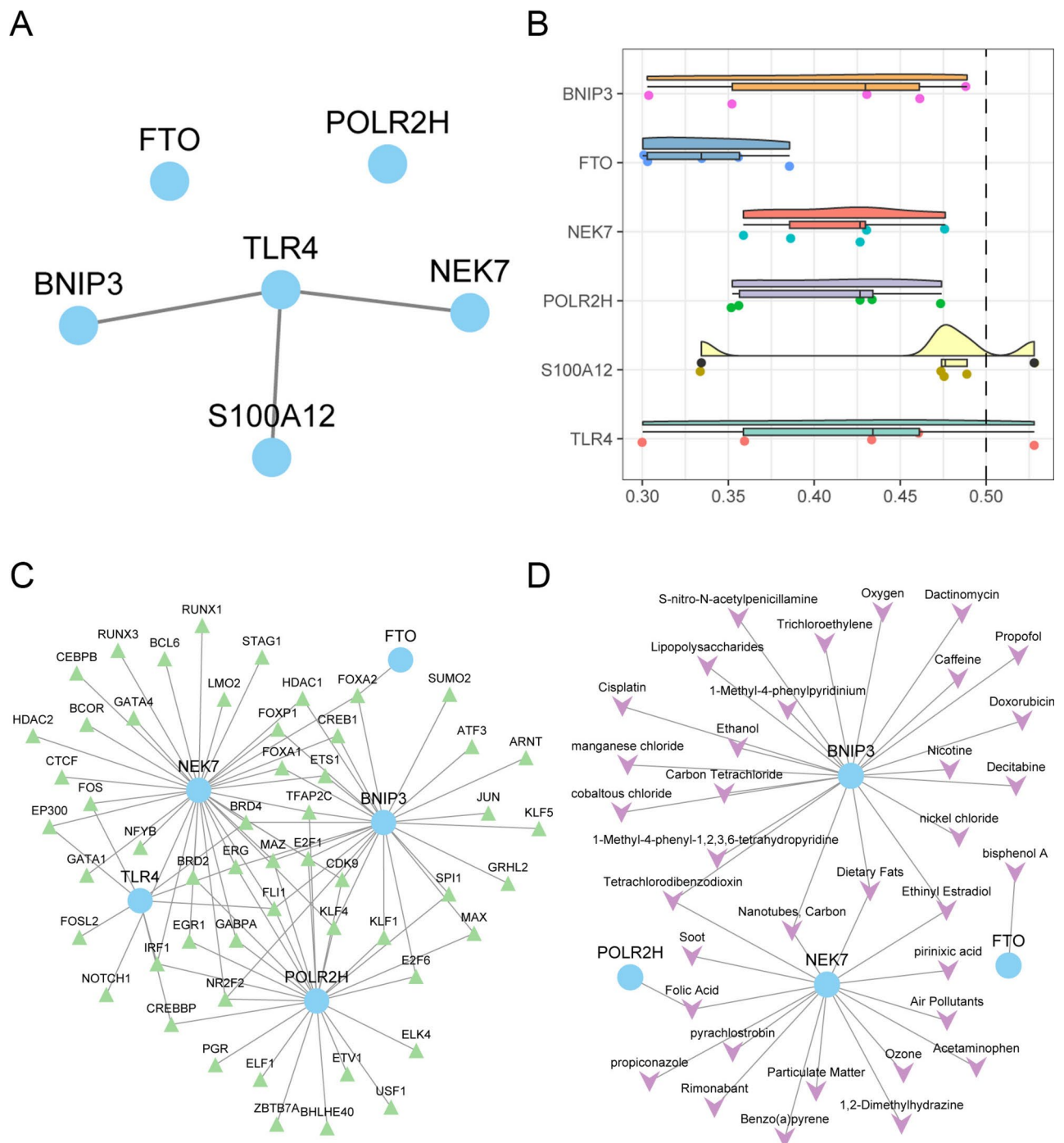


Fig. 6. Development and module analysis of the PPI network **(A)** PPI network of DPRDEGs, illustrating the interactions among the six diagnostic genes. **(B)** Functional similarity analysis of DPRDEGs, presented as a boxplot to display the functional similarity among the genes. **(C)** mRNA- TF interaction network of DPRDEGs, depicting interactions between six DPRDEGs and 53 transcription factors. Light blue circular nodes represent mRNAs, while light green triangular nodes represent transcription factors. **(D)** mRNA-drug interaction network of DPRDEGs, showing the relationships between four DPRDEGs and 35 potential drugs or molecular compounds. Light blue elliptical nodes represent mRNAs, while light purple V-shaped nodes indicate drugs. *PPI* protein-protein interaction, *DPRDEGs* diagnosis pyroptosis related differentially expressed genes, *TF* transcription factor.

tumor cell proliferation and migration in the tumor microenvironment. This finding offers new perspectives on the potential application of S100A12 as a biomarker.

