Original Article

# A new prognostic model for *RHOV*, *ABCC2*, and *CYP4B1* to predict the prognosis and association with immune infiltration of lung adenocarcinoma

**Qiao Li[1#], Xiao-Li Liu[2#], Ni Jiang[3#], Qian-Yun Li[4], Yong-Xiang Song[1], Xi-Xian Ke[1], Hao Han[1], Qian Luo[1], Qiang Guo[1,5], Xiang-Yu Luo[5], Cheng Chen[1]**

[1]Department of Thoracic Surgery, Affiliated Hospital of Zunyi Medical University, Zunyi, China; [2]Department of Ultrasound, The People's Hospital of Jianyang City, Jianyang, China; [3]Department of Obstetrics and Gynecology, Women and Children's Hospital of Chongqing Medical University, Chongqing, China; [4]The Fourth Affiliated Hospital, Zhejiang University School of Medicine, Yiwu, China; [5]Department of Cardiothoracic Surgery, Taihe Hospital, Hubei University of Medicine, Shiyan, China

*Contributions:* (I) Conception and design: C Chen, XY Luo; (II) Administrative support: C Chen, YX Song, XX Ke; (III) Provision of study materials or patients: QY Li, Q Guo; (IV) Collection and assembly of data: Q Li, H Han, Q Luo, Q Guo; (V) Data analysis and interpretation: Q Li, XL Liu, N Jiang, Q Guo; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

*Correspondence to:* Cheng Chen. Department of Thoracic Surgery, Affiliated Hospital of Zunyi Medical University, 149 Dalian Road, Zunyi 563000, Guizhou, China. Email: 29217036@qq.com; Xiang-Yu Luo. Department of Cardiothoracic Surgery, Taihe Hospital, Hubei University of Medicine, 32 Renmin South Road, Shiyan 442012, Hubei, China. Email: 3282174937@qq.com.

**Background:** Lymph node metastasis is one of the important factors affecting the prognosis of lung adenocarcinoma (LUAD) patients. The key molecules in lymph node metastasis have not yet been fully revealed. Therefore, we aimed to construct a prognostic model based on lymph node metastasis-related genes to evaluate the prognosis of LUAD patients.

**Methods:** The differentially expressed genes (DEGs) in the process of LUAD metastasis were identified in The Cancer Genome Atlas (TCGA) database, and the biological roles of the DEGs were depicted using Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), and a protein-protein interaction (PPI) network. Survival analysis and Cox regression analysis were used to identify the genes related to the prognosis of patients with LUAD, and a nomogram and a prognostic model were constructed. The potential prognostic value, immune escape, and regulatory mechanisms of the prognostic model in LUAD progression were explored through survival analysis and gene set enrichment analysis (GSEA).

**Results:** A total of 75 genes were upregulated, and 138 genes were downregulated in tissues of lymph node metastasis. The expression levels of *STC1*, *CYP17A1*, *RHOV*, *GUCA2B*, *TM4SF20*, *DEFB1*, *CRHR2*, *ABCC2*, *CYP4B1*, *KRT16*, and *NTS* were revealed as risk factors for a poor prognosis in LUAD patients. High-risk LUAD patients had a poor prognosis in the prognostic model based on *RHOV*, *ABCC2*, and *CYP4B1*. The clinical stage and the risk score were found to be independent risk factors for a poor prognosis in LUAD patients, and the risk score was associated with the tumor purity, T cell, natural killer (NK) cell, and other immune cells. The prognostic model might affect the progression of LUAD using DNA replication, the cell cycle, P53, and other signaling pathways.

**Conclusions:** Lymph node metastasis-related genes *RHOV*, *ABCC2*, and *CYP4B1* are associated with a poor prognosis in LUAD. A prognostic model based on *RHOV*, *ABCC2*, and *CYP4B1* might predict the prognosis of LUAD patients and be associated with immune infiltration.

**Keywords:** *RHOV*; *ABCC2*; *CYP4B1*; lymph node metastasis; lung adenocarcinoma (LUAD)

Submitted Feb 07, 2023. Accepted for publication Mar 17, 2023. Published online Apr 10, 2023.
doi: 10.21037/jtd-23-265
**View this article at:** https://dx.doi.org/10.21037/jtd-23-265

## Introduction

Lung cancer is one of the most common malignant tumors worldwide, as well as one of the main causes of cancer-related death (1,2). In China, the morbidity and mortality of lung cancer remain high throughout the entire year (3). Lung adenocarcinoma (LUAD) is one of the common subtypes of lung cancer (4). Although there are many effective targeted drugs for treating cancer patients, LUAD remains one of the most common and fatal cancers worldwide.

Early metastasis of LUAD is one of the key factors leading to cancer progression to the middle and advanced stages (5-7). Lymph node metastasis is one of the common modes of LUAD metastasis and is one of the risk factors affecting the long-term survival of patients with LUAD (6,7). Luo *et al.* reported that the metastasis rates of N1 and N2 stages increased with the tumor diameter (6). Dai *et al.* reported that 15% of LUAD patients had lymph node metastasis. Compared with node-negative patients, recurrence-free survival (RFS) and overall survival (OS) have been shown to be significantly decreased in patients with metastasis (7). However, the molecules and the signaling mechanisms of lymph node metastasis in LUAD remain unelucidated. The Cancer Genome Atlas (TCGA) project aims to improve the ability to prevent, diagnose, and treat cancer using high-throughput genome analysis technology. Multiple cancer types and data of genes, microRNAs (miRNAs), long noncoding RNAs (lncRNAs), and others are displayed in the TCGA database (8,9). Lymph node metastasis was a risk factor for poor prognosis in LUAD patients. However, the roles of lymph node metastasis-related genes in the progression of LUAD has not been fully revealed. Therefore, important molecular markers in the process of lymph node metastasis were explored based on the data from the TCGA database. The roles of the established lymph node metastasis-related the prognostic model were investigated in LUAD metastasis to improve the treatment value for LUAD patients. The relationship between the prognostic model and the prognosis, as well as immune cell infiltration of LUAD, were explored to understand the key molecules in LUAD metastasis. We present the following article in accordance with the TRIPOD reporting checklist (available at https://jtd.amegroups.com/article/view/10.21037/jtd-23-265/rc).

### Highlight box

**Key findings**

- Our new prognostic model based on lymph node metastasis-related genes *RHOV*, *ABCC2*, and *CYP4B1* might predict the prognosis of LUAD patients and be associated with immune infiltration.

**What is known and what is new?**

- Lymph node metastasis is related to the poor prognosis of patients with LUAD.
- The molecular mechanisms of lymph node metastasis have not been fully understood in LUAD. Lymph node metastasis-related genes *RHOV*, *ABCC2*, and *CYP4B1* were found to be associated with the prognosis of LUAD patients.

**What is the implication, and what should change now?**

- A prognostic model based on *RHOV*, *ABCC2*, and *CYP4B1* might predict the prognosis of cancer patients, which might become a tool to predict the prognosis of LUAD patients.

## Methods

### *TCGA data download and visualization analysis*

The gene expression data of 594 cases of LUAD high-throughput sequence-fragments per kilobase million (HTSeq-FPKM) and the clinical data of 522 cancer patients were downloaded from the TCGA database. The gene expression data of 535 LUAD tissues were included, and *LUAD* gene expression data with the N0-3 stage were extracted in our study. A total of 330 LUAD patients were lymph node-negative, whereas 171 LUAD patients were lymph node-positive. Differentially expressed genes (DEGs) of LUAD metastasis were screened using the limma package of R software (version 4.0.2; https://www.r-project.org/) with the criteria of a false discovery rate (FDR) <0.05 and |logFC| >1, and the visualization results were displayed using a volcano map. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

### *Biological functions and protein-protein interaction (PPI) network of lymph node metastasis-related genes*

Biological processes, cellular components, molecular functions, and signaling mechanisms involved in lymph node metastasis-related genes were investigated using Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis in the Database for Annotation, Visualization, and Integrated Discovery (DAVID) with the criterion of FDR <0.05. A PPI network of lymph node

metastasis-related genes was visualized in the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database and enriched for analysis using the MCODE method in Cytoscape software (10).

### *Kaplan-Meier (K-M) survival analysis*

Grouped by the median value of DEGs, the relationship between high- and low-expressed DEGs and the OS of LUAD patients were explored via K-M survival analysis and with P<0.05 for filtering and screening. According to the median score, the patients were divided into high-risk and low-risk groups. The prognostic value of LUAD patients in the high-risk and low-risk groups was presented using K-M survival analysis.

### *Construction of nomogram and prognostic model of lymph node metastasis-related genes*

The roles of DEGs in lymph node metastasis were identified in the prognosis of LUAD patients using univariate Cox regression analysis with the criterion of P<0.001. On this basis, multivariate Cox regression analysis and the Akaike information Criterion (AIC) method were performed to screen the factors affecting a poor prognosis of LUAD patients, and a nomogram and a prognostic model were constructed (11).

### *Construction of prognostic model-related nomogram*

The risk score data and the clinicopathological characteristic data of LUAD patients were matched. The relationship between clinicopathological characteristics and the lymph node metastasis prognostic model, and the prognosis of LUAD patients were investigated using univariate and multivariate Cox regression analysis, and a nomogram was constructed based on the multivariate Cox analysis results.

### *Gene set enrichment analysis (GSEA)*

The regulatory mechanisms in which genes might be involved were explored using GSEA (12,13). Grouping of the gene expression data of LUAD was performed according to the median prognostic model score, and the signaling pathways of the lymph node metastasis-related prognostic model were explored. The screening criterion of GSEA: NOM P<0.05.

### *Immune analysis of prognostic model*

Tissue samples from patients with LUAD were scored using Cell-type Identification by Estimating Relative Subsets of RNA Transcripts (CIBERSORT), microenvironment cell populations (MCP)-counter, and estimation methods (14,15). Pearson correlation analysis was used to explore the relationship between the immune score and the levels of tumor purity, immune cells, and immune cell markers, and P<0.05 was considered a significant screening criterion.

### *Statistical analysis*

Perl (https://www.perl.org/) and R were used for data processing and statistical analysis. Cox regression and K-M survival analyses were conducted to filter the risk factors of OS in patients with LUAD, and ROC analysis was carried out to assess the role of a gene-associated nomogram in lymph node metastasis. The relationship between the immune score and the levels of tumor purity, immune cells, and immune cell markers was explored using Pearson correlation analysis. The expression of DEGs in high- and low-risk groups was detected using a *t*-test, and P<0.05 was considered statistically significant.

## Results

### *Identification of DEGs related to metastasis in LUAD*

Comparing the tissues with lymph node-negative LUAD patients, there were 213 DEGs in the tissues of lymph node-positive LUAD patients (Table S1 and Table S2). Among them, 75 DEGs were upregulated (Table S1), and 138 DEGs were downregulated (Table S2). The top 15 DEGs in LUAD tissues were shown by fold changes (*Figure 1*). In detail, the expression levels of *LRRC38*, *TAC1*, *CALB1*, *CGB5*, *KRT20*, *TRIM48*, *TM4SF20*, *NNAT*, *GCG*, *PI3*, *CYP2B6*, *SPAG11B*, *GUCA2B*, *RHCG*, and *MUC2* were increased in the tissues of lymph node-positive LUAD patients (*Figure 1A*), whereas the expression of *ALB*, *NPY*, *AFP*, *SPINK4*, *WFDC12*, *DKK4*, *WFDC5*, *DLK1*, *MSTN*, *FABP7*, *HIST1H4C*, *HIST1H1B*, *GC*, *VTN*, and *FTHL17* was decreased in the tissues of lymph node-positive LUAD patients (*Figure 1B*).

