



Establishment of a Prognostic Model of Lung Adenocarcinoma Based on Tumor Heterogeneity

Pengdou Zheng[†], Huojun Zhang[†], Weiling Jiang, Lingling Wang, Lu Liu, Yuhao Zhou, Ling Zhou* and Huiguo Liu*

Department of Respiratory and Critical Care Medicine, Key Laboratory of Pulmonary Diseases of Health Ministry, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

OPEN ACCESS

Edited by:

Stella Liberman-Aronov,
Ariel University, Israel

Reviewed by:

Pinar Pir,
Gebze Technical University, Turkey
Ravi Kumar Alluri,
Cleveland Clinic, United States
Shilpita Karmakar,
Jackson Laboratory, United States

*Correspondence:

Ling Zhou
471699217@qq.com
Huiguo Liu
huiguol@163.com

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Molecular Diagnostics and
Therapeutics,
a section of the journal
Frontiers in Molecular Biosciences

Received: 02 November 2021

Accepted: 15 March 2022

Published: 11 April 2022

Citation:

Zheng P, Zhang H, Jiang W, Wang L,
Liu L, Zhou Y, Zhou L and Liu H (2022)
Establishment of a Prognostic Model of
Lung Adenocarcinoma Based on
Tumor Heterogeneity.
Front. Mol. Biosci. 9:807497.
doi: 10.3389/fmolb.2022.807497

Lung cancer is one of the main cancer types due to its persistently high incidence and mortality, yet a simple and effective prognostic model is still lacking. This study aimed to identify independent prognostic genes related to the heterogeneity of lung adenocarcinoma (LUAD), generate a prognostic risk score model, and construct a nomogram in combination with other pathological characteristics to predict patients' overall survival (OS). A significant amount of data pertaining to single-cell RNA sequencing (scRNA-seq), RNA sequencing (RNA-seq), and somatic mutation were used for data mining. After statistical analyses, a risk scoring model was established based on eight independent prognostic genes, and the OS of high-risk patients was significantly lower than that of low-risk patients. Interestingly, high-risk patients were more sensitive and effective to immune checkpoint blocking therapy. In addition, it was noteworthy that CCL20 not only affected prognosis and differentiation of LUAD but also led to poor histologic grade of tumor cells. Ultimately, combining risk score, clinicopathological information, and CCL20 mutation status, a nomogram with good predictive performance and high accuracy was established. In short, our research established a prognostic model that could be used to guide clinical practice based on the constantly updated big multi-omics data. Finally, this analysis revealed that CCL20 may become a potential therapeutic target for LUAD.

Keywords: LUAD, prognosis model, tumor microenvironment, tumor differentiation, CCL20

INTRODUCTION

Lung cancer has been one of the most leading causes of cancer death worldwide. According to data from the American Cancer Society in 2021, there are approximately 235,760 new cases and 131,880 deaths, and lung cancer ranks second and first in morbidity and mortality, respectively, among malignant tumors (Siegel et al., 2021). Based on histological subtypes, lung cancer is divided into small cell lung cancer and non-small cell lung cancer (NSCLC). NSCLC represents 85% of all lung cancers and includes lung adenocarcinoma (LUAD), lung squamous cell carcinoma, and large cell carcinoma. Among them, LUAD is the largest NSCLC subgroup, accounting for 40% of all lung tumors (Denisenko et al., 2018). Despite advances in surgical resection, radiotherapy, chemotherapy, immunotherapy, and targeted therapy over the past decades, the prognosis of patients with LUAD is still very poor, with a 5-year survival rate of 15% (Herbst et al., 2018; Zhang et al., 2020a; Isla et al., 2020; Ruiz-Cordero and Devine, 2020).

Therefore, it is necessary to explore this condition at the molecular level to enable accurate diagnosis, treatment, and prognosis assessments.

It is a scientific consensus that tumors are highly heterogeneous (de Sousa and Carvalho, 2018). Heterogeneity makes sensitive cells die as a result of the drug's cytotoxicity, whereas drug-resistant cells survive due to mutations, gene expression changes, alternative splicing, etc., resulting in a poor prognosis of patients and also leading to different sensitivity levels and subsequently different treatment schemes (Aleksakhina et al., 18722019; Dagogo-Jack and Shaw, 2018; Vasan et al., 2019; Marusyk et al., 2020; Narayanan et al., 2020). Therefore, an accurate understanding of tumor heterogeneity is of great significance for assessing the prognosis and for selecting a reasonable and feasible treatment plan. To some extent, the emergence of heterogeneity can be attributed to the evolution of tumors (Lim and Ma, 2019). Single-cell transcriptomics analysis provides novel ideas for solving this problem. The Monocle 2 algorithm uses the reverse graph embedding method to project high-dimensional data into a low-dimensional space, learn cell trajectories, and different cell evolution according to pseudo-time values, thereby successfully constructing differentiation trajectories (Trapnell et al., 2014; Qiu et al., 2017a; Qiu et al., 2017b). Monocle2 has been successfully applied to many types of cancers, such as hepatocellular carcinoma, glioblastoma, multiple myeloma, laryngeal carcinoma, etc., thus verifying the authority, practicability, and reliability of the algorithm (Liang et al., 2020; Ionkina et al., 2021). Unfortunately, monocle2 has not been widely used to screen prognostic factors associated with LUAD.

Considering this issue from a micro perspective, the heterogeneity of tumors can also be partially explained by the tumor microenvironment. Malignant solid tumor tissues include tumor cells but also other types of cells, such as stromal cells, immune cells, vascular cells, etc. (Wu and Dai, 2017). A series of studies have reported that immune cells have a complex impact on tumor growth, disease progression, and drug resistance (Wang et al., 2019a; Greten and Grivennikov, 2019; Hinshaw and Shevde, 2019). Interestingly, recent studies have shown that tumor purity, which represents the proportion of tumor cells in all cells residing in the tumor microenvironment, is related to prognosis. These studies consider that low tumor purity will cause tumor immune escape, which is associated to unfavorable prognosis (Zhang et al., 2017; Gong et al., 2020). In this article, the ESTIMATE (Yoshihara et al., 2013) and CIBERSORT (Newman et al., 2015) algorithms can help us better understand the influence of various immune components and tumor purity on the prognosis of LUAD.

Aiming at the heterogeneity of tumors, many studies had attempted to establish LUAD prognosis prediction models from different aspects such as immune cells, ferroptosis-related genes, non-coding RNAs, and glycolysis genes, among which immune-related prognosis studies accounted for the majority (Liu et al., 2019; Li et al., 2020; Zuo et al., 2020; Gao et al., 2021; Zheng et al., 2021). For example, Yuwei Zhang et al. noted the correlation between tumor invasion environment, especially tumor immune system, and tumor

progression and treatment outcome, and constructed a 27-gene prognostic model based on multiple omics data (Zhang et al., 2020b). Chenghan Luo et al. trained a prognostic model based on two immune genes, ANLN and F2, by using information from 761 LUAD patients (Luo et al., 2020). Both studies noted the effect of tumor immune factors on patient prognosis, and constructed relatively satisfactory models with concordance indexes of 0.73 and 0.66, respectively. However, the analysis of these studies was still incomplete, factors such as tumor purity and the proportion of immune cells had not been taken into account, and the models still needed to be further improved. This study aims to explore the independent prognostic risk genes associated with heterogeneity in LUAD and to construct a prognostic model in combination with other basic clinical information that is capable of predicting the corresponding possible immunotherapy responses of high- and low-risk patients.

MATERIALS AND METHODS

Acquisition, Dimensionality Reduction and Cell Annotation of scRNA-Seq Data

The scRNA-seq data of three human primary LUAD samples with a reading depth of 10× genomics based on Illumina NextSeq 500, accession numbers GSM3304007, GSM3304011 and GSM3304013, were downloaded from the Gene Expression Omnibus (GEO, <http://www.ncbi.nlm.nih.gov/geo/>) database. After that, the scRNA-seq data was preliminarily filtered through the Seurat package in order to obtain higher quality cells. The cells and genes used for subsequent analysis needed to meet the following three quality control standards: 1) the gene was detected in at least three cells, 2) the cell had at least 50 genes detected, and 3) mitochondrial genes accounted for no more than 5%. The FeatureScatter function was run to analyze the correlation between the sequencing depth and total intracellular sequences. The data was normalized by LogNormalize and 1,500 highly variable genes among cells were selected by the default parameter vst for subsequent cell type identification. Principal component analysis (PCA), a multivariate statistical analysis method, was picked for preliminary linear dimensionality reduction (Lall et al., 2018). Each principal component (PC) essentially represented a “meta feature”, which combined relatively concentrated information of related features. Therefore, the PC at the top is more likely to represent the data set. In this study, we chose the first 15 PCs with a p value < 0.05 and further analyzed them through the non-linear dimensionality reduction technique the t-distributed stochastic neighbor embedding (t-SNE) algorithm to obtain the main clusters (Satija et al., 2015). With adjusted p value < 0.05 and $|\log_2[\text{fold change (FC)}]| > 1$ as the cutoff criterion, FindAllMarkers command was used to find the representative genes of each cluster. Finally, SingleR package was run to automatically annotate the corresponding cell type of each cluster according to the marker gene.

Pseudo-Time Analysis and Enrichment Analysis of Lung Adenocarcinoma Differentiation-Related Genes Base on Single-Cell RNA Sequencing Data

We used an unsupervised method to select genes and ran the Monocle2 algorithm to perform non-linear dimensionality reduction, that was, to project the obtained high-dimensional data matrix into a low-dimensional space (Qiu et al., 2017a). The de-embedding machine learning algorithm learned and constructed cell trajectories, and finally projected all cells onto a root and several main branches. Different branches corresponded to different directions of differentiation, and cells on the same branch were considered to have the same differentiation characteristics. Using adjusted p value < 0.05 and $|\log_2FC| > 1$ as the significance standard, the significantly differentially expressed genes among the main branches and the root were defined as LUAD differentiation-related genes (LDRGs). WEB-based GENE SeT AnaLysis Toolkit (WebGestalt, <http://www.webgestalt.org/>) was used for geneontology [GO, including cellular component (CC), biological process (BP), molecular function (MF)] and kyoto encyclopedia of genes and genomes (KEGG) analysis.

