Iran. J. Biotechnol. April 2023;21(2): e3168 DOI: 10.30498/ijb.2022.307202.3168

Research Article

Construction of an *STK11* Mutation and Immune-Related Prognostic Prediction Model in Lung Adenocarcinoma

Bo Tang, Xia Zhao, Hongbing Liu, Qingfeng Zhang, Kui Liu, Xiaoyan Yang, Yun Huang*

Department of Cardio-Thoracic Surgery, Zigong Fourth People's Hospital, Zigong, Sichuan 643099, China.

**Corresponding author*: Yun Huang, Department of Cardio-Thoracic Surgery, Zigong Fourth People's Hospital, Sichuan 643099, China. Tel/Fax: +86-13990025297, E-mail: huangyuuuun@163.com

Received: 2021/09/29 ; **Accepted:** 2022/07/06

Background: *STK11* mutation in LUAD affects immune cell infiltration in tumor tissue, and is associated with tumor prognosis.

Objective: This study aimed to construct a *STK11* mutation and immune-related LUAD prognostic model.

Materials and Methods: The mutation frequency of *STK11* in LUAD was queried via cBioPortal in TCGA and PanCancer Atlas databases. The degree of immune infiltration was analyzed by CIBERSORT analysis. DEGs in *STK11*mut and *STK11*wt samples were analyzed. Metascape, GO and KEGG methods were adopted for functional and signaling pathway enrichment analysis of DEGs. Genes related to immune were overlapped with DEGs to acquire immune-related DEGs, whose Cox regression and LASSO analyses were employed to construct prognostic model. Univariate and multivariate Cox regression analyses verified the independence of riskscore and clinical features. A nomogram was established to predict the OS of patients. Additionally, TIMER was introduced to analyze relationship between infiltration abundance of 6 immune cells and expression of feature genes in LUAD.

Results: The mutation frequency of *STK11* in LUAD was 16%, and the degrees of immune cell infiltration were different between the wild-type and mutant *STK11*. DEGs of *STK11* mutated and unmutated LUAD samples were mainly enriched in immune-related biological functions and signaling pathways. Finally, 6 feature genes were obtained, and a prognostic model was established. Riskscore was an independent immuno-related prognostic factor for LUAD. The nomogram diagram was reliable.

Conclusion: Collectively, genes related to *STK11* mutation and immunity were mined from the public database, and a 6-gene prognostic prediction signature was generated.

Keywords: Cox regression analysis, Immunity, LUAD, LASSO analysis, *STK11*, Signature

1. Background

Lung cancer presents the highest mortality worldwide (1). One of its subtypes, non-small cell lung cancer (NSCLC), accounts for 85% among all NSCLC cases (2). Lung adenocarcinoma (LUAD) as a main subtype of NSCLC accounts for the largest proportion (3). Although technologies for cancer diagnosis and treatment have been improved recently, the overall prognosis of LUAD patients is poor (4). Several studies showed that immune cell infiltration can affect LUAD patient's prognosis, and can be used as a prognostic predictor (5-7). In addition, mutation of the tumor suppressor gene *STK11* not only affects immune cell infiltration in LUAD, but also is linked with poor prognosis of LUAD (7-9). Therefore, combined with the above studies, *STK11* mutation and immune infiltration are both implicated in LUAD patient's prognosis.

At present, the establishment of multi-gene prognostic

prediction signatures in LUAD based on high-throughput expression data has become a research hotspot. In 2021, Xinliang Gao *et al* (10), screened ferroptosis-related genes based on The Cancer Genome Atlas (TCGA) database and established a signature in LUAD with good predictive performance through least absolute shrinkage and selection operator (LASSO) analysis. In 2019, Lei Zhang *et al* (11), mined glycolysis-associated gene sets from TCGA database by Gene Set Enrichment Analysis (GSEA). Combined with Cox regression analysis, a model for predicting metastasis and survival time of LUAD based on glycolysis-related genes was constructed. Similarly, Cheng Yue (12) and his colleagues, based on the gene expression data obtained from Gene Expression Omnibus (GEO) and TCGA databases, established a signature of microenvironment-related genes in LUAD in 2019. However, up to now, no study has constructed a prognostic prediction signature in LUAD based on genes associated with *STK11* mutation and immune.

Herein, gene sets related to *STK11* mutation and immune were screened. Subsequently, Cox regression and LASSO analyses were carried out to establish a 6-gene signature. Finally, receiver operating characteristic (ROC) curves, survival curves, Cox regression analysis, and other methods were introduced to verify the prediction performance of the model. Taken together, the constructed 6-gene signature can effectively predict LUAD patient's prognosis.

2. Objective

Immune cell infiltration can affect the prognosis of LUAD patients and can be served as a prognostic predictor. *STK11* mutations in LUAD can affect the degree of immune cell infiltration in tumor tissues and also have an underlying association with LUAD prognosis. However, the molecular mechanism and prognostic significance of *STK11* remains obscure. Hence, the current study attempts to construct a *STK11* mutation and immune-related LUAD prognostic model.

3. Materials and Methods

3.1. Data Acquisition and Bioinformatics Analysis Process

In TCGA and PanCancer Atlas datasets, cBioPortal (https://www.cbioportal.org/) was utilized to query *STK11* mutation frequency in LUAD (13). The mRNA expression data (FPKM), mutation data (VarScan2 Annotation), and patient clinical features of LUAD samples were accessed from TCGA database (https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga) (**Supplementary Table 1**). LUAD sample dataset (GSE72094) was downloaded from the GEO database (https://www.ncbi.nlm.nih.gov/geo/) as a validation set (44 samples were excluded as missing the clinical data), which contains mRNA expression matrix and clinical information (**Supplementary Table 2**). Bioinformatics analyses were conducted based on the above datasets (**Fig. 1**).

3.2. Analysis of Immune Cell Infiltration

TCGA-LUAD mRNA expression profile (FPKM) data were employed to score abundance of tumor immune cell infiltration in LUAD samples based on CIBERSORT algorithm (14). Permutation test was utilized for the reliability analysis of abundance of immune infiltration, and the analysis results of $p<0.05$ were retained. Immune cells with zero abundance of immune cell infiltration in all samples were deleted. The difference in immune infiltration abundance between *STK11*wt and *STK11*mut groups was analyzed by Wilcox test.

3.3. Differential Expression Analysis and Enrichment Analysis

With *STK11*wt sample as a control, differential analysis was performed on the *STK11*mut sample using limma package (15) to screen differentially expressed genes (DEGs; $|logFC|>1$, $FDR<0.01$). Then, the DEGs obtained were analyzed by DAVID (https://david.ncifcrf.gov/) (16) for GO and KEGG analyses (q value <0.05). GO enrichment analysis includes three categories: molecular function (MF), biological process (BP) and cell component (CC). KEGG enrichment analysis revealed the main enrichment signaling pathways.

