Research article

Construction of a prognostic immune-related lncRNA model and identification of the immune microenvironment in middle- or advanced-stage lung squamous carcinoma patients

Qianqian Xue [a,b,c,1], Yue Wang [a,b,c,1], Qiang Zheng [a,b,c], Lijun Chen [a,b,c], Yan Jin [a,b,c], Xuxia Shen [a,b,c], Yuan Li [a,b,c,*]

[a] Department of Pathology, Fudan University Shanghai Cancer Center, Shanghai, China
[b] Department of Oncology, Shanghai Medical College, Fudan University, Shanghai, China
[c] Institute of Thoracic Oncology, Fudan University, Shanghai, China

ABSTRACT

Background: Globally, non-small-cell lung cancer (NSCLC) has a high incidence, and NSCLC patients have poor prognoses. Lung squamous carcinoma (LUSC) is a major pathological type of NSCLC. LncRNAs play important roles in tumor progression and immune system functions. The aim of this study was to construct a predictive model with immune-related lncRNAs and to assess the immune microenvironment in middle- or advanced-stage LUSC patients.
Methods: RNA sequencing data and corresponding clinical LUSC data were downloaded from The Cancer Genome Atlas. Immune genes were obtained from the Molecular Signatures Database. Immune-related lncRNAs were identified by Pearson correlation analysis in R. The model was constructed using univariate and multivariate Cox regression analyses. Finally, we validated the prognostic immune-related lncRNA model in a cohort from the Fudan University Shanghai Cancer Center.
Results: Our risk model included four immune-related lncRNAs (LINC00944, AL034550.2, AC020907.1 and AC027682.6). Survival analysis revealed that overall and disease-free survival were shorter in the high-risk group than in the low-risk group. Independent prognostic analysis showed that our model could be used as an independent prognostic predictor. The high-risk group was positively associated with CD8+ T cells, B cells, myeloid dendritic cells, macrophages, regulatory T cells (Tregs) and cancer-associated fibroblasts and high expression of PD1 and CTLA4. Additionally, a low-risk score was correlated with lower half maximal inhibitory concentrations (IC_{50}s) of cisplatin, docetaxel, vinorelbine and paclitaxel and a higher IC_{50} of gemcitabine. Gene set enrichment analysis suggested that these lncRNAs may participate in tumor progression and immune processes. Validation with the clinical cancer cohort demonstrated that higher risk scores were associated with a higher, but not statistically significant, likelihood of recurrence.
Conclusion: We established a risk score model including four immune-related lncRNAs. The model accurately predicts the prognosis of middle- or advanced-stage LUSC patients and provides an important reference for individualized treatment.

1. Introduction

Globally, lung cancer is still one of the most commonly diagnosed cancers (11.4%), second only to breast cancer in women (11.7%), and lung cancer remained the leading cause of cancer-related death (18%) in 2020 [1]. There are mainly two pathological categories of lung cancer: non-small-cell lung cancer (NSCLC, approximately 85%) and small cell lung cancer (SCLC, approximately 15%). The World Health Organization classifies NSCLC into three main types: adenocarcinoma, squamous carcinoma and large cell carcinoma, of which squamous carcinoma accounts for approximately 20–30% [2, 3]. Most patients are already at an advanced stage when discovered due to the dearth of classic signs and

symptoms in the early stage, and the 5-year relative survival rate for patients with stage IV disease is only 4% [4]. This outcome shows that the prognoses of middle- or advanced-stage lung squamous carcinoma (LUSC) patients are poor.

In the past decennary, the importance of tumor microenvironment (TME) during the periods of initiation and development of lung cancer has been recognized [5, 6, 7]. The presence of innate immune cells and adaptive immune cells in the TME is conducive to the tumor growth and metastasis process [8]. The introduction of immune checkpoint inhibitors (ICIs), including antibodies against programmed death-1/death-ligand 1 (PD-1/PD-L1) and monoclonal antibodies against cytotoxic T-lymphocyte antigen-4 (CTLA4), ushered in a new era of immunotherapy for lung cancer patients. Many patients with NSCLC have been able to benefit from this therapy [9, 10]. However, in NSCLC, the predictive power of PD-L1, which is currently regarded as the most robust biomarker, is also insufficient in many situations, and the need for further, more complex biomarkers is immense [11]. Therefore, we hoped to select biomarkers appropriate for constructing a prognostic model for LUSC patients with middle- or advanced-stage disease and to assess their immune microenvironment.

Long noncoding RNAs (lncRNAs) are a kind of RNA with lengths longer than 200 bases that cannot be translated into proteins [12]. It has been demonstrated that a comprehensive and complicated position was taken by lncRNAs in regulating cancer initiation and development, and lncRNAs are extremely expected to serve as new biological markers and therapeutic targets [13, 14, 15, 16]. In addition, lncRNAs are essential and crucial in regulating immune gene expression as well as closely associate with the TME [14, 17, 18, 19]. Accordingly, it is worthwhile to construct an immune-related lncRNA model to predict the prognosis of LUSC patients with middle- or advanced-stage disease and to assess their immune microenvironment, which can contribute to the development of individualized treatment plans.

In our study, we comprehensively analyzed immune-related lncRNA data from 175 stage IIB-IV LUSC patients downloaded from The Cancer Genome Atlas (TCGA) to establish a prognostic model and identify immune cell infiltration in tumor tissues.

2. Materials and methods

2.1. Data collection and preprocessing

The RNA sequencing (RNA-seq) data of LUSC patients and the corresponding clinical data were downloaded from The Cancer Genome Atlas (TCGA, https://tcga-data.nci.nih.gov/tcga/) in April 2021. GTF files were downloaded from Ensembl (http://asia.ensembl.org) for gene annotation to classify mRNAs and lncRNAs. We used the "DESeq2" package in R (4.0.5) to screen for differentially expressed mRNAs and lncRNAs (fold change = 1.0, P value <0.05), and the data were normalized by variance-stabilizing transformation (VST) [20]. The list of immune genes was obtained from the Molecular Signatures Database (MSigDB) and cross-referenced with the list of differentially expressed mRNAs. Then, the immune-related lncRNAs were identified by an analysis of the Pearson correlation of immune genes and differentially expressed lncRNAs with R version 4.0.5 (cor-Filter > 0.4 and P value <0.001). Finally, we selected stage IIB-IV patients with survival data >30 days (n = 175) and integrated their clinicopathological characteristics and survival information into the RNA-seq data.

2.2. Development of the immune-related lncRNA risk score model

Immune-related lncRNAs associated with prognosis were screened by univariate Cox regression analysis (P value <0.01). The screened lncRNAs along with clinicopathological features were included in multivariate Cox regression analysis to minimize the Akaike information criterion (AIC) value to obtain the best-performing prognostic model. Cox regression analysis was performed with the package "survival" in R

(4.0.5). Finally, we identified four immune-related lncRNAs. According to the expression level and regression correlation coefficients (β) of these four lncRNAs, risk scores were calculated for each LUSC patient using the following equation: risk score = - 0.136673 × Expression$_{AC020907.1}$ + 0.293362 × Expression$_{AC027682.6}$ + 0.301214 × Expression$_{AL034550.2}$ + 0.120110 × Expression$_{LINC00944}$.