### *Biological functions and PPI network of lymph node metastasis-related genes*

The DEGs of lymph node metastasis were found to be

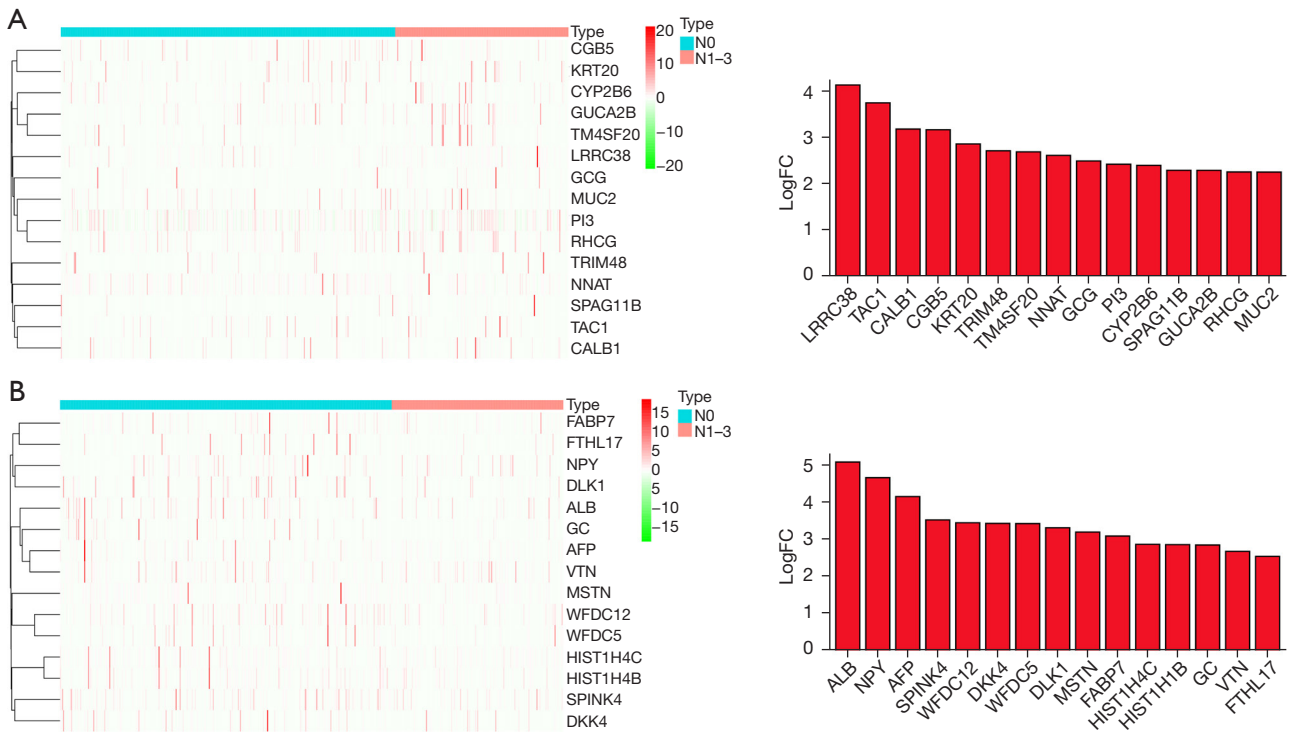


**Figure 1** Fifteen DEGs of lymph node metastasis in LUAD shown using heatmap and histogram. (A) Overexpressed genes; (B) lowly expressed genes. DEGs, differentially expressed genes; LUAD, lung adenocarcinoma; FC, fold change.

involved in the DNA replication-dependent nucleosome assembly, negative regulation of gene expression, cellular protein metabolic process, extracellular exosome, DNA-templated transcription and initiation, positive regulation of cytokine secretion, drug metabolic process, WNT signaling pathway, chemokine production, receptor binding, toll-like receptor 4 binding, and other functions (*Figure 2A-2C* and Table S3). The signaling pathways in which the DEGs of lymph node metastasis were involved were viral carcinogenesis, steroid hormone biosynthesis, drug metabolism-cytochrome P450, chemical carcinogenesis, the PPAR signaling pathway, and others (*Figure 2D* and *Table 1*). The PPI network was presented and was enriched for analysis using the MCODE method (*Figure 3* and Figure S1).

### *Metastasis genes related to prognosis of LUAD*

The results of K-M survival analysis revealed that the expression levels of *PI3*, *CALB1*, *STC1*, *STAR*, *HIST1H4B*, *CYP17A1*, *HIST2H2AB*, *RHOV*, *GUCA2B*, *TM4SF20*, *KRT20*, *HIST1H4A*, *PI15*, *GLP2R*, *KRT78*, *DEFB1*, *CRHR2*, *ABCC2*, *HIST1H2BO*, *CYP4B1*, *LIPF*, *S100G*, *CPB1*, *OTX2*, *KRT16*, *CYP2A6*, and *NTS* were related to the OS of LUAD patients (*Table 2*).

### *Construction of the prognostic model*

Univariate Cox regression analysis revealed that the expression levels of *STC1*, *RHOV*, *GUCA2B*, *ABCC2*, *CYP4B1*, *KRT16*, and *NTS* might be risk factors for the OS of LUAD patients (Figure S2). Multivariate Cox regression analysis and the AIC method revealed that *RHOV*, *ABCC2*, and *CYP4B1* were independent risk factors for a poor prognosis in patients with LUAD, and a prognostic model based on the expression levels of *RHOV*, *ABCC2*, and *CYP4B1* was constructed. In addition, a nomogram based on the expression levels of *RHOV*, *ABCC2*, and *CYP4B1* was constructed (*Figure 4*).

### *Risk score of lymph node metastasis-related genes was associated with poor prognosis in LUAD patients*

*Figure 5A-5C* depict the relationship between the risk score

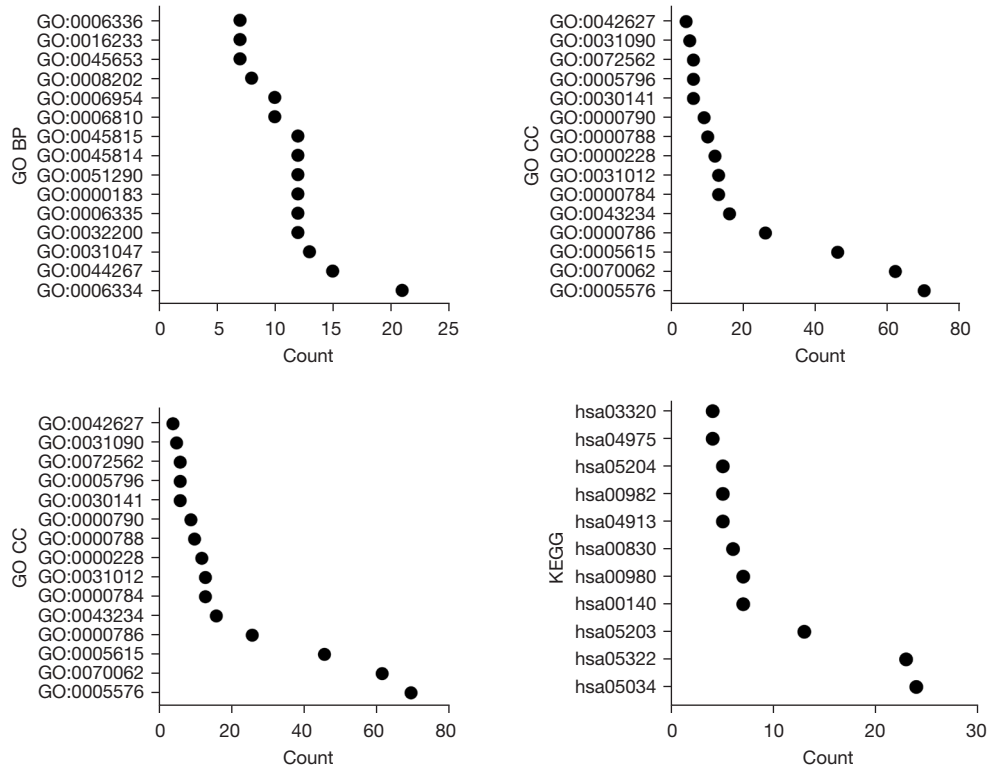


**Figure 2** Functions and mechanisms of LUAD metastasis-related DEGs using GO and KEGG analysis. BP, biological processes; GO, Gene Ontology; CC, cellular components; KEGG, Kyoto Encyclopedia of Genes and Genomes; LUAD, lung adenocarcinoma; DEGs, differentially expressed genes.

**Table 1** Signaling pathways in which DEGs of lymph node metastasis were involved

| Term | Content | Count | P value |
|---|---|---|---|
| Hsa05322 | Systemic lupus erythematosus | 23 | 9.88E-20 |
| Hsa05034 | Alcoholism | 24 | 2.99E-18 |
| Hsa05203 | Viral carcinogenesis | 13 | 5.36E-06 |
| Hsa00140 | Steroid hormone biosynthesis | 7 | 6.31E-05 |
| Hsa00980 | Metabolism of xenobiotics by cytochrome P450 | 7 | 2.48E-04 |
| Hsa00830 | Retinol metabolism | 6 | 9.89E-04 |
| Hsa04913 | Ovarian steroidogenesis | 5 | 0.00276348 |
| Hsa00982 | Drug metabolism-cytochrome P450 | 5 | 0.008937351 |
| Hsa04975 | Fat digestion and absorption | 4 | 0.011324964 |
| Hsa05204 | Chemical carcinogenesis | 5 | 0.015572222 |
| Hsa03320 | PPAR signaling pathway | 4 | 0.046733944 |
| Hsa05202 | Transcriptional misregulation in cancer | 6 | 0.050691025 |

DEGs, differentially expressed genes; PPAR, peroxisome proliferator activated receptor.

