



# Identification of Tumor Microenvironment-Related Prognostic lncRNAs in Lung Adenocarcinoma

Ligong Yuan, Feng Li, Shuaibo Wang, Hang Yi, Fang Li\* and Yousheng Mao\*

Department of Thoracic Surgery, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China

## OPEN ACCESS

### Edited by:

Desi Shang,  
Harbin Medical University, China

### Reviewed by:

Yuxuan Song,  
Tianjin Medical University General  
Hospital, China  
Peng Wang,  
Harbin Medical University, China  
Zhe Wang,  
Liaoning Cancer Hospital, China  
Medical University, China

### \*Correspondence:

Yousheng Mao  
maoysherx@qq.com  
Fang Li  
brucelifang@vip.163.com

### Specialty section:

This article was submitted to  
Cancer Genetics,  
a section of the journal  
Frontiers in Oncology

**Received:** 03 June 2021

**Accepted:** 06 July 2021

**Published:** 02 August 2021

### Citation:

Yuan L, Li F, Wang S, Yi H, Li F and  
Mao Y (2021) Identification of Tumor  
Microenvironment-Related Prognostic  
lncRNAs in Lung Adenocarcinoma.  
*Front. Oncol.* 11:719812.  
doi: 10.3389/fonc.2021.719812

**Background:** Lung adenocarcinoma (LUAD) is the most common type of lung cancer and is a severe threat to human health. Although many therapies have been applied to LUAD, the long-term survival rate of patients remains unsatisfactory. We aim to find reliable immune microenvironment-related lncRNA biomarkers to improve LUAD prognosis.

**Methods:** ESTIMATE analysis was performed to evaluate the degree of immune infiltration of each patient in TAGA LUAD cohort. Correlation analysis was used to identify the immune microenvironment-related lncRNAs. Univariate cox regression analysis, LASSO analysis, and Kaplan Meier analysis were used to construct and validate the prognostic model based on microenvironment-related lncRNAs.

**Results:** We obtained 1,178 immune microenvironment-related lncRNAs after correlation analysis. One hundred and eighty of them are independent prognostic lncRNAs. Sixteen key lncRNAs were selected by LASSO method. This lncRNA-based model successfully predicted patients' prognosis in validation cohort, and the risk score was related to pathological stage. Besides, we also found that TP53 had the highest frequency mutation in LUAD, and the mutation of TP53 in the high-risk group, which was identified by our survival model, has a poor prognosis. lncRNA-mRNA co-expression network further suggested that these lncRNAs play a vital role in the prognosis of LUAD.

**Conclusion:** Here, we filtered 16 key lncRNAs, which could predict the survival of LUAD and may be potential biomarkers and therapeutic targets.

**Keywords:** lung adenocarcinoma, lncRNAs, LASSO regression, immune microenvironment, prognostic biomarkers

## INTRODUCTION

The incidence of lung cancer exceeds 2 million each year, of which approximately 1.8 million ultimately die, making it the leading cause of cancer-related deaths worldwide (1). Of these, 85% of lung cancers are diagnosed as non-small cell lung cancer (NSCLC), and 60% of the patients have locoregional advance or distant metastases (2). There are two major subgroups of NSCLC, LUAD

and LUSC. LUAD and LUSC are distinct at the transcriptome level and in terms of cellular control networks (3, 4). In addition, LUAD shows different genetic drivers and different prognostic profiles compared with LUSC (4, 5). Numerous therapeutic clinical trials in NSCLC have shown that LUAD patients showed different responses compared with LUSC (3, 6). This suggests that LUAD and LUSC are different at pathological and molecular levels. Therefore, the development of new and more effective subtype-specific molecules and associated targeted therapies is of great significance for NSCLC. LUAD is the most common type of lung cancer in nonsmokers, although it can occur in smokers. LUAD morphologic types include glandular alveolar, papillary, solid, micropapillary, and invasive mucinous types (7). In addition, it is more common in women than in men and is more likely to occur in younger people and to present in a more advanced stage (8). LUAD has been the most common histological subtype of lung cancer in the last few decades (9). Lung adenocarcinoma (LUAD) is a severe threat to human health, with more than 1 million deaths per year worldwide (1, 10). Although many therapies have been applied to LUAD, the long-term survival rate of patients remains unsatisfactory, with an average 5-year survival rate of 16% (11, 12).

Recently, there is a growing body of opinion that the immune cell plays an essential role in tumor (13). Cancer is often able to evade different components of the immune system, and the immune microenvironment is a critical factor associated with cancer progression (14). Moreover, many studies have shown that immune-related parameters can predict the prognosis of LUAD patients (15, 16). Therefore, we need a reliable immune microenvironment-related biomarkers to assess LUAD prognosis to guide in the therapeutic management.

lncRNAs are transcripts with non-coding potential and have more than 200 nucleotides (17). The current understanding of the function of lncRNA remains largely unclear. lncRNAs may regulate the expression level of genes by post-transcriptional regulation (18). Meanwhile, lncRNAs may further influence tumor cell migration by regulated target genes (17). However, the immune-related lncRNA signature of lung cancer is still not widely used.

Here, we obtained immune microenvironment-related lncRNAs and evaluated the prognostic efficacy through a mass of bioinformatic analysis. We obtained several lncRNAs that could predict LUAD prognosis, and we also established a lncRNA-mRNA co-expression network to investigate the mechanism of these lncRNAs in LUAD.

## MATERIALS AND METHODS

### Gene Expression Data Sets for Lung Cancer

Data from two publicly available data sets were incorporated into our study. The gene expression data, genomic mutation data, and corresponding clinical information of samples from patients with lung adenocarcinoma (LUAD) were downloaded from the TCGA database. We randomly extracted half of the LUAD

samples as training cohort, and the rest of LUAD samples as validation cohort (**Table S1**).

Gene expression microarray of lung cancer (GSE30219, GSE37745, and GSE31210) with corresponding overall survival (OS) data was downloaded from GEO and served as the validation data set. Gene expression data of all three data sets were normalized by Robust Multichip Average (RMA) method using “affy” package in R.

### Identification of Immune-Related lncRNAs

We evaluated tumor immune infiltration of TCGA LUAD training cohort based on ESTIMATE (Estimation of STromal and Immune cells in Malignant Tumor tissues using Expression data) method by R software with package “estimate” (19). Next, we used Pearson correlation analysis to identify immune-related lncRNAs based on ESTIMATE scores in TCGA LUAD training cohort. A total of 1178 immune-related lncRNAs with  $P < 0.001$  and  $|\text{cor}| > 0.2$  were finally identified.

### Identification of Prognosis-Related lncRNAs and Prognosis Model

A univariate cox regression analysis was performed to select OS-related lncRNAs from abovementioned 1178 lncRNAs in the TCGA LUAD training cohort. A total of 180 immune- and prognosis-related lncRNAs were screened out with  $P < 0.05$ . Next, we performed the LASSO analysis to select the crucial variables and to filter potential survival model, and a model consisting of 16 lncRNAs was identified. Ultimately, the risk score of each patient based on 16 lncRNA expression was calculated by the following formula: risk score =  $(\text{exp}_{\text{lncRNA1}} \times \text{CC}_{\text{lncRNA1}}) + (\text{exp}_{\text{lncRNA2}} \times \text{CC}_{\text{lncRNA2}}) + \dots + (\text{exp}_{\text{lncRNA16}} \times \text{CC}_{\text{lncRNA16}})$ . CC is the coefficient calculated by LASSO. Patients were divided into the high- and low-risk groups based on the median value.

### Pathological, Immune, and Genomic Association Analysis for Risk Group

Wilcoxon rank-sum test was performed to calculate the risk scores' difference between clinical pathological stage groups (TNM Classification of Malignant Tumors). The differences in ESTIMATE scores between high- and low-risk groups in LUAD training cohort were calculated by the same method. Then, we proceed the expression analysis of five immune checkpoint, including PDCD1 (code PD-1), CD274 (code PD-L1), CTLA-4, CD47, and BTLA. Wilcoxon rank-sum test was also used for exploring the risk difference of the five immune checkpoint in TCGA LUAD training cohort. We also analyzed genomic mutation status between high-risk group and low-risk group in TCGA LUAD training cohort and sorted genes according to mutation frequency.

### Establish lncRNA-mRNA Co-Expression Network and Functional Analysis

Pearson correlation analysis was performed for constructing the co-expression network of the 16-lncRNAs, the co-expression

relationships with  $P < 0.001$ ,  $\text{cor} > 0.3$  were retained. We calculated the degree in the network for each lncRNA. We visualized the significant co-expression relationships and retained the mRNAs, which have significant co-expression relationships with more than four lncRNAs. A total of 225 mRNAs were retained for subsequent functional analysis.