2.3. Identification and evaluation of the prognostic immune-related lncRNA model

The median risk score was used to divide LUSC patients into high- and low-risk groups. We assessed the accuracy of the risk score model with survival analysis and receiver operating characteristic (ROC) curves in R (4.0.5), using the "survival" and "survivalROC" packages. Then, we determined whether the model performance was influenced by other clinical characteristics and whether the model could be considered as an independent prognostic indicator with the "survival" package in R (4.0.5). The scatter plot combined with the heatmap shows the survival status and the differential expression of these four lncRNAs in the two risk subgroups. Moreover, we further explored whether the clinical characteristics and the expression of immune checkpoint molecules differed between the two subgroups. The immune subtypes of patients were determined by methods detailed in Vesteinn Thorsson et al [21]. Statistical analyses were performed with the chi-square test in Stata16 and the Wilcoxon test in R (4.0.5). A P value <0.05 was considered statistically significant. Finally, we applied the "scatterplot3d" and "limma" packages in R (4.0.5) for principal component analysis (PCA) and visualization of the prognostic model results.

2.4. GSEA and immune infiltration

We used gene set enrichment analysis (GSEA 4.1.0) to identify biofunctional differences between the two risk subgroups. We considered pathways with normalized enrichment score (NES) absolute values >1, nominal (NOM) p-val < 0.05, and false discovery rate (FDR) q-val < 0.25 to be significantly enriched. Immune cell infiltration estimation data from LUSC patients assessed by several mainstream methods (TIMER [22], CIBERSORT [23], quanTIseq [24], xCell [25], MCP-counter [26] and EPIC [27]) were downloaded from TIMER 2.0 [28] (http://timer.cistrome.org/). We calculated the correlation between the risk scores and the presence of each type of immune cell identified in the aforementioned methods and displayed their differences across the two subgroups. The Wilcox test and Spearman correlation test were performed in R (4.0.5). A P value <0.05 was considered statistically significant.

2.5. Predicting clinical chemotherapeutic response

Responses to clinical chemotherapeutic agents by these 175 LUSC patients, whose data obtained from the TCGA database, were predicted using the *pRRopheticPredict()* function in the "pRRophetic" package [29, 30], and statistical analysis was performed by t test or Wilcox test with a P value less than 0.05 as the threshold using R (4.0.5). The chemotherapeutic agents were selected according to the National Comprehensive Cancer Network (NCCN) guidelines [31].

2.6. Validation of our prognostic model in the Fudan University Shanghai Cancer Center cohort

We collected data from 14 LUSC patients received surgical resection from 2017 to 2018 at the Fudan University Shanghai Cancer Center. All enrolled patients did not undergo any chemotherapy, radiotherapy, or other treatments prior to surgery. Histopathological examination revealed that all enrolled patients had LUSC with stage IIB-IV disease. We detected the expression level of the four identified lncRNAs by quantitative real-time PCR and plotted a disease-free survival (DFS) curve. All samples were obtained from the Department of Biological repositories of

Table 1. qRT–PCR primer sequences.

LncRNA	Sequence
AC020907.1	Forward: TGCTGAAGTAACGAACTAACCTGAA
	Reverse: GGTGGTGATGGTGACGGTAATG
AC027682.6	Forward: TGCCATTTCAGCCAGCCTCAG
	Reverse: TGCCACCACCACCTTCAGGAA
AL034550.2	Forward: TGCTTCAGATTCAATGGATGCTACT
	Reverse: GCCCGAGGATGGAAACAGACT
LINC00944	Forward: CCTCTTAATCCTCTGTCCTCCATC
	Reverse: CTCTCCAGTGTTATGAAGTTCAAGT

the Fudan University Shanghai Cancer Center. All enrolled patients signed informed consent forms for the use of tissue upon admission. The study was supported by the Ethics Committee of Fudan University Shanghai Cancer Center (ethics number 050432-4-2108*).

2.7. RNA extraction and quantification

We extracted total RNA of these 14 patients with TRIzol Reagent (Ambion), then reverse transcribed it into cDNA with RT SuperMix (+gDNA wiper) kit (Vazyme, R323-01). Quantitative real-time PCR was performed with SYBR qPCR Master Mix (Vazyme, Q712) and run on an Applied Biosystems® 7300Plus Real-Time PCR Instrument. Finally, we selected β-actin as an endogenous reference genes to calculate the relative expression level of the four lncRNAs in each sample using the $2^{-\triangle Ct}$ method. All PCR primers were directly synthesized (Sangon Biotech, Shanghai), and the sequences are shown in Table 1.

3. Results

3.1. Acquisition of immune-related lncRNAs

We analyzed the mRNA and lncRNA expression profiles of 502 LUSC samples and 49 normal samples based on set thresholds (fold change >1 and P value <0.05) and obtained 7148 and 4748 differentially expressed genes (DEGs), respectively (Figure 1 and Supplementary Data 1). Subsequently, we downloaded the "immune system process" and "immune response" gene sets from MSigDB and cross-referenced them with the list of differentially expressed mRNAs. Immune-related genes were identified. Finally, we performed Pearson correlation analysis of differentially expressed lncRNAs and immune-related genes (Pearson correction

coefficient >0.4 and P value <0.001), and obtained 755 immune-related lncRNAs.

3.2. Construction and evaluation of an immune-related lncRNA risk score model

We selected stage IIB-IV LUSC patients (n = 175). A total of 8 immune-related lncRNAs were significantly associated with patient prognosis by univariate Cox regression analysis (P value <0.01), and the results are presented in a forest plot (Figure 2A). By including these 8 lncRNAs and patient clinicopathological features in the multivariate Cox regression analysis, a risk model of 4 immune-related lncRNAs was eventually established (Table 2). We calculated the risk score of each LUSC patient according to the following formula: risk score = - 0.136673 × $Expression_{AC020907.1}$ + 0.293362 × $Expression_{AC027682.6}$ + 0.301214 × $Expression_{AL034550.2}$ + 0.120110 × $Expression_{LINC00944}$. We divided the samples into a high-risk group (n = 87) and a low-risk group (n = 88) based on the median risk score.

Not surprisingly, the overall survival (OS) and DFS of patients in the high-risk group were significantly lower than those in the low-risk group (Figure 2B-D). Survival analysis revealed that the survival rate of patients in the high-risk group was 69% after 1 year, 41% after 3 years, and 26% after 5 years, whereas the survival rate of patients in the low-risk group was 88% after 1 year, 60% after 3 years, and 56% after 5 years. We ranked the risk scores of these 175 LUSC patients by combining the survival outcomes and expression of these four lncRNAs (Figure 2E). The findings demonstrated that the expression of AC027682.6, LINC00944 and AL034550.2 was upregulated in the high-risk group, while that of AC020907.1 was downregulated, and the mortality rate rose with the increasing risk score.

To determine whether the model was independent of other clinical factors, such as age, sex, T stage, N stage, M stage and American Joint Committee on Cancer (AJCC) stage, we performed univariate and multivariate independent prognostic analyses. The outcomes (Figure 2F, G) suggested that only the risk score was correlated with OS, with statistically significant differences (P value <0.001). Therefore, the signature of four immune-related lncRNAs we identified can be used as an independent prognostic predictive factor.