Acquisition and Processing of RNA Sequencing Data

RNA-seq data of LUAD samples from GEO and TCGA databases were used in this study. The RNA-seq data matrix composed of 443 LUAD samples was obtained from GSE68465 in the GEO (<http://www.ncbi.nlm.nih.gov/geo/>) database. This cohort was used for 1) molecular classification, 2) analysis of the correlation between prognosis and clinicopathological characteristics, and between prognosis and tumor microenvironment, and 3) verification of prognostic models. The RNA-seq data matrix composed of 535 LUAD samples and 59 normal samples was downloaded from the Cancer Genome Atlas (TCGA, <https://portal.gdc.cancer.gov/>) database for the construction of prognostic models.

Molecular Subtype Classification of Lung Adenocarcinoma Patients Based on Lung Adenocarcinoma Differentiation-Related Genes

After preprocessing, \log_2 -scale transformation and normalization of the RNA-seq matrix, the expression data of LDRGs were extracted for the following consistent clustering analysis. ConsensusClusterPlus, an R package based on resampling methods to verify the rationality of clustering and evaluate the stability of clusters, was used to perform consensus clustering on LUAD samples. The main parameters were 50 resampling, 0.8 sample sampling ratio, max $K = 9$, "pam" clustering algorithm and euclidean distance algorithm. The cumulative distribution function (CDF) curve helps researchers determine the number of subtypes in a visual form. Kaplan-Meier analysis

was chosen to compare survival differences between different clusters.

Tumor Microenvironment Scores, Immune Cell Infiltration and Immune Checkpoint Gene Expression Across Molecular Subtypes

The study calculated the immune/stromal scores and tumor purity of each sample through the ESTIMATE package (Yoshihara et al., 2013). The content of 22 immune cells in each sample from GSE68465 was identified by CIBERSORT package (Newman et al., 2015). Meanwhile, the infiltration density of immune cells and expression of 30 validated ICGs in different LUAD subtypes was compared by the limma package. Kaplan-Meier analysis was used to investigate the prognostic value of immune cells and ICGs.

Risk Scoring Model Generation and Validation

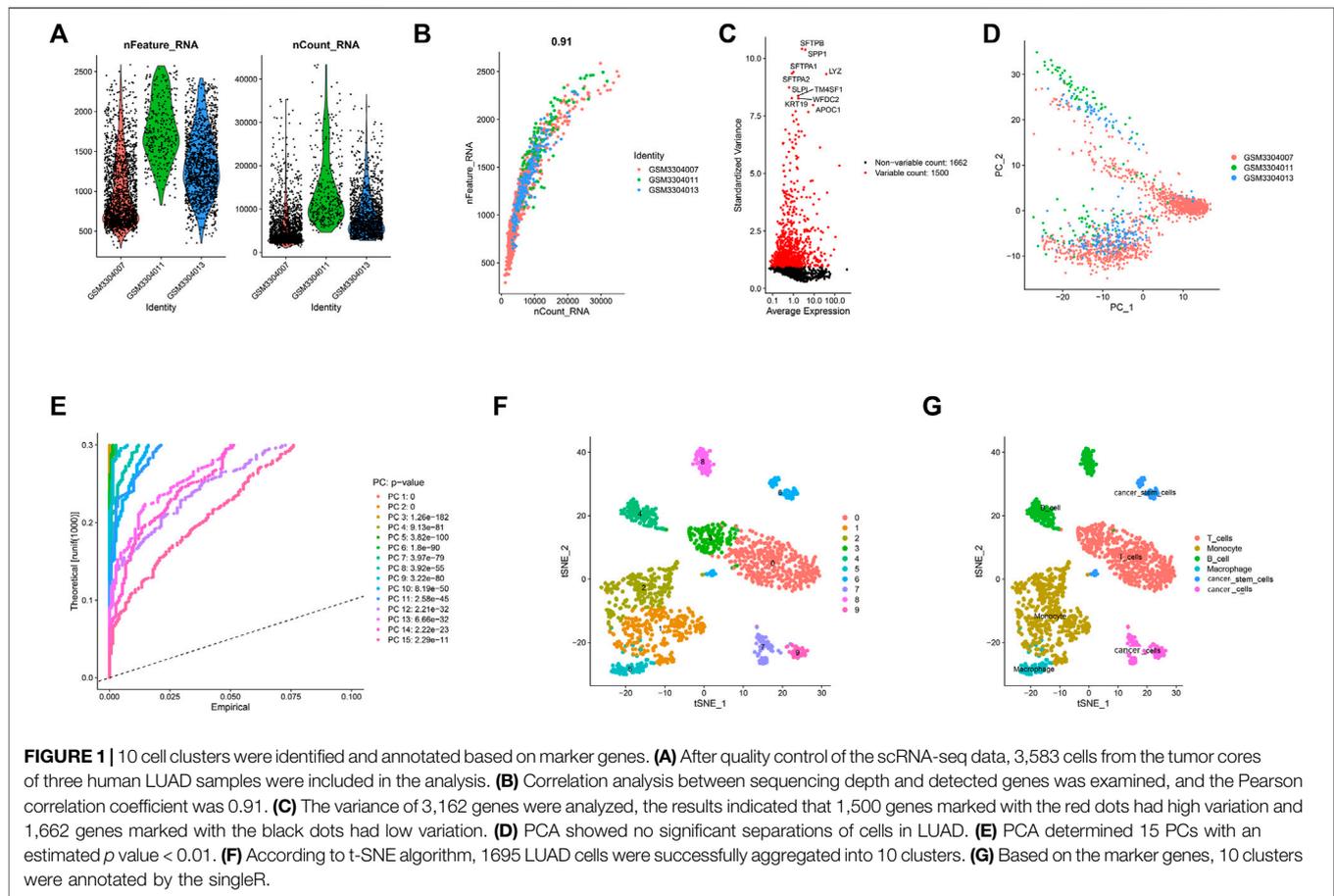
TCGA cohort was used as the training set to generate the risk scoring model, and GSE68465 dataset was used as the test set to verify the model. The expression data of LDRGs, poorly histological grade-related genes and ICGs in the training set and testing set were extracted, and the data were normalized and corrected by \log_2 -scale transformation. Genes with independent prognostic performance were screened by univariate analysis, LASSO regression and multivariate Cox analysis to generate a risk scoring model. The risk scoring could be calculated as the sum of the products of gene expression levels and coefficients. The patients were divided into high-risk groups and low-risk groups based on the mean value of the risk scores. Kaplan-Meier analysis and ROC curves were completed to evaluate model accuracy and predictive performance.

Establishment and Validation of a Nomogram Based on the The Cancer Genome Atlas Cohort

Based on TCGA cohort, age, gender, tumor purity, stage, risk score and CCL20 mutation status were included in univariate and multivariate analyses, respectively. Subsequently, independent prognostic variables were combined into nomograms to predict OS at 1-, 3-, and 5-year. ROC curves and calibration curves were used to evaluate the predictive performance and accuracy of the nomogram.

Immunotherapy Response

The tumor immune dysfunction and elimination (TIDE, <http://tide.dfci.harvard.edu/>) is a computational method that evaluates the prognostic effect of immune checkpoint blocking therapy. It is used in our study to determine the response of high- and low-risk LUAD patients to immunotherapy.



Statistical Analysis and Data Visualization

Statistical analysis and data visualization used R software (version 4.0.4) and Perl (version 5.28.1.0). Packages such as ggplot2, limma, edgeR, survival and venn. diagram in R software are used for data analysis and visualization.

RESULTS

Ten Cell Clusters Were Identified and Annotated Based on Marker Genes

After quality control and normalization of LUAD scRNA-seq data, 1,695 high-quality cells and 3,162 genes from LUAD cores were included for subsequent analysis in this study (Figure 1A). The number of genes detected was obviously positively correlated with the depth of sequencing (Figure 1B). To better distinguish cell types, we used variance analysis to identify the 1,500 highly variable genes (Figure 1C). Furthermore, PCA was used to reduce data dimensionality and screen related genes. The top 20 related genes from PCs are shown as dot plots and heatmaps in the Supplementary Figure S1. However, the PCA results found that there were no clear separations among LUAD cells (Figure 1D). Therefore, we selected 15 primary PCs with significant differences (p values < 0.01) for further analysis (Figure 1E). Based on t-SNE algorithm, 1,695 LUAD cells were successfully divided into ten clusters (Figure 1F), and the

application of differential analysis revealed 1,693 marker genes (Supplementary Figure S2). According to the marker genes, these clusters were annotated by singleR (Figure 1G). A total of 566 cells in clusters 0 and three were annotated as T cells, 598 cells in clusters 1 and 2 as monocytes, 216 cells in clusters 4 and 8 as B cells, 90 cells in cluster 5 as macrophages, 87 cells in cluster 6 as cancer stem cells, and 138 cells in clusters 7 and 9 as cancer cells.