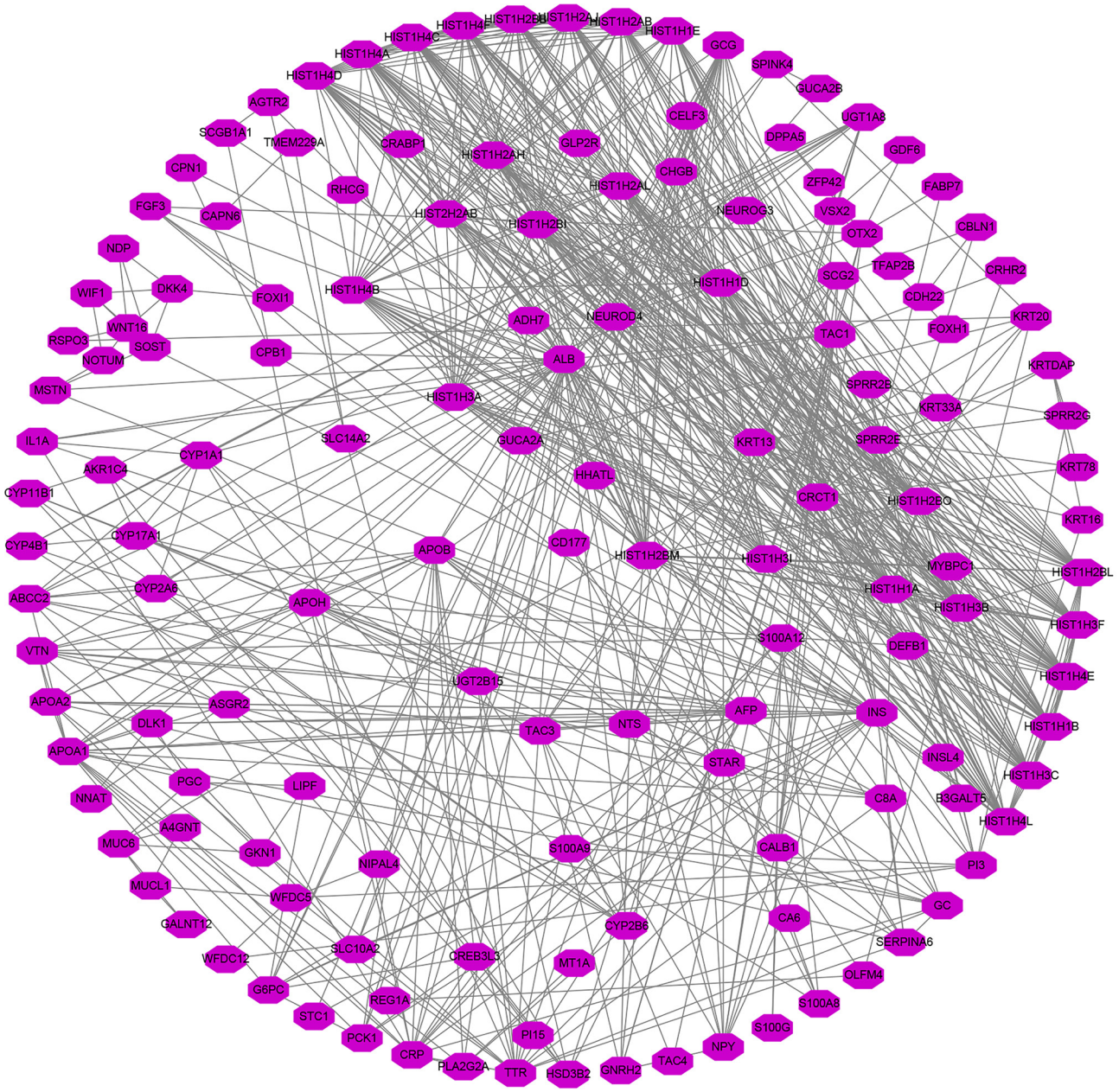


**Figure 3** PPI network of LUAD metastasis-related DEGs. PPI, protein-protein interaction; LUAD, lung adenocarcinoma; DEGs, differentially expressed genes.

and the prognosis of cancer patients. K-M survival analysis illustrated that OS in the low-risk group was significantly higher than in the high-risk group in LUAD patients (*Figure 5D*). Univariate Cox regression analysis demonstrated that the clinical stage, T stage, lymph node metastasis, and risk score were the influencing factors of a poor prognosis in LUAD patients (Figure S3A). Multivariate Cox regression analysis depicted that the clinical stage and prognostic model score were independent risk factors for a poor prognosis in LUAD patients (Figure S3B). A nomogram based on the results of the Cox regression analysis was constructed to assess the prognosis of cancer patients (*Figure 6*).

**Table 2** Screening of prognostic metastasis-related genes using K-M survival analysis

| Gene | P value |
|---|---|
| *ABCC2* | 2.504e-02 |
| *CALB1* | 4.097e-02 |
| *CPB1* | 1.504e-02 |
| *CRHR2* | 2.322e-02 |
| *CYP2A6* | 7.923e-03 |
| *CYP4B1* | 2.655e-03 |
| *CYP17A1* | 9.490e-05 |
| *DEFB1* | 1.287e-02 |
| *GLP2R* | 2.748e-02 |
| *GUCA2B* | 2.236e-03 |
| *HIST1H2BO* | 2.663e-04 |
| *HIST1H4A* | 3.322e-02 |
| *HIST1H4B* | 4.416e-02 |
| *HIST2H2AB* | 1.243e-02 |
| *KRT16* | 6.119e-04 |
| *KRT20* | 4.358e-02 |
| *KRT78* | 1.235e-02 |
| *LIPF* | 3.390e-02 |
| *NTS* | 4.799e-02 |
| *OTX2* | 1.315e-02 |
| *PI15* | 2.613e-02 |
| *PI3* | 3.250e-02 |
| *RHOV* | 7.444e-03 |
| *STAR* | 4.710e-03 |
| *STC1* | 3.702e-03 |
| *TM4SF20* | 4.258e-02 |
| *S100G* | 1.731e-03 |

LUAD, lung adenocarcinoma; K-M, Kaplan-Meier.

### *Prognostic model participation in the signaling mechanisms of LUAD*

In the signaling mechanism module, DNA replication, cell cycle, homologous recombination, mismatch repair, proteasome, pyrimidine metabolism, base excision repair, pentose phosphate pathway, spliceosome, P53 signaling pathway, oocyte meiosis, ubiquitin-mediated proteolysis, basal transcription factors, and other pathways were significantly enriched in the high-risk group (*Table 3*).

### *Prognostic model in LUAD immune microenvironment*

Based on the results of estimation methods concerning LUAD tissues, the risk score was significantly correlated with tumor purity, stromal score, immune score, and estimated score expression levels using correlation analysis (*Figure 7A-7D*). In the high-risk group, the expression of tumor purity was increased, and the expression levels of the stromal score, immune score, and estimate score were decreased (*Figure 7E-7H*). Based on the results of MCP-counter methods concerning LUAD tissues, the risk score was significantly associated with T cells, endothelial cells, neutrophils, and other immune cells (*Table 4*). Based on the results of CIBERSORT methods concerning LUAD tissues, the risk score was significantly associated with the levels of B cell memory, T cell CD8, T cell follicular helper, and other immune cells (*Figure 8* and *Table 5*). In the LUAD tissues, the risk score was significantly associated with the levels of immune cell markers (*Table 6*). More specifically, the risk score was significantly correlated with the expression levels of *BCL6*, *CCR7*, *CCR8*, and other markers of immune cells.

## Discussion

Lung cancer is the most common malignant tumor worldwide and has the highest mortality rate. LUAD accounts for about 40% of lung cancer (1,3,4). Currently, the annual mortality rate of LUAD patients remains high. In the past few decades, the application and mining of big data have constituted one of the important means by which to diagnose, treat, and evaluate the prognosis of patients with cancer. The TCGA database contains high-quality tumor genome data and clinical information of patients. In our study, we found that *PI3*, *CALB1*, *STC1*, *STAR*, *HIST1H4B*, *CYP17A1*, *HIST2H2AB*, *RHOV*, *GUCA2B*, *TM4SF20*, *KRT20*, *HIST1H4A*, *PI15*, *GLP2R*, *KRT78*, *DEFB1*, *CRHR2*, *ABCC2*, *HIST1H2BO*, *CYP4B1*, *LIPF*, *S100G*, *CPB1*, *OTX2*, *KRT16*, CYP2A6, and *NTS* were unusually expressed, and related to the OS of LUAD using the TCGA database. *RHOV*, *ABCC2*, and *CYP4B1* were independent risk factors for a poor prognosis in LUAD patients and were correlated with the prognostic model. Currently, *RHOV*, *ABCC2*, and *CYP4B1* have important biological roles in cancer (16-21). For example, Shepelev

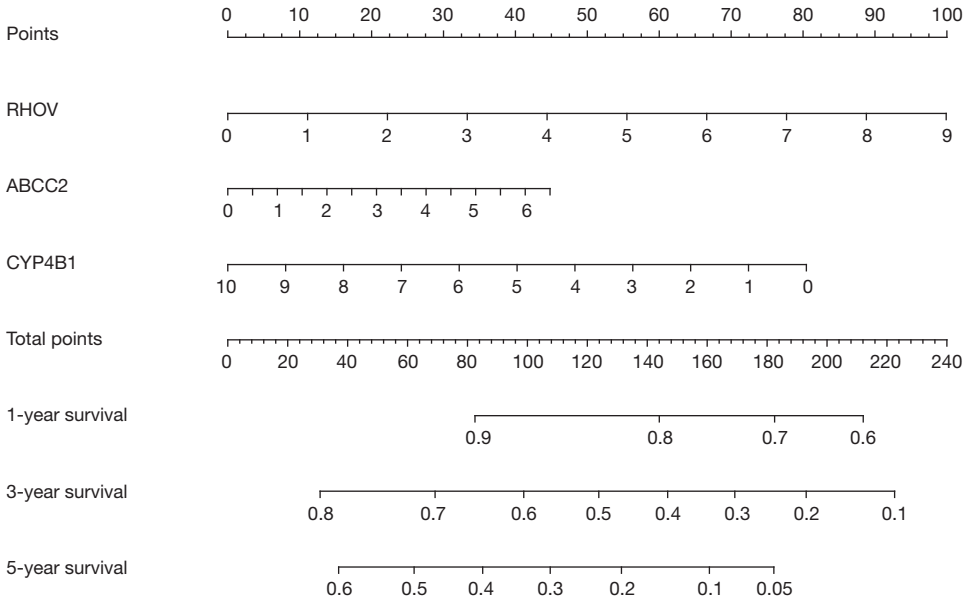


**Figure 4** The nomograms of prognostic genes on overall survival in LUAD. LUAD, lung adenocarcinoma.

*et al.* found that the expression of *RHOV* was increased in lung cancer tissues and cells, and the increased expression of *RHOV* was associated with a poor prognosis of patients (16). Chen *et al.* reported that *ABCC2* was overexpressed in various human cancers (18). The expression of *ABCC2* was upregulated in cisplatin-resistant A549 cells (A549/DDP). Interfering with the expression of *ABCC2* could reverse the resistance of A549/DDP cells to cisplatin *in vitro*, promote G1 phase arrest, and activate the expression of PARP and caspase-3 proteins. The knockout of *ABCC2* expression *in vivo* could enhance the cytotoxicity of cisplatin to subcutaneous transplanted tumors (18). The expression of *CYP4B1* in LUAD decreased, which was related to the history of drug treatment, radiotherapy, and the survival status of cancer patients (21). In addition, we established prognostic model for *RHOV*, *ABCC2*, and *CYP4B1* and demonstrated that *RHOV* and *ABCC2* were overexpressed and *CYP4B1* was underexpressed in the high-risk group, and the OS of LUAD patients in the low-risk group was significantly higher than that in the high-risk group. Univariate Cox regression analysis showed that clinical stage, T stage, lymph node metastasis, and risk score were influencing factors of a poor prognosis in patients with LUAD. Multivariate Cox regression analysis demonstrated that clinical stage and prognostic model score were independent risk factors for a poor prognosis in LUAD patients. This demonstrated that *RHOV*, *ABCC2*, and *CYP4B1* played an important role in lung cancer and indicated that the prognostic model and the nomogram based on the genes *RHOV*, *ABCC2*, and *CYP4B1* of LUAD metastasis have important predictive value.