The area under the ROC curve (AUC) was used to evaluate the accuracy of prediction. As shown in Figure 2H and Figure 2I, the signature of four immune-related lncRNAs we identified performed best (AUC = 0.737) compared with the predictive performances of other clinical characteristics, such as age, sex, T stage, N stage, M stage and AJCC stage.
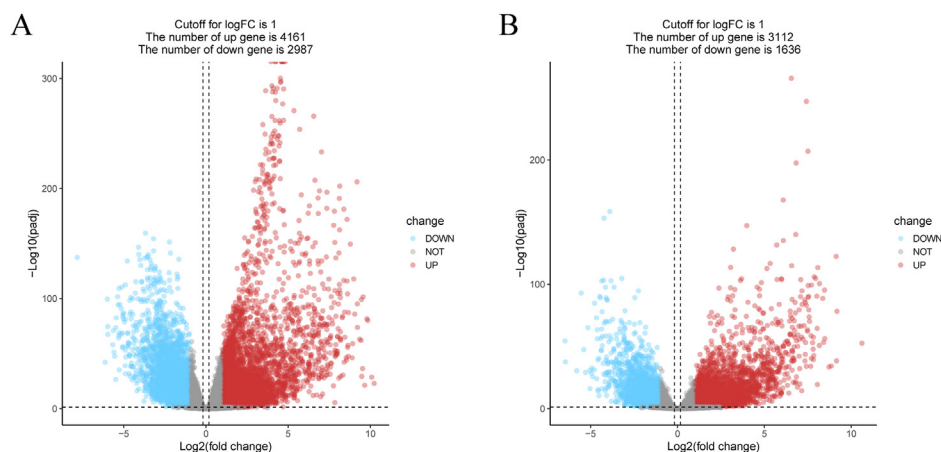
Figure 1. Volcano plots of differentially expressed mRNAs **(A)** and lncRNAs **(B)** in LUSC versus normal samples from the TCGA database. Red dots are upregulated genes, blue dots are downregulated genes, and gray dots are no significantly different genes.
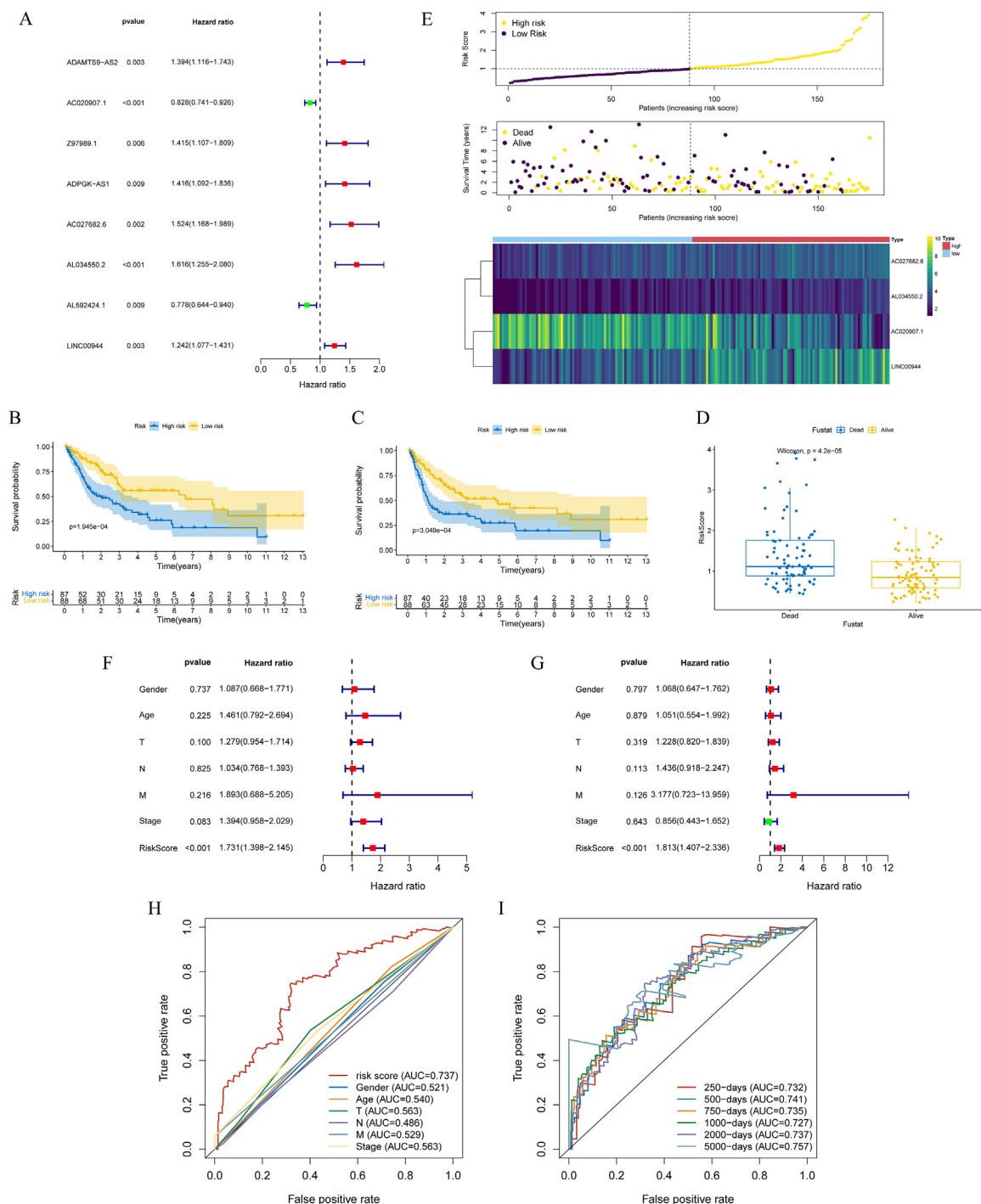
Figure 2. A Forest plot of 8 immune-related lncRNAs associated with prognosis derived from univariate Cox regression analysis. Red and green indicate risk factors and protective factors, respectively. **B-D** Survival analysis showed that the higher the risk score was, the worse the prognosis. **(B)** Kaplan–Meier overall survival (OS) curve (P = 0.00019). **(C)** Kaplan–Meier disease-free survival (DFS) curve (P = 0.00030). **(D)** Scatterplot of survival status and risk score (P = 0.00004). **E** The risk score model of these 4 immune-related lncRNAs in LUSC patients. Risk score (top); survival status (middle); heatmap (bottom). **F, G** Forest plots of univariate independent prognostic analysis **(F)** and multivariate independent prognostic analysis **(G)**. Red and green indicate risk factors and protective factors, respectively. **H, I** ROC curve analysis. **(H)** The area under the ROC curve (AUC) indicated the highest accuracy of our risk model compared to those of other clinical characteristics. **(I)** Time-ROC curve analysis of our risk model at 250, 500, 750, 1000, 2000, and 5000 days.

Time-ROC curve analysis suggested a robust predictive effect of the signature we identified.

3.3. Analysis of clinical features between the different risk subgroups

A series of chi-square tests were carried out to explore the differences in clinical features between the different risk subgroups,

and the results were presented in Figure 3. There were more females and higher mortality in the high-risk group than in the low-risk group, with statistically significant differences (P value <0.05). Regarding immune subtypes, the high-risk group mainly exhibited C1 (wound healing) and C2 (IFN-γ dominant) types, whereas the low-risk group was more likely to exhibit the C1 (wound healing) type (P value <0.05).

Table 2. The coefficients, hazard ratios (HRs) and P values of four immune-related lncRNAs that contributed to the risk model.