Trajectory Analysis Identified Two Branches, and Gene Set Enrichment Analysis Analyzed Biological Functions

Pseudo-time and trajectory analysis suggested that all cells were distributed in one root and in two main branches (branches I and II). The results revealed that clusters 0, 3, 4, and 5 were mainly located in the roots, containing T and B cells; clusters 1, 2, and 6 were mainly located in branch I, containing macrophages and monocytes; and clusters 7, 8, and 9 were mainly located in branch II, containing cancer cells, cancer stem cells, and some B cells (Figure 2A). In addition, this work identified 240 LDRGs in branch I and 213 LDRGs in branch II (Supplementary Tables S1, S2). To understand the correlation between the molecular functions of LDRGs and the different differentiation trajectories, we used the WebGestalt online tool to perform GSEA. The LDRGs in branch I were dramatically associated with repressed immune-related pathway

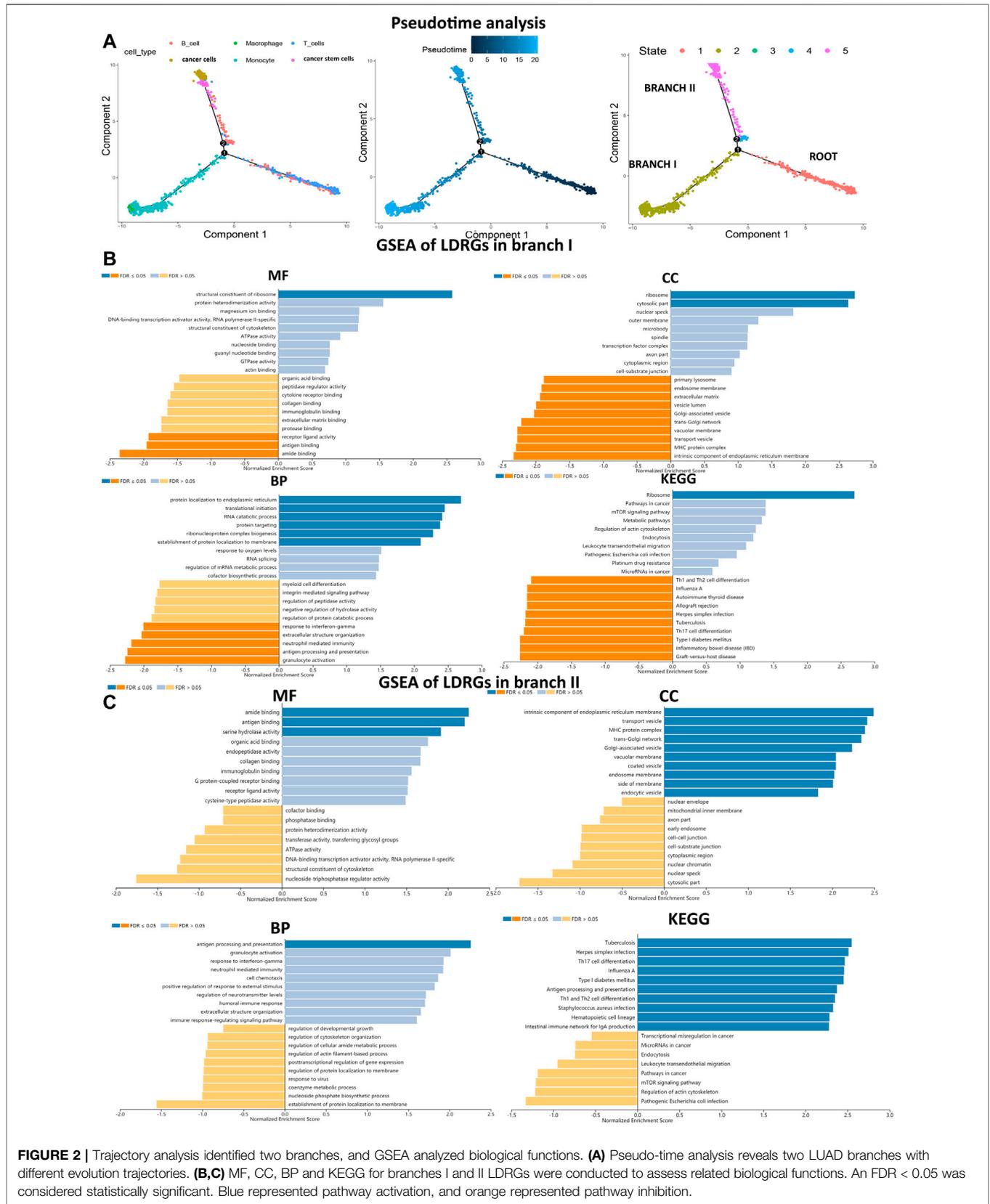


FIGURE 2 | Trajectory analysis identified two branches, and GSEA analyzed biological functions. **(A)** Pseudo-time analysis reveals two LUAD branches with different evolution trajectories. **(B,C)** MF, CC, BP and KEGG for branches I and II LDRGs were conducted to assess related biological functions. An FDR < 0.05 was considered statistically significant. Blue represented pathway activation, and orange represented pathway inhibition.

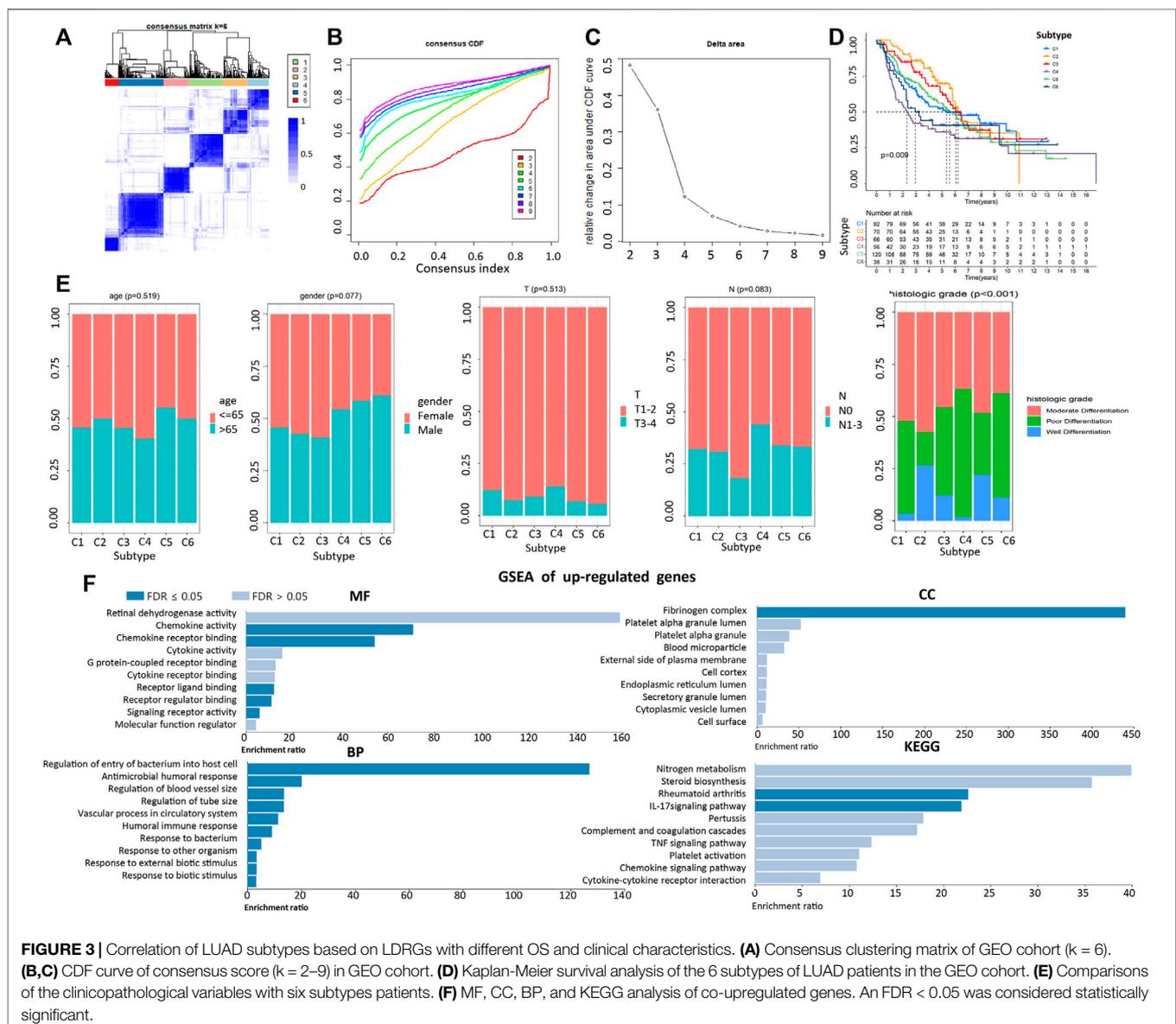
response and active ribosome activity, whereas the LDRGs in branch II were related with vibrant immune-related pathways and susceptibility to bacterial or viral infections. Collectively, these findings demonstrated that LUAD cells in different differentiated directions exhibited different biological tumor characteristics (Figures 2B,C).

Correlation Between Lung Adenocarcinoma Subtypes Based on Lung Adenocarcinoma Differentiation-Related Genes and Different Overall Survival and Clinical Characteristics

A LDRG-based consensus clustering analysis was performed on 443 LUAD patients in GSE68465, and all acquired samples were classified into six subtypes (k value = 6, Figures 3A–C).

Kaplan-Meier survival analysis revealed that subtype 2 (C2) had the highest overall survival (OS), whereas subtype 4 (C4) had the worst OS (log-rank $p = 9 \times 10^{-3}$, Figure 3D). In addition, we compared the clinicopathological characteristics of LUAD among different subtypes of patients in the GSE68465 cohort. As shown in Figure 3E, there were significant differences among the six subtypes pertaining to histologic grades. However, no significant differences were observed in other variables among all subtypes.