The process of lung cancer metastasis involves a variety of biological processes and changes of molecular markers (22-27). The expression of long-chain noncoding RNA NSCLCAT1 was upregulated in the NSCLC tissues. NSCLCAT1 could increase the viability, migration, and invasion of NSCLC cells and reduce apoptosis by inhibiting the expression of *CDH1* and mediating the hippo signaling pathway (22). In NSCLC cells, the inhibition of *PLK1* expression could change the expression of genes related to DNA damage, replication, and repair (23). Rig-G was frequently downregulated in lung cancer tissues and cell lines and associated with a poor prognosis in lung cancer patients. The overexpression of Rig-G has been shown to result in a significant reduction in cell growth and migration inhibition in A549 and NCI-H1944 cells, along with a reduced epithelial-to-mesenchymal transition. Rig-G acted as a tumor suppressor through the p53 signaling mechanism (27). The metastasis genes and the prognostic model had important biological value in the cell cycle, DNA replication, p53 signaling pathway, and other mechanisms using GO, KEGG, and GSEA, whereby proving that our prognostic model had good predictive value in the progression of LUAD.

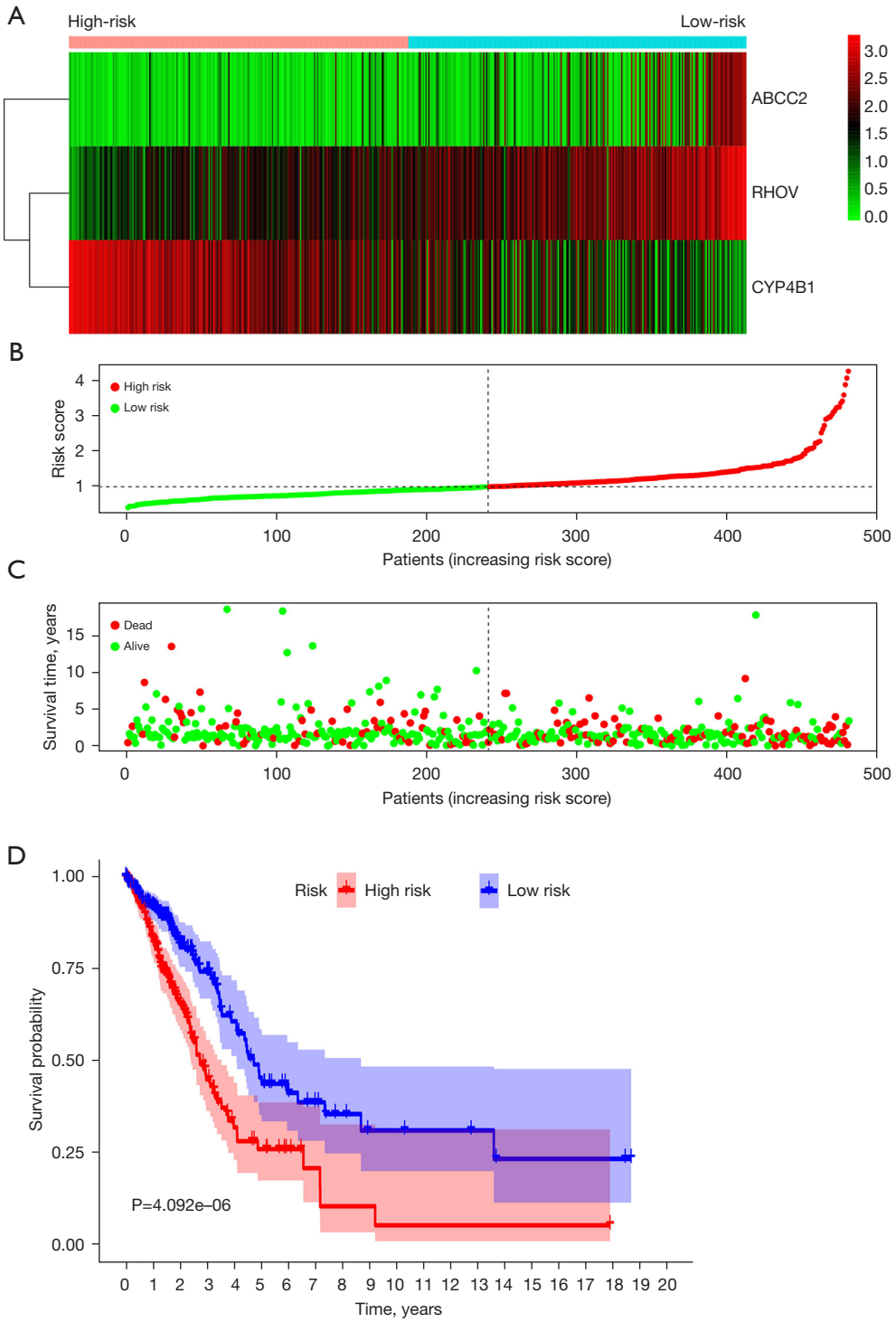


**Figure 5** Evaluation of survival time of patients with LUAD in prognostic model. (A) Prognostic model-related genes showed using heatmap; (B,C) relationship between risk score and prognosis of cancer patients; (D) Kaplan-Meier survival analysis showing OS of LUAD patients in low- and high-risk groups. LUAD, lung adenocarcinoma; OS, overall survival.

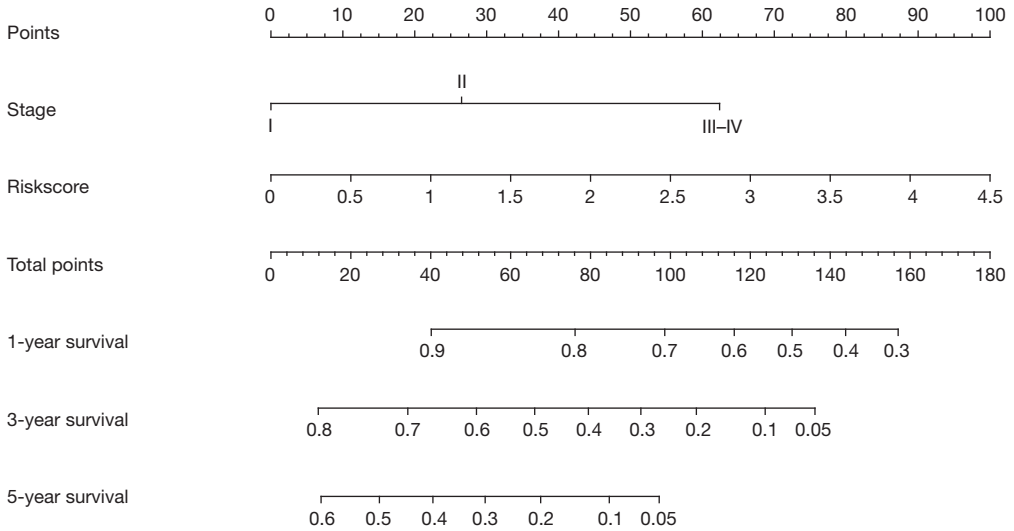


**Figure 6** Nomogram related to prognostic model.

**Table 3** Signaling pathways enriched in high-risk score based on genes of lymph node metastasis

| Name | Size | NES | NOM P value |
|---|---|---|---|
| DNA replication | 36 | 2.188415 | 0 |
| Cell cycle | 124 | 2.101134 | 0 |
| Homologous recombination | 28 | 2.0641458 | 0 |
| Mismatch repair | 23 | 2.0362506 | 0 |
| Proteasome | 44 | 2.0210252 | 0 |
| Pyrimidine metabolism | 97 | 2.0000293 | 0 |
| Base excision repair | 33 | 1.98535 | 0.002083333 |
| Pentose phosphate pathway | 27 | 1.9286441 | 0.004024145 |
| Nucleotide excision repair | 44 | 1.8649865 | 0.005725191 |
| P53 signaling pathway | 68 | 1.7573924 | 0.007905139 |
| Riboflavin metabolism | 15 | 1.6678965 | 0.011494253 |
| Spliceosome | 126 | 1.8515968 | 0.014084507 |
| Pathogenic escherichia coli infection | 55 | 1.7213273 | 0.015717093 |
| Fructose and mannose metabolism | 32 | 1.6948974 | 0.015904572 |
| N glycan biosynthesis | 46 | 1.7191821 | 0.018181818 |
| Protein export | 23 | 1.7690526 | 0.019646365 |
| Oocyte meiosis | 112 | 1.5964646 | 0.023715414 |
| RNA degradation | 57 | 1.628982 | 0.028957529 |
| Glycolysis gluconeogenesis | 61 | 1.6014928 | 0.029166667 |
| Ubiquitin mediated proteolysis | 133 | 1.5566719 | 0.037328094 |
| Purine metabolism | 156 | 1.4549121 | 0.042769857 |

**Table 3** (*continued*)

**Table 3** (*continued*)

| Name | Size | NES | NOM P value |
|---|---|---|---|
| Glyoxylate and dicarboxylate metabolism | 16 | 1.6099515 | 0.042857144 |
| Basal transcription factors | 35 | 1.4993141 | 0.04828974 |
| Galactose metabolism | 25 | 1.5636381 | 0.05 |

NES, normalized enrichment score; NOM, nominal.

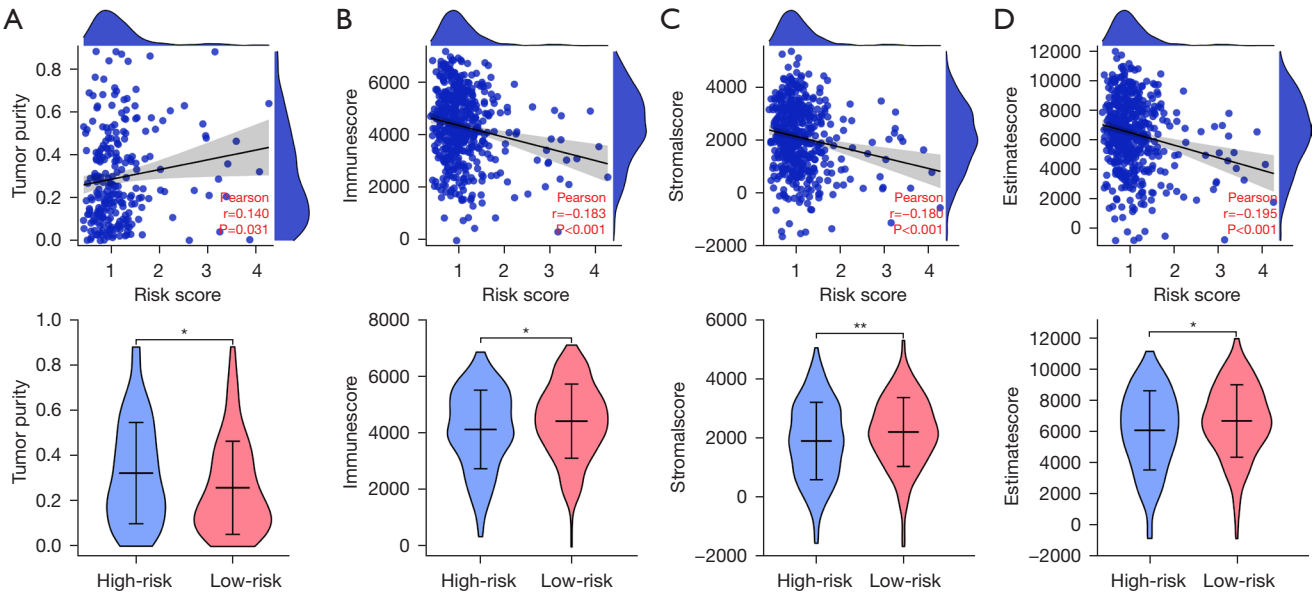


**Figure 7** prognostic model in LUAD immune microenvironment. (A) Tumor purity; (B) immune score; (C) stromal score; (D) estimate score. *, P<0.05; **, P<0.01. LUAD, lung adenocarcinoma.

