



Article **Comprehensive Analysis of a Novel Immune-Related Gene** Signature in Lung Adenocarcinoma

Hongxiang Feng, Chaoyang Liang, Yuhui Shi, Deruo Liu, Jin Zhang and Zhenrong Zhang *

Department of Thoracic Surgery, China-Japan Friendship Hospital, Beijing 100029, China * Correspondence: zhangzhenrong6189@126.com

Abstract: Lung cancer is the major cause of cancer-related deaths around the world. Lung adenocarcinoma (LUAD), the most common subtype of lung cancer, contributed to the majority of mortalities and showed different clinical outcomes in prognosis. Tumor-infiltrated immune cells at the tumor site are associated with better survival and immunotherapy response. Thus, it is essential to further investigate the molecular mechanisms and new prognostic biomarkers of lung adenocarcinoma development and progression. In this study, a six-gene signature (CR2, FGF5, INSL4, RAET1L, AGER, and TNFRSF13C) was established to predict the prognosis of LUAD patients, as well as predictive value. The prognostic risk model was also significantly associated with the infiltration of immune cells in LUAD microenvironments. To sum up, a novel immune-related six-gene signature (CR2, FGF5, INSL4, RAET1L, AGER, and TNFRSF13C) was identified that could predict LUAD survival and is highly related to B cells and dendritic cells, which may provide a theoretical basis of personalized treatment for targeted immunotherapy.

Keywords: lung adenocarcinoma; immune-related gene signature; risk model; tumor infiltration immunity; prognosis

1. Introduction

Lung cancer is currently the most aggressive cancer in terms of morbidity and mortality worldwide [1–3]. There are two main types of lung cancer, small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC). The most common type of lung cancer is nonsmall cell lung cancer (NSCLC), which is diagnosed in 85% of all lung cancers [4,5]. Lung adenocarcinoma (LUAD) is considered to be the most common histologic subtype of NSCLC, which accounts for approximately 40% of patients with lung cancer. Although there have been major advances in lung cancer treatment in recent decades, including surgical resection, radiotherapy, chemotherapy, and immunotherapy, the prognosis and survival of patients are still unsatisfactory [6]. Possible reasons are that patients diagnosed were in advanced stages, or that patients diagnosed early cannot receive targeted therapy yet because they do not carry the common molecular mutations (EGFR, BRAF V600E, MET, or ALK) that could be treated with targeted therapy. It is necessary to further investigate the molecular mechanisms of tumorigenesis and the development of new and reliable biomarkers to improve the survival of LUAD patients.

It is known that a tumor immune microenvironment plays significant roles in tumor development and progression [7–12]. The primary cause is the inhibition of immunosuppressive checkpoints such as CTLA4 or PD-1/PD-L1 or the breakdown in the development of cytotoxic T-cell lymphocytes (CTLs) [13–15]. An increasing number of studies suggest that several tumor prognostic and predictive biomarkers are highly associated with immune response. For example, an immune-related gene signature of 11 genes was able to predict the prognosis and the immunotherapy efficacy of hepatocellular carcinoma (HCC) [16]. A robust six immune-related gene signature played a role in risk stratification and overall survival in patients with lower-grade glioma [17]. An uncovered seven-gene signature was



Citation: Feng, H.; Liang, C.; Shi, Y.; Liu, D.; Zhang, J.; Zhang, Z. Comprehensive Analysis of a Novel Immune-Related Gene Signature in Lung Adenocarcinoma. J. Clin. Med. 2022, 11, 6154. https://doi.org/ 10.3390/jcm11206154

Academic Editor: Rong Na

Received: 5 September 2022 Accepted: 11 October 2022 Published: 19 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/).

verified to predict the prognosis in osteosarcoma [18]. A four-gene signature was developed to predict the prognosis and survival of lung adenocarcinoma patients [19]. These findings have been proven to demonstrate the importance of tumor immune microenvironments in carcinogenesis, progressions, and the development of tumors.

Although many immune-related gene-based markers for LUAD are available, immunotherapy could only benefit a small proportion of patients [19,20]. There is an urgent need for a more comprehensive and reliable indicator that can predict both the survival of LUAD patients and the efficacy of immunotherapy. In this study, we constructed an immune-related prognostic gene signature. Furthermore, we also performed the evaluation of prognostic significance, Gene Set Enrichment Analysis (GSEA) and the possible predictive value in immunotherapy of the gene signatures. The workflow of this study was summarized in Figure 1. We hope the immune-related prognostic gene signature could be used as the predictive biomarker for the prognosis and immunotherapy of LUAD patients.

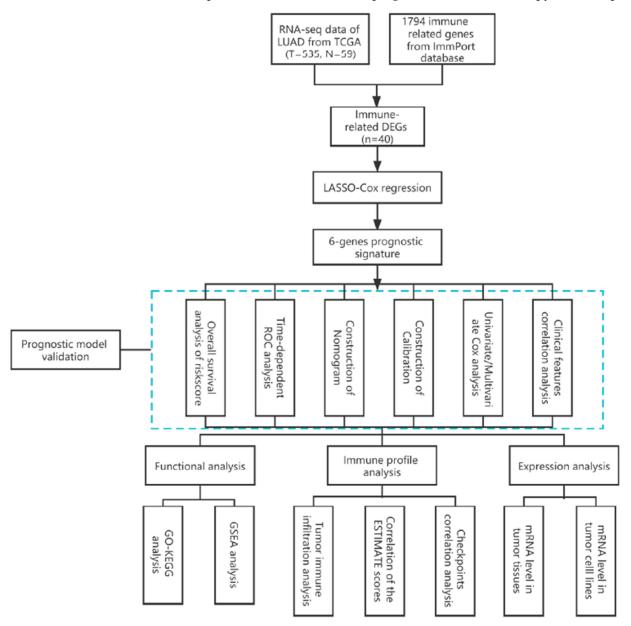


Figure 1. The workflow of this study.

2. Materials and Methods

2.1. Generation of Differentially Expressed Immune-Related Genes (DEIRGs)

Differentially expressed genes (DEGs) between LUAD tumor samples at stage T1 (n = 175) and stage T2-T4 (n = 357) from a TCGA dataset were assessed by using R package "DESeq2 (v 1.26.0)" with the criteria of $|\log 2$ (Fold change) | > 1 and p.adj < 0.05 [21]. The volcano plot of DEGs was drawn by using R package "ggplot2 (v3.3.3)". The 1793 immune-related genes (IRGs) were downloaded from the Immunology Database and Analysis Portal (ImmPort) database (https://www.immport.org/home) [22]. The differentially expressed immune-related genes (DEIRGs) between DEGs and IRGs were analyzed by using R package "ggplot2 (v3.3.3)" and visualized as the Venn diagram.

2.2. Functional Enrichment Analyses

The 40 DEIRGs were obtained from the intersection between DEGs and IRGs. The Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were performed using clusterProfiler (v3.14.3) and org.Hs.eg.db (3.10.0) [23]. The bubble diagram was visualized by R package "ggplot2 (v3.3.3)". Gene set enrichment analysis (GSEA) analysis was conducted by clusterProfiler (3.14.3) and shown as the mountain plot. The reference gene set was c2.cp.v7.2.symbols.gmt [Curated]. Gene set databases are MSigDB Collections. The False Discovery Rate (FDR) < 0.25 and p.Adjust < 0.05 are considered to be significant [23,24].