LncRNAs	Coef	HR (95% CI)	P value
AC020907.1	-0.136673	0.872256 (0.779413–0.976158)	0.0173
AC027682.6	0.293362	1.340928 (1.000794–1.79666)	0.0494
AL034550.2	0.301214	1.351499 (1.028015–1.776772)	0.0309
LINC00944	0.120110	1.12762 (0.966004–1.316277)	0.1281

Coef, regression coefficient; HR, hazard ratio; CI, confidence interval.

3.4. Identification of immune status among the different risk subgroups

First, we conducted PCA of the distribution of the 175 LUSC patients using total gene expression profiles and the expression profiles of the four immune-related lncRNAs of the model (Figure 4A, B). The 175 LUSC patients were better divided into two groups by our risk model. This result suggested that the immune status of LUSC patients with middle- or advanced-stage disease differed in the two risk groups.

Subsequently, to explore the correlation between the risk score and the TME, we adopted several currently acknowledged methods to assess immune cell infiltration in the 175 samples. We discovered that the high-risk group was positively associated with the presence of CD8+ T cells, B cells, myeloid dendritic cells (DCs), macrophages, regulatory T cells (Tregs) and cancer-associated fibroblasts (CAFs), which remained consistent across several methods (Figure 4C).

Finally, since immunotherapy has become a hot spot in the treatment of NSCLC patients, we also evaluated the expression of ICI-related biomarkers in the high- and low-risk groups. The high-risk group highly expressed PD1 (P value <0.001, Figure 4D), CTLA4 (P value <0.001, Figure 4E) and PD-L1 (P value >0.05, Figure 4F); however, PD-L1 expression was not significantly different.

3.5. Exploring the correlation between chemotherapy and risk scores

Chemotherapy is an important clinical treatment for middle- to advanced-stage lung squamous carcinoma patients. We evaluated the correlation between risk scores and the sensitivity of the common chemotherapeutic agents used for lung squamous carcinoma patient treatment. The results are presented in Figure 5 and illustrate that a low-risk score was associated with lower half maximal inhibitory concentrations (IC_{50}s) of cisplatin (P value <0.01), docetaxel (P value <0.001), vinorelbine (P value <0.001), and paclitaxel (P value <0.001), while it was associated with a higher IC_{50} of gemcitabine (P value <0.05). The IC_{50} of etoposide was lower in the low-risk group than in the high-risk group, but the difference was not statistically significant (P value

>0.05). This finding suggested that our risk model may be of value in predicting the efficacy of clinical chemotherapy. Low-risk group is more likely to benefit from the treatment with cisplatin, docetaxel, vinorelbine and paclitaxel, while high-risk group is more likely to benefit from the treatment with gemcitabine.

3.6. Functional enrichment analysis of the risk score model

To identify the biological functions of the genes identified in our immune-related lncRNA risk score model, we used gene set enrichment analysis (GSEA). We displayed the most significant enrichment terms based on Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), Hallmark, and WikiPathways gene sets from the high-risk group (Figure 6A-D). These terms mainly included the Kit receptor signaling pathway (Figure 6E), the KRAS signaling pathway (Figure 6F) and many immune-related pathways. In addition, the GSEA results also showed that the immune response (Figure 6G) and immune system processes (Figure 6H) were enriched in the high-risk group.

3.7. Validation of the immune-related lncRNA risk score model in a clinical cancer cohort

To further validate our prognostic immune-related lncRNA model, we examined the relative expression of the four identified lncRNAs in 14 stage IIB-IV LUSC patients from Fudan University Shanghai Cancer Center. We calculated the risk scores of the 14 patients and divided them into high-risk and low-risk groups. The DFS curve showed a statistically nonsignificant trend: the higher the risk score was, the higher the likelihood of recurrence (P value >0.05, Figure 7).

4. Discussion

Currently, the treatment of lung cancer has entered the era of individualized medicine. Most NSCLC patients with middle- to advanced-stage disease are treated with chemotherapy and/or radiation therapy, and several targeted and immunotherapy drugs are also available for some NSCLC patients [4]. Therefore, it is not sufficient to rely solely on the AJCC TNM staging system for the selection of treatments and the prediction of prognosis for patients. In the era of precision medicine, more and more robust predictive models are urgently needed to enable individualized risk stratification and treatment.