To compare genes related to differential histologic grades in different subtypes, with an adjusted p value < 0.05 and fold change > 2 set as parameters, the EdgeR was conducted to analyze differences between C4 and C2 and between C6 and C2. Taking the intersection, the results showed that there were 18 co-upregulated and 14 co-downregulated genes. GSEA indicated that upregulating genes were closely related to the increased



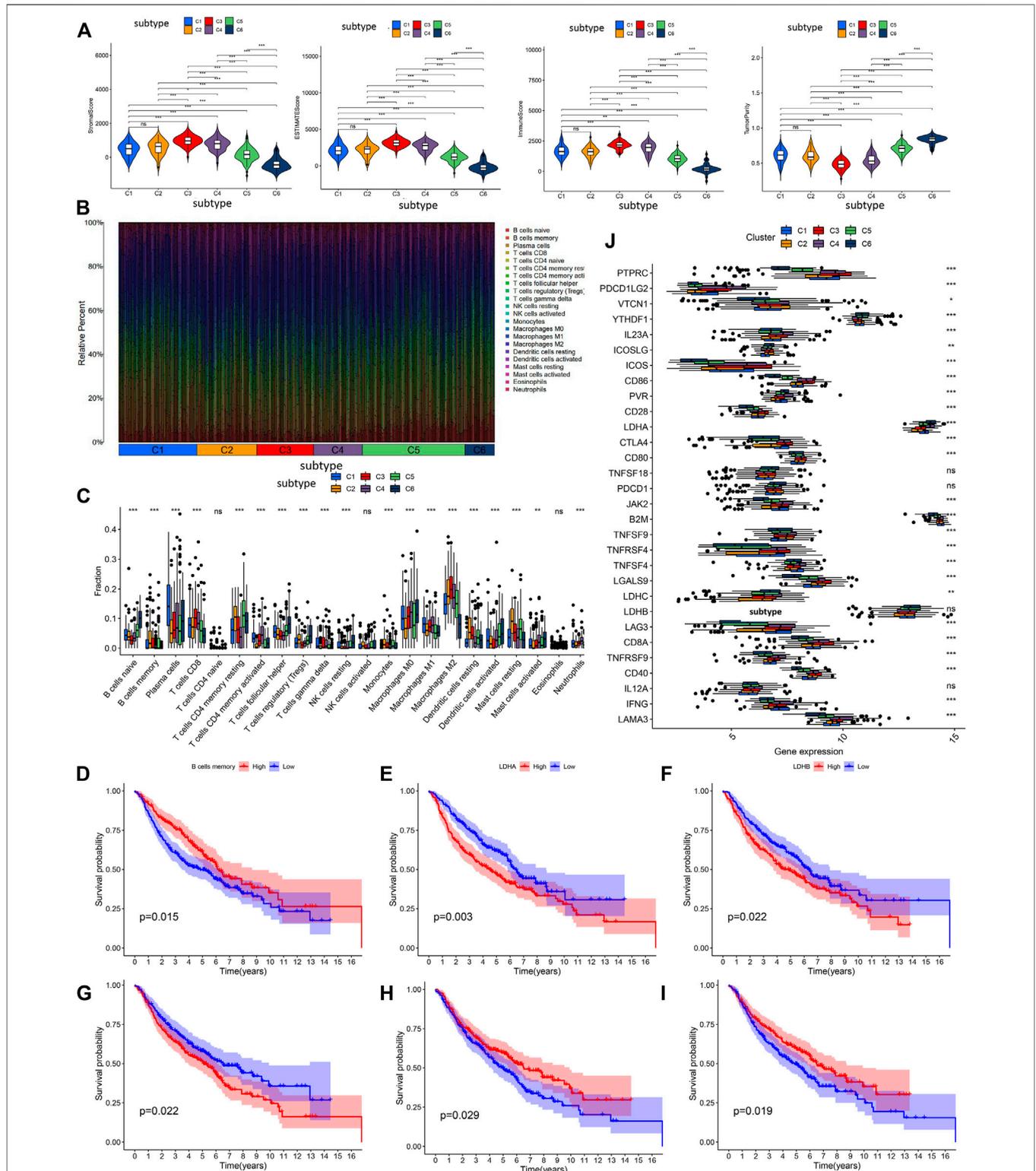


FIGURE 4 | The analysis of tumor microenvironment scores, and prognostic analysis in immune cell infiltration and immune checkpoint genes. **(A)** Difference analysis of immune cell score, stromal cell score, ESTIMATE score and tumor purity among the six subtypes. **(B)** The accumulation percentage map of 22 kinds of immune cells in six subtypes from different LUAD samples. **(C)** Box plots of infiltration density of 22 kinds of immune cells in different subtypes. Kaplan-Meier survival analysis of immune cells **(D)** and immune checkpoints **(E-I)**. **(J)** Differential expression analysis of 30 recognized immune checkpoints in different subtypes. ns $p > 0.05$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

sensitivity of cells to external stimuli, to decreased cell resistance to bacterial infection, and to the widening of the blood vessel diameter (Figure 3F). In contrast, enrichment results pertaining to downregulated genes were not significant (Supplementary Figure S3).

Analysis of Tumor Microenvironment Scores, Prognostic Analysis in Immune Cell Infiltration, and Immune Checkpoint Genes

Based on tumor microenvironment scores, our results revealed that stromal/ESTIMATE/immune scores in C3, C4, C5, and C6 decreased, whereas tumor purity was gradually increased in C3, C4, C5, and C6 (Figure 4A). The content of the 22 immune cells in each LUAD sample was calculated by the CIBERSORT algorithm and was visualized through an accumulation percentage map, where different colors represent different immune cells (Figure 4B). Afterward, the Wilcoxon test was used to compare the infiltration density of the immune cells in different subtypes (Figure 4C), and the results found that the immune cells were significantly different among the six subtypes examined ($p < 0.01$), except for the activated CD4 T cells, activated NK cells, and eosinophils. The Kaplan-Meier survival analysis showed that the infiltration density of memory B cells was positively correlated with patient survival ($p = 0.015$, Figure 4D). Memory B cell infiltration density was found to be the highest in C2, which translates to improved OS, whereas memory B cell infiltration density was the lowest in C6, which means poor OS.

Expression of the 30 immune checkpoints among LUAD subtypes was measured and compared (Figure 4J), and our results found that immune checkpoints were significantly different among the six subtypes ($p < 0.05$), except for interleukin 12A (IL12A), lactate dehydrogenase B (LDHB), programmed cell death 1 (PDCD1), and TNF superfamily member 18 (TNFSF18). It is worth noting that downregulation of LDHB predicted a better OS (Figure 4F). LDHA was downregulated in C2, and the Kaplan-Meier analysis indicated that lowly expressed LDHA predicted better OS (Figure 4E). CD28 and TNFRSF4 were upregulated in C3, and the Kaplan-Meier analysis indicated that highly expressed CD28 and TNFRSF4 predicted better OS (Figures 4H,I). Furthermore, JAK2 was upregulated in C4, and the Kaplan-Meier analysis indicated that highly expressed JAK2 predicted worse OS (Figure 4G). These results provide a molecular basis for understating that the different subtypes have different OS.

Construction of a Prognostic Risk Scoring Model to Predict Lung Adenocarcinoma Patient Survival

Based on previous studies, 313 genes, consisting of 291 LDRGs, 18 significantly upregulated genes related to histologic grade and four immune checkpoint genes were further screened by univariate analysis, LASSO, and multivariate Cox analysis

(Supplementary Tables S3, S4). Ultimately, 14 genes were generated in the TCGA training cohort. According to the 95% confidence interval of these genes, we finally identified eight significant survival-predicting genes: c-type lectin domain containing 7A (CLEC7A, HR = 0.69, $p < 0.001$), baculoviral IAP repeat containing 3 (BIRC3, HR = 1.19, $p = 0.032$), superoxide dismutase 1 (SOD1, HR = 1.62, $p = 0.007$), versican (VCAN, HR = 1.17, $p = 0.039$), kynureninase (KYNU, HR = 1.14, $p = 0.032$), CC motif chemokine ligand 20 (CCL20, HR = 1.10, $p = 0.004$), cancer/testis antigen 1B (CTAG1B, HR = 2.40, $p < 0.001$), and CD69 molecule (CD69, HR = 0.82, $p = 0.022$). Furthermore, the Supplementary Figure S4 demonstrates the expression levels of eight prognostic genes in cell profiles. It is shown that CLEC7A, VCAN, and KYNU were significantly upregulated in monocytes, BIRC3 and SOD1 were mainly distributed in T cells, CCL20 was highly expressed in tumor cells, and CD69 was highly expressed in B cells. CTAG1B was filtered out during quality control.

The prognostic risk scoring model was developed based on the above eight prognostic-related genes, and the formula was as follows:

$$\text{Risk score} = \text{Exp}_{\text{CLEC7A}} \times (-0.377) + \text{Exp}_{\text{BIRC3}} \times 0.170 + \text{Exp}_{\text{SOD1}} \times 0.483 + \text{Exp}_{\text{VCAN}} \times 0.156 + \text{Exp}_{\text{KYNU}} \times 0.132 + \text{Exp}_{\text{CCL20}} \times 0.099 + \text{Exp}_{\text{CTAG1B}} \times 0.877 + \text{Exp}_{\text{CD69}} \times (-0.201).$$