The pyroptotic response mediated by *TLR4* differs between COPD and LUAD. In COPD-related studies, inhibiting the *TLR4* signaling pathway has been shown to alleviate smoking-induced COPD^{49–51}. In LUAD,

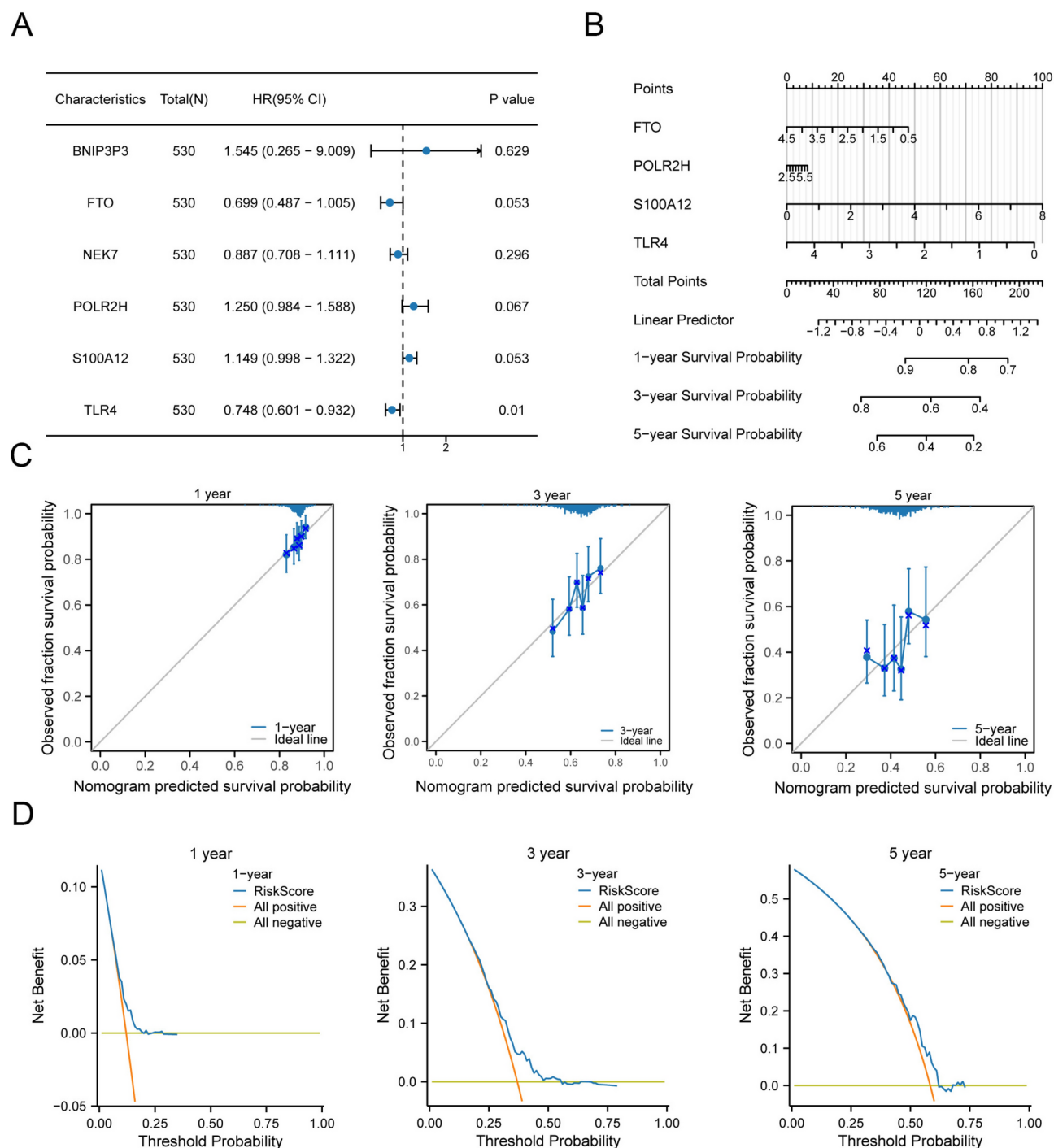


Fig. 7. Development and evaluation of the multi-factor DPRDEGs prognostic model in the TCGA-LUAD dataset (A) Forest plot of univariate Cox regression analysis for DPRDEGs in the TCGA-LUAD dataset, presenting the HR 95% CI and p-values for each gene. (B) Nomogram of the multivariate PRDEG prognostic model in the TCGA-LUAD dataset, illustrating the contribution of DPRDEGs to survival probability prediction. (C) Calibration curve of the multivariate DPRDEGs prognostic model, comparing the predicted survival probabilities at 1, 3, and 5 years with the actual observed survival probabilities. (D) DCA plot of the multivariate DPRDEGs prognostic model, demonstrating the net benefit of the prediction model at different thresholds, with clinical predictive performance ranked as 3-year > 5-year > 1-year. LUAD lung adenocarcinoma, DPRDEGs diagnosis pyroptosis related differentially expressed genes, DCA decision curve analysis.