It has been affirmed that lncRNAs are critical regulators of cancer pathways and are closely associated with most cancer signatures, such as sustained proliferation, replicative immortality, evasion of growth inhibitors, induction of angiogenesis, resistance to cell death and development of metastasis [13, 16]. Recent studies have also indicated the
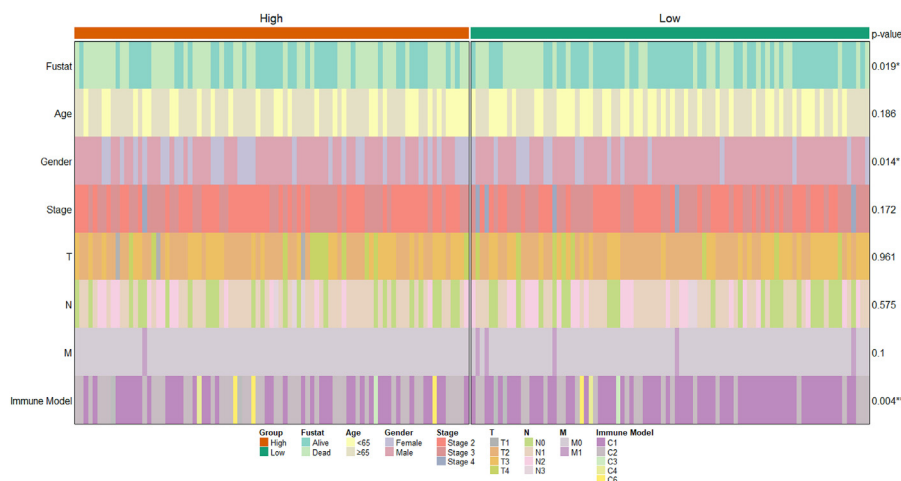
Figure 3. Assessment of clinical characteristics grouped based on risk score. *P < 0.05, **P < 0.01.
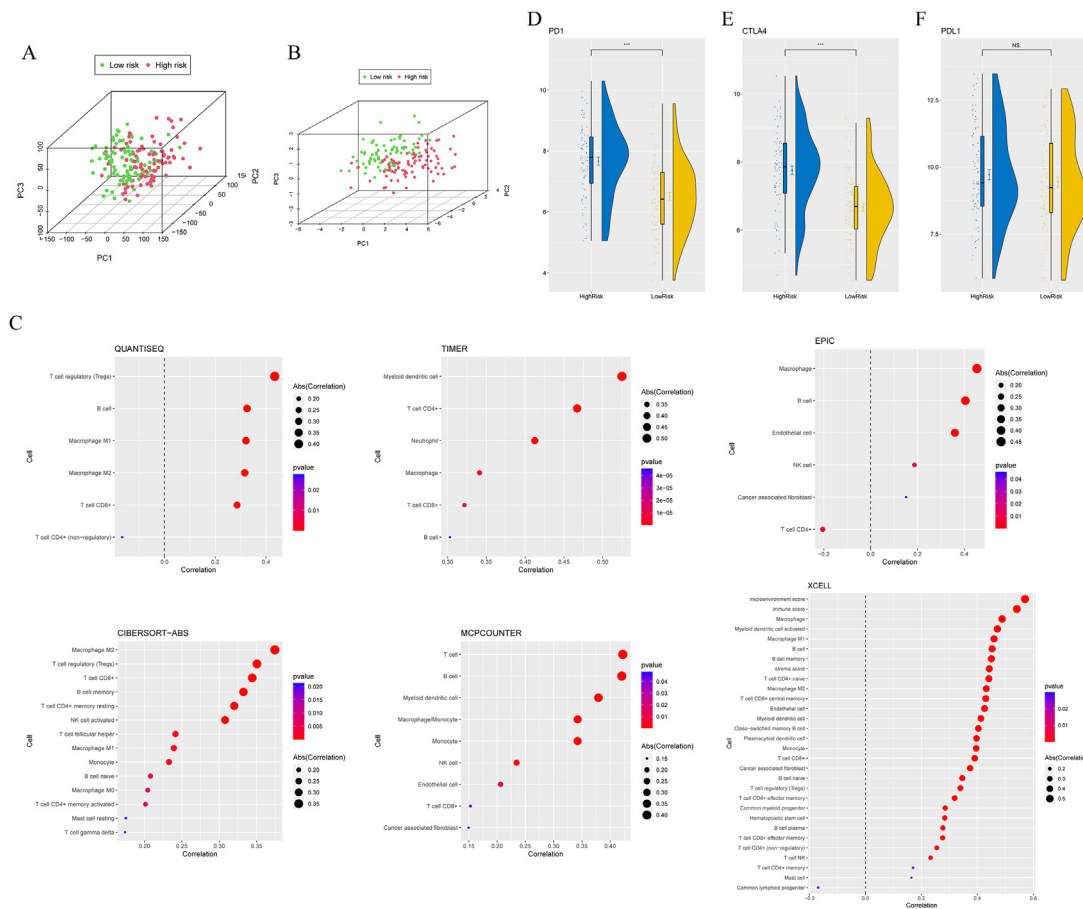
Figure 4. A, B PCA based on the whole gene expression profiles **(A)** and the four immune-related lncRNA signatures **(B)**. Green dots represent the low-risk group, and red dots represent the high-risk group. **C** Correlation of infiltrating immune cells with risk scores assessed by six methods. **D-F** Evaluation of the expression of immune checkpoint molecules between the high- and low-risk groups. NS. -P>0.05, ***P < 0.001.
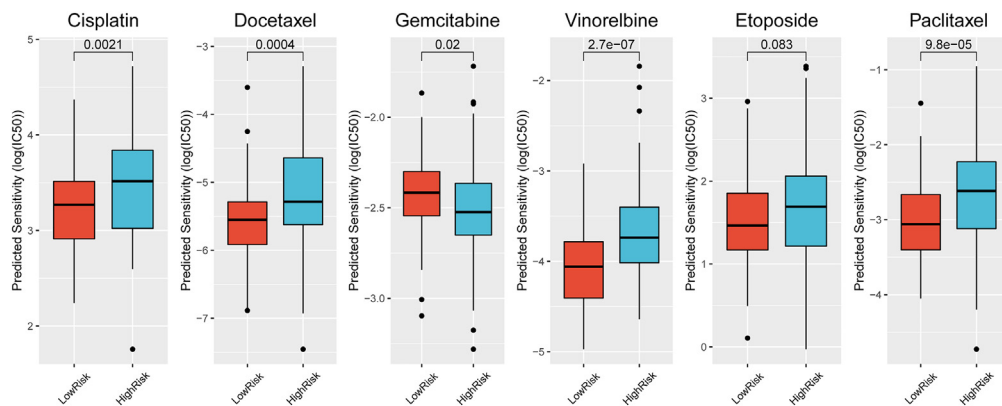

Figure 5. Box plots of the IC_{50} values for several chemotherapeutic agents in the high- and low-risk groups.

potential of lncRNAs as new diagnostic and prognostic biomarkers. LncRNA PCA3 has been used as a noninvasive prediagnostic marker for prostate cancer in clinical practice [32]. In patients with metastatic renal cell carcinoma (RCC), OS was worse in those with high lncARSR expression, and lncARSR could promote resistance to sunitinib [33]. In patients with colon cancer, RAMS11 overexpression was also demonstrated to be related to worse prognosis [34].

Based on the above information, we hoped to build a predictive model for NSCLC patients using immune-related lncRNAs. Immune-related lncRNA risk score models for lung adenocarcinoma have been reported in many studies [35, 36, 37, 38, 39, 40, 41]. Our research is the

first to construct an immune-related lncRNA risk score model in LUSC patients with middle- or advanced-stage disease. Based on the RNA expression profiles and the clinical data of LUSC patients publicly available from TCGA database, we constructed a model using four immune-related lncRNAs assessed by difference analysis, Pearson correlation analysis, and univariate and multivariate Cox regression analysis. The calculated risk scores were interpreted by the following formula: risk score = - 0.136673 × Expression$_{AC020907.1}$ + 0.293362 × Expression$_{AC027682.6}$ + 0.301214 × Expression$_{AL034550.2}$ + 0.120110 × Expression$_{LINC00944}$. Among the lncRNAs identified, in breast cancer, previously LINC00944 was confirmed to be in positive association with
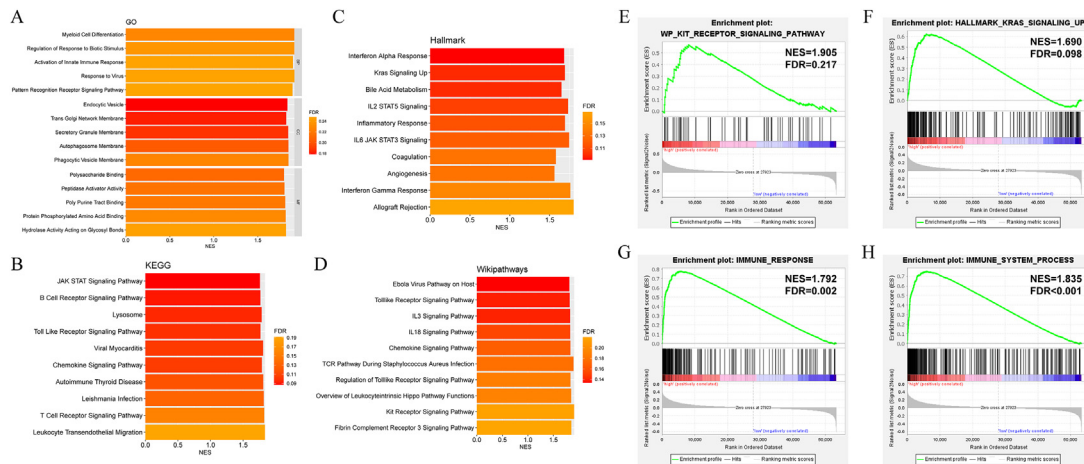
Figure 6. **A-D** The most significantly enriched terms based on GO **(A)**, KEGG **(B)**, hallmark **(C)** and WikiPathways **(D)** gene sets obtained by GSEA. **E-H** GSEA of gene sets for the Kit receptor signaling pathway **(E)**, KRAS signaling pathway **(F)**, immune response **(G)** and immune system processes **(H)**. *NES, normalized enrichment score; FDR, false discovery rate.*
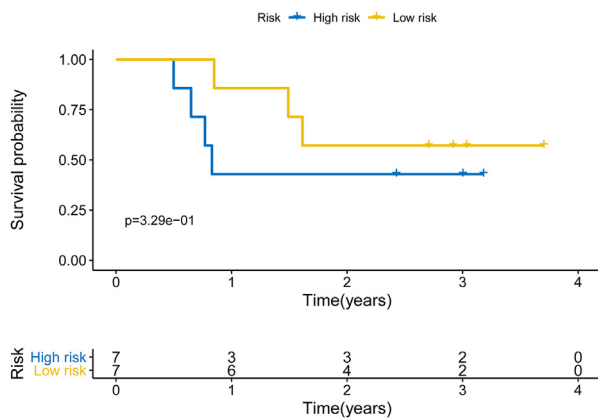
Figure 7. Kaplan–Meier disease-free survival (DFS) curve of our clinical cancer cohort (P = 0.329).