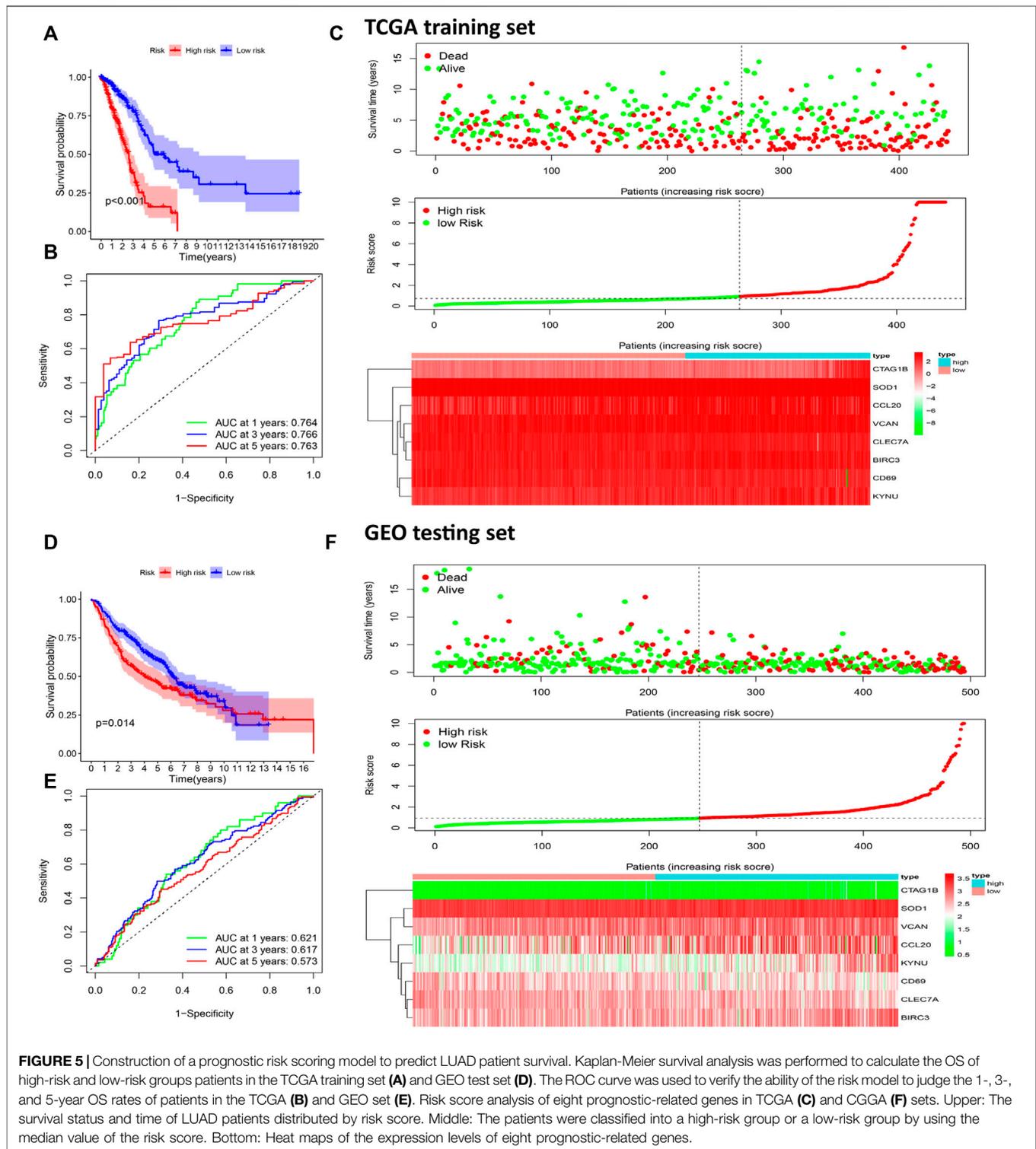
According to this formula, the risk scores of all patients in the TCGA training set and the GEO testing set were calculated, and the patients were divided into high- and low-risk groups based on the median of the risk scores (Figure 5C). The Kaplan-Meier survival analysis showed that OS of patients in the high-risk group was significantly poorer than the OS in the low-risk group both in the training set ($p < 0.001$, Figure 5A) and the testing set ($p = 0.014$, Figure 5D). The areas under the receiver operating characteristic curves (ROC) for predicting 1-, 3-, and 5-year OS were 0.764, 0.766, and 0.760, respectively, in the training set (Figure 5B) and 0.621, 0.617, and 0.573, respectively, in the testing set (Figure 5E). These results showed that this prognostic risk score model could be used as a reliable prognostic predictor for LUAD patients.

The Key Role of CCL20 in Lung Adenocarcinoma

The intersection of 291 LDRGs and 32 genes related to histologic grade showed that they both had CCL20 and C-X-C motif chemokine ligand 8 (CXCL8). Furthermore, CCL20 was also included in the survival-predicting genes (Figure 6) and was mainly expressed in cancer cells of branch II.

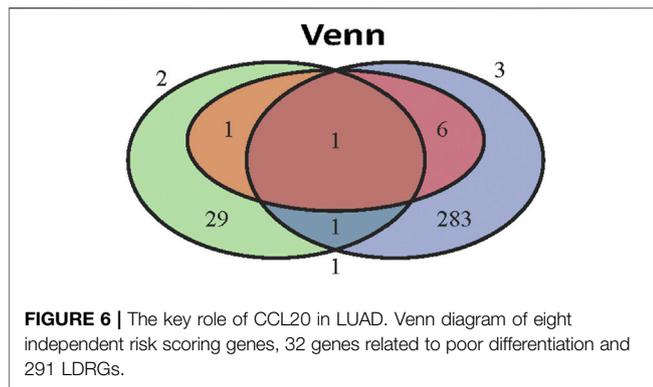
Establishment and Validation of a Nomogram for Predicting Patient OS Based on the The Cancer Genome Atlas Cohort

Univariate and multivariate Cox regression analysis showed that stage, risk score, and CCL20 mutation status could be used as independent factors for the prognosis of LUAD (Supplementary Tables S5, S6). Patients with higher risk



scores and later stage had poorer prognosis. CCL20 mutations were also associated with poor prognosis. Finally, a prognostic nomogram was successfully developed to predict the OS rate of LUAD patients based on the TCGA cohort (Figure 7A). Time-dependent ROC analysis showed that the nomogram could

effectively predict the 1-, 3-, and 5-year OS rates of patients, with AUC values of 0.768, 0.766, and 0.758, respectively (Figure 7B). The calibration chart also revealed that the predicted OS rate was consistent with the actual observations in the TCGA cohort (Figure 7C).

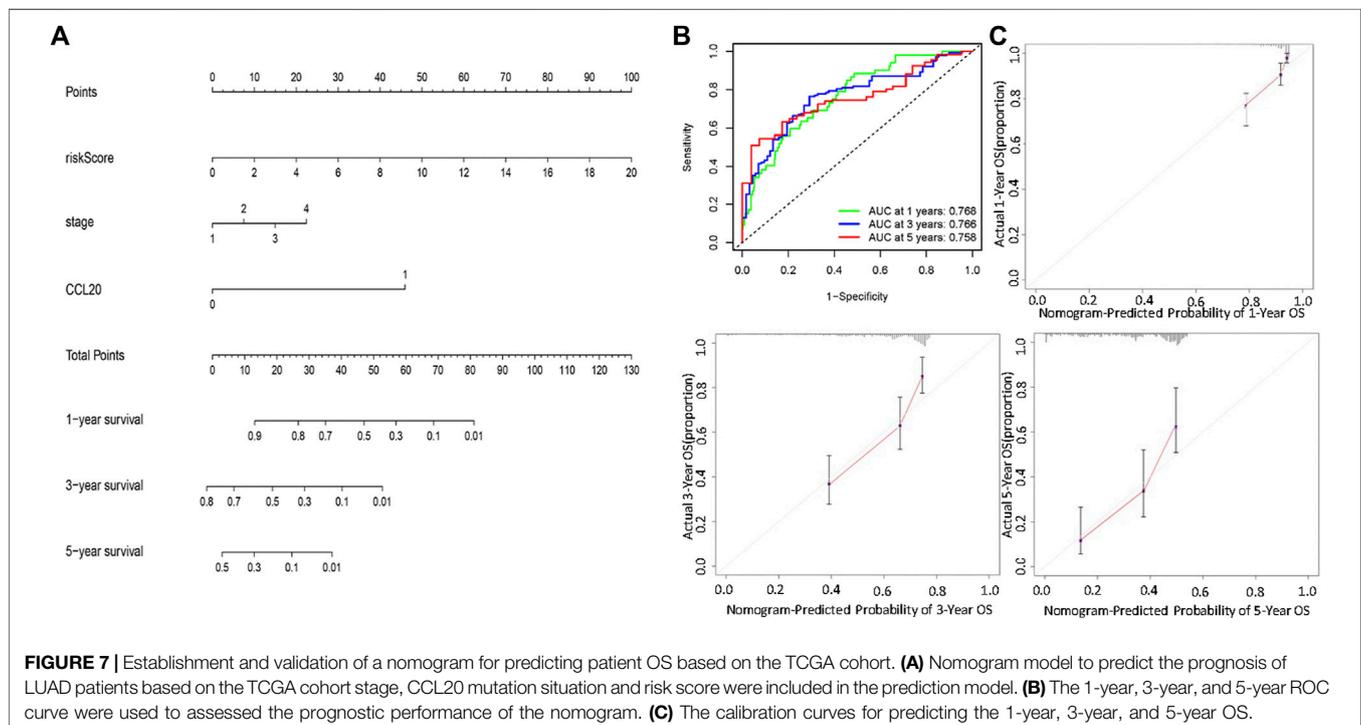


DISCUSSION

Although the incidence of LUAD is consistently reduced year by year, as a result of public education and implementation of etiological preventive measures, such as quitting smoking, it still remains a global public health issue due to its respective high mortality rates (Siegel et al., 2021). To date, there is a lack of effective prediction models that could be used to guide clinical prognosis evident. Most tumors originate from a single cell, and tumor progression is caused by the gradual selection of more aggressive and more resistant dominant cancer cell variant sublines due to acquired genetic instability (Nowell, 1976). The strong drug resistance caused by heterogeneity greatly increases the complexity of tumors and reflects the increasing difficulty of drug design, thus rendering the prognosis of cancer patients unsatisfactory (Fedele et al., 2014). LUAD is the most

common respiratory tumor with extremely high heterogeneity. In this study, we analyzed the factors affecting LUAD heterogeneity based on the RNA-seq and somatic mutation data from TCGA and the scRNA-seq and RNA-seq data from GEO. Single-cell trajectory analysis verified LUAD heterogeneity, and the subsequent enrichment analysis exhibited that this heterogeneity was closely related to the activation and suppression of the immune system function (Satija et al., 2015; Lall et al., 2018). Eventually, CLEC7A, CD69, SOD1, BIRC3, VCAN, KYNU, CCL20, and CTAG1B were identified as the eight most significant, independent predictive genes. Based on these eight genes, we established a risk scoring model and combined the tumor stage and the CCL20 mutation status to successfully develop a 1-, 3-, and 5-year prognostic nomogram. With high predictive power, the nomogram could provide certain evidence support for clinicians to predict the possibility of patient survival time, thus enabling clinicians to formulate individualized optimal treatment plans.