We performed pathway and process enrichment analysis for 225 mRNAs with many ontology sources, such as KEGG pathway, GO biological processes, reactome gene sets, and canonical pathways through Metascape web-based tool. Parameters are selected as  $P < 0.01$ , the terms' minimum count is set at three, and the enrichment factor  $> 1.5$ .

## Survival Analysis

Kaplan-Meier analysis was performed to calculate the difference in the survival time between high-risk and low-risk patients. Survival analysis was used to assess the difference in the survival time between the two groups.  $P < 0.05$  was regarded as statistically significant.

## RESULT

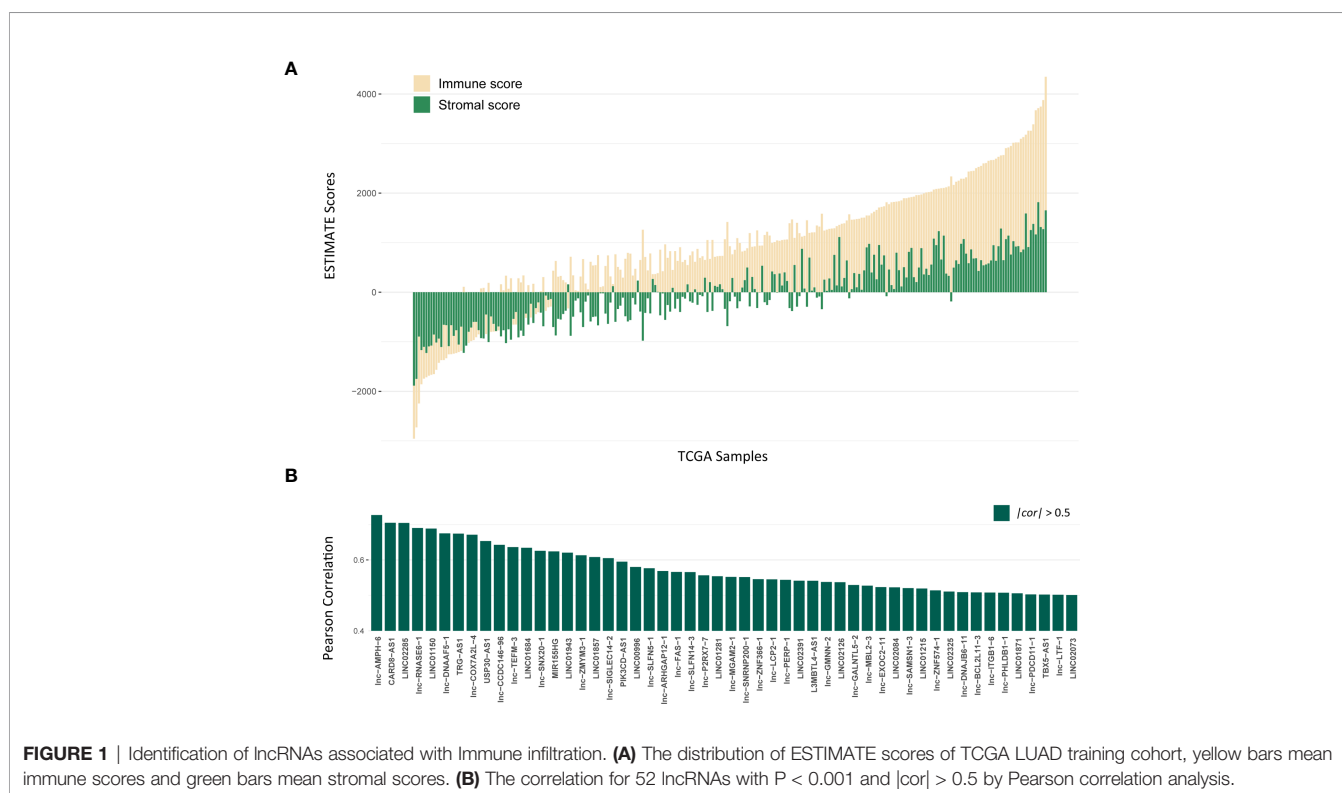
### Identification of lncRNAs Associated With Immune Infiltration

The workflow of this study is shown in **Figure S1**. To explore lncRNAs, which function in tumor immune infiltration, we first used the ESTIMATE method to assess the level of immune cell

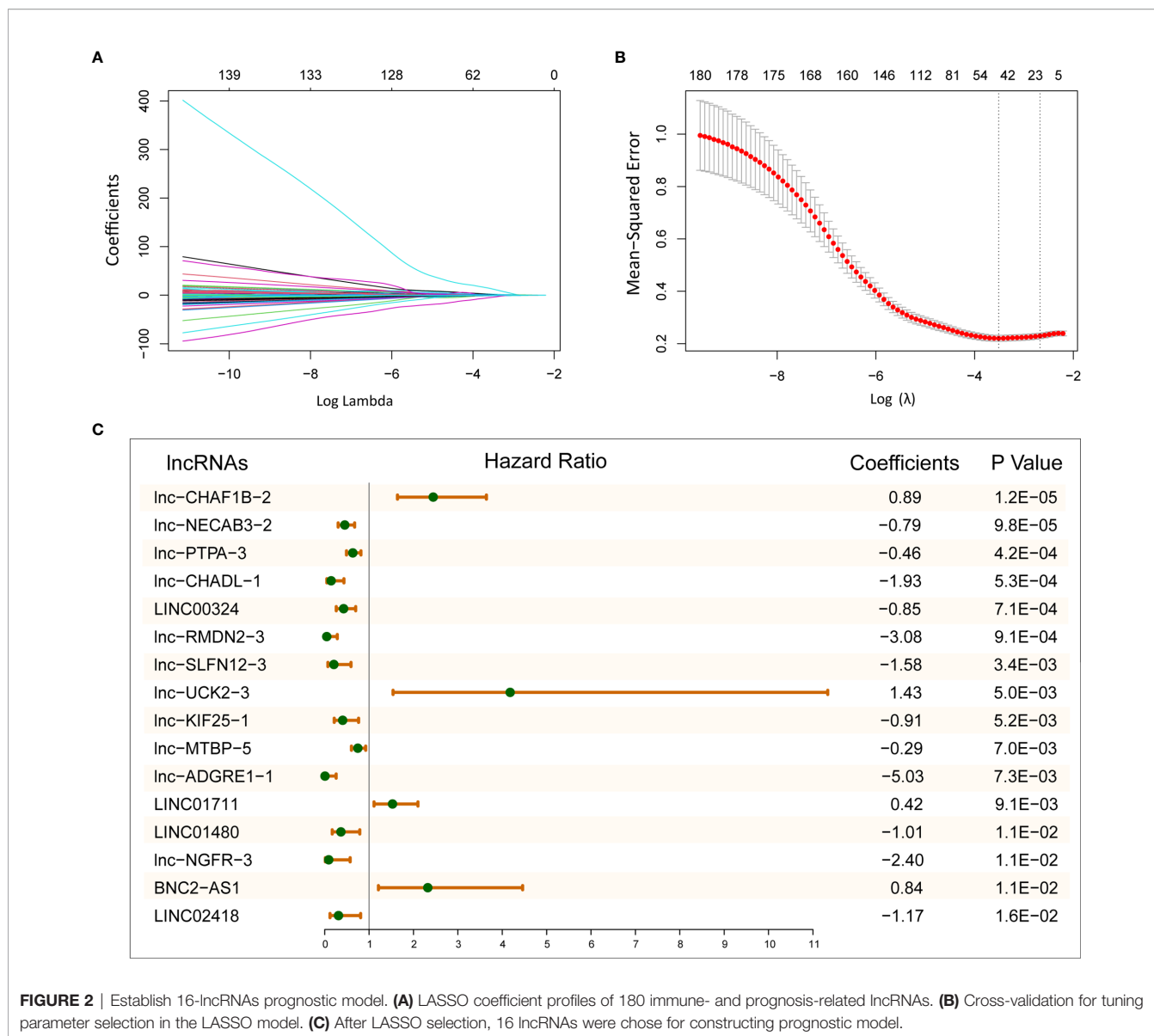
infiltration in the TCGA LUAD training cohort (**Figure 1A**). We analyzed the correlation between ESTIMATE scores and the expression of lncRNAs and identified 1178 lncRNAs significantly associated with ESTIMATE scores. **Figure 1B** showed most significantly associated 52 lncRNAs with  $P < 0.001$  and  $|\text{cor}| > 0.5$  by Pearson correlation analysis.