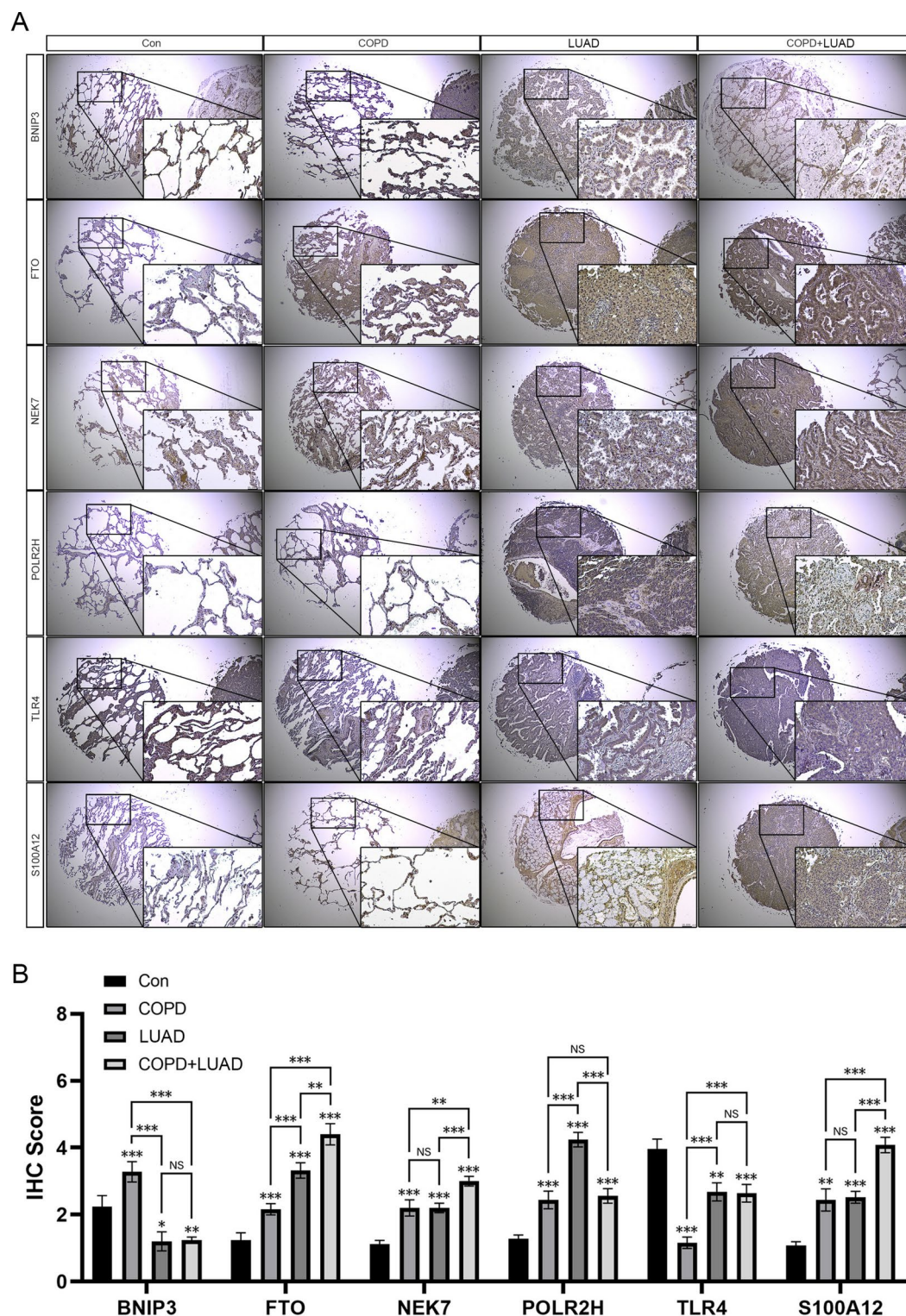


Fig. 8. Immunohistochemical examination of hub genes in clinical samples (A) Expression levels of 6 DPRDEGs detected by immunohistochemistry in normal lung tissue, COPD lung tissue, LUAD lung tissue, and the COPD with LUAD lung tissue. (B) Statistical analysis of immunohistochemical staining scores for each group. “NS” indicates that $P \geq 0.05$, suggesting no statistical significance. “*” means $P < 0.05$, “**” means $P < 0.01$, and “***” means $P < 0.001$, suggesting statistical significance. DPRDEGs diagnosis pyroptosis related differentially expressed genes, COPD chronic obstructive pulmonary disease, LUAD lung adenocarcinoma.