2.3. Establishment and Validation of the Prognostic Immune-Related Gene Signature

Differentially expressed immune-related genes (DEIRGs) were used for the Least Absolute Shrinkage and Selection Operator (LASSO) coefficient screen using the R software glmnet (v4.1-2) and survival (3.2-10). The prognostic immune-related gene signature (risk model) with coefficients were selected based on the optimal lambda value. Overall survival (OS), disease specific survival (DSS) and the progression-free interval (PFI) between high and low risk score groups were conducted using the R software survininer (0.4.9) and survival (3.2-10) [24]. Univariate and multivariate Cox analyses and 1-, 3-, and 5-year receiver operating characteristic (ROC) analysis were used to estimate the prognostic value of the risk model [25]. The predictive risk factors were evaluated by Nomogram and Calibrate curves using the R software rms (6.2-0) and survival (3.2-10) [25].

2.4. Immune Cell Infiltration Analysis

The ssGSEA algorithm was applied for the correlation between the risk score and 24 tumor-infiltrated immune cells by using GSVA (1.34.0) [26,27]. The ESTIMATE method was used for the comparison of the immune score, ESTIMATE score, and stromal score between the low- and high-risk groups [28]. The prognostic value of the 24 tumor-infiltrated immune cells was analyzed by the univariate Cox regression analysis and Kaplan–Meier analysis.

2.5. Gene Expression and Genetic Alterations Analysis

The immunohistochemistry images of six gene signatures in LUAD tumor tissues and normal tissues were obtained from the Human Protein Atlas (HPA) database (https://www.proteinatlas.org). The Genetic alteration analysis was performed using the cBioPortal for Cancer Genomics (https://www.cbioportal.org) [29].

2.6. Reverse Transcription and Real-Time Quantitative PCR (RT-qPCR)

Human lung cancer cell lines (A549, NCI-H1975, Calu-3) and human normal epithelial cells BEAS-2B were purchased from Procell (https://www.procell.com.cn). The A549 cells were cultured in Ham's F-12K Medium (Thermo Fisher, Waltham, MA, USA). The NCI-H1975 cells were cultured in RPMI1640 medium (Hyclone, Logan, UT, USA). The Calu-3 cells were cultured in modified Eagle's medium (MEM, containing NEAA, Procell, Wuhan, China). The BEAS-2B cells were cultured in Dulbecco's modified Eagle's medium (DMEM, HyClone, USA). All cells were cultured in the specific mediums supplemented with 10% fetal bovine serum (Hyclone, USA), penicillin sodium (100 U/mL) and streptomycin (100 mg/mL). The cells were incubated in the thermostatic cell incubator at 37 °C with 5% CO₂. Total RNA was extracted using the RNeasy Plus Mini Kit (QIAGEN, Hilden, Germany). Quantitative PCR (qPCR) was performed using the ABI 7900 qPCR system. Relative RNA expression was normalized to GAPDH and calculated by the $2^{-\Delta\Delta Ct}$ method. The primers were listed in Table 1.

Table 1. Primers for qRT-PCR.

Gene	Forward Sequence (5'-3')	Reverse Sequence (5'-3')
CR2	ACCATGGTCGTCATACAGGTG	AGCCAGGATTGCATCAACA
INSL4	AGCCTGTTCCGGTCCTATCT	ATGATGGCTGCCCTTCAGAC
FGF5	CGCTCACAGTCACCTGGTTT	CACCCTCGTTTGGCTTTTCC
RAET1L	CCATCCCAGCTTTGCTTCTGT	TGACGGGTGTGACTGTCTTG
AGER	GCTTGGAAGGTCCTGTCTCC	CCACCAATTGGACCTCCTCC
TNFRSF13B	GTCAAAGTCCGGCCAAGTCT	CCACTGTCTGGGATGTGTGG
GAPDH	GAGAAGGCTGGGGGCTCATTT	AGTGATGGCATGGACTGTGG

2.7. Statistical Analysis

Data were analyzed using the R software (v3.6.3) and the qPCR analysis was performed by GraphPad Prism. DEGs were analyzed by "ggplot2". GO-KEGG and GSEA analyses were performed by "clusterProfiler". The LASSO analysis was conducted by "glmnet and survival". Univariate and multivariate Cox regression analyses were assessed by using the "survival". The Kaplan–Meier analysis was performed using the "survival" and "survminer" R packages. An immune cell infiltration analysis was applied by "GSVA". A *p* value < 0.05 is statistically significant.

3. Results

3.1. Identification of DEIRGs

The workflow of this study is illustrated in Figure 1. We analyzed a total of 535 LUAD samples and 59 normal samples and obtained 697 DEGs, including 411 upregulated and 286 downregulated genes as shown in the volcano map (Figure 2A). In addition, the forty differentially expressed immune-related genes (DEIRGs) were generated (Figure 2B). The functional analysis of the DEGs was performed. The GO enrichment analysis showed that the DEGs were enriched in "humoral immune response", "defense response to bacterium", and "antimicrobial humoral response" (Figure 2C). The KEGG pathway analysis revealed that the highly relevant signaling pathways were neuroactive ligand—receptor interaction and complement and coagulation cascades (Figure 2D). The functional analysis (GOs and KEGGs) for the remaining genes that were differentially regulated in cancer versus normal tissue were also performed. The enriched functions of either upregulated or downregulated genes were not related to immune responses as shown in Figure 2E,F.

3.2. Construction of a Prognostic Immune-Related Gene Signature

We used a LASSO Cox regression model to perform the overall survival analysis of 40 immune-related potential prognostic genes (Figure 3A). Six genes included in the analysis were identified to be the best predictors for prognostic significance (Figure 3B). Hence, the prognostic model consists of these six genes noted as CR2, FGF5, INSL4, RAET1L, AGER, and TNFRSF13C. The regression coefficient of each gene was displayed in Table 2. The risk scores of this model are formulated as follows: Risk score = -0.01996 * expression level of CR2 + 0.07140 * expression level of INSL4 + 0.03141 * expression level of FGF5 + 0.12065 * expression level of RAET1L + (-0.00451) * expression level of AGER + (-0.00301) * expression level of TNFRSF13B.

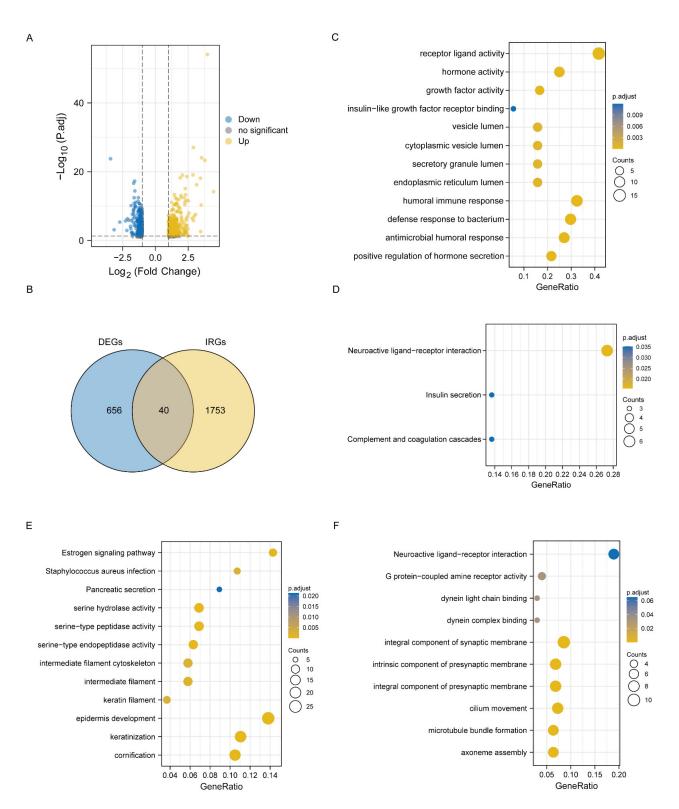


Figure 2. Construction of DEIRGs for LUAD. (**A**) Volcano plot of DEIRGs between LUAD and normal tissues. (**B**) Venn diagram of the intersections of DEGs with IRGs. (**C**) GO enrichment analysis of forty DEIRGs. (**D**) KEGG enrichment analysis of forty DEIRGs. (**E**) GO and KEGG enrichment analysis of the upregulated remaining genes (except the above forty genes). (**F**) GO and KEGG enrichment analysis of the downregulated remaining genes (except the above forty genes).