**Table 4** Risk score was significantly associated with immune cells based on data of MCP-counter analysis

| Immune cells | Correlation coefficient | P value |
|---|---|---|
| T cells | −0.149 | 0.001 |
| CD8 T cells | 0.024 | 0.597 |
| Cytotoxic lymphocytes | 0.03 | 0.505 |
| B lineage | −0.126 | 0.006 |
| NK cells | 0.007 | 0.886 |
| Monocytic lineage | −0.068 | 0.137 |
| Myeloid dendritic cells | −0.290 | <0.001 |
| Neutrophils | −0.109 | 0.017 |
| Endothelial cells | −0.321 | <0.001 |
| Fibroblasts | 0.057 | 0.214 |

MCP, microenvironment cell populations; NK, natural killer.

Recently, immunotherapy has become a dominant therapeutic theme. Immunotherapy could improve long-term survival and the chances of surgery in patients with LUAD (28,29). For example, the CDK5 inhibitor resulted in decreased PD-L1 protein expression in human lung adenocarcinoma (LLC) cells. PD-L1 protein degradation was mediated using the E3 ligase TRIM21 ubiquitination-proteasome pathway (29). *In vitro*, the deletion of CDK5 in the LLC of mice has not been shown to affect cell proliferation. However, the attenuation of CDK5 or binding to anti-PD-L1 was shown to greatly inhibit tumor growth *in vivo* mouse model of LLC implantation. CDK5 disruption caused higher levels of CD3, CD4, and CD8 T cells in the spleen and decreased PD-1 expression in CD4 and CD8 T cells, which provided a potential therapeutic target for LUAD combination immunotherapy (29). In our

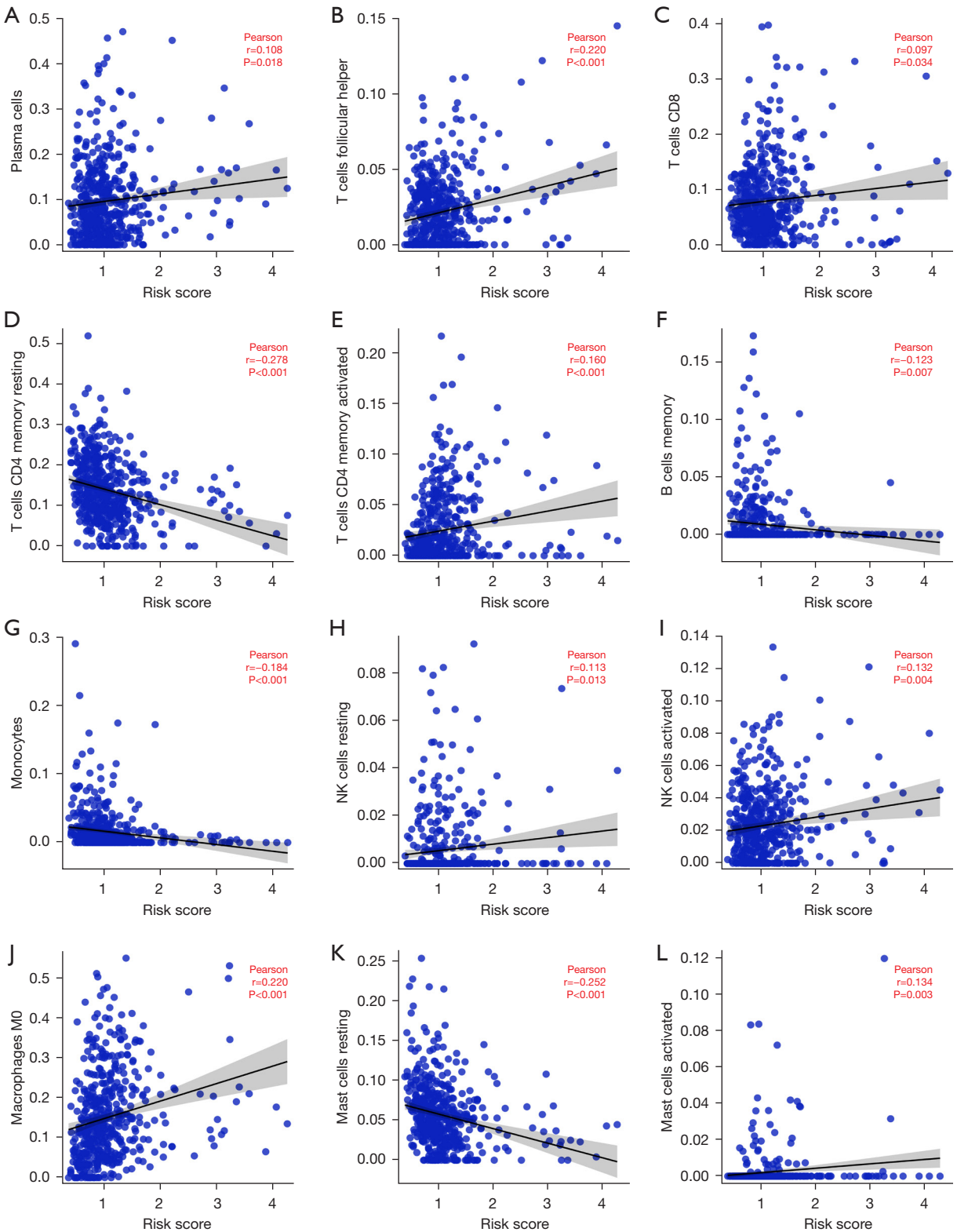


**Figure 8** prognostic model score is significantly associated with immune cell infiltration. (A) Plasma cells. (B) T cells follicular helper. (C) T cells CD8. (D) T cells CD4 memory resting. (E) T cells CD4 memory activated. (F) B cells memory. (G) Monocytes. (H) NK cells resting. (I) NK cells activated. (J) Macrophages M0. (K) Mast cells resting. (L) Mast cells activated. NK, natural killer.

**Table 5** Significant associations of risk score with immune cells based on data of CIBERSORT analysis

| Immune cells | Correlation coefficient | P value |
|---|---|---|
| B cells memory | –0.123 | 0.007 |
| B cells naive | 0.055 | 0.227 |
| Plasma cells | 0.108 | 0.018 |
| T cells CD8 | 0.097 | 0.034 |
| T cells CD4 memory resting | –0.278 | <0.001 |
| T cells CD4 memory activated | 0.160 | <0.001 |
| T cells follicular helper | 0.220 | <0.001 |
| T cells regulatory | –0.021 | 0.652 |
| T cells gamma delta | 0.012 | 0.786 |
| NK cells resting | 0.113 | 0.013 |
| NK cells activated | 0.132 | 0.004 |
| Monocytes | –0.184 | <0.001 |
| Macrophages M0 | 0.220 | <0.001 |
| Macrophages M1 | 0.077 | 0.091 |
| Macrophages M2 | –0.164 | <0.001 |
| Dendritic cells resting | –0.220 | <0.001 |
| Dendritic cells activated | 0.093 | 0.043 |
| Mast cells resting | –0.252 | <0.001 |
| Mast cells activated | 0.134 | 0.003 |
| Eosinophils | 0.019 | 0.672 |
| Neutrophils | 0.095 | 0.037 |

NK, natural killer.

**Table 6** Significant associations of risk score with markers of immune cells

| Cell markers | Correlation coefficient | P value |
|---|---|---|
| *BCL6* | –0.102 | 0.025 |
| *CCR8* | –0.106 | 0.020 |
| *CD2* | –0.100 | 0.028 |
| *CD3E* | –0.101 | 0.027 |
| *CD8B* | 0.024 | 0.597 |
| *CD79A* | –0.081 | 0.076 |
| *CEACAM8* | –0.156 | <0.001 |
| *FOXP3* | –0.091 | 0.046 |
| *GZMB* | 0.170 | <0.001 |
| *HLA-DPA1* | –0.307 | <0.001 |
| *HLA-DQB1* | –0.256 | <0.001 |
| *IFNG* | 0.094 | 0.040 |
| *IL17A* | –0.007 | 0.877 |
| *IRF5* | –0.046 | 0.311 |
| *ITGAX* | –0.154 | <0.001 |
| *MS4A4A* | –0.165 | <0.001 |
| *NRP1* | –0.167 | <0.001 |
| *PTGS2* | 0.039 | 0.392 |
| *STAT3* | –0.072 | 0.114 |
| *STAT5B* | –0.239 | <0.001 |
| *TBX21* | 0.014 | 0.755 |
| *TNF* | –0.002 | 0.971 |
| *CCR7* | –0.180 | <0.001 |
| *CD1C* | –0.268 | <0.001 |
| *CD3D* | –0.058 | 0.208 |
| *CD8A* | 0.022 | 0.626 |
| *CD19* | –0.113 | 0.013 |
| *CD163* | –0.092 | 0.044 |
| *CTLA4* | –0.052 | 0.251 |
| *GATA3* | 0.185 | <0.001 |
| *HAVCR2* | –0.114 | 0.012 |
| *HLA-DPB1* | –0.326 | <0.001 |
| *HLA-DRA* | –0.273 | <0.001 |
| *IL13* | –0.100 | 0.029 |
| *IL21* | –0.012 | 0.785 |
| *ITGAM* | –0.201 | <0.001 |
| *LAG3* | 0.023 | 0.613 |
| *NOS2* | –0.009 | 0.845 |
| *PDCD1* | 0.018 | 0.701 |
| *STAT1* | 0.185 | <0.001 |
| *STAT5A* | –0.182 | <0.001 |
| *STAT6* | –0.180 | <0.001 |
| *TGFB1* | –0.112 | 0.014 |
| *VSIG4* | –0.142 | 0.002 |

research, the prognostic model score based on the *RHOV*, *ABCC2*, and *CYP4B1* was significantly correlated with tumor purity, stromal score, immune score, and estimate score expression levels. The risk score was significantly associated with T cells, endothelial cells, neutrophils, CD8, T cell follicular helper cells, and other immune cells. In LUAD tissues, the risk score was significantly associated with the levels of immune cell markers. In addition, study has confirmed that RHOV is related to the regulation of immune cells (30). Specifically, RHOV expression increased during the differentiation of macrophages into osteoclasts, while a large number of macrophages showed apoptosis. When osteoprotegerin (OPG) inhibits the differentiation of macrophages into osteoclasts, and then OPG can inhibit apoptosis, which is related to the down-regulation of RHOV expression level (30). However, the relationship between *ABCC2*, and *CYP4B1* and immune cell regulation has not been reported in the literature, which will be our research direction in the future.