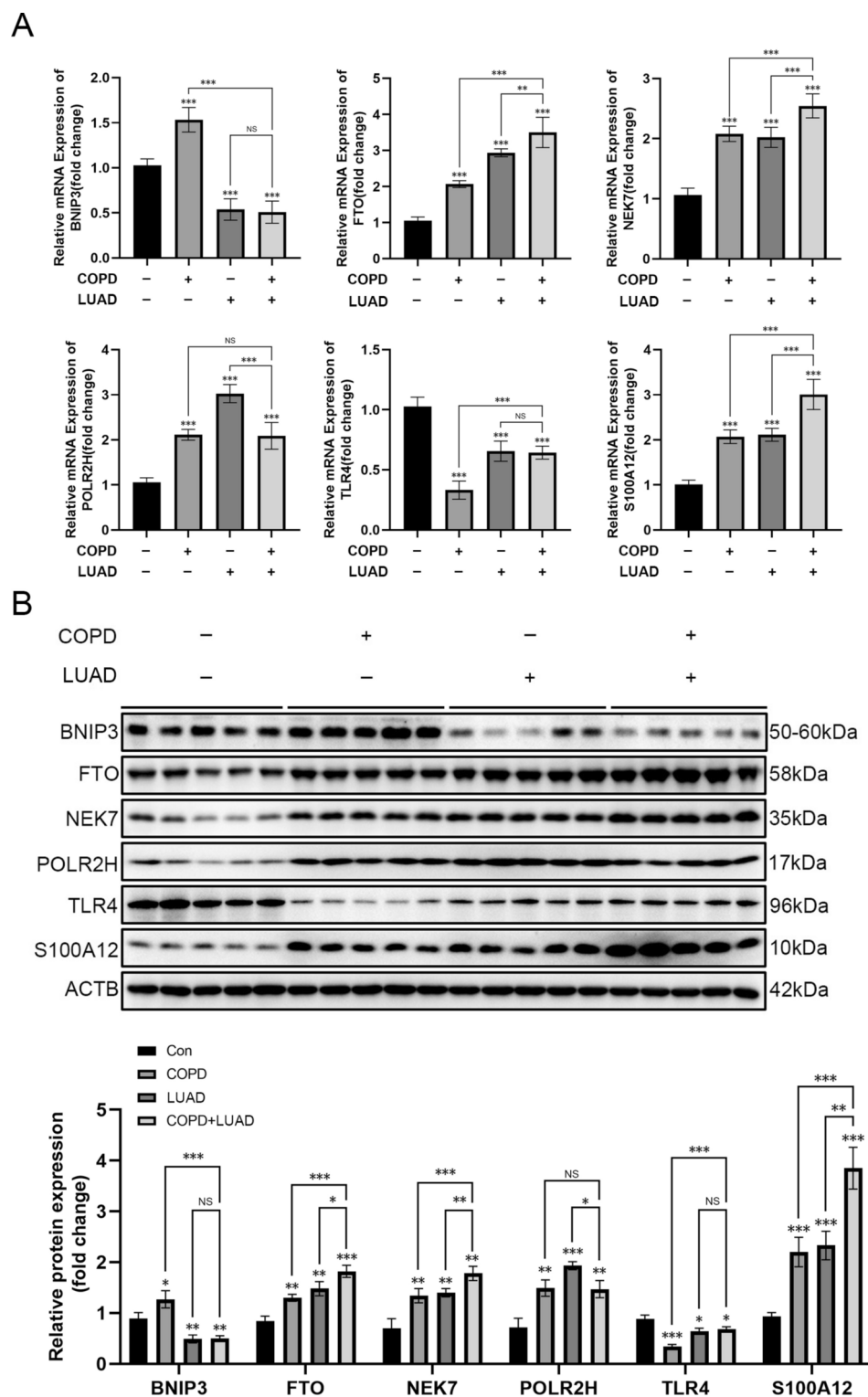


Fig. 9. Expression pattern of hub genes via qPCR and Western blot analyses (**A**) Expression levels of 6 DPRDEGs detected by qPCR in normal lung tissue, COPD lung tissue, LUAD lung tissue, and the COPD with LUAD lung tissue. (**B**) Expression levels of the same 6 DPRDEGs detected through Western blot in normal lung tissue, COPD lung tissue, LUAD lung tissue, and the COPD with LUAD lung tissue. The original blots are presented in (Supplementary Figure S1–S6). “NS” means $P \geq 0.05$, suggesting no statistical significance. “*” means $P < 0.05$, “**” means $P < 0.01$, and “***” means $P < 0.001$, suggesting statistical significance. DPRDEGs diagnosis pyroptosis related differentially expressed genes, COPD chronic obstructive pulmonary disease, LUAD lung adenocarcinoma.

however, activating the *TLR4/NLRP3/caspase-1/GSDMD*-dependent pyroptosis pathway has been reported to inhibit the progression of LUAD⁵². This discrepancy may arise from differences in the types of cells involved in the pyroptotic response. In COPD, the goal is to reduce pyroptotic damage in macrophages, thereby decreasing inflammation and mitigating potential fibrosis⁵³. In LUAD, by contrast, the objective is to directly target tumor cells with drugs to exert inhibitory and cytotoxic effects⁵⁴. Furthermore, analysis of the CTD database revealed putative interactions between *BNIP3*, *FTO*, *NEK7*, and *POLR2H* and existing drugs or small molecule compounds, providing novel insights for future clinical translational research. In summary, pyroptosis may constitute a common pathogenic mechanism underlying the comorbidity of COPD and LUAD, thereby warranting further investigation into its specific mechanisms.