### Identification of Prognosis-Related lncRNAs and Construction of 16-lncRNAs Prognostic Model

To explore the prognostic efficacy of selected immune-related lncRNAs, we performed univariate Cox proportional regression analysis and obtained 180 immune-related lncRNAs, which were significantly related to OS. Next, LASSO analysis was used to filter the potential survival model. A model with 16 immune- and prognosis-related lncRNAs was constructed (lnc-CHAF1B-2, lnc-NECAB3-2, lnc-PTPA-3, lnc-CHADL-1, LINC00324, lnc-RMDN2-3, lnc-SLFN12-3, lnc-UCK2-3, lnc-KIF25-1, lnc-MTBP-5, lnc-ADGRE1-1, LINC01711, LINC01480, lnc-NGFR-3, BNC2-AS1, and LINC02418) (**Figures 2A, B**). Then, based on the TCGA LUAD training set, we established a predictive model: risk score =  $(0.89 \times \text{lnc-CHAF1B-2 exp}) + (-0.79 \times \text{lnc-NECAB3-2 exp}) + (-0.46 \times \text{lnc-PTPA-3 exp}) + (-1.93 \times \text{lnc-CHADL-1 exp}) + (-0.85 \times \text{LINC00324 exp}) + (-3.08 \times \text{lnc-RMDN2-3 exp}) + (-1.58 \times \text{lnc-SLFN12-3 exp}) + (1.43 \times \text{lnc-UCK2-3 exp}) + (-0.91 \times \text{lnc-KIF25-1 exp}) + (-0.29 \times \text{lnc-MTBP-5 exp}) + (-5.03 \times \text{lnc-ADGRE1-1 exp}) + (0.42 \times \text{LINC01711 exp}) + (-1.01 \times \text{LINC01480 exp}) + (-2.40 \times \text{lnc-NGFR-3 exp}) + (0.84 \times \text{BNC2-AS1 exp}) + (-1.17 \times \text{LINC02418 exp})$  (**Figure 2C**).



**FIGURE 1** | Identification of lncRNAs associated with Immune infiltration. **(A)** The distribution of ESTIMATE scores of TCGA LUAD training cohort, yellow bars mean immune scores and green bars mean stromal scores. **(B)** The correlation for 52 lncRNAs with  $P < 0.001$  and  $|\text{cor}| > 0.5$  by Pearson correlation analysis.



### The Relationship Between Risk Scores With Clinical Pathological Features

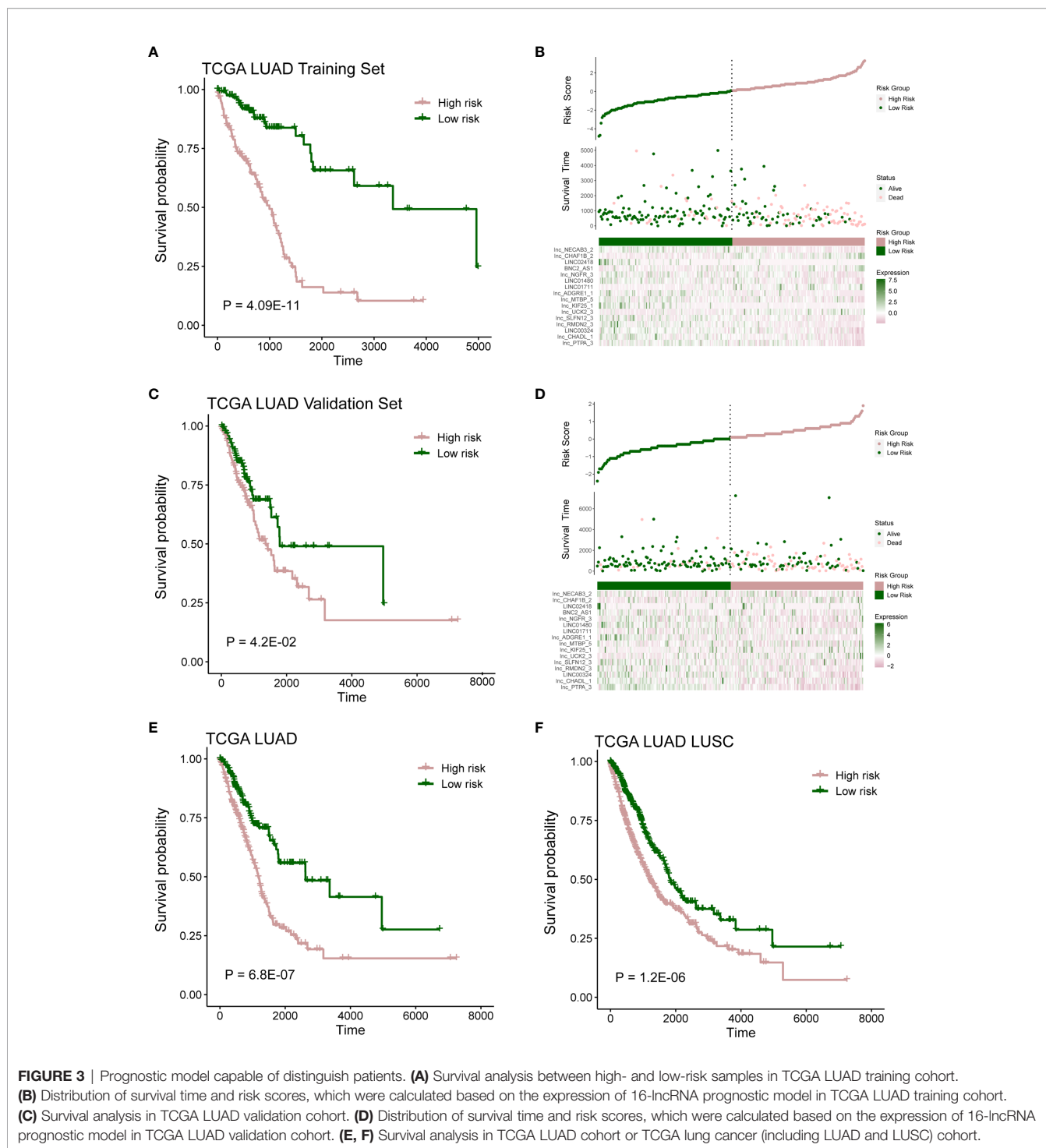
According to the risk score of LASSO analysis, we distinguished the patients into low-risk and high-risk groups. In training cohort, patients in the high-risk group had a poor OS (Figure 3A; P = 4.09E-11; log-rank test), and there are more alive patients in the low-risk group (Figure 3B). In validation cohort, the high-risk group also had a poor prognosis (Figure 3C; P = 4.2E-02; log-rank test), and the alive patients in the high-risk group were less than the low-risk group (Figure 3D). The same observation was also found in the whole TCGA LUAD (training set and validation set) and TCGA lung cancer (LUAD and LUSC) cohorts (Figures 3E, F and Figures S2A, B; P = 6.8E-07, P = 1.2E-02; log-rank test).

Next, further investigation was conducted to determine whether the risk scores could indicate prognosis in different

subgroups of clinical features. In the gender subgroup, high-risk patients had a poor prognosis (Figures 4A, B; P < 0.05; log-rank test). Similarly, in the T subgroups (T2 and T3), M1 subgroup, N subgroups (N0 and N1), and pathological stage (stage I and stage III), high-risk score patients had a significantly poor survival (Figures 4C-H, J; P < 0.05; log-rank test), and stage II subgroup has the same trend but without statistical significance (Figure 4I; P = 9.2E-02; log-rank test).

We also analyzed the difference of risk score among tumor grade groups. The risk scores in stage T2 and T3 were significantly higher than those in stage T1, and the risk scores in stage T3 and T4 were higher than stage T2 (Figure 4K; P < 0.05). Besides, the risk scores in stage N1 and N2 were higher than stage N0 (Figure 4L; P < 0.05). Moreover, the risk scores in stage II, stage III, and stage IV were significantly higher than stage I, and the risk scores in stage III were also higher than stage II (Figure 4M; P < 0.05).