3.4. Construction of a Signature

Immune-related genes (17) were obtained from Import database (https://www.immport.org/shared/home), then the genes were taken intersection with DEGs to get immune-related DEGs. Subsequently, Metascape (http://metascape.org/gp/index.html#/main/step1) was adopted to perform functional enrichment analysis on immune-related DEGs (18). Univariate Cox regression

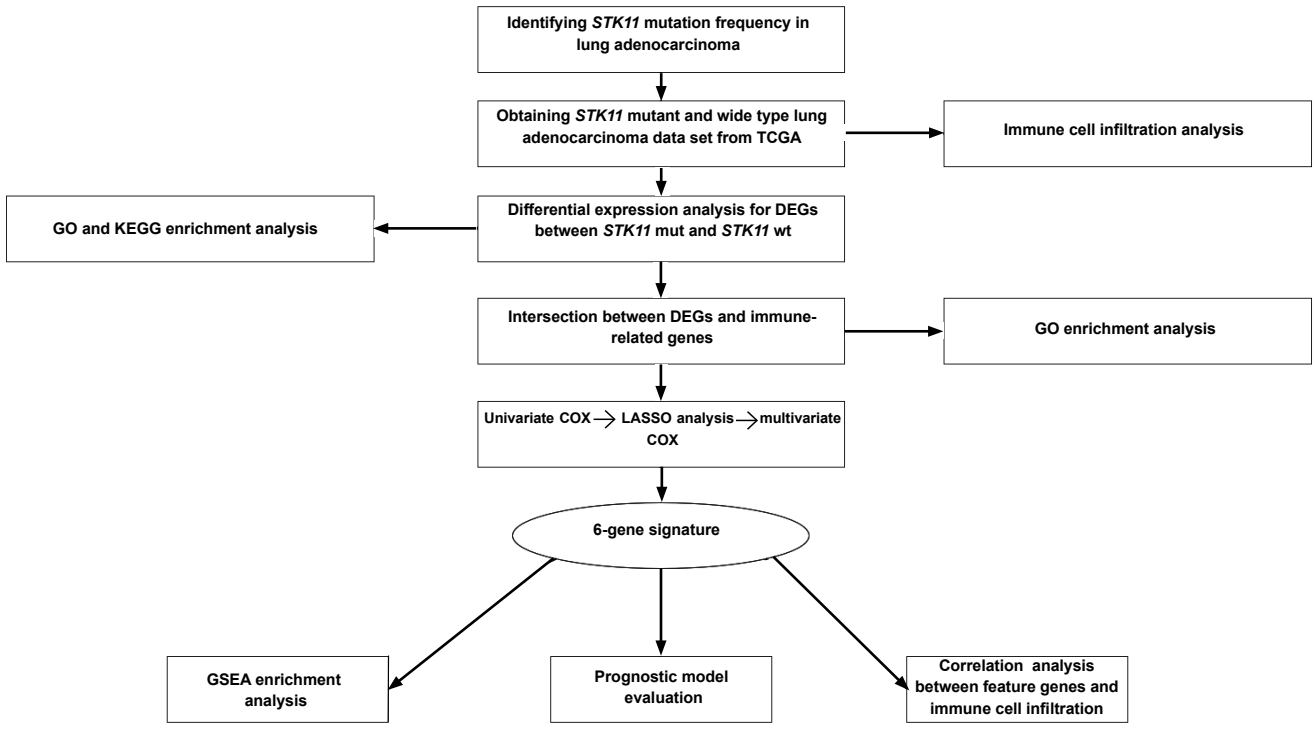

Figure 1. Flow chart of the analysis.

analysis was done using survival package (p<0.01). Then, glmnet (19) was applied to conduct LASSO analysis on the obtained genes. The penalty parameter lambda (λ) was selected by cross-validation method to remove the genes with strong correlation. Survival package was employed for multivariate Cox regression analysis of the genes screened by LASSO analysis, and a signature was established finally.

3.5. Signature Evaluation and GSEA

The ROC curve was plotted using survival ROC package (20) with TCGA-LUAD dataset as training set and GSE72094 dataset as validation set. The area under the ROC curve (AUC) for 1, 2 and 3 years was analyzed. The risk score of TCGA-LUAD samples was computed by the signature, and patients were classified into high- and low-risk groups with the median risk score as cut-off value. Survival curves of patients were plotted using the survival package based on TCGA-LUAD and GSE72094 datasets. GSEA of DEGs in patients of the two risk groups was performed using GSEA software (21).

3.6. Analysis of Tumor Immune Cell Infiltration Based on TIMER Algorithm

Based on TCGA-LUAD dataset, TIMER tool (https://cistrome.shinyapps.io/timer/)(22) was employed to analyze correlation between abundance of tumor immune cell infiltration and patients' overall survival (OS), feature gene expression and OS, the abundance of tumor immune cell infiltration and feature gene expression.

3.7. Independent Analysis of Signature and Construction of the Nomogram

Univariate and multivariate Cox regression analyses were performed to the risk score calculated by the prognostic model and the clinical information (age, gender, TMN, and stage). Through the rms (https://cran.r-project.org/web/packages/rms/index.html) package and combining with different clinical information and risk scores, the nomogram for 3-year and 5-year survival probabilities was plotted. Foreign package (https://cran.r-project.org/web/packages/foreign/index.html) was applied to draw the correction curve of the nomogram.

4. Results

4.1. Frequency of STK11 Mutation in LUAD

Multiple studies indicated that *STK11* is a common mutation site of LUAD and affects immune infiltration of tumor tissue (7,23). We used cBioPortal to detect the mutation frequency of *STK11* in LUAD, and results displayed that the mutation frequency of *STK11* in LUAD was 16% (**Supplementary Fig. 1**).

4.2. Analysis of Immune Cell Infiltration Between STK11wt and STK11mut in LUAD

TCGA-LUAD mutation data were analyzed and then 434 *STK11*wt and 69 *STK11*mut LUAD samples were obtained. Then, CIBERSORT algorithm was applied to score the abundance of immune cell infiltration in TCGA-derived dataset (**Fig. 2A**). Then, 402 *STK11*wt samples and 62 *STK11*mut samples were retained, and the correlation between the abundance of immune cell infiltration was analyzed (**Fig. 2B**). Subsequently, abundance of immune infiltration was compared between *STK11*mut group and *STK11*wt group (T cells CD4 naive abundance was 0 in each sample and therefore was not shown in the results). As demonstrated in the results, compared with *STK11*wt group, *STK11*mut group presents higher T cells follicular helper, Plasma cells, NK cells activated, and Neutrophils infiltration abundance, while lower Macrophages M1, Macrophages M2, and Dendritic cells resting abundance (**Fig. 2C**).

4.3. DEGs and Enrichment Analysis of STK11mut and STK11wt LUAD Samples

Based on TCGA-LUAD dataset, *STK11*wt group was set as the control group, differential analysis was performed on *STK11*mut group. 823 DEGs were screened out, wherein expression of 455 genes was prominently increased and that of 368 genes was down-regulated (**Fig. 3A**). To detect DEGs-related biological functions and signaling pathways, GO and KEGG analyses were carried out. As indicated by GO enrichment analysis result, in the BP module, DEGs presented enrichment in antigen processing, immune response, and presentation of peptide or polysaccharide antigen via MHC class II (**Fig. 3B**). In the CC module, DEGs were mainly gathered in plasma membrane, extracellular space, and cell surface (**Fig. 3C**). In the MF module, DEGs were mainly enriched in calcium ion binding, cytokine activity, and MHC class II receptor activity (**Fig. 3D**). KEGG enrichment analysis result revealed that DEGs showed enrichment in signaling pathways such as rheumatoid arthritis, hematopoietic cell lineage, autoimmune thyroid disease, etc. (**Fig. 3E**). In sum, DEGs of *STK11* mutated and *STK11* unmutated LUAD samples were mainly gathered in immune-related biological functions and signaling pathways.