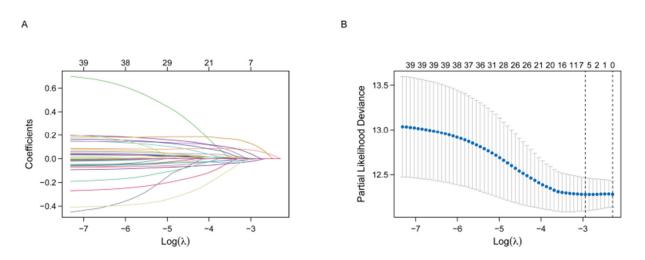


Figure 3. Establishment of the prognostic gene signature using LASSO regression analysis. (A) LASSO coefficient profiles of 40 immune-related potential prognostic genes. Each curve corresponds to a gene. (B) Tenfold cross-validation for tuning parameter selection in the LASSO model. The partial likelihood deviance is plotted against log (Lambda), where Lambda is the tuning parameter. Partial likelihood deviance values are shown, with error bars representing SE. The dotted vertical lines are drawn at the optimal values by minimum criteria and 1-SE criteria. LASSO—the least absolute shrinkage and selection operator Cox regression model.

Gene Symbol	Description	Risk Coefficient	
CR2	Complement C3d receptor 2	-0.019957508	
INSL4	Insulin-like 4	0.071400283	
FGF5	Fibroblast growth factor 5	0.031415586	
RAET1L	Retinoic acid early transcript 1L	0.120647299	
AGER	Advanced glycosylation end-product specific receptor	-0.004511221	
TNFRSF13B	TNF receptor superfamily member 13B	-0.003010621	

Table 2. Six immune-related prognostic genes obtained from LASSO Cox regression model.

3.3. Evaluation of the Prognostic Value of the Six-Gene Signature

The risk score of each LUAD patient was a linear combination of each six-gene signature expression and its risk coefficient. High- and low-risk groups of LUAD patients were divided according to the median risk score. The distribution plots of risk scores and the outcome status, including the Overall Survival (OS), Disease Specific Survival (DSS), and Progression-Free Interval (PFI) of the gene signature, are shown in Figure 4A. Patients in the high-risk group showed a lower survival probability compared with those in the low-risk group. Consistently, Kaplan–Meier plotter analysis showed that the OS, DSS and PFIs of the patients in the high-risk group exhibited remarkably worse outcomes than those in the corresponding low-risk group (Figure 4B). Next, we performed the univariate analysis to validate the six immune-related genes with prognosis significance in LUAD (Figure 4C). In addition, we generated a nomogram tool to predict the 1-year, 3-year and 5-year overall survival risk based on the selected factors (Figure 4D). Moreover, the calibration curves for the prognostic nomograms presented a better accuracy and consistency between prediction and actual 1-, 3-, and 5-year survival in the TCGA-LUAD cohort (Figure 4E).

7 of 18

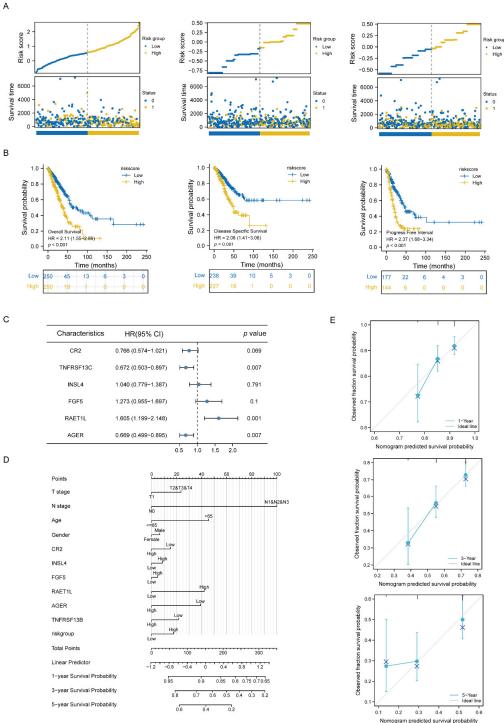


Figure 4. Evaluation of prognostic significance of the six-gene signature risk score. **(A)** Risk score distribution and outcome status, including OS, DSS, and PFI of LUAD patients in low- and high-risk groups. **(B)** Kaplan–Meier plotter curves of the six-gene signature risk score. **(C)** Forest plot for the multivariable Cox model results of each gene in the six-gene signature risk score. **(D)** Nomogram for predicting 1-, 3-, and 5-year OS in TCGA-LUAD cohort. **(E)** Calibration curves of prognostic nomogram on consistency between prediction and observed 1-, 3-, and 5-year survival in TCGA-LUAD cohort. Dashed line at 45° implicated a perfect prediction, and the actual performances of our nomogram were shown in blue lines. OS—overall survival. DSS—disease specific survival. PFI—progression-free interval.

We performed the time-dependent receiver operating characteristic (ROC) plotter and found that the values of the area under the curve (AUC) for the six-gene signature risk score at 1, 3, 5, and 10 years were all above 0.6 (Figure 5A). Next, we performed the correlation between the six genes individually and the risk score analysis using the Pearson correlation analysis as shown in Figure 5B–H. TNFRSF13B was significantly related with most of the genes, except for RAET1L, among the six-gene signature model (Figure 5B). There were negative correlations between CR2, AGER, TNFRSF13B and the risk score (Figure 5C,E,H). There was a positive relationship between RAET1L, FGF5, INSL4 and the risk score (Figure 5D,F,G). These results indicated that the six-gene signature risk score could be used as an independent prognostic factor for LUAD.

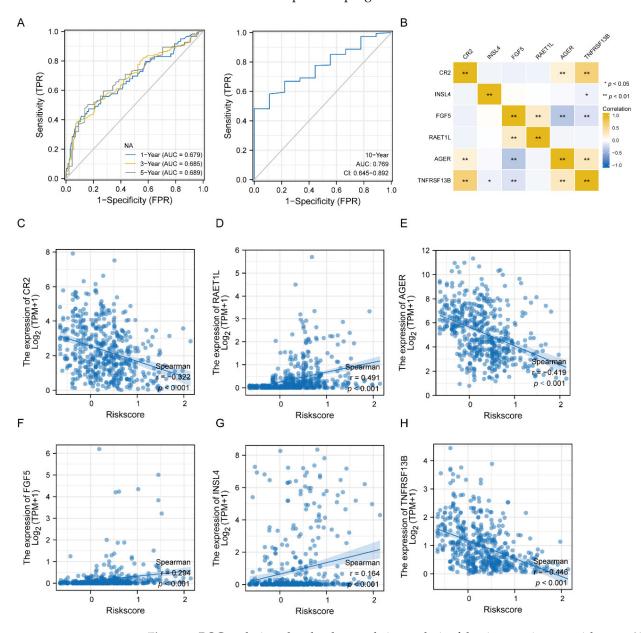


Figure 5. ROC analysis and molecular correlation analysis of the six-gene signature risk score. (**A**) The 1-year, 3-year, 5-year and 10-year ROC curve plot of LUAD patients. (**B**) Heatmap shows the correlation between the six-gene signature molecules. (**C**–**H**) Scatter diagram shows correlation between risk score and CR2 (**C**), RAET1L (**D**), AGER (**E**), FGF5 (**F**), INSL4 (**G**), and TNFRSF13B (**H**). p < 0.001. ROC—receiver operating characteristic.