CLEC7A, also known as Dectin-1 and CD369, is a pattern recognition receptor expressed by myeloid phagocytes (macrophages, dendritic cells, and neutrophils) and plays a pivotal role in antifungal response, antimicrobial reaction, and autoimmunity through phagocytosis, respiratory burst, cytokines, and others (Goodridge et al., 2011; Mattioli et al., 2019). Besides, accumulated evidence has suggested that CLEC7A had dual effects on cancer, namely, inhibiting and promoting cancer (Tone et al., 2019). CLEC7A has been widely studied in pancreatic cancer, kidney cancer, and liver cancer, etc., but the true role of this gene remains controversial (Seifert et al., 2015; Xia et al., 2016; Daley et al., 2017). CLEC7A, which is highly expressed in M2-like tumor-associated macrophages, is linked to



galectin 9, thus leading to tolerogenic macrophage programming and adaptive immunosuppression that contribute to tumor exacerbation (Daley et al., 2017; Zhou et al., 2021a). Conversely, activation of CLEC7A can exert effective anti-tumor immunity by enhancing the expression of interleukin-33 and interleukin-9 in dendritic cells and by inducing differentiation of naive CD4⁺T cells into T-helper 9 cells (Zhao et al., 2016; Chen et al., 2018). CLEC7A could also stimulate the secretion of cytokines, such as interleukin-13, interleukin-1, and TNF- α in macrophages to activate macrophage tumoricidal response (Deng et al., 2018; Alaeddine et al., 2019). This opposite effect may be caused by the activation of different types of immune cells by CLEC7A in different pathways. So far, little is known regarding the relationship between CLEC7A and lung cancer. Chiba et al. found that CLEC7A could promote natural killer cells to secrete cytotoxic factors and thus fight lung cancer through cell-to-cell contact (Chiba et al., 2014). In this study, CLEC7A had a positive effect on the prognosis of LUAD patients, which was consistent with the findings of previous studies (Zou et al., 2020; Zhou et al., 2021b; Lin et al., 2021).

In addition, the key role of CCL20 in LUAD was identified in this study. CCL20, also called liver and activation-regulated chemokine, macrophage inflammatory protein 3 alpha, and small-inducible cytokine A20, is highly expressed in many types of cancer, including NSCLC, ovarian cancer, colorectal cancer, penile cancer, breast cancer, gastric cancer, and pancreatic cancer (Han et al., 2015; Liu et al., 2016; Lee et al., 2017; Wang et al., 2019b; Liu et al., 2020a; Mo et al., 2020). This ligand has been found to enhance the migration and proliferation of cancer cells through the CCL20/CCR6 signaling pathway and to inhibit cell apoptosis, which directly promotes cancer progression. In addition, CCL20 can indirectly promote cancer progression through immune cell control to reshape tumor microenvironment. For instance, Wei et al. concluded that the high expression of CCL20 in LUAD could enhance CCL20/CCR6/ERK signal transduction, thereby promoting the proliferation and migration of tumor cells (Wei et al., 2019). Furthermore, Wang et al. believed that NSCLC cells produce CCL20 in an autocrine and paracrine manner. Overexpression of CCL20 overactivates the MAPK-PI3K signaling pathway, thus leading to high chemoresistance, accelerated proliferation, or a more stable ability to reside in new locations of cancer cells (Wang et al., 2016). In this work, CCL20 was found to be highly expressed mainly in branch II and concentrated in tumor cells. Enrichment analysis showed that the significant CCL20 upregulation can increase the sensitivity of tumors to extracellular stimuli and promote bacterial invasion of host cells, leading to instability of the intracellular environment, but it can also increase the diameter of blood vessels in the tumor environment to meet the needs of tumor proliferation and migration. Furthermore, our findings showed that CCL20 was significantly negatively correlated with the patient's OS. Consistent with previous research, CCL20 was confirmed to be an important gene involved in the evolution of LUAD tumors, poor histologic grade, and independent prognosis of patients in our work.

CD69 upregulation could induce LUAD cell apoptosis and inhibit tumor proliferation, invasion, and metastasis, thus

suggesting a good prognosis. BIRC3, VCAN, CTAG1B, SOD1, and KYNU promoted the survival and proliferation of cancer cells, inhibited cancer cell apoptosis, and were related to poor prognosis and treatment resistance (Jin et al., 2018; Liu et al., 2020b; Al-Mansoob et al., 2021; Frazzi, 2021; Li et al., 2021). In this work, these genes could also be used as potential prognostic genes for LUAD, and their impact on LUAD was consistent with the findings of previous reports.

Tumor immunotherapy has proven to be very effective because it can manipulate the individual's immune system to identify and attack cancer cells, thereby inducing a long-lasting and strong anti-tumor response (Gotwals et al., 2017; Kennedy and Salama, 2020). Immunotherapy has achieved amazing effects in solid tumors, such as melanoma, bladder cancer, and breast cancer, but the effective survival benefit of lung cancer has only recently been confirmed (van den Bulk et al., 2018). However, it should not be ignored that immunotherapy is not applicable to all patients with LUAD (Zhao et al., 2019; Hegde and Chen, 2020). Unfortunately, some patients have poor sensitivity to immunotherapy, but also treatment side effects, such as pneumonia, can further damage the already poor lung reserve, leading to accelerated disease deterioration (Steven et al., 2016). Therefore, it is very important to distinguish patients who are sensitive to immunotherapy. The tumor immune dysfunction and elimination (TIDE) algorithm came into being. The results of this study showed that the TIDE score of low-risk patients was significantly higher than that of high-risk patients, indicating that high-risk patients are bound to benefit more from immunotherapy and thus get a better prognosis compared to low-risk patients (**Supplementary Figure S5**). In general, a risk scoring model composed of eight independent genes could provide a preliminary estimate of patients' responsiveness to immunotherapy, thereby minimizing the negative effects of drug toxicity.

Our research also has certain limitations. In this research, we only analyzed the molecular phenotypes and existing clinical data that we were interested in. Phenotypes such as gene methylation, gene mutations, long non-coding RNA, competing endogenous RNA, and clinicopathological characteristics, such as smoking history, chemotherapy and radiotherapy history, history of underlying diseases, history of disease progression, etc., were not considered due to the lack of unified standard data. Consequently, survival prediction models necessitate continuous improvement.

CONCLUSION

The current research analyzed the genes that affected tumor heterogeneity and established a risk scoring model consisting of eight independent prognostic genes through data dimensionality reduction. The nomogram that combined risk scores, clinicopathological variables, and CCL20 mutation status provided a visual and accurate method for predicting the 1-, 3-, and 5-year survival rates of patients. In conclusion, this study emphasizes the impact of tumor heterogeneity on predicting

patient OS and immunotherapy response and underlines a potential therapeutic target for the treatment of LUAD.

acquisition, HL. All authors have read and agreed to the published version of the manuscript.

DATA AVAILABILITY STATEMENT

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

FUNDING

This study was supported by the National key Research and Development Program of China (Project No. 2016YFC1304203) and the National Natural Science Foundation of China (No. 81770088).

AUTHOR CONTRIBUTIONS

Conceptualization, PZ and HL; methodology, YZ, LW, WJ, and LL.; writing—original draft preparation, HZ and LZ; writing—review and editing, HZ and HL; funding

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2022.807497/full#supplementary-material>