The limitations of this study are as follows. First, the study relies on publicly available data from databases, which makes the results exploratory in nature. Second, although DPRDEGs were rigorously validated using techniques such as IHC, qPCR, and WB, the sample size ($n=5$ per group) may limit the statistical power of the results. A small sample size can introduce variability and restrict the generalizability of the findings. Additionally, the molecular mechanisms by which these six DPRDEGs contribute to COPD and LUAD were not explored in depth. To overcome these limitations, future studies will focus on increasing the sample size and further investigating the molecular pathways involved, thereby enhancing the clinical relevance and broader applicability of the findings.

Conclusions

In conclusion, this study identified six DPRDEGs and developed a prognostic model for DPRDEGs. The expression levels of these six DPRDEGs were verified through IHC, qPCR, and WB. The results of this study will offer new and valuable information for understanding and treating the comorbidity of COPD and LUAD.

Data availability

All sequencing raw data used in this work can be obtained from the Gene Expression Omnibus (GEO; <https://www.ncbi.nlm.nih.gov/geo/>), the Cancer Genome Atlas (TCGA) datasets (<https://xenabrowser.net/>), and the GeneCards (<https://www.genecards.org/>). Additionally, to ensure the reproducibility of our findings, all scripts used for data preprocessing, differential expression analysis, and model construction have been made publicly available. Further inquiries can be directed to the corresponding author.