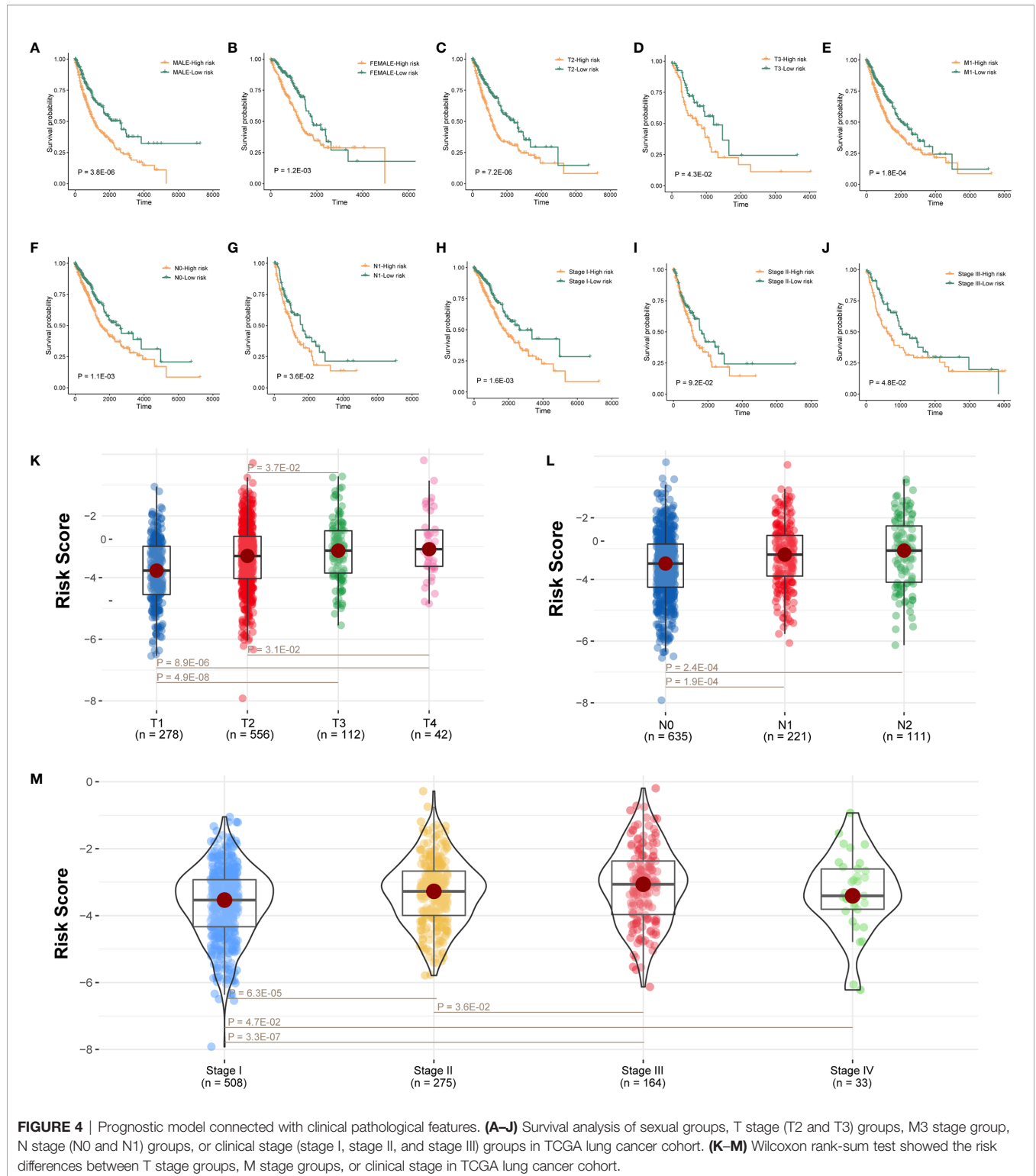


**FIGURE 3 |** Prognostic model capable of distinguish patients. **(A)** Survival analysis between high- and low-risk samples in TCGA LUAD training cohort. **(B)** Distribution of survival time and risk scores, which were calculated based on the expression of 16-lncRNA prognostic model in TCGA LUAD training cohort. **(C)** Survival analysis in TCGA LUAD validation cohort. **(D)** Distribution of survival time and risk scores, which were calculated based on the expression of 16-lncRNA prognostic model in TCGA LUAD validation cohort. **(E, F)** Survival analysis in TCGA LUAD cohort or TCGA lung cancer (including LUAD and LUSC) cohort.

### Immune Infiltration and Genomic Mutation Discrepancy in Different Risk Groups

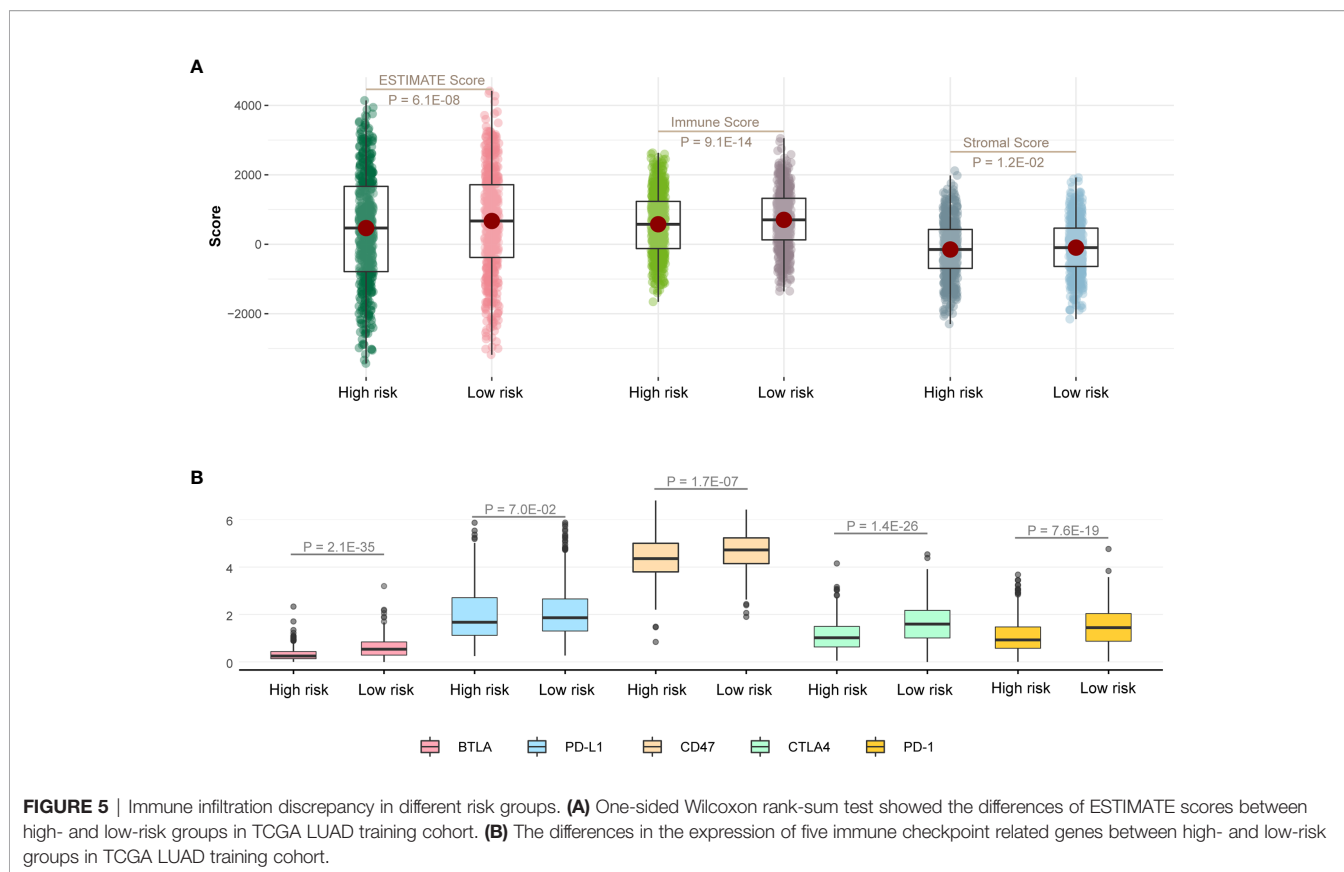
At first, we calculated ESTIMATE scores (immune scores and stromal scores) in high-risk and low-risk groups. We found that the high-risk score group had lower immune scores, stromal scores, and ESTIMATE scores compared with the low-risk score group (**Figure 5A**;  $P < 0.05$ ). Next, we evaluated the expression

difference of five immune checkpoint between high-risk score group and low-risk score group. The expressions of BTLA ( $P = 2.1E-35$ ), CD47 ( $P = 1.7E-07$ ), CTLA4 ( $P = 1.4E-26$ ), and PD-1 ( $P = 7.6E-19$ ) in low-risk group were significantly greater than those in the high-risk group (**Figure 5B**), and expression of PD-L1 has the same trend but no statistical significance (**Figure 5B**;  $P = 7.0E-02$ ).