To further understand the biological functions of the six-gene signature risk score in LUAD, we performed enrichment analyses of GO and KEGG pathways. Results showed that the six-gene signature was enriched in lymphocyte mediated immunity, leukocyte proliferation, mononuclear cell proliferation, and lymphocyte proliferation (Figure 6A). Consistently, multiple GSEA analyses revealed that the six-gene signature was highly enriched in immune-related signaling pathways such as antigen activates B cell receptor (BCR) leading to the generation of second messengers, reactome immunoregulatory interactions between a lymphoid and a non-lymphoid cell, reactome interleukin 20 family signaling, type II interferon signaling, and CD22 mediated BCR regulation (Figure 6B–G).

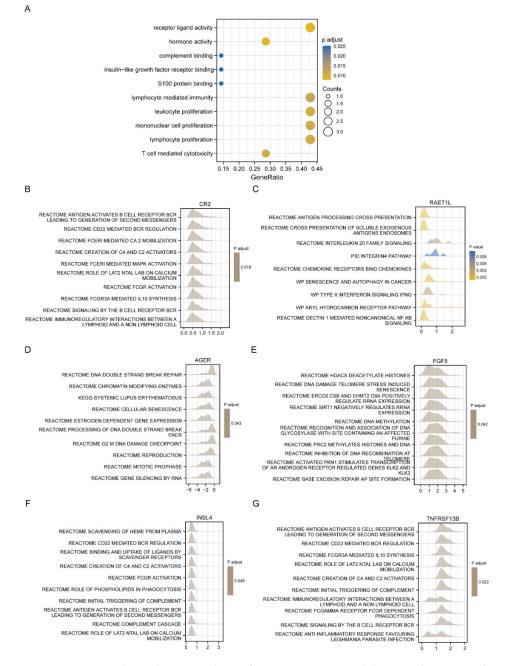


Figure 6. Functional enrichment analysis of six-gene signature. (**A**) GO and KEGG pathway analysis indicating the enriched pathways with six-gene signature risk score. (**B**–**G**) GSEA analysis of differential genes associated with CR2, RAET1L, AGER, FGF5, INSL4, and TNFRSF13B, respectively.

3.5. Clinicopathological Characteristics Correlation Analysis

Next, we investigated the possible relationship between multiple clinicopathological features and the six-gene signature risk score. The correlation analysis indicated that a higher risk score was significantly associated with age, gender, pathologic stages, T/N/M stages, and OS/DSS/PFI events (Figure 7A–I). Moreover, we also did univariate and multivariate Cox analyses with risk scores as shown in Table 3. The results validated that higher age, T/N/M and pathologic stage, as well as higher risk scores were risk factors for LUAD patients with HRs > 1, p < 0.01.

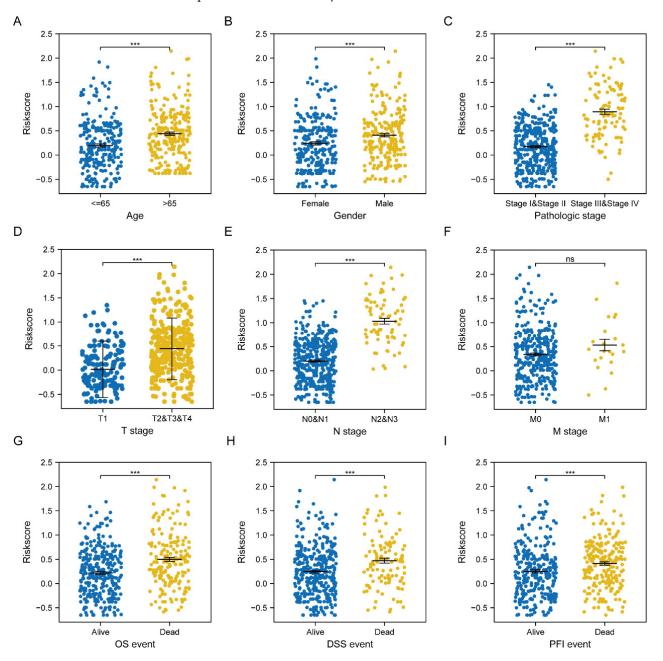


Figure 7. Clinicopathological features correlation analysis. (A–I) The relationships between risk score and clinicopathological factors (Age, Gender, Pathologic stage, T stage, N stage, M stage, OS, DSS, and PFI events) in TCGA-LUAD cohort, respectively. *** p < 0.001.

	Total (N)	Univariate Analysis		Multivariate Analysis	
Characteristics		Hazard Ratio (95% CI)	p Value	Hazard Ratio (95% CI)	p Value
T stage	501				
T1	168	Reference			
T2&T3&T4	333	1.668 (1.184-2.349)	0.003	1.591 (1.004-2.520)	0.048
N stage	492				
NO	325	Reference			
N1&N2&N3	167	2.606 (1.939-3.503)	< 0.001	1.790 (1.004-3.191)	0.048
M stage	360				
MO	335	Reference			
M1	25	2.111 (1.232-3.616)	0.007	1.371 (0.744-2.523)	0.311
Gender	504				
Female	270	Reference			
Male	234	1.060 (0.792-1.418)	0.694		
Age	494				
≤ 65	238	Reference			
>65	256	1.228 (0.915-1.649)	0.171		
Pathologic stage	496				
Stage I	270	Reference			
Stage II & Stage III & Stage IV	226	2.975 (2.188-4.045)	< 0.001	1.177 (0.617-2.244)	0.621
Smoker	490				
No	71	Reference			
Yes	419	0.887 (0.587-1.339)	0.568		
Risk group	482	· · · · · ·			
Low	244	Reference			
High	238	2.217 (1.625-3.025)	< 0.001	1.612 (1.081-2.402)	0.019

Table 3. Univariate and multivariate Cox regression analyses of risk score and other clinicopathologic factors for OS in the entire TCGA cohort.

3.6. Evaluation of Immune Cell Infiltration Characterization

To further explore the association between the six-gene signature risk score and immunity, we firstly evaluated the correlation with 24 types of tumor-infiltrating immune cells in the high- or low-risk groups using the ssGSEA algorithm (Figure 8A). Results showed that T cells, plasmacytoid DC (pDC), mast cells, DC, cytotoxic cells, B cells, and Th17 cells showed a lower expression in the high-risk group (p < 0.05), whereas T gamma delta (Tgd) and Th2 cells were highly expressed in the high-risk group. Furthermore, we investigated the correlation with tumor lymph cells by Spearman's correlation analysis. The lollipop chart indicated that risk score was positively related to Th2 cells and Tgd (Figure 8B). Moreover, a negative correlation was observed between risk score and TFH, B cells, Mast cells, T cells, and CD8 T cells (Figure 8B).

In addition, we also investigated the prognostic significance of 22 types of tumorinfiltrating immune cells. The Kaplan–Meier analysis demonstrated that the lower infiltrating abundance of B cells and dendritic cells were significantly associated with better OS (p < 0.05, Figure 8C,D). The univariate Cox analysis revealed that B cells and CD4 T cells remarkably influenced prognosis (Figure 8E). These findings indicate that B cells and CD4 T cells may play a meaningful role in the prognostic ability of the six-gene signature in LUAD.