the presence of tumor-infiltrating T cells and proapoptotic markers, and low expression of LINC00944 showed correlation with poorer OS and relapse-free survival (RFS) [42]. However, an oncogenic role of LINC00944 in RCC has also been reported [43]. In a pancancer study of immune-related lncRNAs, LINC00944 was validated as a cancer-associated lncRNA whose expression was significantly perturbed in breast, kidney, lung and colorectal cancers and was influenced by the T-cell receptor (TCR) signaling pathway in 16 cancer categories [44]. It has been proven that the expression of AL034550.2 is remarkably higher in CD8+ T cells from patients with polymyositis (PM) than in patients with dermatomyositis (DM) [45]. In addition, AC020907.1 is considered to be one of the possible oncogenes of lung adenocarcinoma (LUAD) [46]. However, we could not retrieve any relevant reports for AC027682.6. In terms of survival analysis, both the OS and DFS of high-risk patients were worse than those of low-risk patients, suggesting a poorer prognosis. The subsequent independent prognostic analysis, PCA, and AUC data all showed that our model could serve as an independent predictor with good accuracy.

Platinum-based chemotherapy is an important treatment for middle-to advanced-stage LUSC patients. We evaluated the sensitivity to common chemotherapy drugs in high- and low-risk patients. Low-risk patients had higher sensitivity to cisplatin, docetaxel, vinorelbine and paclitaxel, while high-risk patients were more sensitive to gemcitabine. This finding implies that our model might be potentially helpful in the therapeutic decision-making process. Immunotherapy has also been proven to improve the outcome of NSCLC patients [47]. Our results

suggested that high-risk patients with a high expression of PD1 and CTLA4 were probably more likely to benefit from immunotherapy, but this finding remains to be validated by further studies. The important role of the TME in the development of primary lung cancer has been recognized over the past decade, and the use of the TME indicators as predictive biomarkers has been extensively studied [48]. Therefore, we further assessed the relationship between the risk score and the immune microenvironment, revealing that the risk score was positively correlated with the presence of CD8+ T cells, B cells, myeloid DCs, macrophages, Tregs and CAFs. CAFs are considered as a factor of bad prognosis in NSCLC patients [49, 50, 51, 52, 53, 54]. Tregs have been demonstrated to be associated with worse RFS in NSCLC patients without lymph node metastases [55]. O'Callaghan DS et al. have also reported that the infiltration of tumor islet Foxp3 (+) T-cells indicates worse prognosis among NSCLC patients [56]. Recently, a large study revealed that LUAD patients with high expression of immune-infiltrating Treg-related genes had worse OS [57]. A high density of mature DCs in tertiary lymphoid structures suggests a better prognosis [58]. In contrast, the prognostic role of tumor-associated macrophages (TAMs) has been controversial [59, 60], and more studies are required to further our understanding of the differences in local TAM phenotypes [48].

The GSEA results suggested possible biological functions of our model, indicating that these lncRNAs may have critical and important effects in the progression of LUSC as well as in the diagnosis and treatment of LUSC patients. c-Kit is a tyrosine kinase receptor that is deregulated and then activates downstream signaling in diseases such as cancer [61]. Activated signaling cascades include the Ras/Raf/MEK/-MAPK and PI3K/AKT/RPS6K pathways. KIT stimulation is as well known to activate the JAK/STAT and PLC/PKC signaling pathways [62]. c-Kit activation has been demonstrated to be associated with a variety of human cancers, including hematopoietic malignancies, malignant melanomas, gastrointestinal stromal tumors, and small cell lung cancers [61]. KRAS belongs to the rat sarcoma virus (Ras) gene family, and Ras genes are key players in human tumor pathogenesis [63]. KRAS is a commonly mutated gene in NSCLC patients, occurring in approximately 30% of LUAD [64] and 6% of LUSC patients [65]. Mutations in Ras genes induce sustained activation of downstream pathways, leading to tumor growth, proliferation and survival [66].

However, there are still some shortcomings of our study. For our clinical cancer cohort, we were unable to obtain OS data limited by the relatively short follow-up period. Additionally, the DFS curve was not statistically significant, which we speculated might be related to the insufficient sample size. Currently, we are unable to obtain additional fresh tissue samples from LUSC patients eligible for enrollment. Moreover, we failed to retrieve publicly available datasets containing the

expression profiles of these four lncRNAs, the clinicopathological characteristics and the survival outcomes of LUSC patients with middle- or advanced-stage disease. In the future, we will collect more clinical samples, expand the sample size and extend the follow-up time to further validate and evaluate our risk model.

In conclusion, we first constructed a risk score model of immune-related lncRNAs in LUSC patients with middle- or advanced-stage disease. The results demonstrated that our four immune-related lncRNA model performed well in individualized risk stratification and prognosis prediction.

Declarations

Author contribution statement

Qianqian Xue & Yue Wang: Conceived and designed the experiment; Performed the experiments; Analyzed and interpreted the data; Wrote the paper.

Qiang Zheng: Conceived and designed the experiment; Contributed reagents, materials, analysis tools or data.

Lijun Chen: Performed the experiments.

Yan Jin & Xuxia Shen: Contributed reagents, materials, analysis tools or data.

Yuan Li: Conceived and designed the experiment; Contributed reagents, materials, analysis tools or data.

Data availability statement

Data included in article/supplementary material/referenced in article.

Declaration of interests statement

The authors declare no conflict of interest.

Additional information

Supplementary content related to this article has been published online at https://doi.org/10.1016/j.heliyon.2022.e09521.

References

[1] H. Sung, J. Ferlay, R.L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, et al., Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA A Cancer J. Clin. 71 (3) (2021) 209–249.

[2] WHO Classification of Tumours Editorial Board, Thoracic Tumours, International Agency for Research on Cancer, Lyon, France, 2021.

[3] L. Osmani, F. Askin, E. Gabrielson, Q.K. Li, Current WHO guidelines and the critical role of immunohistochemical markers in the subclassification of non-small cell lung carcinoma (NSCLC): moving from targeted therapy to immunotherapy, Semin. Cancer Biol. 52 (Pt 1) (2018) 103–109.

[4] K.D. Miller, L. Nogueira, A.B. Mariotto, J.H. Rowland, K.R. Yabroff, C.M. Alfano, et al., Cancer treatment and survivorship statistics, CA A Cancer J. Clin. 69 (5) (2019) 363–385.

[5] Z. Chen, C.M. Fillmore, P.S. Hammerman, C.F. Kim, K.K. Wong, Non-small-cell lung cancers: a heterogeneous set of diseases, Nat. Rev. Cancer 14 (8) (2014) 535–546.