We then explored genomic mutation status between high- and low-risk groups. **Figures 6A, B** showed the top 20 mutations in TCGA cohort. TP53 is the most frequent mutation gene in both groups. We next investigated relationship between the mutation of TP53 and OS in the high- and low-risk groups

and found the mutation of TP53 indicated poor prognosis in the high-risk group (**Figure 6C**;  $P = 3.1E-02$ ), but there was no distinction in the low-risk group (**Figure 6D**;  $P = 9.8E-01$ ), and there has a trend but no statistical significance in whole TCGA LUAD cohort (**Figure 6E**;  $P = 6.3E-02$ ).



## Hub lncRNAs in Co-Expression Network Were Associated With Prognosis

To identify mRNAs that are associated with 16 lncRNAs in prognostic model, we constructed lncRNAs and mRNAs co-expression network. **Figure 7A** showed degrees of 16-lncRNAs in the co-expression network. We found that three lncRNAs were related to prognosis in top 5 degrees lncRNAs in TCGA LUAD cohort. Patients with high expression of lnc-SLFN12-3, lnc-NECAB3-2, or lnc-CHADL-1 had a significantly better OS than those in the low expression (**Figures 7B–D**;  $P < 0.05$ ; log-rank test). We visualized mRNA-lncRNA co-expression network in **Figure 7E** in which mRNAs has more than four co-expressed lncRNA partners.

## The Function of mRNAs in Co-Expression Network Involved in Immune Pathways

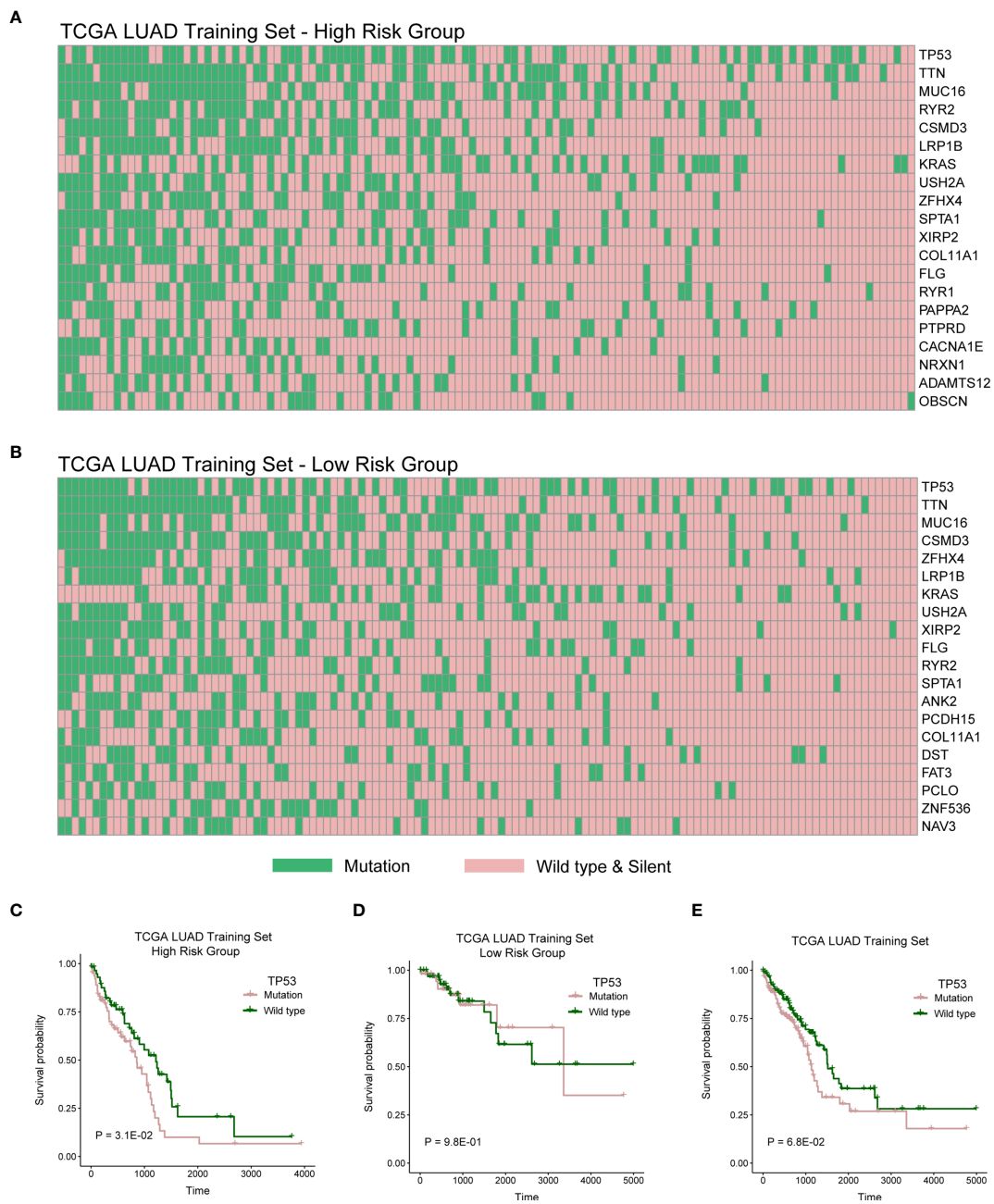
To identify functional processes regulated by co-expression network comprehensively, we performed pathway and process enrichment analysis for 225 mRNAs by Metascape, including 10 pathway resources. Significant terms in biological processes were “lymphocyte activation,” “TYROBP Causal Network,” “leukocyte activation involved in immune response,” “immune response-regulating signaling pathway,” “regulation of cytokine production,” and so on (**Figure 8A**). Network plot showed the subset of enriched terms, only the terms with a similarity  $> 0.3$  are connected by edges (**Figure 8B**).

We found that patients with mRNA BCL2 high expression had a better prognosis than low expression in TCGA LUAD, TCGA lung cancer, GSE37745, and GSE37745 cohorts (**Figures 8C–F**;  $P < 0.05$ ). BCL2 is a member of the Bcl-2 protein family that regulates cell death (apoptosis) by inhibiting (anti-apoptotic) or inducing (pro-apoptotic) apoptosis (20, 21). It was the first apoptosis regulator identified in any organism.

In addition, patients with mRNA IKZF1 high expression had a better prognosis in TCGA LUAD, TCGA lung cancer, GSE37745, GSE37745, and GSE30219 cohorts (**Figure S3A–E**;  $P < 0.05$ ). IKZF1 encodes the DNA-binding protein Ikaros (also known as Ikaros family zinc finger protein 1), which plays an important function in the hematopoietic system and is a regulator of early B-cell, CD4+ T-cell, and other immune cell development. It is closely associated with the development of chronic lymphocytic leukemia (22, 23).