4.4. Construction of a STK11 Mutation and Immune-Related Prognostic Prediction Signature in LUAD

For the purpose of constructing the *STK11* mutation and immune related signature for LUAD, the following steps were performed. First, 1,811 immune-related genes were overlapped with 827 DEGs obtained above to acquire 117 immune-related DEGs (**Fig. 4A**) (**Supplementary Table 3**). Subsequently, enrichment analysis was completed on immune-related DEGs. Result exhibited that the enrichment of genes laid in regulation of leukocyte activation, antigen processing and presentation and endogenous lipid antigen via MHC class Ib (**Fig. 4B-C**). Hence, it could be inferred that immune-related DEGs played a part in immune regulation. In order to construct a signature related to *STK11* mutation and immune, a univariate Cox regression analysis was first performed on 117 immune-related DEGs, obtaining 17 genes evidently correlated with prognosis ($p<0.01$) (**Supplementary Table 4**). To prevent model overfitting, LASSO Cox regression analysis was conducted on 17 prognostic genes, and 9 candidate feature genes were acquired (**Fig. 4D-E**). Finally, was performed on the results of LASSO analysis, and 6 optimal feature genes (*CCL20, PGC, RAET1L, CD1E, FURIN, KL*) were gained (**Fig. 4A**). A signature was constructed: riskscore=0.09935 * CCL20 - 0.03346 * PGC + 0.04047 * RAET1L - 0.06405 * CD1E + 0.13980 * FURIN - 0.08910 * KL. Among the features, *CCL20, CD1E, FURIN* seemed to significantly correlate to the OS. Therefore, we subsequently analyzed expression of these genes in LUAD, finding that *CCL20* and *FURIN* were evidently upregulated in the LUAD tumor tissues, while *CD1E* was downregulated in the LUAD tissues (**Supplementary Fig. 2**). As our regression analysis indicated *CCL20* and *FURIN* as the risk factors, *CD1E* as a protecting factor, the results of our expression support our prognostic model.

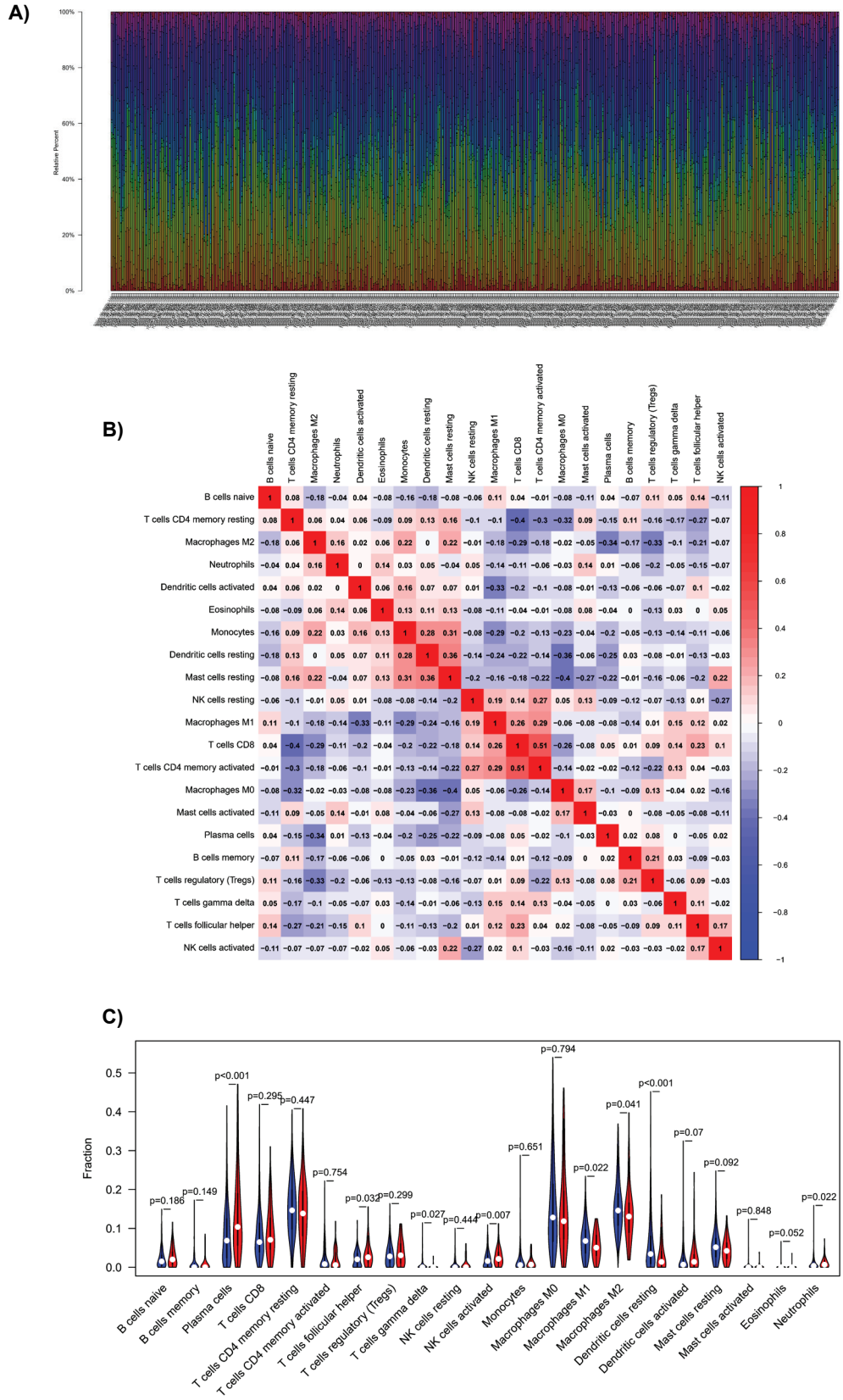

Figure 2. Analysis of immune cell infiltration. A) Bar graph of the abundance of immune cell infiltration in *STK11*wt and *STK11*mut LUAD samples (different colors represent varying types of immune cells). **B)** Heat map of infiltration abundance of immune cells in *STK11*wt group (blue) and *STK11*mut group (pink). **C)** Violin diagram of the abundance of immune cell infiltration in *STK11*mut group (red) and *STK11*wt group (blue).