[6] D. Hanahan, L.M. Coussens, Accessories to the crime: functions of cells recruited to the tumor microenvironment, Cancer Cell 21 (3) (2012) 309–322.

[7] D.F. Quail, J.A. Joyce, Microenvironmental regulation of tumor progression and metastasis, Nat. Med. 19 (11) (2013) 1423–1437.

[8] D.C. Hinshaw, L.A. Shevde, The tumor microenvironment innately modulates cancer progression, Cancer Res. 79 (18) (2019) 4557–4566.

[9] F. Zhou, M. Qiao, C. Zhou, The cutting-edge progress of immune-checkpoint blockade in lung cancer, Cell. Mol. Immunol. 18 (2) (2021) 279–293.

[10] R.S. Herbst, D. Morgensztern, C. Boshoff, The biology and management of non-small cell lung cancer, Nature 553 (7689) (2018) 446–454.

[11] L. Horvath, B. Thienpont, L. Zhao, D. Wolf, A. Pircher, Overcoming immunotherapy resistance in non-small cell lung cancer (NSCLC) - novel approaches and future outlook, Mol. Cancer 19 (1) (2020) 141.

[12] F. Kopp, J.T. Mendell, Functional classification and experimental dissection of long noncoding RNAs, Cell 172 (3) (2018) 393–407.

[13] A.M. Schmitt, H.Y. Chang, Long noncoding RNAs in cancer pathways, Cancer Cell 29 (4) (2016) 452–463.

[14] L. Statello, C.J. Guo, L.L. Chen, M. Huarte, Gene regulation by long non-coding RNAs and its biological functions, Nat. Rev. Mol. Cell Biol. 22 (2) (2021) 96–118.

[15] Y. Chen, E. Zitello, R. Guo, Y. Deng, The function of LncRNAs and their role in the prediction, diagnosis, and prognosis of lung cancer, Clin. Transl. Med. 11 (4) (2021) e367.

[16] S.J. Liu, H.X. Dang, D.A. Lim, F.Y. Feng, C.A. Maher, Long noncoding RNAs in cancer metastasis, Nat. Rev. Cancer 21 (7) (2021) 446–460.

[17] Y.G. Chen, A.T. Satpathy, H.Y. Chang, Gene regulation in the immune system by long noncoding RNAs, Nat. Immunol. 18 (9) (2017) 962–972.

[18] M.K. Atianand, D.R. Caffrey, K.A. Fitzgerald, Immunobiology of long noncoding RNAs, Annu. Rev. Immunol. 35 (2017) 177–198.

[19] Q. Hu, S.D. Egranov, C. Lin, L. Yang, Long noncoding RNA loss in immune suppression in cancer, Pharmacol. Ther. 213 (2020) 107591.

[20] S.M. Lin, P. Du, W. Huber, W.A. Kibbe, Model-based variance-stabilizing transformation for Illumina microarray data, Nucleic Acids Res. 36 (2) (2008) e11.

[21] V. Thorsson, D.L. Gibbs, S.D. Brown, D. Wolf, D.S. Bortone, T.H. Ou Yang, et al., The immune landscape of cancer, Immunity 48 (4) (2018) 812–830, e814.

[22] B. Li, E. Severson, J.C. Pignon, H. Zhao, T. Li, J. Novak, et al., Comprehensive analyses of tumor immunity: implications for cancer immunotherapy, Genome Biol. 17 (1) (2016) 174.

[23] A.M. Newman, C.L. Liu, M.R. Green, A.J. Gentles, W. Feng, Y. Xu, et al., Robust enumeration of cell subsets from tissue expression profiles, Nat. Methods 12 (5) (2015) 453–457.

[24] F. Finotello, C. Mayer, C. Plattner, G. Laschober, D. Rieder, H. Hackl, et al., Molecular and pharmacological modulators of the tumor immune contexture revealed by deconvolution of RNA-seq data, Genome Med. 11 (1) (2019) 34.

[25] D. Aran, Z. Hu, A.J. Butte, xCell: digitally portraying the tissue cellular heterogeneity landscape, Genome Biol. 18 (1) (2017) 220.

[26] E. Becht, N.A. Giraldo, L. Lacroix, B. Buttard, N. Elarouci, F. Petitprez, et al., Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression, Genome Biol. 17 (1) (2016) 218.

[27] J. Racle, K. de Jonge, P. Baumgaertner, D.E. Speiser, D. Gfeller, Simultaneous enumeration of cancer and immune cell types from bulk tumor gene expression data, Elife 6 (2017).

[28] T. Li, J. Fu, Z. Zeng, D. Cohen, J. Li, Q. Chen, et al., TIMER2.0 for analysis of tumor-infiltrating immune cells, Nucleic Acids Res. 48 (W1) (2020) W509–W514.

[29] P. Geeleher, N.J. Cox, R.S. Huang, Clinical drug response can be predicted using baseline gene expression levels and in vitro drug sensitivity in cell lines, Genome Biol. 15 (3) (2014) R47.

[30] P. Geeleher, N. Cox, R.S. Huang, pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels, PLoS One 9 (9) (2014), e107468.

[31] National Comprehensive Cancer Network, NCCN clinical practice guidelines in oncology (NCCN Guidelines®). Non-small cell lung cancer, Version 4.2021 [cited 2021 March 3]. Available from: https://www.nccn.org/guidelines/.

[32] M.J. Roobol, F.H. Schroder, P. van Leeuwen, T. Wolters, R.C. van den Bergh, G.J. van Leenders, et al., Performance of the prostate cancer antigen 3 (PCA3) gene and prostate-specific antigen in prescreened men: exploring the value of PCA3 for a first-line diagnostic test, Eur. Urol. 58 (4) (2010) 475–481.

[33] L. Qu, J. Ding, C. Chen, Z.J. Wu, B. Liu, Y. Gao, et al., Exosome-transmitted lncARSR promotes sunitinib resistance in renal cancer by acting as a competing endogenous RNA, Cancer Cell 29 (5) (2016) 653–668.

[34] J.M. Silva-Fisher, H.X. Dang, N.M. White, M.S. Strand, B.A. Krasnick, E.B. Rozycki, et al., Long non-coding RNA RAMS11 promotes metastatic colorectal cancer progression, Nat. Commun. 11 (1) (2020) 2156.

[35] Y. Li, R. Shen, A. Wang, J. Zhao, J. Zhou, W. Zhang, et al., Construction of a prognostic immune-related LncRNA risk model for lung adenocarcinoma, Front. Cell Dev. Biol. 9 (2021) 648806.

[36] B. Zhang, R. Wang, K. Li, Z. Peng, D. Liu, Y. Zhang, et al., An immune-related lncRNA expression profile to improve prognosis prediction for lung adenocarcinoma: from bioinformatics to clinical word, Front. Oncol. 11 (2021) 671341.

[37] J.P. Li, R. Li, X. Liu, C. Huo, T.T. Liu, J. Yao, et al., A seven immune-related lncRNAs model to increase the predicted value of lung adenocarcinoma, Front. Oncol. 10 (2020) 560779.

[38] J. Li, C. Zhang, C. Zhang, H. Wang, Construction of immune-related and prognostic lncRNA clusters and identification of their immune and genomic alterations characteristics in lung adenocarcinoma samples, Aging (Albany NY) 12 (10) (2020) 9868–9881.