The molecular mechanisms of lymph node metastasis have not been fully understood in LUAD. Lymph node metastasis-related genes *RHOV*, *ABCC2*, and *CYP4B1* were found to be associated with the prognosis of LUAD patients. In addition, this study used big data samples to provide new candidate biomarkers related to LUAD metastasis for the prognosis of LUAD patients, which has the advantage of high reliability. However, our study also had some limitations. First, the expression levels of *RHOV*, *ABCC2*, and *CYP4B1* in clinical LUAD samples need to be verified. Moreover, the expression levels of *RHOV*, *ABCC2*, and *CYP4B1* and the values of their constructed prognostic model in the prognosis of LUAD patients need to be explored. The roles and the underlying signaling mechanisms of our constructed prognostic model were explored in LUAD using basic research in the future.

## Conclusions

There are many molecular DEGs in the process of LUAD metastasis. *RHOV*, *ABCC2*, and *CYP4B1* are influencing factors of a poor prognosis, and clinical stage and risk score are independent risk factors for a poor prognosis in patients with LUAD. The prognostic model might be involved in the progression of LUAD through the cell cycle, DNA replication, p53 signaling pathway, and others. A prognostic model based on *RHOV*, *ABCC2*, and *CYP4B1* might predict the prognosis of LUAD patients and be associated with immune infiltration.

## Acknowledgments

Thanks to the TCGA database for providing the data used for analysis.
*Funding:* None.

## Footnote

*Reporting Checklist:* The authors have completed the TRIPOD reporting checklist. Available at https://jtd.amegroups.com/article/view/10.21037/jtd-23-265/rc

*Conflicts of Interest:* All authors have completed the ICMJE uniform disclosure form (available at https://jtd.amegroups.com/article/view/10.21037/jtd-23-265/coif). The authors have no conflicts of interest to declare.

*Ethical Statement:* The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).



## References

1. Chen PW, Huang SK, Chou WC, et al. Severinia buxifolia-isolated acridones inhibit lung cancer invasion and decrease HIFα protein synthesis involving 5'UTR-mediated translation inhibition. Phytomedicine 2023;109:154570.
2. Ding DX, Li Q, Shi K, et al. LncRNA NEAT1-miR-101-3p/miR-335-5p/miR-374a-3p/miR-628-5p-TRIM6 axis identified as the prognostic biomarker for lung adenocarcinoma via bioinformatics and meta-analysis. Transl Cancer Res 2021;10:4870-83.
3. Feng RM, Zong YN, Cao SM, et al. Current cancer situation in China: good or bad news from the 2018 Global

 

Cancer Statistics? Cancer Commun (Lond) 2019;39:22.
4. Kleczko EK, Kwak JW, Schenk EL, et al. Targeting the Complement Pathway as a Therapeutic Strategy in Lung Cancer. Front Immunol 2019;10:954.
5. Yang J, Peng A, Wang B, et al. The prognostic impact of lymph node metastasis in patients with non-small cell lung cancer and distant organ metastasis. Clin Exp Metastasis 2019;36:457-66.
6. Luo T, Chen Q, Zeng J. Analysis of lymph node metastasis in 200 patients with non-small cell lung cancer. Transl Cancer Res 2020;9:1577-83.
7. Dai C, Xie H, Kadeer X, et al. Relationship of Lymph Node Micrometastasis and Micropapillary Component and Their Joint Influence on Prognosis of Patients With Stage I Lung Adenocarcinoma. Am J Surg Pathol 2017;41:1212-20.
8. Guo Q, Wu CY, Jiang N, et al. Downregulation of T-cell cytotoxic marker IL18R1 promotes cancer proliferation and migration and is associated with dismal prognosis and immunity in lung squamous cell carcinoma. Front Immunol 2022;13:986447.
9. Gong WJ, Zhou T, Wu SL, et al. A novel immune-related ceRNA network that predicts prognosis and immunotherapy response in lung adenocarcinoma. Ann Transl Med 2021;9:1484.
10. Du J, Zheng L, Chen S, et al. NFIL3 and its immunoregulatory role in rheumatoid arthritis patients. Front Immunol 2022;13:950144.
11. Guo Q, Liu XL, Liu HS, et al. The Risk Model Based on the Three Oxidative Stress-Related Genes Evaluates the Prognosis of LAC Patients. Oxid Med Cell Longev 2022;2022:4022896.
12. Huang H, Wu W, Lu Y, et al. The development and validation of a m6A-lncRNAs based prognostic model for overall survival in lung squamous cell carcinoma. J Thorac Dis 2022;14:4055-72.
13. Zhang YQ, Li K, Guo Q, et al. A New Risk Model Based on 7 Quercetin-Related Target Genes for Predicting the Prognosis of Patients With Lung Adenocarcinoma. Front Genet 2022;13:890079.
14. Guo Q, Xiao XY, Wu CY, et al. Clinical Roles of Risk Model Based on Differentially Expressed Genes in Mesenchymal Stem Cells in Prognosis and Immunity of Non-small Cell Lung Cancer. Front Genet 2022;13:823075.
15. Shen A, Ye Y, Chen F, et al. Integrated multi-omics analysis identifies CD73 as a prognostic biomarker and immunotherapy response predictor in head and neck squamous cell carcinoma. Front Immunol 2022;13:969034.
16. Shepelev MV, Korobko IV. The RHOV gene is overexpressed in human non-small cell lung cancer. Cancer Genet 2013;206:393-7.
17. Zhang D, Jiang Q, Ge X, et al. RHOV promotes lung adenocarcinoma cell growth and metastasis through JNK/c-Jun pathway. Int J Biol Sci 2021;17:2622-32.
18. Chen Y, Zhou H, Yang S, et al. Increased ABCC2 expression predicts cisplatin resistance in non-small cell lung cancer. Cell Biochem Funct 2021;39:277-86.
19. Qian CY, Zheng Y, Wang Y, et al. Associations of genetic polymorphisms of the transporters organic cation transporter 2 (OCT2), multidrug and toxin extrusion 1 (MATE1), and ATP-binding cassette subfamily C member 2 (ABCC2) with platinum-based chemotherapy response and toxicity in non-small cell lung cancer patients. Chin J Cancer 2016;35:85.
20. Lim S, Alshagga M, Ong CE, et al. Cytochrome P450 4B1 (CYP4B1) as a target in cancer treatment. Hum Exp Toxicol 2020;39:785-96.
21. Liu X, Jia Y, Shi C, et al. CYP4B1 is a prognostic biomarker and potential therapeutic target in lung adenocarcinoma. PLoS One 2021;16:e0247020.
22. Zhao W, Zhang LN, Wang XL, et al. Long noncoding RNA NSCLCAT1 increases non-small cell lung cancer cell invasion and migration through the Hippo signaling pathway by interacting with CDH1. FASEB J 2019;33:1151-66.
23. Yao D, Gu P, Wang Y, et al. Inhibiting polo-like kinase 1 enhances radiosensitization via modulating DNA repair proteins in non-small-cell lung cancer. Biochem Cell Biol 2018;96:317-25.
24. Wang Y, Zhang J, Li YJ, et al. MEST promotes lung cancer invasion and metastasis by interacting with VCP to activate NF-κB signaling. J Exp Clin Cancer Res 2021;40:301.
25. Zhang Q, Liu H, Zhang J, et al. MiR-142-5p Suppresses Lung Cancer Cell Metastasis by Targeting Yin Yang 1 to Regulate Epithelial-Mesenchymal Transition. Cell Reprogram 2020;22:328-36.
26. Chao CC, Lee WF, Wang SW, et al. CXC chemokine ligand-13 promotes metastasis via CXCR5-dependent signaling pathway in non-small cell lung cancer. J Cell Mol Med 2021;25:9128-40.
27. Sun J, Wang X, Liu W, et al. Novel evidence for retinoic acid-induced G (Rig-G) as a tumor suppressor by activating p53 signaling pathway in lung cancer. FASEB J 2020;34:11900-12.

28. Zhang YQ, Yuan Y, Zhang J, et al. Evaluation of the roles and regulatory mechanisms of PD-1 target molecules in NSCLC progression. Ann Transl Med 2021;9:1168.
29. Gao L, Xia L, Ji W, et al. Knockdown of CDK5 down-regulates PD-L1 via the ubiquitination-proteasome pathway and improves antitumor immunity in lung adenocarcinoma. Transl Oncol 2021;14:101148.
30. Song R, Liu X, Zhu J, et al. RhoV mediates apoptosis of RAW264.7 macrophages caused by osteoclast differentiation. Mol Med Rep 2015;11:1153-9.

(English Language Editor: J. Jones)

**Cite this article as:** Li Q, Liu XL, Jiang N, Li QY, Song YX, Ke XX, Han H, Luo Q, Guo Q, Luo XY, Chen C. A new prognostic model for *RHOV*, *ABCC2*, and *CYP4B1* to predict the prognosis and association with immune infiltration of lung adenocarcinoma. J Thorac Dis 2023;15(4):1919-1934. doi: 10.21037/jtd-23-265

## Supplementary

**Table S1** Highly expressed genes associated with LUAD metastasis

| Gene | logFC | Gene | logFC | Gene | logFC |
|---|---|---|---|---|---|
| *LRRC38* | 4.115872102 | *GUCA2B* | 2.276642003 | *ABCC2* | 1.374776363 |
| *NNAT* | 2.598832901 | *TM4SF20* | 2.67559724 | *CHGB* | 1.670343784 |
| *S100A8* | 2.03031514 | *KRT20* | 2.845900777 | *NDP* | 1.17157837 |
| *MT1A* | 1.876167804 | *SPAG11B* | 2.277371446 | *NEUROD4* | 1.175330904 |
| *PI3* | 2.408197028 | *RSPO3* | 1.257348989 | *FAM228A* | 1.021207733 |
| *S100A12* | 1.540121983 | *GLP2R* | 1.419295928 | *ADH7* | 1.401616715 |
| *CYP2B6* | 2.384631964 | *KRT78* | 1.242914841 | *PAGE1* | 1.839951996 |
| *RHCG* | 2.241716309 | *RTP1* | 1.268393463 | *EPHA5* | 1.130400739 |
| *TAC1* | 3.731489817 | *SOST* | 1.696655357 | *KRT16* | 1.071459263 |
| *CALB1* | 3.16661865 | *CXorf67* | 1.465135742 | *KRTDAP* | 1.159388837 |
| *S100A9* | 1.466930019 | *GDPD2* | 1.312368616 | *SCG2* | 1.144724598 |
| *CGB5* | 3.152939448 | *ZACN* | 1.177117745 | *HTR3B* | 1.123094007 |
| *C1QTNF3* | 1.124963406 | *TRIM48* | 2.697396009 | *GUCA2A* | 1.039662581 |
| *STC1* | 1.053503933 | *BPIFB4* | 1.681609102 | *EIF4E1B* | 1.133032617 |
| *NKX2-3* | 1.865700882 | *AC187653.1* | 1.378861741 | *KHDC1L* | 1.130389461 |
| *COLEC10* | 1.514190244 | *PRH2* | 1.347665707 | *SPAG11A* | 1.481473314 |
| *SPX* | 1.788358216 | *DEFB1* | 1.259832041 | *SLC10A2* | 1.408515173 |
| *MUC2* | 2.238366652 | *NIPAL4* | 1.123020494 | *HHATL* | 1.149929728 |
| *VSX2* | 1.572115559 | *GSG1L2* | 1.24992093 | *TEX19* | 1.015740755 |
| *IL1A* | 1.347787178 | *A2ML1* | 1.72894757 | *NTS* | 1.467268835 |
| *PRB3* | 1.499905792 | *CRCT1* | 1.564500616 | *CRP* | 1.056388369 |
| *C8A* | 1.775486777 | *MYBPH* | 1.268795033 | *KRT13* | 1.090032304 |
| *MUCL1* | 1.942169109 | *ECEL1* | 1.196685891 | *NFE4* | 1.004469025 |
| *ACTL8* | 2.033243312 | *GCG* | 2.47824742 | *PRB1* | 1.112467601 |
| *RHOV* | 1.106145485 | *DAPL1* | 1.175622384 | *INSL4* | 1.10861026 |