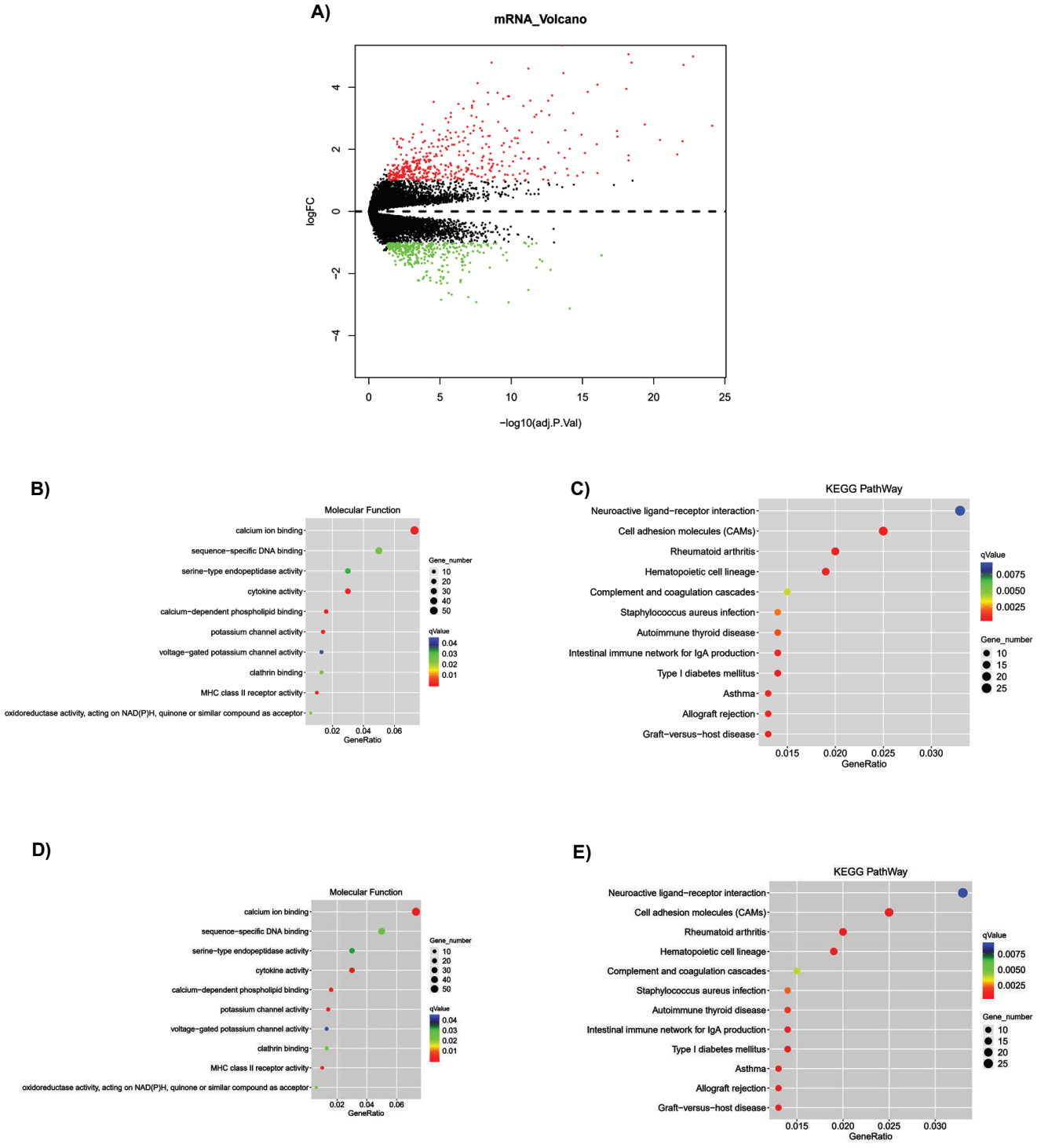

Figure 3. DEGs and enrichment analysis of *STK11*mutand *STK11*wt LUAD samples. A) Volcano map of DEGs between *STK11*mut and *STK11*wt groups (red: remarkably upregulated genes, green: remarkably downregulated genes). **B)** GO enrichment analysis shows that the main BPs of DEGs enrichment (q value decreases as the color tends to red. The larger the circle, the more genes are enriched). **C)** GO enrichment analysis shows the CCs where DEGs are mainly enriched. **D)** GO enrichment analysis exhibits the MFs where DEGs are mainly enriched; E. KEGG enrichment analysis displays the signaling pathways where DEGs are mainly enriched.

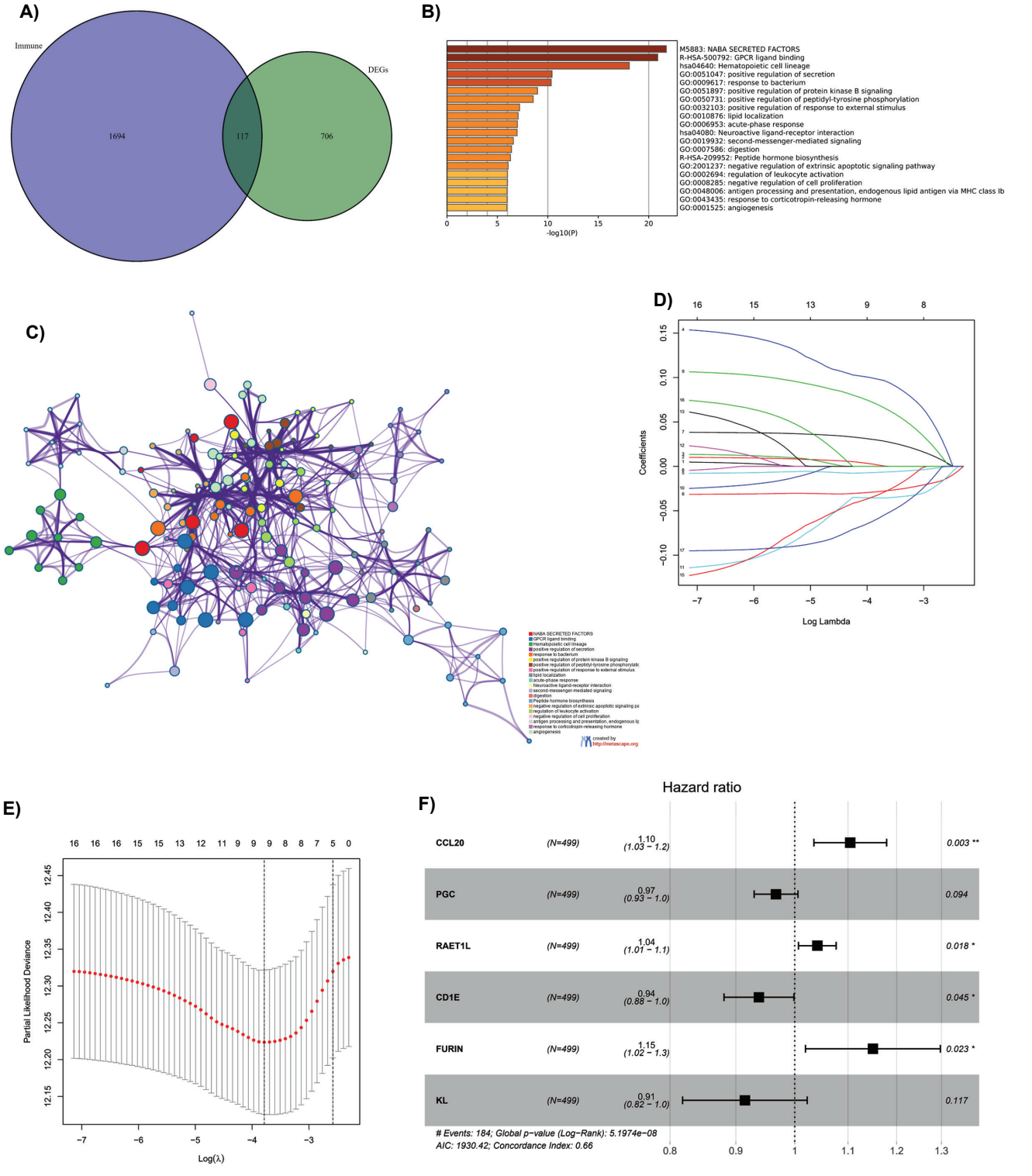

Figure 4. Construction of a signature in LUAD related to *STK11* mutation and immune. A) Venn diagram of immune-related genes and DEGs. **B)** Functional enrichment analysis of immune-related DEGs and sequencing the enriched modules according to the *P* value. **C)** Enrichment analysis network of immune-related DEGs displayed according to the enriched functional modules. Nodes of the same color correspond to function modules of the corresponding color. The larger the size of nodes indicates that more genes are contained. **D)** The LASSO regression model coefficients of 17 prognostic genes, with different color curves representing the variation trajectories of gene coefficients of different characteristics with the penalty parameter λ. **E)** The optimal penalty parameter λ selected from the LASSO regression model. **F)** Forest plot of the relationship between 6 feature genes in multivariate Cox regression model and OS rate of patients.