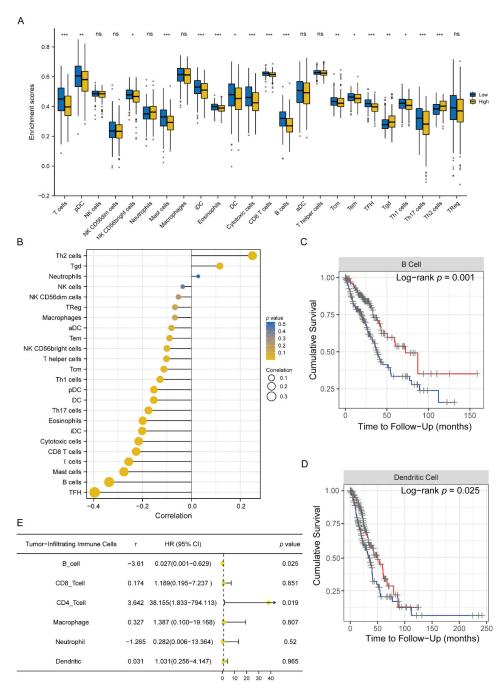


Figure 8. Relationship between immune cell infiltration and six-gene signature risk score and prognostic value evaluation. (**A**) Correlations of risk score with immune cell infiltration. The blue and yellow boxes represented the six-gene signature low- and high-risk group, respectively. The medium line inside the box represented the mean value. Wilcoxon rank-sum was applied for the significance test. * p < 0.05; ** p < 0.01; *** p < 0.001. (**B**) Lollipop chart showing the Spearman analysis of risk score with tumor-infiltrating immune cells. (**C**) Relationship between overall survival and B cells. High level (Red line) vs. low level (black line). (**D**) Relationship between overall survival and dendritic cells. High level (Red line) vs. low level (black line). (**E**) Univariate Cox regression model built for six tumor-infiltrating immune cell types (B cells, CD8 cells, CD4 cells, Macrophage, Neutrophil, Dendritic) based on overall survival. T cells; pDC [Plasmacytoid DC]; NK cells; NK CD56dim cells; NK CD56bright cells; Neutrophils; Mast cells; Macrophages; iDC [immature DC]; Eosinophils; DC; Cytotoxic cells; CD8 T cells; B cells; aDC [activated DC]; T helper cells; Tcm [T central memory]; Tem [T effector memory]; Tfh [T follicular helper]; Tgd [T gamma delta]; Th1 cells; Th17 cells; Th2 cells; and Treg.

3.7. Estimation of the Immunity Response

We also evaluated the association between the six-gene signature and tumor immune infiltration to calculate the immune score, ESTIMATE score, and stromal score using the ESTIMATE package. We observed that there was a significant difference in the immune score and stromal score between the low-risk group and the high-risk group, except for the ESTIMATE score (Figure 9A–C), suggesting that the six-gene signature risk score was significantly associated with the immune score and stromal score. Next, we also did the correlation analysis of immune checkpoints as shown in Figure 9D–I. There was a significant relationship between CD27, TNFSF15, TNFRSF25, TIGIT, VSIR, BTNL2 and risk score, suggesting that the six-gene signature risk score might have an immune response and respond to immunotherapy in LUAD.

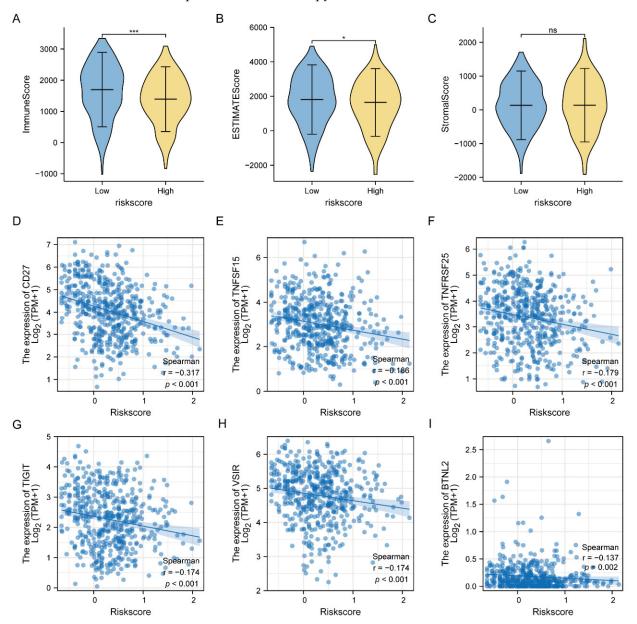


Figure 9. Correlation analysis of six-gene signature risk score with the ESTIMATE score and immune checkpoints. (**A**–**C**) Immune score, ESTIMATE score, and stromal score between low risk-group and high-risk group of six-gene signature, respectively. * p < 0.05; *** p < 0.001. (**D**–**I**) Correlation analysis of immune checkpoints (CD27, TNFSF15, TNFRSF25, TIGIT, VSIR, and BTNL2) with risk score, respectively. p < 0.001.

3.8. Verification of Hub Genes

The mRNA level of CR2, INSL4, FGF5, RAET1L and TNFRSF13B was upregulated in LUAD tumor tissues compared with paired normal tissues whereas AGER was downregulated (Figure 10A). The protein distribution from the HPA database showed the consistency at protein level of CR2, TNFRSF13B, and AGER in LUAD tissues without the other three other genes present for the immunohistochemistry staining as they were not available in the HPA (Figure 10B). In addition, we also measured the mRNA levels of the six hub genes in LUAD cell lines, including A549, NCI-H1975, and Calu-3 as shown in Figure 10C, of which the patterns were consistent with the RNA level of tumor tissues. We also analyzed genetic alterations of the hub genes in LUAD via cBioportal. There were 16% (186/1144) genetic alterations in total, and CR2 was found to have the highest alteration (8%) in LUAD patients (Figure 10D).

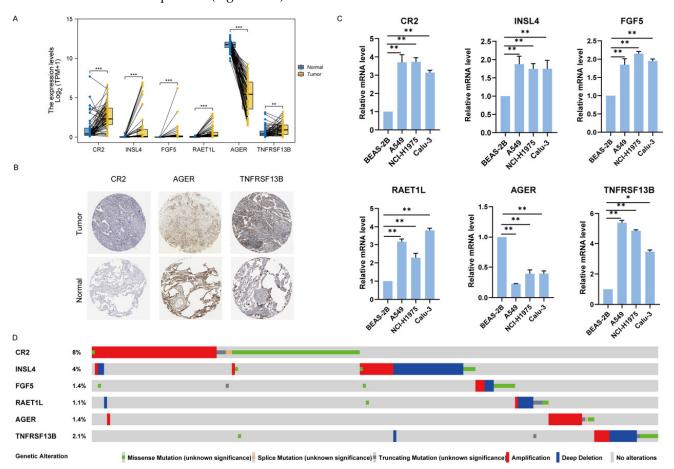


Figure 10. Verification of the expression and genetic alterations of 6 hub genes in tumor and normal tissues. (**A**) The boxplot showed the expression of 6 hub genes between paired LUAD tumor and normal tissues in TCGA database. (**B**) The immunohistochemical staining of CR2, AFER, TNFRSF13B was obtained from HPA database. (**C**) The mRNA levels of hub genes were measured by qRT-PCR in lung adenocarcinoma cell lines (A549, NCI-H1975, and Calu-3) and compared with human normal lung epithelial cells BEAS-2B. (**D**) Genetic alterations of hub genes in LUAD in TCGA datasets. * p < 0.05, ** p < 0.01, *** p < 0.001.