## DISCUSSION

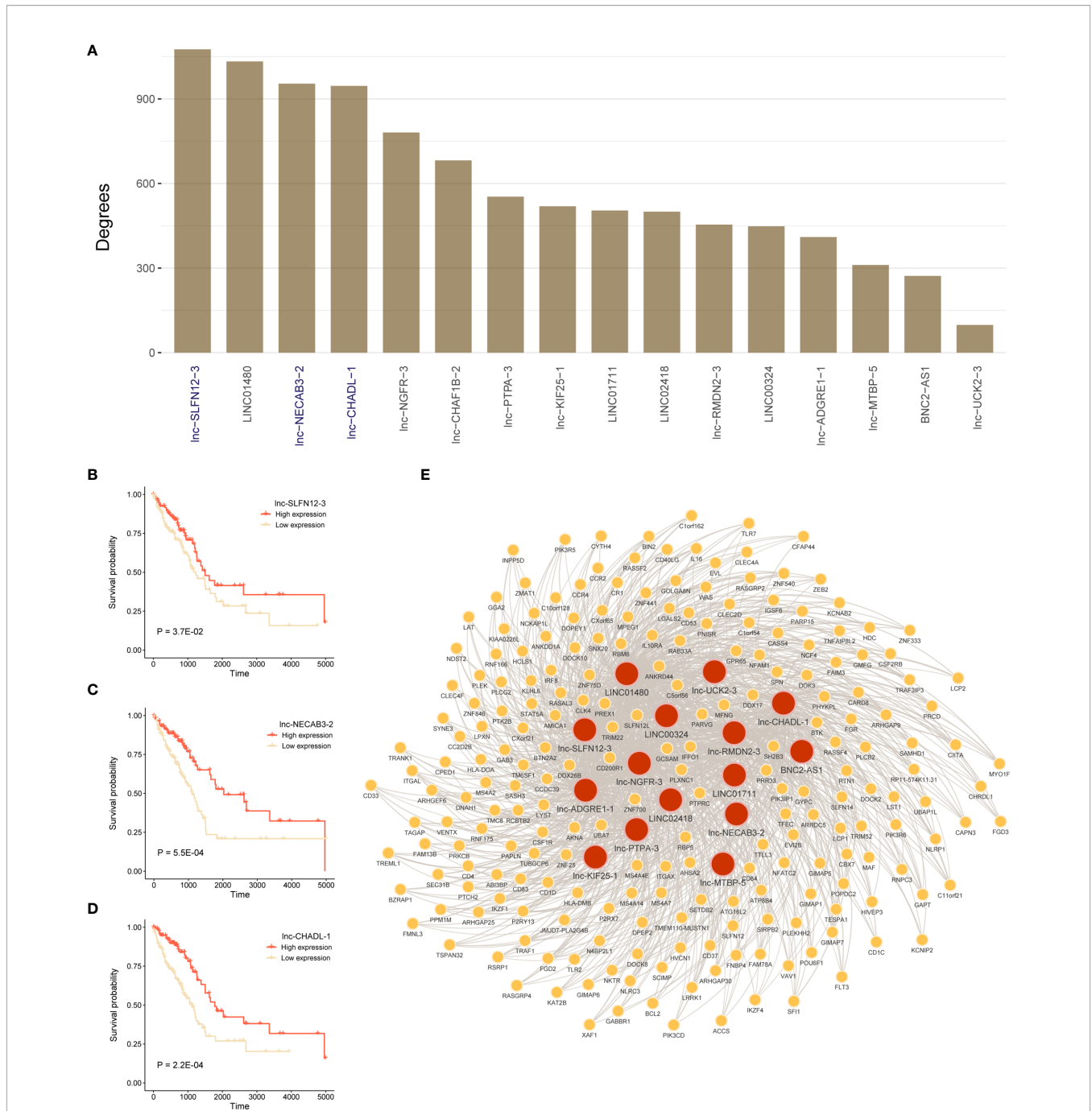
For the past few years, bioinformatics analysis was widely preformed in cancer research (24). The immune microenvironment has also been shown to play an essential part in various cancers. For example, in invasive breast cancer, CD2 was an immune-related prognostic biomarker regulating the tumor microenvironment (25). MNK1/2-eIF4E regulatory axis can affect immunosuppression and metastasis



**FIGURE 6** | High- and low-risk groups with different distribution of mutation. **(A)** The distribution of top 20 genes with high frequency in high-risk group of TCGA LUAD training cohort. **(B)** The distribution of top 20 genes with high frequency in low-risk group of TCGA LUAD training cohort. **(C–E)** Log-rank test was used to assess the difference in OS between TP53 mutation and wild type samples in TCGA LUAD training cohort with high risk, TCGA LUAD training cohort with low risk, or TCGA LUAD training cohort.

in breast cancer (26). lncRNAs have also been reported to influence the prognosis of many tumors. For instance, lncRNA AGAP2-AS1 can enhance lung cancer radiotherapy immunity by regulating the expression of microRNA-296 and NOTCH2 (27). lncRNA TP53TG1 inhibits hepatocellular carcinoma growth and metastasis by affecting the PRDX4/ $\beta$ -catenin axis (28). The lncRNA BBOX1-

AS1 is closely associated with the malignant cellular phenotype of non-small cell lung cancer. It can regulate miR-27a-5p through the ceRNA network, thereby up-regulating MELK to activate the FAK signaling pathway (29). In summary, lncRNA is non-negligible in LUAD. Nevertheless, there are few studies on immune microenvironment-related lncRNAs in LUAD.

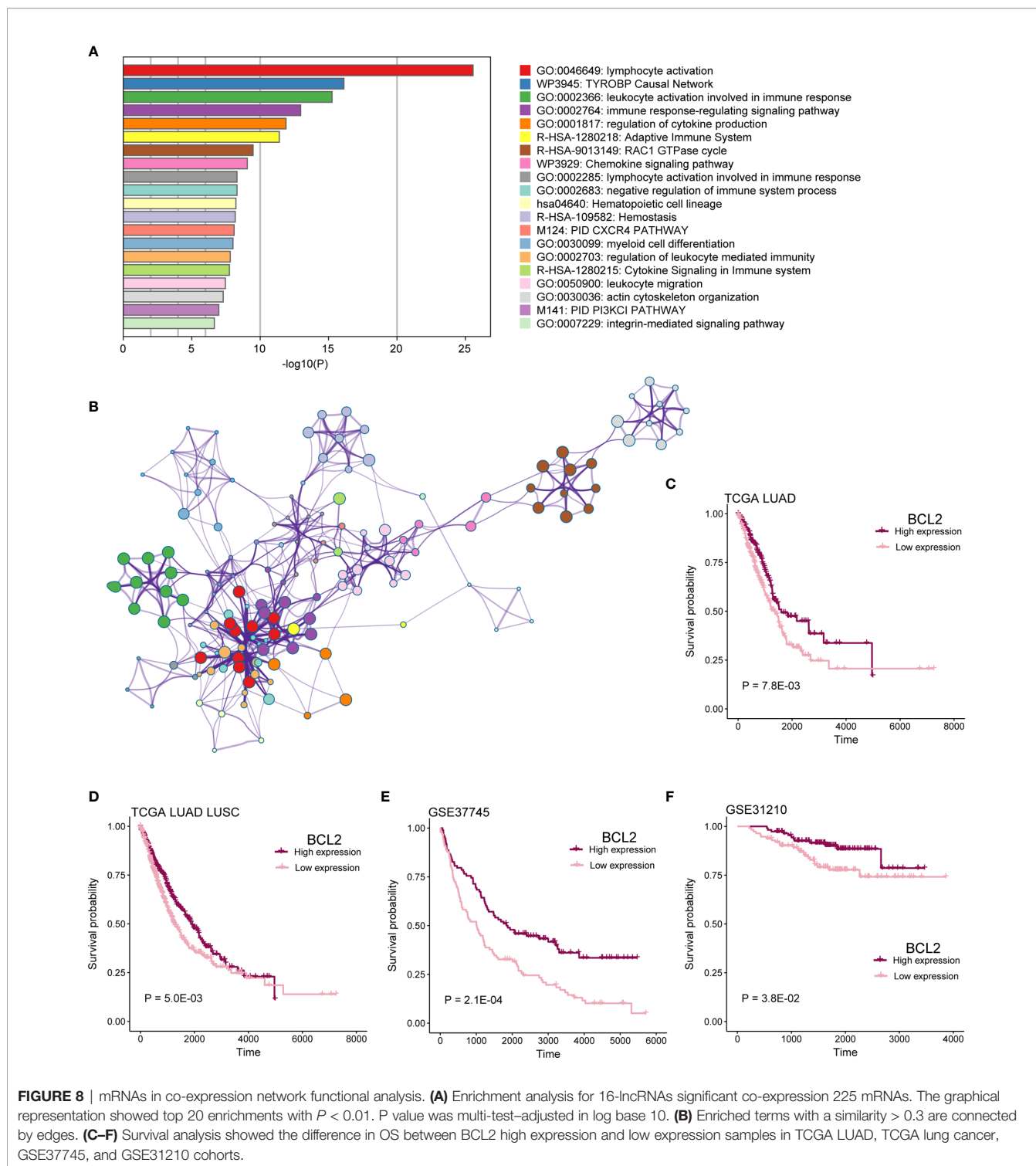


**FIGURE 7 |** Construct lncRNA-mRNA co-expression network and survival analysis for hub lncRNAs. **(A)** Degrees of 16-lncRNAs in lncRNA-mRNA co-expression network. **(B–D)** Log-rank test was used to assess the difference in OS between lnc-SLFN12-3, lnc-NECAB3-2, or lnc-CHADL-1 high expression and low expression samples in TCGA LUAD training cohort. **(E)** Significant co-expression network between 16-lncRNAs and 225 mRNAs (keep significant co-expression relationships more than quarter 16-lncRNAs). Red nodes mean lncRNAs, and yellow nodes mean mRNAs. The width of the edge represents the Pearson correlation.