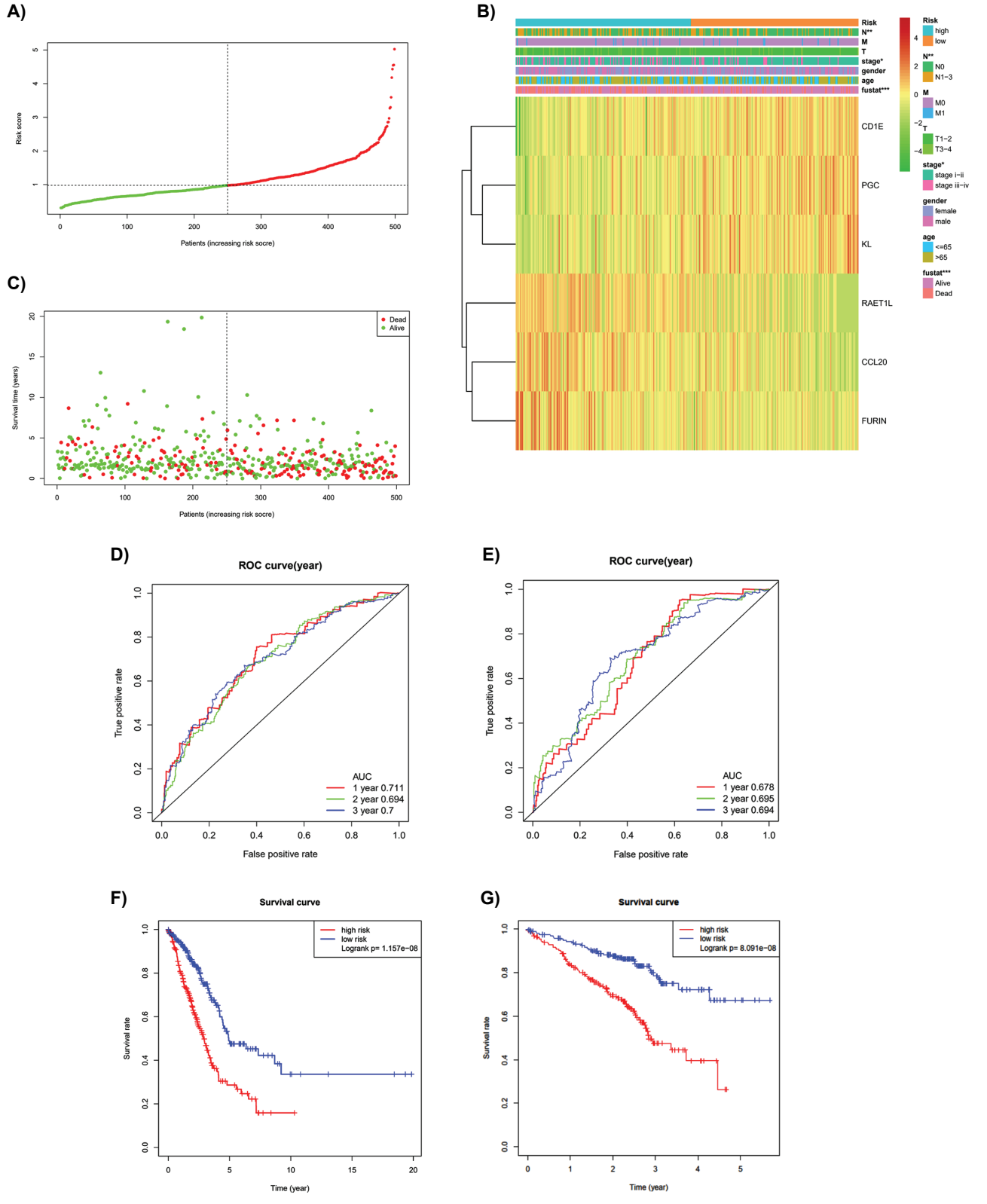

Figure 5. Evaluation of performance of signature. A) Distribution of patient risk scores in TCGA-LUAD dataset. The median value of risk score is utilized as a cutoff line to classify patients into two risk groups. **B)** Heat map of 6 feature genes and clinical features in the two risk groups (*p<0.05, ** p<0.01, and *** p<0.001). **C)** Relationship between patient risk scores and survival in TCGA-LUAD dataset. **(D,E)** ROC analysis of the 6-gene signature in TCGA-LUAD and GSE72094 datasets. **(F,G)** Survival curves of the two risk groups in TCGA-LUAD and GSE72094 datasets.

4.5. Evaluation of the Performance of the Signature

To evaluate performance of the 6-gene signature, riskscore of each sample was firstly c omputed b y the model, and the median value of the riskscore was utilized as the threshold to sort patients into two risk groups (**Fig. 5A**). Subsequently, the relationship between survival time and risk score was analyzed, revealing that patients' survival deteriorated as risk score increased (**Fig. 5C**). Concomitantly, the expression of 6 feature genes in different clinical features and the two risk groups was analyzed. The results suggested that *CD1E, PGC* and *KL* were lowly expressed in the high-risk group while *RAET1L*, *CCL20* and *FURIN* expression was the opposite. In addition, the chi-square test revealed that the distribution of N stage, staging and survival status also had prominent differences between the two risk groups (**Fig. 5B**). Next, performance of the 6-gene signature was evaluated by drawing ROC curves and survival curves. As demonstrated in the results, in the training set, the AUC of 1, 2 and 3 years of OS was 0.711, 0.694 and 0.7, respectively (**Fig. 5D**). In the validation set, the AUC of 1, 2 and 3 years of OS was 0.678, 0.695 and 0.694, respectively (**Fig. 5E**). The survival curves of both TCGA-LUAD and GSE72094 datasets indicated a lower survival rate in the high-risk group (**Fig. 5F-5G**). To examine whether this immune-related prognostic model could be used to assess immunotherapy efficacy, the correlations between the risk score and several immune checkpoint genes (*PD-1, PDL-1, LAG3, CTLA-4*) were analyzed. However, the analyses results did not show any significant associations between them (**Supplementary Fig. 3**), indicating this model was inappropriate to assess immunotherapy efficacy.

4.6. GSEA Results

For the purpose of exploring the signaling pathways of marked enrichment in the two risk groups, GSEA results suggested that the high-risk group was remarkably enriched in the p53 and riboflavin metabolism-related signaling pathways (**Fig. 6A-6B**), indicating that the regulation of these two signaling pathways was prominently different in the two risk groups.

4.7. Analysis of Correlation Between Feature Genes, Immune Cell Infiltration and Prognosis of LUAD

By using TIMER, the survival curves of patients in high and low immune infiltration groups, and high and low expression groups of 6 optimal feature genes were analyzed in the LUAD dataset. The results indicated that patients in the B-cell high infiltration group had a higher survival rate (**Fig. 7A**). Patients in the high *KL* expression group, and low *REAT1L* had a higher survival rate, compared with the corresponding control groups, respectively (**Fig. 7B**). Besides, the correlation between the expression of 6 optimal feature genes and immune cell infiltration degree was analyzed. The results uncovered that *CD1E* level was markedly positively linked with the infiltration degree of CD4+T cells, B cells, macrophages and dendritic cells (**Fig. 7C**).

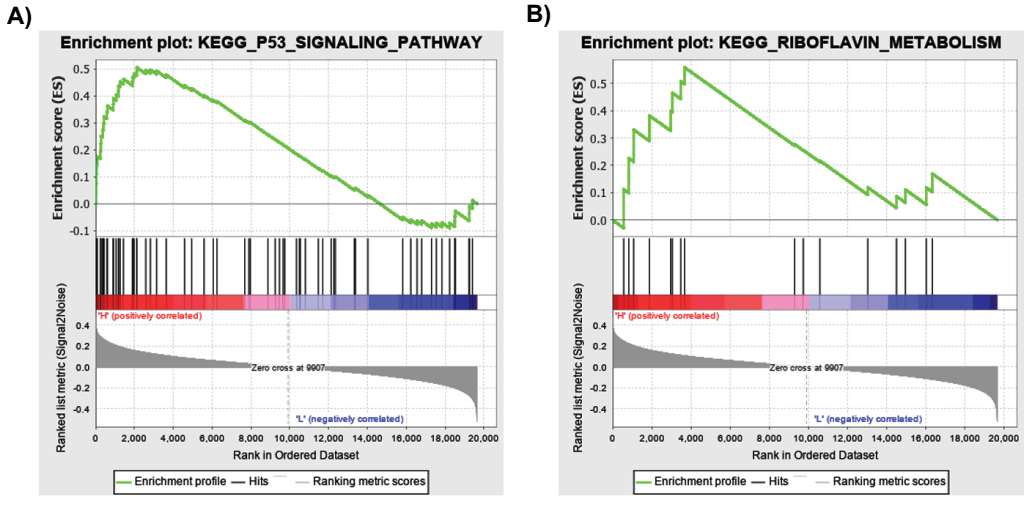

Figure 6. GSEA analyzed pathway enrichment in low- and high-risk groups. A) Enrichment of DEGs in the p53 signaling pathway in the two risk groups. **B)** Enrichment of DEGs in the signaling pathways related to riboflavin metabolism in the two risk groups.