4. Discussions

In recent years, the development of cancer treatment strategies has been toward immunotherapy [30,31]. Cell-mediated immunity has been shown to play a role in both lung cancer monitoring and development. As studies on first-line immunotherapy for NSCLC proceeds, it is more and more challenging to choose the most suitable therapeutic treatment for patients with different clinicopathological features from among such options as immunotherapy alone, or chemotherapy combined with immunotherapy, or immunotherapy combined with other treatment [32]. Thus, it is essential and urgent to develop an immune-related prognostic model to better predict the prognosis and evaluate the efficacy of immunotherapy in LUAD patients. In this study, we identified 40 significantly different immune- related genes between LUAD tumors and normal tissues. We established a prognostic risk model that included six hub genes as determined by LASSO, univariate, and multivariate regression analyses. Moreover, the gene signature was found to be highly associated with tumor-infiltrated immune cells, which might be used as an independent factor. The prognostic risk model consisting of CR2, INSL4, FGF5, RAET1L, AGER, and TNFRSF13C acts as a novel potential biomarker for evaluating the prognosis and the efficacy of immunotherapy in LUAD.

Nowadays, it has become a growing focus for research to construct prognostic models based on lncRNA, miRNA, and mRNA to predict the prognostic or diagnostic value in tumors [33]. In this study, the gene signature was established by taking the cross from LUAD transcriptome differential expressed genes and immune-related genes. We found a total of 40 genes that were involved in the "humoral immune response" and "defense response to bacterium", suggesting that the gene signature was closely related to immune response. Next, LASSO and regression analyses were performed to identify 6 out of these 40 immune-related genes to construct a prognostic risk model, which included CR2, INSL4, FGF5, RAET1L, AGER, and TNFRSF13C. Therein, the complement receptor type 2 (CR2) was reported to combine with the B-Cell receptor to inhibit the activation, proliferation, and antibody production of human B cells [34]. CR2 was also verified to be involved in the nine-gene signature with the prognostic and predicted values for LUAD [35]. Insulin-like 4 (INSL4) has been reported to be required for the growth and viability of LKB1-inactivated lung cancer [36]. A recent study showed that INSL4 might be a prognostic marker for proliferation and invasiveness in non-small cell lung cancer (NSCLC) [37]. It was reported that FGF5 could promote cell proliferation via activation of the MAPK signaling pathway in osteosarcoma [38]. Moreover, high FGF5 expression was found to be significantly associated with poor overall survival and relapse-free survival in LUAD [39]. A report showed that RAET1L together with ULBP1, ULBP2, ULBP3 had the high diagnostic values in colon adenocarcinoma (COAD) [40]. AGER overexpression was found to suppress the cell proliferation, invasion and migration of H1299 cells [41]. TNFRSF13C has been demonstrated to be involved in a five-gene signature (CD40LG, TNFRSF6B, TNFSF13, TNFRSF13C, and TNFRSF19) which was helpful for the prognosis and immunotherapy response prediction in LUAD [42]. These previous studies have revealed that several genes were closely related to lung adenocarcinoma, which also further confirmed the validity and reliability of our prognostic six-gene risk model.

Studies have shown that a tumor microenvironment plays a key role in NSCLC development and is closely related to immunotherapy response [43,44]. Given the importance of tumor immune infiltration, we performed the correlation analysis of immune infiltrates using ssGSEA and stromal score algorithms. Recent studies have shown that Th2 cells have tumor-promoting effects in lung cancer and even human primary NSCLC tumors [45–47]. Existing evidence shows that tumor-infiltrating B cells play a role in almost all stages of lung cancer [48,49]. Consistent with the above findings, our work showed that the risk score (six-gene signature) was highly correlated with Th2 and B cells, suggesting a potential role in immune response. Moreover, we identified that higher infiltrations of B cells and dendritic cells were associated with a better cumulative survival of LUAD patients, findings which were well in line with recent research showing that a gene expression signature associated with B cells was correlated with survival using immunotherapy in lung adenocarcinoma [50]. Our findings were identical to a clinical trial where personalized neoantigen pulsed dendritic cell vaccines were administered for advanced lung cancer [51]. Above all, the six-gene signature risk model constructed in this study may have a better prognosis value and clinical significance in LUAD. Furthermore, we built a nomogram to assess the prediction accuracy by linking the risk model with T stage, N stage, age, and

gender. Similarly, it was reported that a nomogram was also used to enhance the prognostic prediction accuracy by combining the risk index in HCC patients with clinicopathological features [16]. Thus, it is of significance to provide and improve the reliability and efficacy of survival risk prediction for LUAD patients by using the novel prognostic model.

The present study identified an original immune-related risk model for LUAD. There are some highlights in this study. First, we did a comprehensive analysis of the six-gene signature in LUAD, including the functional enrichment, the correlation with clinical features, the tumor-infiltrated immune cells, the immune checkpoints, and the expression profiles in tissues and cell lines. Second, a nomogram and calibration were used to further validate the accuracy for clinical outcomes for the risk model in LUAD. Nevertheless, there are some limitations in this study. First, clinical LUAD samples need to be collected to verify the validity and accuracy of the prognostic model. Second, relevant functional and mechanical experiments should be carried out to further explore the effect of this risk model in vitro and in vivo.

5. Conclusions

In conclusion, we established an immune-related prognostic gene signature which was significantly associated with the tumor immune microenvironment. The six-gene signature model could predict the prognosis and survival of LUAD patients, as well as imply a better response to immune-based therapies, which may provide a theoretical basis for the prognosis and immunotherapy of LUAD.

Author Contributions: H.F. and Z.Z. designed the project. C.L. and Y.S. performed the experiments. C.L., Y.S. and J.Z. performed the data analysis. H.F. wrote the manuscript. D.L. and Z.Z. revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Elite Medical Professionals Project of China-Japan Friendship Hospital (No. ZRJY2021-TD04) and National High Level Hospital Clinical Research Funding (NO.2022-NHLHCRF-YS-04).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets used in the current study are available, and can be found in the article.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

GO	Gene Ontology
BP	Biological process
MF	Molecular function
CC	Cellular component
KEGG	Kyoto Encyclopedia of Genes and Genomes
HRs	Hazard ratios
DEGs	Differentially expressed genes
IRGs	Immune-related gene
DEIRGs	Differently expressed immune-related genes
OS	Overall Survival
DSS	Disease Specific Survival
PFI	Progression-Free Interval

References

- Bade, B.C.; dela Cruz, C.S. Lung Cancer 2020: Epidemiology, Etiology, and Prevention. *Clin. Chest Med.* 2020, 41, 1–24. [CrossRef] [PubMed]
- 2. Wu, F.; Wang, L.; Zhou, C. Lung Cancer in China: Current and Prospect. Curr. Opin. Oncol. 2021, 33, 40–46. [CrossRef] [PubMed]
- 3. Nooreldeen, R.; Bach, H. Current and Future Development in Lung Cancer Diagnosis. Int. J. Mol. Sci. 2021, 22, 8861. [CrossRef]