In this study, we constructed a survival model containing 16 immune microenvironment-related lncRNAs in LUAD through ESTIMATE method and other bioinformatics analysis. This method has been shown to be feasible in cancer research and has been widely used (30, 31). A model with 16 immune- and prognosis-related lncRNAs was constructed and successfully

predicted patients' OS (Figures 2, 3). Of 16 lncRNAs, *lnc-CHAF1B-2*, *lnc-NECAB3-2*, *lnc-PTPA-3*, *lnc-CHADL-1*, *lnc-RMDN2-3*, *lnc-SLFN12-3*, *lnc-UCK2-3*, *lnc-KIF25-1*, *lnc-MTBP-5*, *lnc-ADGRE1-1*, *lnc-NGFR-3* have not been reported in LUAD and other cancers. They may be novel prognostic biomarkers of LUAD. *LINC00324* have been reported to





influence cell proliferation and invasion of several tumors and could regulate the IGF1R to affect non-small cell lung cancer cell invasion (32–34). LINC01711 also have prognostic ability in esophageal squamous cell carcinoma (35). *LINC01480* is a known biomarker of endometrial cancer (36). BNC2-AS1 could influence the proliferation and invasion of gastric cancer

(37). LINC02418 has also been recovered to promote colon cancer progression (38).

Next, further investigation was conducted to determine whether the risk scores could indicate prognosis in different subgroups of clinical features. We found that this survival model was able to distinguish high- and low-risk groups, irrespective of

gender (**Figures 4A, B**). Meanwhile, patients with early stage in TNM stage and pathological stage had higher risk scores and lower survival rates (**Figure 4**), suggesting that our survival model may have an equivalent or better efficacy in early-stage patients. In addition, we found significant differences in the expression of five immune checkpoint-related genes among the high-risk and low-risk groups of the survival model (**Figure 5**), suggesting that the model we constructed could distinguish groups that are more suitable for immunotherapy.

We then explored genomic mutation status between high- and low-risk groups. **Figures 6A, B** showed the top 20 mutations in TCGA cohort. TP53 is the most frequent mutation gene in both groups. We next investigated relationship between the mutation of TP53 and OS in high- and low-risk group and found that the mutation of TP53 indicated poor prognosis in high-risk group. TP53 mutation significantly increased the expression of immune checkpoints and activated T-effector (39–41). In this study, we also explored mutation status between high- and low-risk groups and found that TP53 was the gene with the highest mutation frequency in both risk groups (**Figures 6A, B**). Then, we further explored the influence of TP53 mutation on patients' survival. We found that patients with TP53 mutations had poor survival rates in the high-risk group, suggesting that the survival model also associated with TP53 mutation. Some studies indicated that TP53 mutation significantly increased the expression of immune checkpoints and activated T-effector and associated with poor survival (39–41). Meanwhile, we found lower immune infiltration in the high-risk group (**Figure 5**). These results indicated that patients with TP53 mutations in the high-risk group might have the potential of better immunotherapy efficacy.

Finally, our model-based co-expression network also enriched many immune-related pathways, such as immune response-regulating signaling pathway and cytokine signaling in immune system. It also suggests that lncRNAs in this model are closely related to immunity. These results showed that the biomarkers we obtained have a large research potential in LUAD.

In conclusion, we hope that the results of this study will help identify immune-related potential prognostic lncRNAs and thus

provide new molecular biomarkers for improving the poor prognosis of LUAD.

## DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

## AUTHOR CONTRIBUTIONS

YM contributed to the conception and design of the review. LY performed the research and wrote the manuscript. FeL and SW revised the manuscript. All authors contributed to the article and approved the submitted version.

## FUNDING

This study was supported by Wu Jie Ping Medical Foundation (320.6750.2020-19-31).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2021.719812/full#supplementary-material>

**Supplementary Figure 1** | Workflow of this study.

**Supplementary Figure 2** | Distribution of survival time and risk scores in TCGA cohort. Distribution of survival time and risk scores which was calculated based on the expression of 16-lncRNA prognostic model in **(A)** TCGA LUAD and **(B)** TCGA lung cancer cohorts.

**Supplementary Figure 3** | Survival analysis for IKZF1. **(A–E)** The difference in OS between IKZF1 high expression and low expression samples in TCGA LUAD, TCGA lung cancer, GSE37745, GSE31210, and GSE30219 cohorts.

## REFERENCES

- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* (2018) 68(6):394–424. doi: 10.3322/caac.21492
- Tun AM, Thein KZ, Thein WL, Guevara E. Checkpoint Inhibitors Plus Chemotherapy for First-Line Treatment of Advanced non-Small Cell Lung Cancer: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. *Future Sci OA* (2019) 5(9):FSO421. doi: 10.2144/fsoa-2019-0081
- Liu J, Yang XY, Shi WJ. Identifying Differentially Expressed Genes and Pathways in Two Types of Non-Small Cell Lung Cancer: Adenocarcinoma and Squamous Cell Carcinoma. *Genet Mol Res* (2014) 13(1):95–102. doi: 10.4238/2014.January.8.8
- Herbst RS, Morgensztern D, Boshoff C. The Biology and Management of Non-Small Cell Lung Cancer. *Nature* (2018) 553(7689):446–54. doi: 10.1038/nature25183
- Jamal-Hanjani M, Wilson GA, McGranahan N, Birkbak NJ, Watkins TBK, Veeriah S, et al. Tracking the Evolution of Non-Small-Cell Lung Cancer. *N Engl J Med* (2017) 376(22):2109–21. doi: 10.1056/NEJMoa1616288
- Charkiewicz R, Niklinski J, Claesen J, Sulewska A, Kozłowski M, Michalska-Falkowska A, et al. Gene Expression Signature Differentiates Histology But Not Progression Status of Early-Stage NSCLC. *Transl Oncol* (2017) 10(3):450–8. doi: 10.1016/j.tranon.2017.01.015
- Tang ER, Schreiner AM, Pua BB. Advances in Lung Adenocarcinoma Classification: A Summary of the New International Multidisciplinary Classification System (IASLC/ATS/ERS). *J Thorac Dis* (2014) 6(Suppl 5):S489–501. doi: 10.3978/j.issn.2072-1439.2014.09.12
- de Groot P, Munden RF. Lung Cancer Epidemiology, Risk Factors, and Prevention. *Radiol Clin North Am* (2012) 50(5):863–76. doi: 10.1016/j.rcl.2012.06.006
- Relli V, Trerotola M, Guerra E, Alberti S. Abandoning the Notion of Non-Small Cell Lung Cancer. *Trends Mol Med* (2019) 25(7):585–94. doi: 10.1016/j.molmed.2019.04.012
- Cheng TY, Cramb SM, Baade PD, Youlden DR, Nwogu C, Reid ME. The International Epidemiology of Lung Cancer: Latest Trends, Disparities, and Tumor Characteristics. *J Thorac Oncol* (2016) 11(10):1653–71. doi: 10.1016/j.jtho.2016.05.021
- Ettinger DS, Aisner DL, Wood DE, Akerley W, Bauman J, Chang JY, et al. Nccn Guidelines Insights: Non-Small Cell Lung Cancer, Version 5.2018. *J Natl Compr Canc Netw* (2018) 16(7):807–21. doi: 10.6004/jnccn.2018.0062