Received: 22 November 2024; Accepted: 7 April 2025

Published online: 30 April 2025

References

- Celli, B. et al. Definition and nomenclature of chronic obstructive pulmonary disease: Time for its revision. *Am. J. Respir. Crit. Care Med.* **206**, 1317–1325. <https://doi.org/10.1164/rccm.202204-0671PP> (2022).
- Adeloye, D. et al. Global and regional estimates of COPD prevalence: Systematic review and meta-analysis. *J. Glob. Health* <https://doi.org/10.7189/jogh.05.020415> (2015).
- Agusti, A. et al. GOLD COPD DOCUMENT 2023: a brief update for practicing cardiologists. *Clin. Res. Cardiol.* **113**, 020415. <https://doi.org/10.1007/s00392-023-02217-0> (2023).
- Siegel, R. L., Miller, K. D., Fuchs, H. E. & Jemal, A. Cancer statistics, 2021. *CA Cancer J. Clin.* **71**, 7–33. <https://doi.org/10.3322/caac.21654> (2021).
- Forder, A. et al. Mechanisms contributing to the comorbidity of COPD and lung cancer. *Int. J. Mol. Sci.* **24**, 2859. <https://doi.org/10.3390/ijms24032859> (2023).
- Takiguchi, Y. Chronic obstructive pulmonary disease as a risk factor for lung cancer. *World J. Clin. Oncol.* **5**, 660–666. <https://doi.org/10.5306/wjco.v5.i4.660> (2014).
- Liu, J., Lin, P. & Zhou, B. Inflammation fuels tumor progress and metastasis. *Curr Pharm. Des.* **21**, 3032–3040. <https://doi.org/10.2174/1381612821666150514105741> (2015).
- Jing, B. et al. IL6/STAT3 signaling orchestrates premetastatic niche formation and immunosuppressive traits in lung. *Cancer Res.* **80**, 784–797. <https://doi.org/10.1158/0008-5472.CAN-19-2013> (2020).
- Frank, D. & Vince, J. E. Pyroptosis versus necroptosis: similarities, differences, and crosstalk. *Cell Death Differ.* **26**, 99–114. <https://doi.org/10.1038/s41418-018-0212-6> (2018).
- Feng, Y. et al. Pyroptosis in inflammation-related respiratory disease. *J. Physiol. Biochem.* **78**, 721–737. <https://doi.org/10.1007/s13105-022-00909-1> (2022).
- Zhang, M.-Y. et al. Cigarette smoke extract induces pyroptosis in human bronchial epithelial cells through the ROS/NLRP3/caspase-1 pathway. *Life Sci.* **269**, 119090. <https://doi.org/10.1016/j.lfs.2021.119090> (2021).
- Xia, X. et al. The role of pyroptosis in cancer: pro-cancer or pro-‘host’?. *Cell Death Dis.* <https://doi.org/10.1038/s41419-019-1883-8> (2019).
- Leek, J. T., Johnson, W. E., Parker, H. S., Jaffe, A. E. & Storey, J. D. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* **28**, 882–883. <https://doi.org/10.1093/bioinformatics/bts034> (2012).
- Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47–e47. <https://doi.org/10.1093/nar/gkv007> (2019).
- Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550. <https://doi.org/10.1186/s13059-014-0550-8> (2014).
- Li, S., Gao, K. & Yao, D. Comprehensive Analysis of angiogenesis associated genes and tumor microenvironment infiltration characterization in cervical cancer. *Heliyon* **10**, e33277. <https://doi.org/10.1016/j.heliyon.2024.e33277> (2024).
- Yu, G. Gene ontology semantic similarity analysis using GOSemSim. *Methods Mol. Biol.* **2117**, 207–215. https://doi.org/10.1007/978-1-0716-0301-7_11 (2020).
- Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30. <https://doi.org/10.1093/nar/28.1.27> (2000).
- Yu, G., Wang, L.-G., Han, Y. & He, Q.-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287. <https://doi.org/10.1089/omi.2011.0118> (2012).
- Sanz, H., Valim, C., Vegas, E., Oller, J. M. & Reverter, F. SVM-RFE: selection and visualization of the most relevant features through non-linear kernels. *BMC Bioinform.* **19**, 432. <https://doi.org/10.1186/s12859-018-2451-4> (2018).
- Liu, Y. & Zhao, H. Variable importance-weighted random forests. *Quant. Biol.* **5**, 338–351 (2017).