- 4. The Lancet. Lung Cancer: Some Progress, but Still a Lot More to Do. Lancet 2019, 394, 1880. [CrossRef]
- Herbst, R.S.; Morgensztern, D.; Boshoff, C. The Biology and Management of Non-Small Cell Lung Cancer. *Nature* 2018, 553, 446–454. [CrossRef] [PubMed]
- Remark, R.; Becker, C.; Gomez, J.E.; Damotte, D.; Dieu-Nosjean, M.C.; Sautès-Fridman, C.; Fridman, W.H.; Powell, C.A.; Altorki, N.K.; Merad, M.; et al. The Non-Small Cell Lung Cancer Immune Contexture: A Major Determinant of Tumor Characteristics and Patient Outcome. Am. J. Respir. Crit. Care Med. 2015, 191, 446–454. [CrossRef]
- Kadara, H.; Choi, M.; Zhang, J.; Parra, E.R.; Rodriguez-Canales, J.; Gaffney, S.G.; Zhao, Z.; Behrens, C.; Fujimoto, J.; Chow, C.; et al. Whole-Exome Sequencing and Immune Profiling of Early-Stage Lung Adenocarcinoma with Fully Annotated Clinical Follow-Up. *Ann. Oncol.* 2017, 28, 377–390. [CrossRef]
- Marzagalli, M.; Ebelt, N.D.; Manuel, E.R. Unraveling the Crosstalk between Melanoma and Immune Cells in the Tumor Microenvironment. Semin. Cancer. Biol. 2019, 59, 75–82. [CrossRef]
- 9. Pitt, J.M.; Marabelle, A.; Eggermont, A.; Soria, J.C.; Kroemer, G.; Zitvogel, L. Targeting the Tumor Microenvironment: Removing Obstruction to Anticancer Immune Responses and Immunotherapy. *Ann. Oncol.* **2016**, *27*, 236–250. [CrossRef]
- Stankovic, B.; Bjørhovde, H.A.K.; Skarshaug, R.; Aamodt, H.; Frafjord, A.; Müller, E.; Hammarström, C.; Beraki, K.; Bækkevold, E.S.; Woldbæk, P.R.; et al. Immune Cell Composition in Human Non-Small Cell Lung Cancer. *Front. Immunol.* 2018, 9, 1482–1492. [CrossRef]
- Usó, M.; Jantus-Lewintre, E.; Bremnes, R.M.; Calabuig, S.; Blasco, A.; Pastor, E.; Borreda, I.; Molina-Pinelo, S.; Paz-Ares, L.; Guijarro, R.; et al. Analysis of the Immune Microenvironment in Resected Non-Small Cell Lung Cancer: The Prognostic Value of Different T Lymphocyte Markers. *Oncotarget* 2016, 7, 3101. [CrossRef] [PubMed]
- Bremnes, R.M.; Busund, L.T.; Kilver, T.L.; Andersen, S.; Richardsen, E.; Paulsen, E.E.; Hald, S.; Khanehkenari, M.R.; Cooper, W.A.; Kao, S.C.; et al. The Role of Tumor-Infiltrating Lymphocytes in Development, Progression, and Prognosis of Non-Small Cell Lung Cancer. J. Thorac. Oncol. 2016, 11, 52849–52861. [CrossRef] [PubMed]
- Wang, H.; Xia, L.; Yao, C.C.; Dong, H.; Yang, Y.; Li, C.; Ji, W.X.; Sun, R.M.; Duan, H.Q.; Mengzhou, W.; et al. NLRP4 Negatively Regulates Type I Interferon Response and Influences the Outcome in Anti-Programmed Cell Death Protein (PD)-1/PD-Ligand 1 Therapy. *Cancer Sci.* 2022, 113, 789–800. [CrossRef] [PubMed]
- Perez-Ruiz, E.; Minute, L.; Otano, I.; Alvarez, M.; Ochoa, M.C.; Belsue, V.; de Andrea, C.; Rodriguez-Ruiz, M.E.; Perez-Gracia, J.L.; Marquez-Rodas, I.; et al. Prophylactic TNF Blockade Uncouples Efficacy and Toxicity in Dual CTLA-4 and PD-1 Immunotherapy. *Nature* 2019, 569, 838–851. [CrossRef]
- 15. Naidoo, J.; Page, D.B.; Wolchok, J.D. Immune Checkpoint Blockade. Hematol. Oncol. Clin. N. Am. 2014, 28, 428–432. [CrossRef]
- 16. Dai, Y.; Qiang, W.; Lin, K.; Gui, Y.; Lan, X.; Wang, D. An Immune-Related Gene Signature for Predicting Survival and Immunotherapy Efficacy in Hepatocellular Carcinoma. *Cancer Immunol. Immunother.* **2021**, *70*, 967–979. [CrossRef]
- 17. Zhang, M.; Wang, X.; Chen, X.; Zhang, Q.; Hong, J. Novel Immune-Related Gene Signature for Risk Stratification and Prognosis of Survival in Lower-Grade Glioma. *Front. Genet.* **2020**, *11*, 363. [CrossRef]
- Xiao, B.; Liu, L.; Li, A.; Xiang, C.; Wang, P.; Li, H.; Xiao, T. Identification and Verification of Immune-Related Gene Prognostic Signature Based on SsGSEA for Osteosarcoma. *Front. Oncol.* 2020, 10, 607622. [CrossRef]
- 19. Sun, S.; Guo, W.; Wang, Z.; Wang, X.; Zhang, G.; Zhang, H.; Li, R.; Gao, Y.; Qiu, B.; Tan, F.; et al. Development and Validation of an Immune-Related Prognostic Signature in Lung Adenocarcinoma. *Cancer Med.* **2020**, *9*, 5960–5975. [CrossRef]
- Dranoff, G. Immunotherapy at Large: Balancing Tumor Immunity and Inflammatory Pathology. Nat. Med. 2013, 19, 1100–1101. [CrossRef]
- Love, M.I.; Huber, W.; Anders, S. Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2. *Genome Biol.* 2014, 15, 550. [CrossRef] [PubMed]
- Bhattacharya, S.; Dunn, P.; Thomas, C.G.; Smith, B.; Schaefer, H.; Chen, J.; Hu, Z.; Zalocusky, K.A.; Shankar, R.D.; Shen-Orr, S.S.; et al. ImmPort, toward Repurposing of Open Access Immunological Assay Data for Translational and Clinical Research. *Sci. Data* 2018, *5*, 180015. [CrossRef] [PubMed]
- Yu, G.; Wang, L.G.; Han, Y.; He, Q.Y. ClusterProfiler: An R Package for Comparing Biological Themes among Gene Clusters. Omics J. Integr. Biol. 2012, 16, 284–287. [CrossRef] [PubMed]
- Subramanian, A.; Tamayo, P.; Mootha, V.K.; Mukherjee, S.; Ebert, B.L.; Gillette, M.A.; Paulovich, A.; Pomeroy, S.L.; Golub, T.R.; Lander, E.S.; et al. Gene Set Enrichment Analysis: A Knowledge-Based Approach for Interpreting Genome-Wide Expression Profiles. *Proc. Natl. Acad. Sci. USA* 2005, 102, 15545–15550. [CrossRef]
- Liu, J.; Lichtenberg, T.; Hoadley, K.A.; Poisson, L.M.; Lazar, A.J.; Cherniack, A.D.; Kovatich, A.J.; Benz, C.C.; Levine, D.A.; Lee, A.V.; et al. An Integrated TCGA Pan-Cancer Clinical Data Resource to Drive High-Quality Survival Outcome Analytics. *Cell* 2018, 173, 400–416.e411. [CrossRef]
- Hänzelmann, S.; Castelo, R.; Guinney, J. GSVA: Gene Set Variation Analysis for Microarray and RNA-Seq Data. BMC Bioinform. 2013, 14, 7. [CrossRef]
- Bindea, G.; Mlecnik, B.; Tosolini, M.; Kirilovsky, A.; Waldner, M.; Obenauf, A.C.; Angell, H.; Fredriksen, T.; Lafontaine, L.; Berger, A.; et al. Spatiotemporal Dynamics of Intratumoral Immune Cells Reveal the Immune Landscape in Human Cancer. *Immunity* 2013, 39, 782–795. [CrossRef]