12. Miller KD, Nogueira L, Mariotto AB, Rowland JH, Yabroff KR, Alfano CM, et al. Cancer Treatment and Survivorship Statistics, 2019. *CA Cancer J Clin* (2019) 69(5):363–85. doi: 10.3322/caac.21565
13. Hanahan D, Weinberg RA. Hallmarks of Cancer: The Next Generation. *Cell* (2011) 144(5):646–74. doi: 10.1016/j.cell.2011.02.013
14. Garner H, de Visser KE. Immune Crosstalk in Cancer Progression and Metastatic Spread: A Complex Conversation. *Nat Rev Immunol* (2020) 20(8):483–97. doi: 10.1038/s41577-019-0271-z
15. Mao S, Li Y, Lu Z, Che Y, Huang J, Lei Y, et al. Systematic Profiling of Immune Signatures Identifies Prognostic Predictors in Lung Adenocarcinoma. *Cell Oncol (Dordr)* (2020) 43(4):681–94. doi: 10.1007/s13402-020-00515-7
16. Zhang C, Zhang Z, Zhang G, Zhang Z, Luo Y, Wang F, et al. Clinical Significance and Inflammatory Landscapes of a Novel Recurrence-Associated Immune Signature in Early-Stage Lung Adenocarcinoma. *Cancer Lett* (2020) 479:31–41. doi: 10.1016/j.canlet.2020.03.016
17. Ramilowski JA, Yip CW, Agrawal S, Chang JC, Ciani Y, Kulakovskiy IV, et al. Functional Annotation of Human Long Noncoding RNAs Via Molecular Phenotyping. *Genome Res* (2020) 30(7):1060–72. doi: 10.1101/gr.254219.119
18. Huang Q, Yan J, Agami R. Long Non-Coding RNAs in Metastasis. *Cancer Metastasis Rev* (2018) 37(1):75–81. doi: 10.1007/s10555-017-9713-x
19. Yoshihara K, Shahmoradgoli M, Martinez E, Vegesna R, Kim H, Torres-Garcia W, et al. Inferring Tumour Purity and Stromal and Immune Cell Admixture From Expression Data. *Nat Commun* (2013) 4:2612. doi: 10.1038/ncomms3612
20. Tsujimoto Y, Finger LR, Yunis J, Nowell PC, Croce CM. Cloning of the Chromosome Breakpoint of Neoplastic B Cells With the T (14,18) Chromosome Translocation. *Science* (1984) 226(4678):1097–9. doi: 10.1126/science.6093263
21. Cleary ML, Smith SD, Sklar J. Cloning and Structural Analysis of cDNAs for Bcl-2 and a Hybrid bcl-2/immunoglobulin Transcript Resulting From the T(14;18) Translocation. *Cell* (1986) 47(1):19–28. doi: 10.1016/0092-8674(86)90362-4
22. Kastner P, Chan S. Role of Ikaros in T-cell Acute Lymphoblastic Leukemia. *World J Biol Chem* (2011) 2(6):108–14. doi: 10.4331/wjbc.v2.i6.108
23. Oliveira VC, Lacerda MP, Moraes BBM, Gomes CP, Maricato JT, Souza OF, et al. Deregulation of Ikaros Expression in B-1 Cells: New Insights in the Malignant Transformation to Chronic Lymphocytic Leukemia. *J Leukoc Biol* (2019) 106(3):581–94. doi: 10.1002/JLB.MA1118-454R
24. Chen Y, Liao LD, Wu ZY, Yang Q, Guo JC, He JZ, et al. Identification of Key Genes by Integrating DNA Methylation and Next-Generation Transcriptome Sequencing for Esophageal Squamous Cell Carcinoma. *Aging (Albany NY)* (2020) 12(2):1332–65. doi: 10.18632/aging.102686
25. Chen Y, Meng Z, Zhang L, Liu F. Cd2 Is a Novel Immune-Related Prognostic Biomarker of Invasive Breast Carcinoma That Modulates the Tumor Microenvironment. *Front Immunol* (2021) 12:664845. doi: 10.3389/fimmu.2021.664845
26. Guo Q, Bartish M, Goncalves C, Huang F, Smith-Voudouris J, Krisna SS, et al. The MNK1/2-eIF4E Axis Supports Immune Suppression and Metastasis in Postpartum Breast Cancer. *Cancer Res* (2021) 81(14):3876–89. doi: 10.1158/0008-5472.CAN-20-3143
27. Zhang F, Sang Y, Chen D, Wu X, Wang X, Yang W, et al. M2 Macrophage-Derived Exosomal Long Non-Coding RNA Agap2-AS1 Enhances Radiotherapy Immunity in Lung Cancer by Reducing microRNA-296 and Elevating NOTCH2. *Cell Death Dis* (2021) 12(5):467. doi: 10.1038/s41419-021-03700-0
28. Chen B, Lan J, Xiao Y, Liu P, Guo D, Gu Y, et al. Long Noncoding RNA TP53TG1 Suppresses the Growth and Metastasis of Hepatocellular Carcinoma by Regulating the PRDX4/beta-Catenin Pathway. *Cancer Lett* (2021) 513:75–89. doi: 10.1016/j.canlet.2021.04.022
29. Shi J, Yang C, An J, Hao D, Liu C, Liu J, et al. KLF5-Induced BBOX1-AS1 Contributes to Cell Malignant Phenotypes in Non-Small Cell Lung Cancer Via Sponging miR-27a-5p to Up-Regulate MELK and Activate FAK Signaling Pathway. *J Exp Clin Cancer Res* (2021) 40(1):148. doi: 10.1186/s13046-021-01943-5
30. Meng Z, Ren D, Zhang K, Zhao J, Jin X, Wu H. Using ESTIMATE Algorithm to Establish an 8-mRNA Signature Prognosis Prediction System and Identify Immuncyte Infiltration-Related Genes in Pancreatic Adenocarcinoma. *Aging (Albany NY)* (2020) 12(6):5048–70. doi: 10.18632/aging.102931
31. Qu Y, Cheng B, Shao N, Jia Y, Song Q, Tan B, et al. Prognostic Value of Immune-Related Genes in the Tumor Microenvironment of Lung Adenocarcinoma and Lung Squamous Cell Carcinoma. *Aging (Albany NY)* (2020) 12(6):4757–77. doi: 10.18632/aging.102871
32. Xu J, Li Z, Su Q, Zhao J, Ma J. Suppression of Long Noncoding RNA LINC00324 Restricts Cell Proliferation and Invasion of Papillary Thyroid Carcinoma Through Downregulation of TRIM29 Via Upregulating MicroRNA-195-5p. *Aging (Albany NY)* (2020) 12(24):26000–11. doi: 10.18632/aging.202219
33. Chen M, Zhang M, Xie L, Wu S, Zhong Y. LINC00324 Facilitates Cell Proliferation Through Competing for miR2145p in Immature Ovarian Teratocarcinoma. *Int J Mol Med* (2021) 47(1):397–407. doi: 10.3892/ijmm.2020.4800
34. Zhang M, Lin B, Liu Y, Huang T, Chen M, Lian D, et al. LINC00324 Affects Non-Small Cell Lung Cancer Cell Proliferation and Invasion Through Regulation of the miR-139-5p/IGF1R Axis. *Mol Cell Biochem* (2020) 473(1–2):193–202. doi: 10.1007/s11010-020-03819-2
35. Li W, Liu J, Zhao H. Identification of a Nomogram Based on Long Non-Coding RNA to Improve Prognosis Prediction of Esophageal Squamous Cell Carcinoma. *Aging (Albany NY)* (2020) 12(2):1512–26. doi: 10.18632/aging.102697
36. Chen BJ, Byrne FL, Takenaka K, Modesitt SC, Olzomer EM, Mills JD, et al. Transcriptome Landscape of Long Intergenic Non-Coding RNAs in Endometrial Cancer. *Gynecol Oncol* (2017) 147(3):654–62. doi: 10.1016/j.ygyno.2017.10.006
37. Liu C, Yang G, Liu N, Zhou Z, Cao B, Zhou P, et al. Effect of Lncrna BNC2-AS1 on the Proliferation, Migration and Invasion of Gastric Cancer Cells. *Clin Lab* (2018) 64(11):1901–8. doi: 10.7754/Clin.Lab.2018.180537
38. Tian J, Cui P, Li Y, Yao X, Wu X, Wang Z, et al. LINC02418 Promotes Colon Cancer Progression by Suppressing Apoptosis Via Interaction With miR-34b-5p/BCL2 Axis. *Cancer Cell Int* (2020) 20:460. doi: 10.1186/s12935-020-01530-2
39. Dong ZY, Zhong WZ, Zhang XC, Su J, Xie Z, Liu SY, et al. Potential Predictive Value of TP53 and KRAS Mutation Status for Response to PD-1 Blockade Immunotherapy in Lung Adenocarcinoma. *Clin Cancer Res: Off J Am Assoc Cancer Res* (2017) 23(12):3012–24. doi: 10.1158/1078-0432.CCR-16-2554
40. Aisner DL, Sholl LM, Berry LD, Rossi MR, Chen H, Fujimoto J, et al. The Impact of Smoking and TP53 Mutations in Lung Adenocarcinoma Patients With Targetable Mutations-the Lung Cancer Mutation Consortium (Lcm2). *Clin Cancer Res: Off J Am Assoc Cancer Res* (2018) 24(5):1038–47. doi: 10.1158/1078-0432.CCR-17-2289
41. Biton J, Mansuet-Lupo A, Pecuchet N, Alifano M, Ouakrim H, Arrondeau J, et al. Tp53, STK11, and EGFR Mutations Predict Tumor Immune Profile and the Response to Anti-PD-1 in Lung Adenocarcinoma. *Clin Cancer Res: Off J Am Assoc Cancer Res* (2018) 24(22):5710–23. doi: 10.1158/1078-0432.CCR-18-0163

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

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Yuan, Li, Wang, Yi, Li and Mao. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.