22. Engebretsen, S. & Bohlén, J. Statistical predictions with glmnet. *Clin. Epigenet.* **11**, 123. <https://doi.org/10.1186/s13148-019-0730-1> (2019).
23. Park, S. Y. Nomogram: An analogue tool to deliver digital knowledge. *J. Thorac. Cardiovasc. Surg.* **155**, 1793. <https://doi.org/10.1016/j.jtcvs.2017.12.107> (2018).
24. Van Calster, B. et al. Reporting and interpreting decision curve analysis: A guide for investigators. *Eur. Urol.* **74**, 796–804. <https://doi.org/10.1016/j.eururo.2018.08.038> (2018).
25. Szklarczyk, D. et al. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **47**, D607–D613. <https://doi.org/10.1093/nar/gky1131> (2019).
26. Shannon, P. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504. <https://doi.org/10.1101/gr.1239303> (2003).
27. Zhou, K. et al. ChIPBase v2.0: decoding transcriptional regulatory networks of non-coding RNAs and protein-coding genes from ChIP-seq data. *Nucleic Acids Res.* **45**, D43–D50. <https://doi.org/10.1093/nar/gkw965> (2017).
28. Zhang, Q. et al. hTFtarget: A comprehensive database for regulations of human transcription factors and their targets. *Genom. Proteom. Bioinform.* **18**, 120–128. <https://doi.org/10.1016/j.gpb.2019.09.006> (2020).
29. Davis, A. P. et al. Comparative toxicogenomics database (CTD): update 2021. *Nucleic Acids Res.* **49**, D1138–D1143. <https://doi.org/10.1093/nar/gkaa891> (2020).
30. Perrotta, F. et al. Evolving concepts in COPD and lung cancer: a narrative review. *Minerva Med.* **113**, 436–448. <https://doi.org/10.23736/S0026-4806.22.07962-9> (2022).
31. Wang, H. et al. Genetic variant in the 3'-untranslated region of VEGFR1 gene influences chronic obstructive pulmonary disease and lung cancer development in Chinese population. *Mutagenesis* **29**, 311–317. <https://doi.org/10.1093/mutage/geu020> (2017).
32. Dimitrakopoulos, F. I. D., Kottorou, A. E., Melpomeni, K. & Kalofonos, H. P. The fire within: NF- κ B involvement in non-small cell lung cancer. *Cancer Res.* **80**, 4025–4036. <https://doi.org/10.1158/0008-5472.CAN-19-3578> (2020).
33. Zhao, G. et al. Prevalence of lung cancer in chronic obstructive pulmonary disease: A systematic review and meta-analysis. *Front. Oncol.* **12**, 947981. <https://doi.org/10.3389/fonc.2022.947981> (2022).
34. Brody, J. S. & Spira, A. State of the art. Chronic obstructive pulmonary disease, inflammation, and lung cancer. *Proc. Am. Thorac. Soc.* **3**, 535–537. <https://doi.org/10.1513/pats.200603-089MS> (2006).
35. Aoshiba, K., Zhou, F., Tsuji, T. & Nagai, A. DNA damage as a molecular link in the pathogenesis of COPD in smokers. *Eur. Respir. J.* **39**, 1368–1376. <https://doi.org/10.1183/09031936.00050211> (2012).
36. Gohy, S. T. et al. Imprinting of the COPD airway epithelium for dedifferentiation and mesenchymal transition. *Eur. Respir. J.* **45**, 1258–1272. <https://doi.org/10.1183/09031936.00135814> (2015).
37. Zaynagetdinov, R. et al. Chronic NF- κ B activation links COPD and lung cancer through generation of an immunosuppressive microenvironment in the lungs. *Oncotarget* **7**, 5470–5482. <https://doi.org/10.18632/oncotarget.6562> (2015).
38. Scrimini, S. et al. Expansion of myeloid-derived suppressor cells in chronic obstructive pulmonary disease and lung cancer: potential link between inflammation and cancer. *Cancer Immunol. Immunother.* **64**, 1261–1270. <https://doi.org/10.1007/s00262-015-1737-x> (2015).
39. Mao, Q. et al. Interplay between the lung microbiome and lung cancer. *Cancer Lett.* **415**, 40–48. <https://doi.org/10.1016/j.canlet.2017.11.036> (2018).
40. Chiba, Y. et al. Caspase-4 promotes metastasis and interferon- γ -induced pyroptosis in lung adenocarcinoma. *Commun. Biol.* **7**, 699. <https://doi.org/10.1038/s42003-024-06402-3> (2024).
41. Zhao, Z. et al. Melatonin downregulates angiogenesis and lymphangiogenesis by regulating tumor-associated macrophages via NLRP3 inflammasomes in lung adenocarcinoma. *Aging* **16**, 12225–12238. <https://doi.org/10.18632/aging.206057> (2024).
42. Su, L. et al. Targeting Src reactivates pyroptosis to reverse chemoresistance in lung and pancreatic cancer models. *Sci. Transl. Med.* **15**, eab17895. <https://doi.org/10.1126/scitranslmed.abl7895> (2023).
43. Christy, et al. Smoking-associated exposure of human primary bronchial epithelial cells to aldehydes: Impact on molecular mechanisms controlling mitochondrial content and function. *Cells* **11**, 3481–3481. <https://doi.org/10.3390/cells11213481> (2022).