- Yoshihara, K.; Shahmoradgoli, M.; Martínez, E.; Vegesna, R.; Kim, H.; Torres-Garcia, W.; Treviño, V.; Shen, H.; Laird, P.W.; Levine, D.A.; et al. Inferring Tumour Purity and Stromal and Immune Cell Admixture from Expression Data. *Nat. Commun.* 2013, 4, 2612. [CrossRef]
- 29. Gao, J.; Aksoy, B.A.; Dogrusoz, U.; Dresdner, G.; Gross, B.; Sumer, S.O.; Sun, Y.; Jacobsen, A.; Sinha, R.; Larsson, E.; et al. Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the CBioPortal. *Sci. Signal.* **2013**, *6*, pl1. [CrossRef]
- Abbott, M.; Ustoyev, Y. Cancer and the Immune System: The History and Background of Immunotherapy. *Semin. Oncol. Nurs.* 2019, 35, 150923. [CrossRef] [PubMed]
- 31. Steven, A.; Fisher, S.A.; Robinson, B.W. Immunotherapy for Lung Cancer. Respirology 2016, 21, 821–833. [CrossRef] [PubMed]
- Zhou, Y.; Zhang, Y.; Guo, G.; Cai, X.; Yu, H.; Cai, Y.; Zhang, B.; Hong, S.; Zhang, L. Nivolumab plus Ipilimumab versus Pembrolizumab as Chemotherapy-free, First-line Treatment for PD-L1-positive Non-small Cell Lung Cancer. *Clin. Transl. Med.* 2020, 10, 107–115. [CrossRef] [PubMed]
- 33. Pajouheshnia, R.; Damen, J.A.A.G.; Groenwold, R.H.H.; Moons, K.G.M.; Peelen, L.M. Treatment Use in Prognostic Model Research: A Systematic Review of Cardiovascular Prognostic Studies. *Diagn. Progn. Res.* **2017**, *1*, 15. [CrossRef] [PubMed]
- Kovács, K.G.; Mácsik-Valent, B.; Matkó, J.; Bajtay, Z.; Erdei, A. Revisiting the Coreceptor Function of Complement Receptor Type 2 (CR2, CD21); Coengagement with the B-Cell Receptor Inhibits the Activation, Proliferation, and Antibody Production of Human B Cells. *Front. Immunol.* 2021, 12, 620427. [CrossRef] [PubMed]
- 35. Zhong, H.; Wang, J.; Zhu, Y.; Shen, Y. Comprehensive Analysis of a Nine-Gene Signature Related to Tumor Microenvironment in Lung Adenocarcinoma. *Front. Cell Dev. Biol.* **2021**, *9*, 700607. [CrossRef]
- 36. Yang, R.; Li, S.W.; Chen, Z.; Zhou, X.; Ni, W.; Fu, D.A.; Lu, J.; Kaye, F.J.; Wu, L. Role of INSL4 Signaling in Sustaining the Growth and Viability of LKB1-Inactivated Lung Cancer. *J. Natl. Cancer Inst.* **2018**, *111*, 664–674. [CrossRef]
- Scopetti, D.; Piobbico, D.; Brunacci, C.; Pieroni, S.; Bellezza, G.; Castelli, M.; Ludovini, V.; Tofanetti, F.R.; Cagini, L.; Sidoni, A.; et al. INSL4 as Prognostic Marker for Proliferation and Invasiveness in Non-Small-Cell Lung Cancer. J. Cancer 2021, 12, 3781–3795. [CrossRef]
- Han, D.; Wang, M.; Yu, Z.; Yin, L.; Liu, C.; Wang, J.; Liu, Y.; Jiang, S.; Ren, Z.; Yin, J. FGF5 Promotes Osteosarcoma Cells Proliferation via Activating MAPK Signaling Pathway. *Cancer Manag. Res.* 2019, 11, 6457–6466. [CrossRef]
- 39. Zhao, T.; Qian, K.; Zhang, Y. High Expression of FGF5 Is an Independent Prognostic Factor for Poor Overall Survival and Relapse-Free Survival in Lung Adenocarcinoma. *J. Comput. Biol.* **2020**, *27*, 948–957. [CrossRef]
- Ruan, G.T.; Wang, S.; Zhu, L.C.; Liao, X.W.; Wang, X.K.; Liao, C.; Yan, L.; Xie, H.L.; Gong, Y.Z.; Gan, J.L.; et al. Investigation and Verification of the Clinical Significance and Perspective of Natural Killer Group 2 Member D Ligands in Colon Adenocarcinoma. *Aging* 2021, 13, 12565–12586. [CrossRef]
- Wang, Q.; Zhu, W.; Xiao, G.; Ding, M.; Chang, J.; Liao, H. Effect of AGER on the Biological Behavior of Non-Small Cell Lung Cancer H1299 Cells. *Mol. Med. Rep.* 2020, 22, 810–818. [CrossRef] [PubMed]
- Zhang, C.; Zhang, Z.; Sun, N.; Zhang, Z.; Zhang, G.; Wang, F.; Luo, Y.; Che, Y.; He, J. Identification of a Costimulatory Molecule-Based Signature for Predicting Prognosis Risk and Immunotherapy Response in Patients with Lung Adenocarcinoma. *OncoImmunology* 2020, *9*, 1824641. [CrossRef] [PubMed]
- 43. Wood, S.L.; Pernemalm, M.; Crosbie, P.A.; Whetton, A.D. The Role of the Tumor-Microenvironment in Lung Cancer-Metastasis and Its Relationship to Potential Therapeutic Targets. *Cancer Treat. Rev.* **2014**, *40*, 558–566. [CrossRef] [PubMed]
- 44. Saab, S.; Zalzale, H.; Rahal, Z.; Khalifeh, Y.; Sinjab, A.; Kadara, H. Insights into Lung Cancer Immune-Based Biology, Prevention, and Treatment. *Front. Immunol.* **2020**, *11*, 159. [CrossRef] [PubMed]
- Zhong, R.; Zhang, Y.; Chen, D.; Cao, S.; Han, B.; Zhong, H. Single-Cell RNA Sequencing Reveals Cellular and Molecular Immune Profile in a Pembrolizumab-Responsive PD-L1-Negative Lung Cancer Patient. *Cancer Immunol. Immunother.* 2021, 70, 2261–2274. [CrossRef]
- Duan, M.C.; Zhong, X.N.; Liu, G.N.; Wei, J.R. The Treg/Th17 Paradigm in Lung Cancer. J. Immunol. Res. 2014, 2014, 730380. [CrossRef]
- Frafjord, A.; Buer, L.; Hammarström, C.; Aamodt, H.; Woldbæk, P.R.; Brustugun, O.T.; Helland, Å.; Øynebråten, I.; Corthay, A. The Immune Landscape of Human Primary Lung Tumors Is Th2 Skewed. *Front. Immunol.* 2021, 12, 764596. [CrossRef]
- Wang, S.S.; Liu, W.; Ly, D.; Xu, H.; Qu, L.; Zhang, L. Tumor-Infiltrating B Cells: Their Role and Application in Anti-Tumor Immunity in Lung Cancer. Cell. Mol. Immunol. 2018, 16, 6–18. [CrossRef]
- Chen, J.; Tan, Y.; Sun, F.; Hou, L.; Zhang, C.; Ge, T.; Yu, H.; Wu, C.; Zhu, Y.; Duan, L.; et al. Single-Cell Transcriptome and Antigen-Immunoglobin Analysis Reveals the Diversity of B Cells in Non-Small Cell Lung Cancer. *Genome Biol.* 2020, 21, 152. [CrossRef]
- Budczies, J.; Kirchner, M.; Kluck, K.; Kazdal, D.; Glade, J.; Allgäuer, M.; Kriegsmann, M.; Heußel, C.P.; Herth, F.J.; Winter, H.; et al. A Gene Expression Signature Associated with B Cells Predicts Benefit from Immune Checkpoint Blockade in Lung Adenocarcinoma. Oncoimmunology 2021, 10, 1860586. [CrossRef]
- 51. Ding, Z.; Li, Q.; Zhang, R.; Xie, L.; Shu, Y.; Gao, S.; Wang, P.; Su, X.; Qin, Y.; Wang, Y.; et al. Personalized Neoantigen Pulsed Dendritic Cell Vaccine for Advanced Lung Cancer. *Signal Transduct. Target. Ther.* **2021**, *6*, 26. [CrossRef] [PubMed]