FC, fold change; LUAD, lung adenocarcinoma.


https://dx.doi.org/10.21037/jtd-23-265

**Table S2** Lowly expressed genes associated with LUAD metastasis

| Gene | logFC | Gene | logFC | Gene | logFC |
|---|---|---|---|---|---|
| *ALB* | –5.079728232 | *HIST2H2AB* | –1.95079741 | *APOB* | –1.48668674 |
| *NPY* | –4.660537685 | *HIST1H2AH* | –1.945305733 | *CYP1A1* | –1.479514228 |
| *AFP* | –4.150027541 | *FER1L6* | –1.930156354 | *CPB1* | –1.446707372 |
| *SPINK4* | –3.520266681 | *HIST1H3C* | –1.927950184 | *FGF3* | –1.44132808 |
| *WFDC12* | –3.442749412 | *HIST1H2BB* | –1.918138428 | *TTR* | –1.431390359 |
| *DKK4* | –3.426817639 | *CRABP1* | –1.916379679 | *LIPF* | –1.492784384 |
| *WFDC5* | –3.421041261 | *HIST1H2BI* | –1.88901368 | *COX8C* | –1.413522106 |
| *DLK1* | –3.309753519 | *HIST1H3I* | –1.863873343 | *ZFP42* | –1.400084091 |
| *MSTN* | –3.192568886 | *RERGL* | –1.834274109 | *CYP2A6* | –1.387772556 |
| *FABP7* | –3.084941905 | *HIST1H3F* | –1.805336967 | *PSG3* | –1.38172679 |
| *GKN1* | –1.801620515 | *HIST1H4E* | –1.787070398 | *S100G* | –1.349125942 |
| *BHMT* | –1.796430282 | *UGT2B15* | –1.779261119 | *TUBA3E* | –1.32599556 |
| *GC* | –2.843840534 | *HIST1H2AB* | –1.768817617 | *PRSS33* | –1.312170884 |
| *VTN* | –2.672501666 | *HIST1H2AJ* | –1.646772648 | *VCX* | –1.134784664 |
| *FTHL17* | –2.536548085 | *CYP11B1* | –1.626906727 | *NPTX1* | –1.132641041 |
| *MYBPC1* | –2.478767731 | *HIST1H4C* | –2.861322629 | *RBM46* | –1.135815572 |
| *PSG4* | –2.390784252 | *HIST1H1B* | –2.853524819 | *ASGR2* | –1.109988044 |
| *OLFM4* | –2.376884767 | *HIST1H3B* | –2.295596172 | *MAEL* | –1.303017648 |
| *VGLL2* | –2.376288645 | *HIST1H4L* | –2.267483191 | *AKR1C4* | –1.30261952 |
| *APOA1* | –2.365204735 | *HIST1H4D* | –2.168005302 | *NEUROG3* | –1.296631195 |
| *KIR2DL1* | –2.312066188 | *HIST1H4F* | –2.152486674 | *LILRA2* | –1.295379798 |
| *PSG5* | –2.311257398 | *HIST1H1E* | –2.149380269 | *CAPN6* | –1.294482763 |
| *FOXI1* | –1.788148655 | *HIST1H1D* | –2.08342783 | *BMX* | –1.043823741 |
| *GNRH2* | –1.557652305 | *SPRR2G* | –2.01057535 | *CA6* | –1.25855064 |
| *G6PC* | –2.190878152 | *HIST1H4B* | –1.991030038 | *DPPA5* | –1.249194272 |
| *SPIC* | –1.606336263 | *HSD3B2* | –1.554002906 | *B3GALT5* | –1.248282703 |
| *STAR* | –1.58316233 | *HIST1H2AL* | –1.427435476 | *SLC7A3* | –1.245150828 |
| *TFAP2B* | –1.568999384 | *PLA2G2A* | –1.410032331 | *DHRS2* | –1.226463753 |
| *INS* | –2.095832241 | *HIST1H4A* | –1.403263725 | *CPN1* | –1.222939832 |
| *MUC6* | –1.743521454 | *TMEM229A* | –1.35988277 | *PCK1* | –1.217873797 |
| *PGC* | –2.057997158 | *HIST1H3A* | –1.308476052 | *PI15* | –1.214958866 |
| *TAC3* | –2.03204835 | *HIST1H2BM* | –1.282864769 | *COL11A2* | –1.208569681 |
| *ZP2* | –1.611639323 | *HIST1H1A* | –1.175265584 | *TAC4* | –1.199695777 |
| *WIF1* | –1.637147552 | *SERPINA6* | –1.159766806 | *CDH16* | –1.197352159 |
| *XAGE3* | –1.545787281 | *HIST1H2BO* | –1.15219093 | *FXYD4* | –1.082317459 |
| *CRHR2* | –1.197145579 | *APOA2* | –1.517103224 | *SLC5A8* | –1.071277875 |
| *CBLN1* | –1.194660601 | *TUBA3C* | –1.489488203 | *SCGB1A1* | –1.056351964 |
| *WNT16* | –1.190775724 | *CYP17A1* | –1.958874326 | *SPRR2B* | –1.05593709 |
| *OTX2* | –1.189934074 | *PIK3C2G* | –1.03113146 | *CD177* | –1.05557762 |
| *CREB3L3* | –1.509649185 | *SPANXD* | –1.108219388 | *NOTUM* | –1.049113799 |
| *CDH22* | –1.168108863 | *CDC20B* | –1.105517671 | *PRAC1* | –1.048660802 |
| *PAGE2* | –1.539126019 | *NELL1* | –1.104602759 | *CYP4B1* | –1.04420733 |
| *A4GNT* | –1.528089966 | *UGT1A8* | –1.102632454 | *FOXH1* | –1.043004968 |
| *SPRR2E* | –1.52807188 | *CELF3* | –1.087787494 | *KRT33A* | –1.142192724 |
| *REG1A* | –1.526407309 | *AGTR2* | –1.020513278 | *HIST1H2BL* | –1.136379748 |
| *SLC14A2* | –1.00392211 | *APOH* | –1.015314806 | *GDF6* | –1.021729806 |

FC, fold change; LUAD, lung adenocarcinoma.