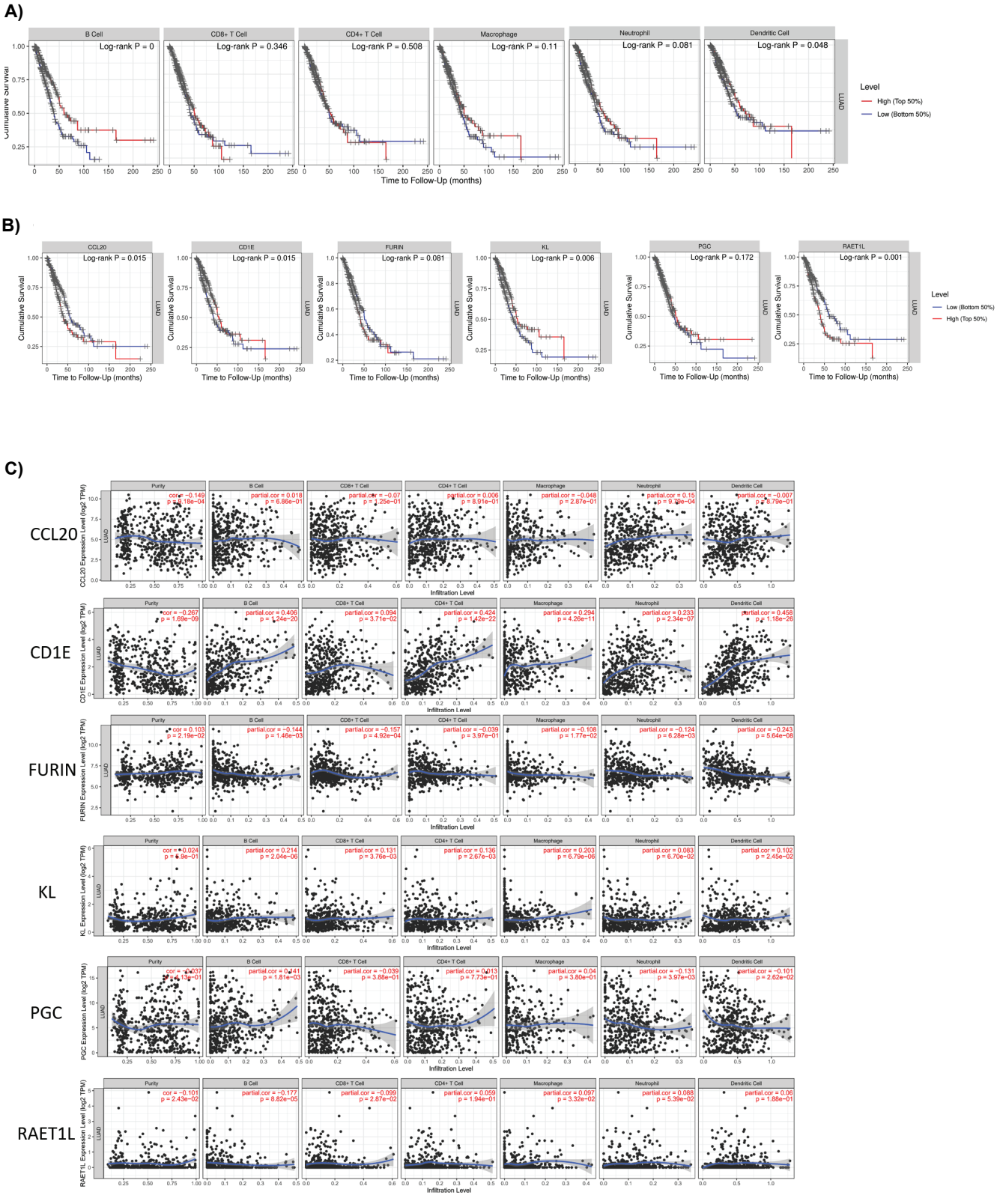

Figure 7. Immune cell infiltration analysis of feature genes. A) Survival curves of the high and low immune cell infiltration groups. **B)** Survival curves of the 6 optimal feature genes in the high and the low expression groups. **C)** Analysis of the correlation between the expression of the optimal feature genes and the degree of tumor immune cell infiltration.