REFERENCES

- Al-Mansoob, M., Gupta, I., Stefan Rusyniak, R., and Ouhtit, A. (2021). KYN^U, a Novel Potential Target that Underpins CD44-promoted Breast Tumour Cell Invasion. *J. Cell. Mol. Med.* 25 (5), 2309–2314. doi:10.1111/jcmm.16296
- Alaeddine, M., Prat, M., Poinot, V., Gouazé-Andersson, V., Authier, H., Meunier, E., et al. (2019). IL13-Mediated Dectin-1 and Mannose Receptor Overexpression Promotes Macrophage Antitumor Activities through Recognition of Sialylated Tumor Cells. *Cancer Immunol. Res.* 7 (2), 321–334. doi:10.1158/2326-6066.CIR-18-0213
- Aleksakhina, S. N., Kashyap, A., and Imyanotov, E. N. (18722019). Mechanisms of Acquired Tumor Drug Resistance. *Biochim. Biophys. Acta (Bba) - Rev. Cancer* 1872 (2), 188310. doi:10.1016/j.bbcan.2019.188310
- Chen, J., Zhao, Y., Jiang, Y., Gao, S., Wang, Y., Wang, D., et al. (2018). Interleukin-33 Contributes to the Induction of Th9 Cells and Antitumor Efficacy by Dectin-1-Activated Dendritic Cells. *Front. Immunol.* 9, 1787. doi:10.3389/fimmu.2018.01787
- Chiba, S., Ikushima, H., Ueki, H., Yanai, H., Kimura, Y., Hangai, S., et al. (2014). Recognition of Tumor Cells by Dectin-1 Orchestrates Innate Immune Cells for Anti-tumor Responses. *Elife* 3, e04177. doi:10.7554/eLife.04177
- Dagogo-Jack, I., and Shaw, A. T. (2018). Tumour Heterogeneity and Resistance to Cancer Therapies. *Nat. Rev. Clin. Oncol.* 15 (2), 81–94. doi:10.1038/nrclinonc.2017.166
- Daley, D., Mani, V. R., Mohan, N., Akkad, N., Ochi, A., Heindel, D. W., et al. (2017). Dectin 1 Activation on Macrophages by Galectin 9 Promotes Pancreatic Carcinoma and Peritumoral Immune Tolerance. *Nat. Med.* 23 (5), 556–567. doi:10.1038/nm.4314
- Deng, C., Fu, H., Shang, J., Chen, J., and Xu, X. (2018). Dectin-1 Mediates the Immunoenhancement Effect of the Polysaccharide from Dictyophora Indusiata. *Int. J. Biol. Macromolecules* 109, 369–374. doi:10.1016/j.ijbiomac.2017.12.113
- Denisenko, T. V., Budkevich, I. N., and Zhivotovsky, B. (2018). Cell Death-Based Treatment of Lung Adenocarcinoma. *Cell Death Dis* 9 (2), 117. doi:10.1038/s41419-017-0063-y
- de Sousa, V. M. L., and Carvalho, L. (2018). Heterogeneity in Lung Cancer. *Pathobiology* 85 (1–2), 96–107. doi:10.1159/000487440
- Fedele, C., Tothill, R. W., and McArthur, G. A. (2014). Navigating the challenge of Tumor Heterogeneity in Cancer Therapy. *Cancer Discov.* 4 (2), 146–148. doi:10.1158/2159-8290.CD-13-1042
- Frazzi, R. (2021). BIRC3 and BIRC5: Multi-faceted Inhibitors in Cancer. *Cell Biosci* 11 (1), 8. doi:10.1186/s13578-020-00521-0
- Gao, X., Tang, M., Tian, S., Li, J., and Liu, W. (2021). A Ferroptosis-Related Gene Signature Predicts Overall Survival in Patients with Lung Adenocarcinoma. *Future Oncol.* 17 (12), 1533–1544. doi:10.2217/fon-2020-1113
- Gong, Z., Zhang, J., and Guo, W. (2020). Tumor Purity as a Prognosis and Immunotherapy Relevant Feature in Gastric Cancer. *Cancer Med.* 9 (23), 9052–9063. doi:10.1002/cam4.3505
- Goodridge, H. S., Reyes, C. N., Becker, C. A., Katsumoto, T. R., Ma, J., Wolf, A. J., et al. (2011). Activation of the Innate Immune Receptor Dectin-1 upon Formation of a 'phagocytic Synapse'. *Nature* 472 (7344), 471–475. doi:10.1038/nature10071
- Gotwals, P., Cameron, S., Cipolletta, D., Cremasco, V., Crystal, A., Hewes, B., et al. (2017). Prospects for Combining Targeted and Conventional Cancer Therapy with Immunotherapy. *Nat. Rev. Cancer* 17 (5), 286–301. doi:10.1038/nrc.2017.17
- Greten, F. R., and Grivnenkov, S. I. (2019). Inflammation and Cancer: Triggers, Mechanisms, and Consequences. *Immunity* 51 (1), 27–41. doi:10.1016/j.immuni.2019.06.025
- Han, G., Wu, D., Yang, Y., Li, Z., Zhang, J., and Li, C. (2015). Crkl Mediates CCL20/CCR6-Induced EMT in Gastric Cancer. *Cytokine* 76 (2), 163–169. doi:10.1016/j.cyto.2015.05.009
- Hegde, P. S., and Chen, D. S. (2020). Top 10 Challenges in Cancer Immunotherapy. *Immunity* 52 (1), 17–35. doi:10.1016/j.immuni.2019.12.011
- Herbst, R. S., Morgensztern, D., and Boshoff, C. (2018). The Biology and Management of Non-small Cell Lung Cancer. *Nature* 553 (7689), 446–454. doi:10.1038/nature25183
- Hinshaw, D. C., and Shevde, L. A. (2019). The Tumor Microenvironment Innately Modulates Cancer Progression. *Cancer Res.* 79 (18), 4557–4566. doi:10.1158/0008-5472.CAN-18-3962
- Ionkina, A. A., Balderrama-Gutierrez, G., Ibanez, K. J., Phan, S. H. D., Cortez, A. N., Mortazavi, A., et al. (2021). Transcriptome Analysis of Heterogeneity in Mouse Model of Metastatic Breast Cancer. *Breast Cancer Res.* 23 (1), 93. doi:10.1186/s13058-021-01468-x
- Isla, D., de Castro, J., García-Campelo, R., Lianes, P., Felip, E., Garrido, P., et al. (2020). Treatment Options beyond Immunotherapy in Patients with Wild-type Lung Adenocarcinoma: a Delphi Consensus. *Clin. Transl Oncol.* 22 (5), 759–771. doi:10.1007/s12094-019-02191-y
- Jin, S., Cao, S., Li, J., Meng, Q., Wang, C., Yao, L., et al. (2018). Cancer/testis Antigens (CTAs) Expression in Resected Lung Cancer. *Ott Vol.* 11, 4491–4499. doi:10.2147/OTT.S159491
- Kennedy, L. B., and Salama, A. K. S. (2020). A Review of Cancer Immunotherapy Toxicity. *CA A. Cancer J. Clin.* 70 (2), 86–104. doi:10.3322/caac.21596
- Lall, S., Sinha, D., Bandyopadhyay, S., and Sengupta, D. (2018). Structure-Aware Principal Component Analysis for Single-Cell RNA-Seq Data. *J. Comput. Biol.* 25, 1365–1373. doi:10.1089/cmb.2018.0027
- Lee, S. K., Park, K.-K., Kim, H.-J., Park, J., Son, S. H., Kim, K. R., et al. (2017). Human Antigen R-Regulated CCL20 Contributes to Osteolytic Breast Cancer Bone Metastasis. *Sci. Rep.* 7 (1), 9610. doi:10.1038/s41598-017-09040-4
- Li, J.-P., Li, R., Liu, X., Huo, C., Liu, T.-T., Yao, J., et al. (2020). A Seven Immune-Related lncRNAs Model to Increase the Predicted Value of Lung Adenocarcinoma. *Front. Oncol.* 10, 560779. doi:10.3389/fonc.2020.560779