44. Huang, X., Lv, D., Yang, X., Li, M. & Zhang, H. m6A RNA methylation regulators could contribute to the occurrence of chronic obstructive pulmonary disease. *J. Cell. Mol. Med.* **24**, 12706–12715. <https://doi.org/10.1111/jcmm.15848> (2020).
45. Wu, G. et al. ALKBH1-8 and FTO: Potential therapeutic targets and prognostic biomarkers in lung adenocarcinoma pathogenesis. *Front. Cell. Dev. Biol.* **9**, 633927. <https://doi.org/10.3389/fcell.2021.633927> (2021).
46. Chen, X. et al. NEK7 interacts with NLRP3 to modulate the pyroptosis in inflammatory bowel disease via NF- κ B signaling. *Cell Death Dis.* **10**, 906. <https://doi.org/10.1038/s41419-019-2157-1> (2019).
47. Dong, A., Wang, Z.-W., Ni, N., Li, L. & Kong, X.-Y. Similarity and difference of pathogenesis among lung cancer subtypes suggested by expression profile data. *Pathol. Res. Pract.* **220**, 153365–153365. <https://doi.org/10.1016/j.prp.2021.153365> (2021).
48. Kim, K., Kim, H. J., Binas, B., Kang, J. H. & Chung, I. Y. Inflammatory mediators ATP and S100A12 activate the NLRP3 inflammasome to induce MUC5AC production in airway epithelial cells. *Biochem. Biophys. Res. Commun.* **503**, 657–664. <https://doi.org/10.1016/j.bbrc.2018.06.057> (2018).
49. Yuan, J. et al. Perilla leaf extract (PLE) attenuates COPD airway inflammation via the TLR4/Syk/PKC/NF- κ B pathway in vivo and in vitro. *Front. Pharmacol.* **12**, 763624. <https://doi.org/10.3389/fphar.2021.763624> (2022).
50. Qiu, X. Y. et al. Eucalyptol, limonene and pinene enteric capsules attenuate airway inflammation and obstruction in lipopolysaccharide-induced chronic bronchitis rat model via TLR4 signaling inhibition. *Int. Immunopharmacol.* **129**, 111571–111571. <https://doi.org/10.1016/j.intimp.2024.111571> (2024).
51. Wang, L. et al. Hydrogen sulfide attenuates cigarette smoke-induced pyroptosis through the TLR4/NF- κ B signaling pathway. *Int. J. Mol. Med.* **49**, 56. <https://doi.org/10.3892/ijmm.2022.5112> (2022).
52. Zhu, H. et al. Reniformin A suppresses non-small cell lung cancer progression by inducing TLR4/NLRP3/caspase-1/GSDMD-dependent pyroptosis. *Int. Immunopharmacol.* **133**, 112068–112068. <https://doi.org/10.1016/j.intimp.2024.112068> (2024).
53. Long, L. et al. Mefenidone alleviates silica-induced inflammation and fibrosis by inhibiting the TLR4-NF- κ B/MAPK pathway and attenuating pyroptosis in murine macrophages. *Biomed. Pharmacother.* **178**, 117216–117216. <https://doi.org/10.1016/j.biopha.2024.117216> (2024).
54. Yuan, R. et al. Cucurbitacin B inhibits non-small cell lung cancer in vivo and in vitro by triggering TLR4/NLRP3/GSDMD-dependent pyroptosis. *Pharmacol. Res.* **170**, 105748. <https://doi.org/10.1016/j.phrs.2021.105748> (2021).

Author contributions

C.C. and S.X. conceived this project. C.C. and Z.Z. took charge of the project administration. C.C. and Y.Y. performed experiments, data analysis and interpretation. X.X. coordinated patient enrollment and/or sample collection. C.C. and B.W. wrote the original draft. S.X., L.K., and X.X. conducted critically review and editing of the manuscript. All authors have read and approved the final version of the manuscript.

Funding

This study was supported by National Natural Science Foundation of China (82072008); Liaoning Provincial Department of Education Scientific Research Platform Construction Project (LJ232410164021); Liaoning Pro-

vincial Department of Education Independent Research Topic Project (LJ212410164020); Provincial key clinical specialty construction project in Liaoning Province (Inslczdzk2024); Scientific research project of Shenyang Health Commission (202316).

Declarations

Competing interests

The authors declare no competing interests.

Ethics approval and consent to participate

The study was approved by the ethic committee of the Second Affiliated Hospital of Shenyang Medical College. This research conformed to the principles of the Helsinki Declaration. Written informed content was obtained from every participant.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-97727-4>.

Correspondence and requests for materials should be addressed to S.X. or G.X.

Reprints and permissions information is available at www.nature.com/reprints.

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

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

© The Author(s) 2025