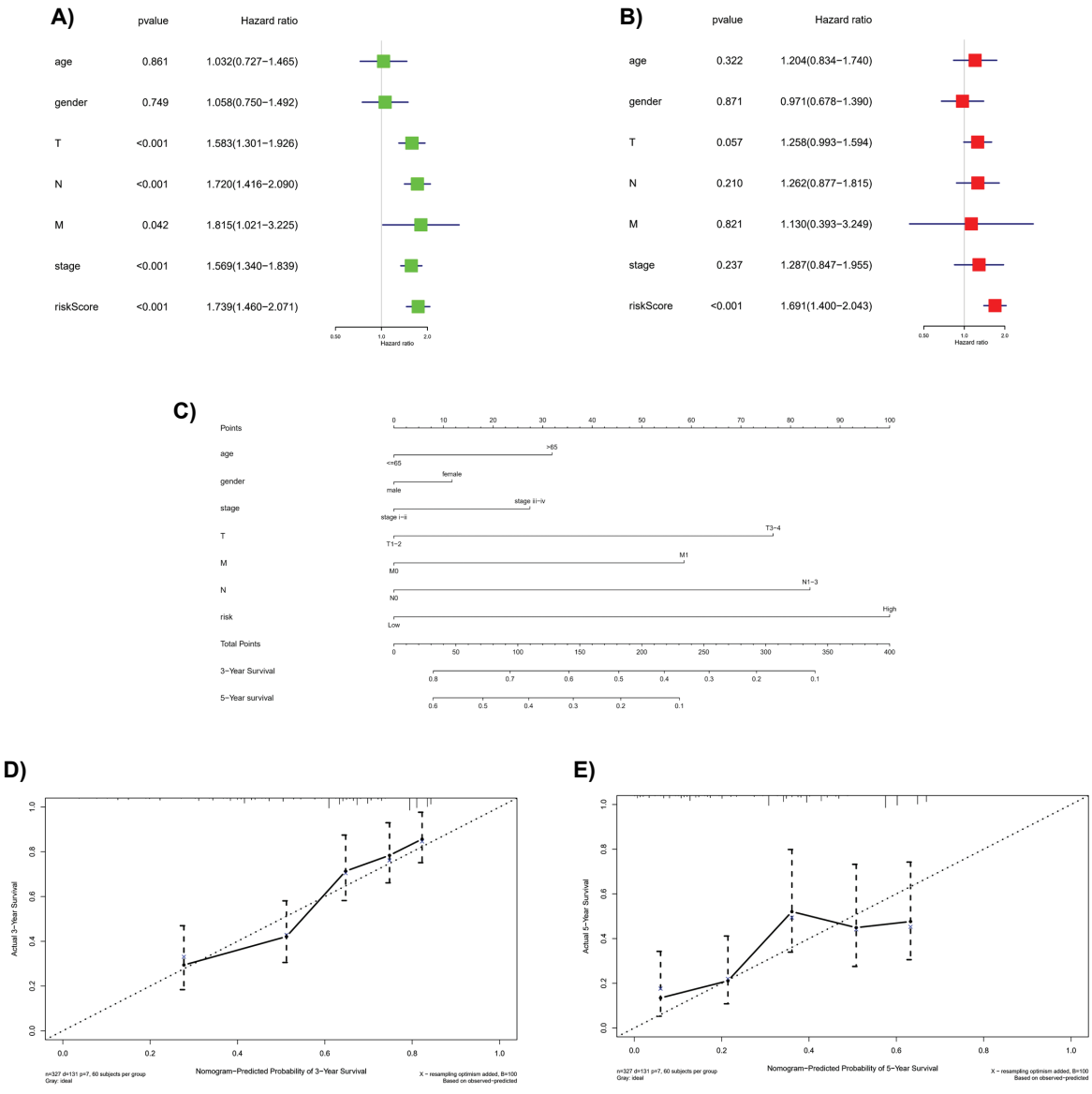

Figure 8. Independence of the signature evaluated with clinical information. (A,B) Univariate and multivariate Cox regression analyses for varying clinical features and risk scores. **C)** Nomogram of predicted 3-year and 5-year survival drawn in combination with different clinical features and risk groups. **(D,E)** Calibration curves of the 3-year and 5-year survival prediction nomogram.

4.8. Independence of the Signature Evaluated with Clinical Information and Construction of the Nomogram

To determine the independence of the signature, univariate Cox regression analysis was conducted on independent prognostic factors like age, gender, TMN, stage, and risk score, uncovering that TNM, stage, and risk score were evidently correlated with LUAD patient's prognosis (**Fig. 8A**). Further, we performed multivariate Cox regression analysis of the above clinical features and risk scores. The result suggested that only risk score was markedly associated with the prognosis of LUAD patients (**Fig. 8B**). Finally, the risk scores and clinical features were combined to draw a nomogram to predict the 3-year and 5-year OS probabilities of patients (**Fig. 8C**). At the same time, the prediction performance of the nomogram was measured by calibration curves, and the results suggested that the predicted 3-year and 5-year survival rates presented a high degree of fit with the actual ones (**Fig. 8D,E**).

5. Discussion

In the past 20 years, with the in-depth understanding of lung cancer genome atlas and the development of new drugs, the treatment regimens for NSCLC have developed from a simple combination of targeted therapy and chemotherapy to a combination of chemotherapy, targeted therapy and immunotherapy based on gene diagnosis (24). Meanwhile, the selection of treatment plan for lung cancer surgery also refers to the prognosis assessment of lung cancer, and clinically it is widely believed that the factors influencing the prognosis of lung cancer include pathological and clinical features of lung cancer (25). In recent years, to evaluate the prognosis of lung cancer more accurately, a number of studies generated prognostic risk model according to on the gene expression profile of lung cancer to evaluate the prognosis of lung cancer (26). Similarly, based on the public database data set of LUAD, this study constructed a 6-gene signature with genes related to *STK11* mutation and immune as features.

STK11 is a commonly mutated gene in LUAD. Based on TCGA data, this study uncovered that the mutation frequency of *STK11* in LUAD was 16%, which was consistent with the results of several published studies (27). In a literature review published in *Nature Reviews Caner*, Ferdinandos Skoulidis and John V. Hepmach(28) concluded that mutation of *STK11*, a tumor suppressor gene, is central to the heterogeneity of lung cancer, but also affect immune cell infiltration in LUAD tumor tissue(29,30). Here, the abundance of immune cell infiltration in *STK11*mut and *STK11*wt LUAD datasets was analyzed, finding that there were marked differences in the abundance of immune cell infiltration between the two groups. Moreover, Weijing Cai *et al.* (31), study in 2018 mentioned that MHC I and II neoantigens often appear in *STK11* mutated LUAD tissue, which was in accordance with the results of functional enrichment analysis of immune-related DEGs in this work.

Qian Song *et al*. (32), designed a 30-gene prognostic risk prediction model in LUAD on the basis of immune-related genes, and revealed that the prediction performance of the model is good, and there are evident differences in immune infiltration degree among the two risk groups predicted by the model. Similarly, this study identified the DEGs of *STK11* mutated and *STK11* non-mutated LUAD samples, selected immune-related DEGs, further screened OS-related feature genes from them, and then constructed a 6-gene signature of LUAD. Compared with the research of Qian Song *et al.*, this study not only employed ROC curve and survival curve to evaluate prediction performance of the model, but also drew a nomogram based on risk grouping and clinical features to predict probability of 3-year and 5-year survival. Therefore, this paper provides more reference data for prognosis prediction of LUAD. Interestingly, in a similar study, survival curves between *STK11* mutation and non-mutation LUAD cases presented differently depending on their immune infiltration degree (33), indicating that *STK11* mutation correlated highly with immune infiltration, being consistent with the conclusion in section 1.1. Based on the understandings above, we first exhibited an *STK11*-immune-related model to effectively assess LUAD prognosis.

Among the 6 genes, *CCL20* and *FURIN* seem to contribute much to LUAD patients` prognosis. CCL20, understood as an inflammatory chemokine, could specifically bind to *CCR6*, promoting cancer progression in various tumors, which consists with our prediction. The diverse pathways triggered by *CCL20-CCR6* interaction were associated with cell migration, invasion, angiogenesis, as well as immune infiltration in cancers (34). Also, furin was considered as a well-understood tumorigenesis factor reported to promote tumor growth in various cancers (35). According to the previous studies, activation or upregulation of furin could mediate several tumor-promoting signaling pathways, like IGF1R/STAT3, Hippo-YAP, and NICD/PTEN, causing aggressive phenotypes in different cancers (36-38).

Although 6 prognostic-related feature genes in LUAD were finally identified and a 6-gene signature was constructed, there are still deficiencies. The analysis demonstrated that *CD1E* gene, as a protective factor that could evidently suppress LUAD, was also markedly positively correlated with the abundance of different immune cell infiltration. Therefore, we speculated that this gene is pivotal in LUAD development. However, the mechanism of this gene in LUAD has not been further investigated. Hence, in the next step, we try to further investigate the regulatory mechanism of *CD1E* in LUAD by combining bioinformatics analysis, molecular experiments and cell experiments.