- Li, R., Hou, S., Zou, M., Ye, K., and Xiang, L. (2021). miR-543 Impairs Cell Proliferation, Migration, and Invasion in Breast Cancer by Suppressing VCAN. *Biochem. Biophysical Res. Commun.* 570, 191–198. doi:10.1016/j.bbrc.2021.07.005
- Liang, C., Yang, K. Y., Chan, V. W., Li, X., Fung, T. H. W., Wu, Y., et al. (2020). CD8+ T-Cell Plasticity Regulates Vascular Regeneration in Type-2 Diabetes. *Theranostics* 10 (9), 4217–4232. doi:10.7150/thno.40663
- Lim, Z.-F., and Ma, P. C. (2019). Emerging Insights of Tumor Heterogeneity and Drug Resistance Mechanisms in Lung Cancer Targeted Therapy. *J. Hematol. Oncol.* 12 (1), 134. doi:10.1186/s13045-019-0818-2
- Lin, J., Weng, J., Lin, S., Lin, C., Huang, J., Zhang, C., et al. (2021). Development and Validation of a Novel mRNA Signature for Predicting Early Relapse in Non-small Cell Lung Cancer. *Jpn. J. Clin. Oncol.* 51 (8), 1277–1286. doi:10.1093/jjco/hyab075
- Liu, B., Jia, Y., Ma, J., Wu, S., Jiang, H., Cao, Y., et al. (2016). Tumor-associated Macrophage-Derived CCL20 Enhances the Growth and Metastasis of Pancreatic Cancer. *Acta Biochim. Biophys. Sin* 48 (12), 1067–1074. doi:10.1093/abbs/gmw101
- Liu, C., Li, Y., Wei, M., Zhao, L., Yu, Y., and Li, G. (2019). Identification of a Novel Glycolysis-Related Gene Signature that Can Predict the Survival of Patients with Lung Adenocarcinoma. *Cell Cycle* 18 (5), 568–579. doi:10.1080/15384101.2019.1578146
- Liu, S., Li, B., Xu, J., Hu, S., Zhan, N., Wang, H., et al. (2020). SOD1 Promotes Cell Proliferation and Metastasis in Non-small Cell Lung Cancer via an miR-409-3p/SOD1/SETDB1 Epigenetic Regulatory Feedforward Loop. *Front. Cell Dev. Biol.* 8, 213. doi:10.3389/fcell.2020.00213
- Liu, W., Wang, W., Wang, X., Xu, C., Zhang, N., and Di, W. (2020). Cisplatin-stimulated Macrophages Promote Ovarian Cancer Migration via the CCL20-CCR6 axis. *Cancer Lett.* 472, 59–69. doi:10.1016/j.canlet.2019.12.024
- Luo, C., Lei, M., Zhang, Y., Zhang, Q., Li, L., Lian, J., et al. (2020). Systematic Construction and Validation of an Immune Prognostic Model for Lung Adenocarcinoma. *J. Cel Mol Med* 24 (2), 1233–1244. doi:10.1111/jcmm.14719
- Marusyk, A., Janiszewska, M., and Polyak, K. (2020). Intratumor Heterogeneity: The Rosetta Stone of Therapy Resistance. *Cancer Cell* 37 (4), 471–484. doi:10.1016/j.ccell.2020.03.007
- Mattiola, I., Tomay, F., De Pizzol, M., Silva-Gomes, R., Savino, B., Gulic, T., et al. (2019). The Macrophage Tetraspan MS4A4A Enhances Dectin-1-dependent NK Cell-Mediated Resistance to Metastasis. *Nat. Immunol.* 20 (8), 1012–1022. doi:10.1038/s41590-019-0417-y
- Mo, M., Tong, S., Huang, W., Cai, Y., Zu, X., and Hu, X. (2020). High Serum CCL20 Is Associated with Tumor Progression in Penile Cancer. *J. Cancer* 11 (23), 6812–6822. doi:10.7150/jca.48939
- Narayanan, S., Cai, C.-Y., Assaraf, Y. G., Guo, H.-Q., Cui, Q., Wei, L., et al. (2020). Targeting the Ubiquitin-Proteasome Pathway to Overcome Anti-cancer Drug Resistance. *Drug Resist. Updates* 48, 100663. doi:10.1016/j.drug.2019.100663
- Newman, A. M., Liu, C. L., Green, M. R., Gentles, A. J., Feng, W., Xu, Y., et al. (2015). Robust Enumeration of Cell Subsets from Tissue Expression Profiles. *Nat. Methods* 12 (5), 453–457. doi:10.1038/nmeth.3337
- Nowell, P. C. (1976). The Clonal Evolution of Tumor Cell Populations. *Science* 194 (4260), 23–28. doi:10.1126/science.959840
- Qiu, X., Hill, A., Packer, J., Lin, D., Ma, Y.-A., and Trapnell, C. (2017). Single-cell mRNA Quantification and Differential Analysis with Census. *Nat. Methods* 14 (3), 309–315. doi:10.1038/nmeth.4150
- Qiu, X., Mao, Q., Tang, Y., Wang, L., Chawla, R., Pliner, H. A., et al. (2017). Reversed Graph Embedding Resolves Complex Single-Cell Trajectories. *Nat. Methods* 14 (10), 979–982. doi:10.1038/nmeth.4402
- Ruiz-Cordero, R., and Devine, W. P. (2020). Targeted Therapy and Checkpoint Immunotherapy in Lung Cancer. *Surg. Pathol. Clin.* 13 (1), 17–33. doi:10.1016/j.path.2019.11.002
- Satija, R., Farrell, J. A., Gennert, D., Schier, A. F., and Regev, A. (2015). Spatial Reconstruction of Single-Cell Gene Expression Data. *Nat. Biotechnol.* 33 (5), 495–502. doi:10.1038/nbt.3192
- Seifert, L., Deutsch, M., Allothman, S., Alqunaibit, D., Werba, G., Pansari, M., et al. (2015). Dectin-1 Regulates Hepatic Fibrosis and Hepatocarcinogenesis by Suppressing TLR4 Signaling Pathways. *Cel Rep.* 13 (9), 1909–1921. doi:10.1016/j.celrep.2015.10.058
- Siegel, R. L., Miller, K. D., Fuchs, H. E., and Jemal, A. (2021). Cancer Statistics, 2021. *CA A. Cancer J. Clin.* 71 (1), 7–33. doi:10.3322/caac.21654
- Steven, A., Fisher, S. A., and Robinson, B. W. (2016). Immunotherapy for Lung Cancer. *Respirology* 21 (5), 821–833. doi:10.1111/resp.12789
- Tone, K., Stappers, M. H. T., Willment, J. A., and Brown, G. D. (2019). C-type Lectin Receptors of the Dectin-1 Cluster: Physiological Roles and Involvement in Disease. *Eur. J. Immunol.* 49 (12), 2127–2133. doi:10.1002/eji.201847536
- Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., et al. (2014). The Dynamics and Regulators of Cell Fate Decisions Are Revealed by Pseudotemporal Ordering of Single Cells. *Nat. Biotechnol.* 32 (4), 381–386. doi:10.1038/nbt.2859
- van den Bulk, J., Verdegaal, E. M., and de Miranda, N. F. (2018). Cancer Immunotherapy: Broadening the Scope of Targetable Tumours. *Open Biol.* 8 (6), 180037. doi:10.1098/rsob.180037
- Vasan, N., Baselga, J., and Hyman, D. M. (2019). A View on Drug Resistance in Cancer. *Nature* 575 (7782), 299–309. doi:10.1038/s41586-019-1730-1
- Wang, B., Shi, L., Sun, X., Wang, L., Wang, X., and Chen, C. (2016). Production of CCL20 from Lung Cancer Cells Induces the Cell Migration and Proliferation through PI 3K Pathway. *J. Cel. Mol. Med.* 20 (5), 920–929. doi:10.1111/jcmm.12781
- Wang, D., Yang, L., Yu, W., Wu, Q., Lian, J., Li, F., et al. (2019). Colorectal Cancer Cell-Derived CCL20 Recruits Regulatory T Cells to Promote Chemoresistance via FOXO1/CEBPB/NF- κ B Signaling. *J. Immunotherapy Cancer* 7 (1), 215. doi:10.1186/s40425-019-0701-2
- Wang, S.-s., Liu, W., Ly, D., Xu, H., Qu, L., and Zhang, L. (2019). Tumor-infiltrating B Cells: Their Role and Application in Anti-tumor Immunity in Lung Cancer. *Cell Mol Immunol* 16 (1), 6–18. doi:10.1038/s41423-018-0027-x
- Wei, W., Zhao, X., Zhu, J., Zhang, L., Chen, Y., Zhang, B., et al. (2019). lncRNA-u50355 Promotes the Progression of Lung Cancer by Activating CCL20/ERK Signaling. *Oncol. Rep.* 42 (5), 1946–1956. doi:10.3892/or.2019.7302
- Wu, T., and Dai, Y. (2017). Tumor Microenvironment and Therapeutic Response. *Cancer Lett.* 387, 61–68. doi:10.1016/j.canlet.2016.01.043
- Xia, Y., Liu, L., Bai, Q., Wang, J., Xi, W., Qu, Y., et al. (2016). Dectin-1 Predicts Adverse Postoperative Prognosis of Patients with clear Cell Renal Cell Carcinoma. *Sci. Rep.* 6, 32657. doi:10.1038/srep32657
- Yoshihara, K., Shahmoradgoli, M., Martínez, E., Vegesna, R., Kim, H., Torres-García, W., et al. (2013). Inferring Tumour Purity and Stromal and Immune Cell Admixture from Expression Data. *Nat. Commun.* 4, 2612. doi:10.1038/ncomms3612
- Zhang, C., Cheng, W., Ren, X., Wang, Z., Liu, X., Li, G., et al. (2017). Tumor Purity as an Underlying Key Factor in Glioma. *Clin. Cancer Res.* 23 (20), 6279–6291. doi:10.1158/1078-0432.CCR-16-2598
- Zhang, C., Zhang, G., Sun, N., Zhang, Z., Zhang, Z., Luo, Y., et al. (2020). Comprehensive Molecular Analyses of a TNF Family-Based Signature with Regard to Prognosis, Immune Features, and Biomarkers for Immunotherapy in Lung Adenocarcinoma. *EBioMedicine* 59, 102959. doi:10.1016/j.ebiom.2020.102959
- Zhang, Y., Yang, M., Ng, D. M., Haleem, M., Yi, T., Hu, S., et al. (2020). Multi-omics Data Analyses Construct TME and Identify the Immune-Related Prognosis Signatures in Human LUAD. *Mol. Ther. - Nucleic Acids* 21, 860–873. doi:10.1016/j.omtn.2020.07.024
- Zhao, Y., Chu, X., Chen, J., Wang, Y., Gao, S., Jiang, Y., et al. (2016). Dectin-1-activated Dendritic Cells Trigger Potent Antitumor Immunity through the Induction of Th9 Cells. *Nat. Commun.* 7, 12368. doi:10.1038/ncomms12368
- Zhao, Z., Zheng, L., Chen, W., Weng, W., Song, J., and Ji, J. (2019). Delivery Strategies of Cancer Immunotherapy: Recent Advances and Future Perspectives. *J. Hematol. Oncol.* 12 (1), 126. doi:10.1186/s13045-019-0817-3
- Zheng, Y., Tian, H., Zhou, Z., Xiao, C., Liu, H., Liu, Y., et al. (2021). A Novel Immune-Related Prognostic Model for Response to Immunotherapy and Survival in Patients with Lung Adenocarcinoma. *Front. Cell Dev. Biol.* 9, 651406. doi:10.3389/fcell.2021.651406
- Zhou, T., Yang, P., Tang, S., Zhu, Z., Li, X., Yang, Z., et al. (2021). Classification of Lung Adenocarcinoma Based on Immune Checkpoint and Screening of Related Genes. *J. Oncol.* 2021, 1–12. doi:10.1155/2021/5512325
- Zhou, W., Zhou, Y., Chen, X., Ning, T., Chen, H., Guo, Q., et al. (2021). Pancreatic Cancer-Targeting Exosomes for Enhancing Immunotherapy and Reprogramming Tumor Microenvironment. *Biomaterials* 268, 120546. doi:10.1016/j.biomaterials.2020.120546
- Zou, X., Hu, Z., Huang, C., and Chang, J. (2020). A Seven-Gene Signature with Close Immune Correlation Was Identified for Survival Prediction of Lung Adenocarcinoma. *Med. Sci. Monit.* 26, e924269. doi:10.12659/MSM.924269

Zuo, S., Wei, M., Wang, S., Dong, J., and Wei, J. (2020). Pan-Cancer Analysis of Immune Cell Infiltration Identifies a Prognostic Immune-Cell Characteristic Score (ICCS) in Lung Adenocarcinoma. *Front. Immunol.* 11, 1218. doi:10.3389/fimmu.2020.01218

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

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of

the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Zheng, Zhang, Jiang, Wang, Liu, Zhou, Zhou and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

GLOSSARY

BIRC3	Baculoviral IAP repeat containing 3	LDRGs	LUAD differentiation-related genes
BP	Biological process	LUAD	Lung adenocarcinoma
CC	Component	LUSC	Lung squamous cell carcinoma
CCL20	CC motif chemokine ligand 20	MF	Molecular function
CD69	CD69 molecule	NSCLC	Non-small cell lung cancer
CDF	The cumulative distribution function	OS	Overall survival
CLEC7A	c-type lectin domain containing 7A	PC	Principal components
CTAG1B	Cancer/testis antigen 1B	PCA	Principal component analysis
GEO	Gene Expression Omnibus	ROC	Receiver operating characteristic curves
GO	Geneontology	SCLC	Small cell lung cancer
ICG	Immune checkpoint gene	SOD1	Superoxide dismutase one
KEGG	Kyoto encyclopedia of genes and genomes	TCGA	The Cancer Genome Atlas
KYNU	Kynureninase	TF	Transcription factors
LASSO	Least absolute shrinkage selection operator	TIDE	The tumor immune dysfunction and elimination
		t-SNE	The t-distributed stochastic neighbor embedding
		VCAN	Versican