[39] J. Wang, X. Yin, Y.Q. Zhang, X. Ji, Identification and validation of a novel immune-related four-lncRNA signature for lung adenocarcinoma, Front. Genet. 12 (2021) 639254.

[40] H. Miao, D. Chen, R. Li, J. Hu, Y. Chen, C. Xu, et al., Identification of an immune-related six-long noncoding RNA signature as a novel prognosis biomarker for adenocarcinoma of lung, Biosci. Rep. 41 (1) (2021).

[41] D. Jin, Y. Song, Y. Chen, P. Zhang, Identification of a seven-lncRNA immune risk signature and construction of a predictive nomogram for lung adenocarcinoma, BioMed Res. Int. 2020 (2020) 7929132.

[42] P.R. de Santiago, A. Blanco, F. Morales, K. Marcelain, O. Harismendy, M. Sjoberg Herrera, et al., Immune-related lncRNA LINC00944 responds to variations in ADAR1 levels and it is associated with breast cancer prognosis, Life Sci. 268 (2021) 118956.

[43] C. Chen, H. Zheng, LncRNA LINC00944 promotes tumorigenesis but suppresses Akt phosphorylation in renal cell carcinoma, Front. Mol. Biosci. 8 (2021) 697962.

[44] Y. Li, T. Jiang, W. Zhou, J. Li, X. Li, Q. Wang, et al., Pan-cancer characterization of immune-related lncRNAs identifies potential oncogenic biomarkers, Nat. Commun. 11 (1) (2020) 1000.

[45] M. Houtman, L. Ekholm, E. Hesselberg, K. Chemin, V. Malmstrom, A.M. Reed, et al., T-cell transcriptomics from peripheral blood highlights differences between polymyositis and dermatomyositis patients, Arthritis Res. Ther. 20 (1) (2018) 188.

[46] H. Feng, Exploration of Molecular Mechanism and Prognosis Risk Model of Lung Adenocarcinoma Based on Bioinformatics, Lanzhou University, 2018.

[47] J.W. Carlisle, C.E. Steuer, T.K. Owonikoko, N.F. Saba, An update on the immune landscape in lung and head and neck cancers, CA A Cancer J. Clin. 70 (6) (2020) 505–517.

[48] N.K. Altorki, G.J. Markowitz, D. Gao, J.L. Port, A. Saxena, B. Stiles, et al., The lung microenvironment: an important regulator of tumour growth and metastasis, Nat. Rev. Cancer 19 (1) (2019) 9–31.

[49] R. Navab, D. Strumpf, B. Bandarchi, C.Q. Zhu, M. Pintilie, V.R. Ramnarine, et al., Prognostic gene-expression signature of carcinoma-associated fibroblasts in non-small cell lung cancer, Proc. Natl. Acad. Sci. U. S. A. 108 (17) (2011) 7160–7165.

[50] M. Ito, G. Ishii, K. Nagai, R. Maeda, Y. Nakano, A. Ochiai, Prognostic impact of cancer-associated stromal cells in patients with stage I lung adenocarcinoma, Chest 142 (1) (2012) 151–158.

[51] J. Alcaraz, J.L. Carrasco, L. Millares, I.C. Luis, F.J. Fernandez-Porras, A. Martinez-Romero, et al., Stromal markers of activated tumor associated fibroblasts predict poor survival and are associated with necrosis in non-small cell lung cancer, Lung Cancer 135 (2019) 151–160.

[52] K. Shimizu, K. Kirita, K. Aokage, M. Kojima, T. Hishida, T. Kuwata, et al., Clinicopathological significance of caveolin-1 expression by cancer-associated fibroblasts in lung adenocarcinoma, J. Cancer Res. Clin. Oncol. 143 (2) (2017) 321–328.

[53] T. Miyamura, N. Sakamoto, K. Ishida, T. Kakugawa, H. Taniguchi, Y. Akiyama, et al., Presence of heat shock protein 47-positive fibroblasts in cancer stroma is associated with increased risk of postoperative recurrence in patients with lung cancer, Respir. Res. 21 (1) (2020) 234.

[54] H. Zhang, H. Jiang, L. Zhu, J. Li, S. Ma, Cancer-associated fibroblasts in non-small cell lung cancer: recent advances and future perspectives, Cancer Lett. 514 (2021) 38–47.

[55] K. Shimizu, M. Nakata, Y. Hirami, T. Yukawa, A. Maeda, K. Tanemoto, Tumor-infiltrating Foxp3+ regulatory T cells are correlated with cyclooxygenase-2 expression and are associated with recurrence in resected non-small cell lung cancer, J. Thorac. Oncol. 5 (5) (2010) 585–590.

[56] D.S. O'Callaghan, E. Rexhepaj, K. Gately, L. Coate, D. Delaney, D.M. O'Donnell, et al., Tumour islet Foxp3+ T-cell infiltration predicts poor outcome in nonsmall cell lung cancer, Eur. Respir. J. 46 (6) (2015) 1762–1772.

[57] X. Wang, Z. Xiao, J. Gong, Z. Liu, M. Zhang, Z. Zhang, A prognostic nomogram for lung adenocarcinoma based on immune-infiltrating Treg-related genes: from bench to bedside, Transl. Lung Cancer Res. 10 (1) (2021) 167–182.

[58] J. Goc, C. Germain, T.K. Vo-Bourgais, A. Lupo, C. Klein, S. Knockaert, et al., Dendritic cells in tumor-associated tertiary lymphoid structures signal a Th1 cytotoxic immune contexture and license the positive prognostic value of infiltrating CD8+ T cells, Cancer Res. 74 (3) (2014) 705–715.

[59] E.M. Conway, L.A. Pikor, S.H. Kung, M.J. Hamilton, S. Lam, W.L. Lam, et al., Macrophages, inflammation, and lung cancer, Am. J. Respir. Crit. Care Med. 193 (2) (2016) 116–130.

[60] A. Mantovani, F. Marchesi, S. Jaillon, C. Garlanda, P. Allavena, Tumor-associated myeloid cells: diversity and therapeutic targeting, Cell. Mol. Immunol. 18 (3) (2021) 566–578.

[61] J. Lennartsson, L. Ronnstrand, Stem cell factor receptor/c-Kit: from basic science to clinical implications, Physiol. Rev. 92 (4) (2012) 1619–1649.

[62] K. Kandasamy, S.S. Mohan, R. Raju, S. Keerthikumar, G.S. Kumar, A.K. Venugopal, et al., NetPath: a public resource of curated signal transduction pathways, Genome Biol. 11 (1) (2010) R3.

[63] A.E. Karnoub, R.A. Weinberg, Ras oncogenes: split personalities, Nat. Rev. Mol. Cell Biol. 9 (7) (2008) 517–531.

[64] L. Ding, G. Getz, D.A. Wheeler, E.R. Mardis, M.D. McLellan, K. Cibulskis, et al., Somatic mutations affect key pathways in lung adenocarcinoma, Nature 455 (7216) (2008) 1069–1075.

[65] M.L. Sos, R.K. Thomas, Genetic insight and therapeutic targets in squamous-cell lung cancer, Oncogene 31 (46) (2012) 4811–4814.

[66] P. Tomasini, P. Walia, C. Labbe, K. Jao, N.B. Leighl, Targeting the KRAS pathway in non-small cell lung cancer, Oncol. 21 (12) (2016) 1450–1460.