https://dx.doi.org/10.21037/jtd-23-265

**Table S3** Functions of LUAD metastasis-related DEGs using GO analysis

| GO | Category | Count | P |
|---|---|---|---|
| CC | Nucleosome | 26 | 7.61E–28 |
| CC | Extracellular region | 70 | 8.70E–24 |
| BP | Nucleosome assembly | 21 | 4.67E–18 |
| BP | Telomere organization | 12 | 3.59E–15 |
| BP | DNA replication-dependent nucleosome assembly | 12 | 3.43E–14 |
| BP | Chromatin silencing at rDNA | 12 | 2.16E–13 |
| BP | Protein heterotetramerization | 12 | 1.03E–12 |
| BP | Negative regulation of gene expression, epigenetic | 12 | 8.30E–12 |
| CC | Nuclear chromosome | 12 | 1.33E–11 |
| CC | Extracellular space | 46 | 2.23E–11 |
| MF | Protein heterodimerization activity | 26 | 6.45E–11 |
| BP | Cellular protein metabolic process | 15 | 7.48E–11 |
| BP | Positive regulation of gene expression, epigenetic | 12 | 1.00E–10 |
| CC | Nuclear nucleosome | 10 | 1.17E–09 |
| BP | Gene silencing by RNA | 13 | 5.20E–09 |
| CC | Nuclear chromosome, telomeric region | 13 | 2.54E–08 |
| BP | Negative regulation of megakaryocyte differentiation | 7 | 3.42E–08 |
| MF | Histone binding | 12 | 1.11E–07 |
| CC | Extracellular exosome | 62 | 1.15E–07 |
| BP | Telomere capping | 7 | 1.77E–07 |
| BP | DNA replication-independent nucleosome assembly | 7 | 3.93E–07 |
| BP | Steroid metabolic process | 8 | 5.21E–07 |
| BP | DNA-templated transcription, initiation | 7 | 3.03E–06 |
| BP | Regulation of gene silencing | 5 | 5.13E–06 |
| BP | Beta-catenin-TCF complex assembly | 7 | 8.86E–06 |
| BP | CENP-A containing nucleosome assembly | 7 | 8.86E–06 |
| CC | Protein complex | 16 | 6.53E–05 |
| BP | Double-strand break repair via nonhomologous end joining | 7 | 8.17E–05 |
| MF | Hormone activity | 8 | 8.31E–05 |
| CC | Extracellular matrix | 13 | 1.32E–04 |
| CC | Chylomicron | 4 | 4.59E–04 |
| MF | Cholesterol transporter activity | 4 | 5.52E–04 |
| BP | Retinoid metabolic process | 6 | 6.57E–04 |
| BP | Triglyceride metabolic process | 5 | 6.57E–04 |
| BP | Glucocorticoid metabolic process | 3 | 7.69E–04 |
| MF | Vitamin D binding | 3 | 0.001194091 |
| CC | Very-low-density lipoprotein particle | 4 | 0.001369119 |
| CC | Secretory granule | 6 | 0.001441816 |
| CC | Nuclear chromatin | 9 | 0.001519683 |
| MF | Nucleosomal DNA binding | 5 | 0.00165476 |
| BP | Chromatin silencing | 5 | 0.001711258 |
| BP | Keratinocyte differentiation | 6 | 0.001777404 |
| MF | Oxygen binding | 5 | 0.00179322 |
| BP | Coumarin metabolic process | 3 | 0.001892821 |
| MF | Protein domain specific binding | 9 | 0.002258561 |
| BP | Negative regulation of endopeptidase activity | 7 | 0.002678439 |
| BP | Positive regulation of cytokine secretion | 4 | 0.002812071 |
| BP | Defense response to Gram-positive bacterium | 6 | 0.00290604 |
| BP | Tachykinin receptor signaling pathway | 3 | 0.003480372 |
| BP | Drug metabolic process | 4 | 0.003516806 |
| BP | Female pregnancy | 6 | 0.003545304 |
| MF | Chromatin DNA binding | 5 | 0.003881638 |
| BP | Negative regulation of heart rate | 3 | 0.00444119 |
| BP | Glucocorticoid biosynthetic process | 3 | 0.00444119 |
| BP | Lipoprotein biosynthetic process | 3 | 0.00444119 |
| MF | Serine-type endopeptidase inhibitor activity | 6 | 0.004479843 |
| CC | Golgi lumen | 6 | 0.004491116 |
| MF | RAGE receptor binding | 3 | 0.006286122 |
| BP | Regulation of blood pressure | 5 | 0.006515875 |
| BP | Response to estrogen | 5 | 0.006515875 |
| BP | Low-density lipoprotein particle remodeling | 3 | 0.006683856 |
| MF | Iron ion binding | 7 | 0.007220785 |
| BP | Cholesterol metabolic process | 5 | 0.007633451 |
| CC | Secretory granule lumen | 3 | 0.007644016 |
| BP | Exogenous drug catabolic process | 3 | 0.007960614 |
| BP | Lipoprotein metabolic process | 4 | 0.009253098 |
| BP | Acute inflammatory response | 3 | 0.009337671 |
| MF | Oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen, reduced flavin or flavoprotein as one donor, and incorporation of one atom of oxygen | 3 | 0.011657032 |
| BP | Antibacterial humoral response | 4 | 0.0138206 |
| CC | Cornified envelope | 4 | 0.014739111 |
| MF | Phospholipid binding | 5 | 0.015309052 |
| CC | Organelle membrane | 5 | 0.016493902 |
| BP | Keratinization | 4 | 0.017460907 |
| BP | Epoxygenase P450 pathway | 3 | 0.017643624 |
| MF | Heme binding | 6 | 0.01827361 |
| BP | Transport | 10 | 0.018975122 |
| BP | Peptide cross-linking | 4 | 0.019462409 |
| BP | Regulation of cytoskeleton organization | 3 | 0.019572763 |
| MF | Structural molecule activity | 8 | 0.020333057 |
| BP | WNT signaling pathway | 7 | 0.020713935 |
| BP | Insecticide metabolic process | 2 | 0.022737969 |
| BP | Neutrophil aggregation | 2 | 0.022737969 |
| CC | High-density lipoprotein particle | 3 | 0.024865479 |
| BP | Defense response to Gram-negative bacterium | 4 | 0.024998124 |
| MF | Oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen | 4 | 0.025282421 |
| MF | Monooxygenase activity | 4 | 0.026449736 |
| MF | Retinoic acid binding | 3 | 0.026510249 |
| MF | Lipid binding | 6 | 0.026517937 |
| CC | Blood microparticle | 6 | 0.02841997 |
| BP | Glucose homeostasis | 5 | 0.028683462 |
| BP | Inflammatory response | 10 | 0.030153755 |
| BP | O-glycan processing | 4 | 0.031292909 |
| BP | Triglyceride catabolic process | 3 | 0.03283976 |
| BP | Cholesterol efflux | 3 | 0.03283976 |
| MF | Apolipoprotein receptor binding | 2 | 0.032867777 |
| MF | High-density lipoprotein particle receptor binding | 2 | 0.032867777 |
| MF | Heparin binding | 6 | 0.032881847 |
| MF | Steroid hydroxylase activity | 3 | 0.033328039 |
| BP | Protein oxidation | 2 | 0.03391333 |
| BP | Positive regulation of saliva secretion | 2 | 0.03391333 |
| BP | Biphenyl metabolic process | 2 | 0.03391333 |
| BP | Chemokine production | 2 | 0.03391333 |
| BP | Negative regulation of very-low-density lipoprotein particle remodeling | 2 | 0.03391333 |
| MF | Aromatase activity | 3 | 0.035736913 |
| BP | Cholesterol homeostasis | 4 | 0.036870385 |
| BP | Bile acid and bile salt transport | 3 | 0.037857464 |
| BP | Defense response to fungus | 3 | 0.037857464 |
| BP | Response to lipopolysaccharide | 6 | 0.04031827 |
| BP | Defense response | 4 | 0.041365056 |
| BP | Cellular response to dexamethasone stimulus | 3 | 0.043146799 |
| MF | Receptor binding | 9 | 0.043246864 |
| MF | Neuropeptide hormone activity | 3 | 0.043347817 |
| MF | Toll-like receptor 4 binding | 2 | 0.04358313 |
| BP | Positive regulation of peptide secretion | 2 | 0.044961554 |
| BP | Sequestering of zinc ion | 2 | 0.044961554 |
| BP | Negative regulation of lipase activity | 2 | 0.044961554 |

GO, Gene Ontology; MF, molecular function; BP: Biological processes; CC: Cellular components; DEGs, differentially expressed genes; LUAD, lung adenocarcinoma.


https://dx.doi.org/10.21037/jtd-23-265

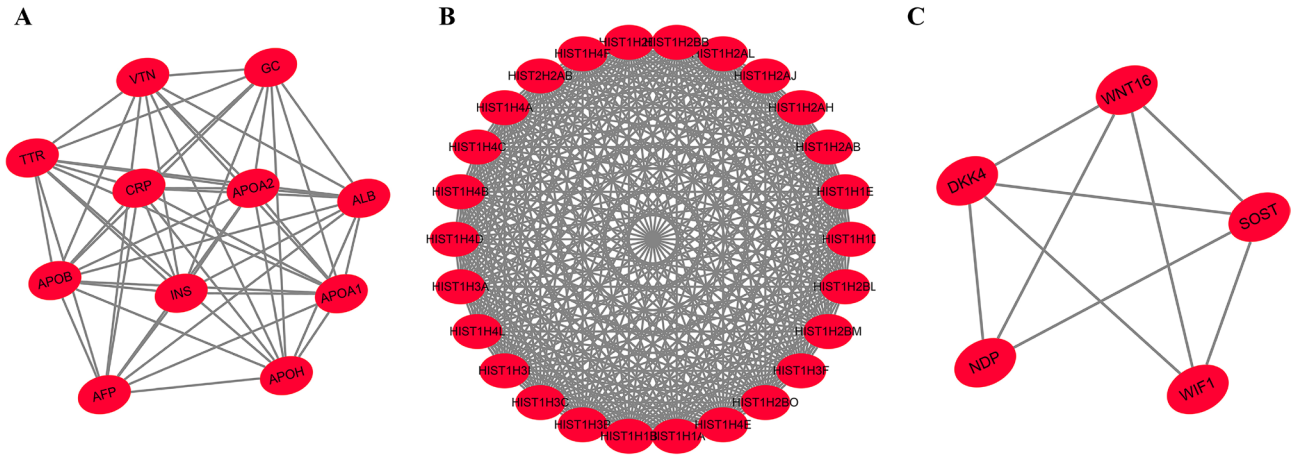


**Figure S1** PPI network of LUAD metastasis-related DEGs using MCODE method. LUAD, lung adenocarcinoma; DEGs, differentially expressed genes.

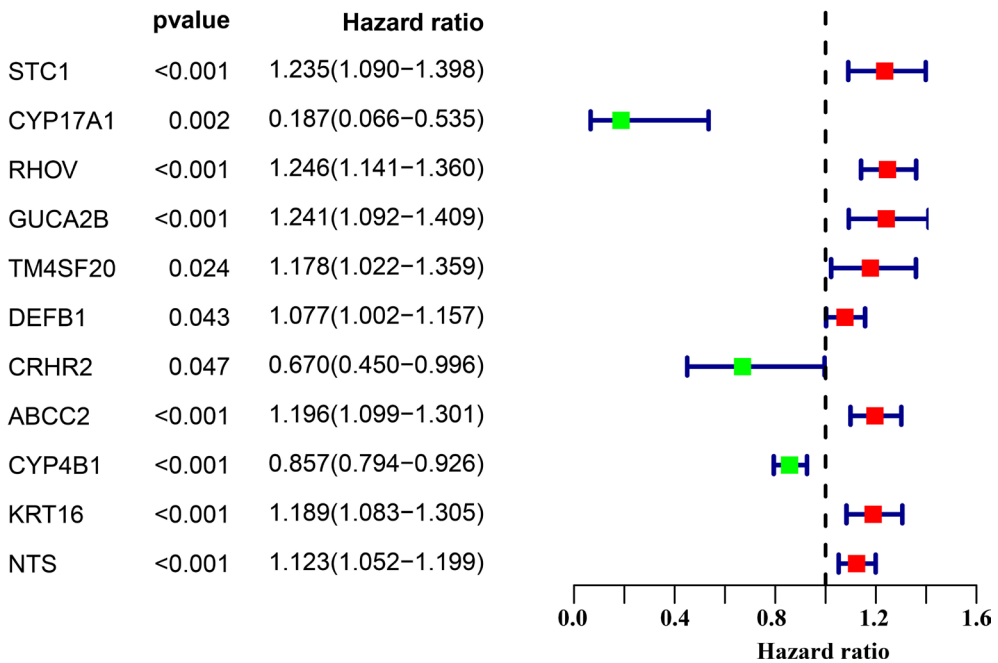


**Figure S2** Lymph node metastasis-related genes in prognosis of LUAD patients using univariate Cox analysis. LUAD, lung adenocarcinoma.


https://dx.doi.org/10.21037/jtd-23-265

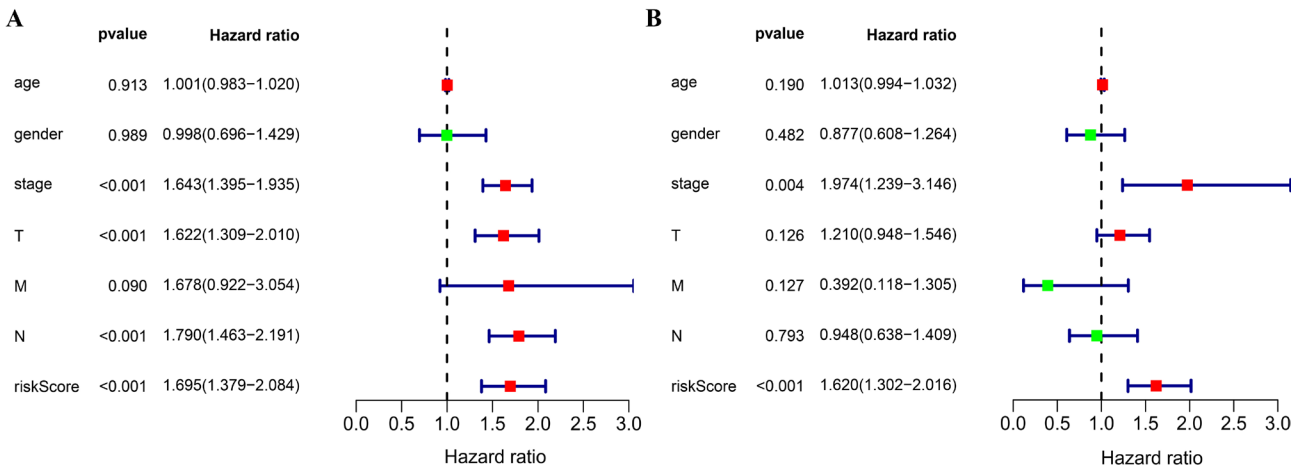


**Figure S3** Clinicopathological characteristic factors of prognosis in LUAD using Cox analysis. LUAD, lung adenocarcinoma.


https://dx.doi.org/10.21037/jtd-23-265