6. conclusion

To sum up, we mined the genes closely correlated with *STK11* mutation and immune via a series of bioinformatics analysis, disclosed the immunoregulatory function of genes, and established a risk assessment model that can accurately predict the prognosis of LUAD patients. It underlay the exploration of the prognostic value of *STK11* in regulating the immune microenvironment, and provided an essential reference for the diagnosis and treatment of clinical *STK11* mutation in LUAD.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: Cancer j clinic.*2018;**68**:394-424. doi:10.3322/caac.21492
2. Gridelli C, Rossi A, Carbone DP, Guarize J, Karachaliou N., Mok T, *et al.* Non-small-cell lung cancer. *Nat Rev Dis Primers*. 2015;**1**:15009. doi:10.1038/nrdp.2015.9
3. Cancer Genome Atlas Research, N. Comprehensive molecular profiling of lung adenocarcinoma. *Nature*.2014;**511**:543-550. doi:10.1038/nature13385
4. Denisenko T V, Budkevich I N, Zhivotovsky B. Cell death-based treatment of lung adenocarcinoma. *Cell Death Dis.*2018;**9**:117. doi:10.1038/s41419-017-0063-y
5. Guo L, Li X, Liu R, Chen Y, Ren C, Du S, *et al.* TOX correlates with prognosis, immune infiltration, and T cells exhaustion in lung adenocarcinoma. *Cancer Med*. 2020;**9**:6694-6709. doi:10.1002/cam4.3324
6. Zuo S, Wei M, Wang S, Dong J, Wei J. Pan-Cancer Analysis of Immune Cell Infiltration Identifies a Prognostic Immune-Cell Characteristic Score (ICCS) in Lung Adenocarcinoma. *Front Immunol*.2020;**11**:1218. doi:10.3389/fimmu.2020.01218
7. Wang H, Guo J, Shang X, Wang Z. Less immune cell infiltration and worse prognosis after immunotherapy for patients with lung adenocarcinoma who harbored STK11 mutation. *Int Immunopharmacol*. 2020;**84**:106574. doi:10.1016/j.intimp.2020.106574
8. Skoulidis F, Goldberg ME, Greenawalt DM, Hellmann MD, Awad MM, Gainor J F, *et al.* STK11/LKB1 Mutations and PD-1 Inhibitor Resistance in KRAS-Mutant Lung Adenocarcinoma. *Cancer Discov*.2018;**8**:822-835. doi:10.1158/2159-8290.CD-18-0099
9. 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*.2018;**24**:5710-5723. doi:10.1158/1078-0432.CCR-18-0163
10. Gao X, Tang M, Tian S, Li J, Liu W. A ferroptosis-related gene signature predicts overall survival in patients with lung adenocarcinoma. *Future Oncol*.2021;**17**:1533-1544. doi:10.2217/fon-2020-1113
11. Zhang L, Zhang Z, Yu Z. Identification of a novel glycolysis-related gene signature for predicting metastasis and survival in patients with lung adenocarcinoma. *J Transl Med*. 2019;**17**:423. doi:10.1186/s12967-019-02173-2
12. Yue C, Ma H, Zhou Y. Identification of prognostic gene signature associated with microenvironment of lung adenocarcinoma. *PeerJ*. 2019;**7**:e8128. doi:10.7717/peerj.8128
13. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, *et al.* The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov*.2012;**2**:401-404. doi:10.1158/2159-8290.CD-12-0095
14. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, *et al.* Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*.2015;**12**:453-457. doi:10.1038/nmeth.3337
15. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.*2015;**43**:e47. doi:10.1093/nar/gkv007
16. Huang DW, Sherman BT, Tan Q, Kir J, Liu D, Bryant D, *et al.* DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. *Nucleic Acids Res*.2007;**35**:169-175. doi:10.1093/nar/gkm415
17. Bhattacharya S, Dunn P, Thomas CG, Smith B, Schaefer H, Chen J, *et al.* ImmPort, toward repurposing of open access immunological assay data for translational and clinical research. *Sci Data*.2018;**5**:180015.doi:10.1038/sdata.2018.15
18. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun*.2019;**10**:1523. doi:10.1038/s41467-019-09234-6
19. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J Stat Softw*.2010;**33**:1-22
20. Heagerty PJ, Lumley T, Pepe MS. Time-dependent ROC curves for censored survival data and a diagnostic marker. *Biometrics*.2000;**56**:337-344. doi:10.1111/j.0006-341x.2000.00337.x
21. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;**102**:15545-15550. doi:10.1073/pnas.0506580102
22. Li T, Fan J, Wang B, Traugh N , Chen Q, Liu JS, *et al.* TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells. *Cancer Res*.2017;**77**:e108-e110. doi:10.1158/0008-5472.CAN-17-0307
23. Koyama S, Akbay EA, Li YY, Aref AR, Skoulidis F, Herter-Sprie GS, *et al.* STK11/LKB1 Deficiency Promotes Neutrophil Recruitment and Proinflammatory Cytokine Production to Suppress T-cell Activity in the Lung Tumor Microenvironment. *Cancer Res*.2016;**76**:999-1008. doi:10.1158/0008-5472.CAN-15-1439
24. Herbst RS, Morgensztern D, Boshoff C. The biology and management of non-small cell lung cancer. *Nature.* 2018;**553**:446-454. doi:10.1038/nature25183
25. Rami-Porta R, Call S, Dooms C, Obiols C, Sanchez M, Travis WD, *et al.* Lung cancer staging: a concise update. *Eur Respir J.* 2018;**51**. doi:10.1183/13993003.00190-2018
26. Tang H, Wang S, Xiao G, Schiller J, Papadimitrakopoulou V, Minna J, *et al.* Comprehensive evaluation of published gene expression prognostic signatures for biomarker-based lung cancer clinical studies. *Ann Oncol.* 2017;**28**:733-740. doi:10.1093/annonc/mdw683
27. Devarakonda S, Morgensztern D, Govindan R. Genomic

alterations in lung adenocarcinoma. *The Lancet.* Oncology. 2015;**16**:e342-e351.doi:10.1016/s1470 2045(15)000 77-7

28. Skoulidis F, Heymach J V. Co-occurring genomic alterations in non-small-cell lung cancer biology and therapy. *Nat Rev Cancer*. 2019;**19**:495-509. doi:10.1038/s41568-019-0179-8
29. Kadara H, Choi M, Zhang J, Parra ER, Rodriguez-Canales J, Gaffney SG, *et al.* Whole-exome sequencing and immune profiling of early-stage lung adenocarcinoma with fully annotated clinical follow-up. *Ann Oncol*. 2017;**28**:75-82. doi:10.1093/annonc/mdw436
30. Cristescu R, Mogg R, Ayers M, Albright A, Murphy E, Yearley J, *et al.* Pan-tumor genomic biomarkers for PD-1 checkpoint blockade-based immunotherapy. *Science.* 2018;**362**. doi:10.1126/science.aar3593
31. Cai W, Zhou D, Wu W, Tan WL, Wang J, Zhou C, *et al.* MHC class II restricted neoantigen peptides predicted by clonal mutation analysis in lung adenocarcinoma patients: implications on prognostic immunological biomarker and vaccine design. *BMC Genomics*.2018;**19**:582. doi:10.1186/s12864-018-4958-5
32. Song Q, Shang J, Yang Z, Zhang L, Zhang C, Chen J, *et al.* Identification of an immune signature predicting prognosis risk of patients in lung adenocarcinoma. *J Transl Med.2019;***17**:70. doi:10.1186/s12967-019-1824-4
33. Zhang Y, Yang M, Ng DM, Haleem M, Yi T, Hu S, *et al.* Multi-omics Data Analyses Construct TME and Identify the Immune-Related Prognosis Signatures in Human LUAD. *Molecular therapy. Nucleic acids*. 2020;**21**:860-873.doi:10.1016/j.omtn.2020.07.024
34. Kadomoto S, Izumi K, Mizokami A. The CCL20-CCR6 Axis in Cancer Progression. *Int J Mol Sci*. 2020;**21**. doi:10.3390/ijms21155186
35. Jaaks P, Bernasconi M. The proprotein convertase furin in tumour progression. *Int j cancer*. 2017;**141**:654-663. doi:10.1002/ijc.30714
36. Chen C, Gupta P, Parashar D, Nair GG, George J, Geethadevi A, *et al.* ERBB3-induced furin promotes the progression and metastasis of ovarian cancer via the IGF1R/STAT3 signaling axis. *OncoGene*. 2020;**39**:2921-2933. doi:10.1038/s41388-020-1194-7
37. Zhang Y, Zhou M, Wei H, Zhou H, He J, Lu Y, *et al.* Furin promotes epithelial-mesenchymal transition in pancreatic cancer cells via Hippo-YAP pathway. *Int joncolog*. 2017;**50**:1352-1362. doi:10.3892/ijo.2017.3896
38. Li Y, Chu J, Li J, Feng W, Yang F, Wang Y, *et al.* Cancer/testis antigen-Plac1 promotes invasion and metastasis of breast cancer through Furin/NICD/PTEN signaling pathway. *Mol Oncolog*. 2018;**12**:1233-1248. doi:10. 1002/1878-0261.12311