

OPEN

# Myeloid cell differentiation-related gene signature for predicting clinical outcome, immune microenvironment, and treatment response in lung adenocarcinoma

Di Wu<sup>1,3</sup>, Yibing Liu<sup>1,3</sup>, Jian Liu<sup>2</sup>, Li Ma<sup>1</sup> & Xiaoxia Tong<sup>1</sup>✉

Considering the key role of myeloid cell differentiation—related genes in the tumor microenvironment (TME), we aimed to build a prognostic risk model using these genes for Lung adenocarcinoma (LUAD). The mRNA gene expression profiles of LUAD patients from The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) databases were downloaded as the training and validation sets. Then, “edgeR” R package was applied to screen out the differentially expressed genes (DEGs) and univariate cox regression, backward stepwise selection analyses were performed to construct a prognostic model for LUAD. ESTIMATE, TIMER, XCELL, CIBERSORT abs, QUANTISEQ, MCPOUNTER, EPIC, and CIBERSORT algorithms were conducted to access the association of risk levels with the stromal and immune cell infiltration levels in LUAD. Six genes (F2RL1, PRKDC, TNFSF11, INHA, PLA2G3 and TUBB1) were utilized to construct the prognostic model. The risk model showed excellent prognostic performance for LUAD in both TCGA and GEO datasets. Also, compared to the low-risk patients, the high-risk patients had higher expression of immune checkpoint molecules and showed a lower IC50 value to the chemotherapy agents. Our findings provided a myeloid cell differentiation—related gene signature that could effectively predict prognosis and guide treatment strategies for LUAD patients.

**Keywords** Lung adenocarcinoma, Myeloid cell differentiation, Tumor microenvironment, Immunotherapy, Drug response

Lung cancer is the most prevalent type of cancer that threatens human health and life quality, and it is the leading cause of cancer-related deaths worldwide<sup>1</sup>. Approximately 85% of lung cancer cases are non-small cell lung cancer (NSCLC), and Lung adenocarcinoma (LUAD) is the most common histological subtype of NSCLC<sup>2</sup>. Despite new medical treatments that have prolonged the survival of LUAD patients, the overall treatment outcomes of LUAD remain not ideal due to distant metastasis, recurrence, or advanced stage at diagnosis<sup>3,4</sup>. Therefore, continued efforts to identify novel and sensitive therapeutic targets for prognosis prediction and providing individualized therapy for LUAD patients are urgently required.

Immune checkpoint inhibitors (ICIs), a class of therapeutic agents that function by activating the immune system and modulating T lymphocyte activities, have achieved unprecedented results in NSCLC<sup>5–8</sup>. However, it is important to note that ICIs are associated with a specific set of treatment-related adverse events<sup>9–12</sup>. Growing studies focused on tumor microenvironments have reported that the tumor immune microenvironment (TME) determined by tumor-infiltrating stromal and immune cells, is closely related to tumor progression, metastasis, immune escape, and affects the sensitivity of immunotherapy and the survival of cancer patients<sup>13</sup>. For instance, during LUAD progression, cancer-associated fibroblasts (CAFs) regulate the synthesis and remodeling of the extracellular matrix, which could increase the flintiness of tumor tissue and tumor growth<sup>14,15</sup>. CAFs can paracrine transforming growth factor- $\beta$  (TGF- $\beta$ ) and lead to tumor heterogeneity in LUAD<sup>16</sup>. Also, CAFs can secrete chemokine CXCL12 to mediate M2-polarized macrophages, thereby promoting LUAD progression both in vitro

<sup>1</sup>Experimental Research Center, Qingpu Branch of Zhongshan Hospital Affiliated to Fudan University, Shanghai 201700, China. <sup>2</sup>Department of Otolaryngology-Head and Neck Surgery, Qingpu Branch of Zhongshan Hospital Affiliated to Fudan University, Shanghai, China. <sup>3</sup>These authors contributed equally: Di Wu and Yibing Liu. ✉email: xiaoxia\_tong@163.com

and *in vivo*<sup>17</sup>. Tumor-associated immune cells, primarily T cells, can produce cytokines that promote tumor cell migration and metastatic spreading<sup>18</sup>.

Growing studies had reported that pre-metastatic tumor microenvironment, recruiting various myeloid cells, to facilitate cancer cell recruitment and colonization<sup>19,20</sup>. Stefanescu et al found that myeloid cells can suppress CD8+ T cell activity to exert antitumor immune responses during lung metastasis by activating TGF- $\beta$  signaling pathway<sup>21</sup>. Another study showed Gr-1 + CD11b + myeloid cells can upregulating metalloproteinases MMP9 to promote lung metastasis and vascular remodeling<sup>22</sup>. Since the above studies distinctly demonstrated the important roles of myeloid cells in driving lung cancer cell metastases, we believe myeloid cells may act as prognostic markers for LUAD patient overall survival.

Herein, we collected the clinical and mRNA expression data of LUAD patients from TCGA and GEO databases in order to identify the association of myeloid cell differentiation—related genes expression with prognosis, immune cell infiltration, therapeutic benefits via integrative bioinformatics.

## Method and material

### Data download and myeloid cell differentiation—related genes selection

The RNA-sequencing (RNA-seq) data and clinicopathological features of TCGA LUAD patients were downloaded from the TCGA website (<https://portal.gdc.cancer.gov/>). Among them, 58 normal samples and 501 LUAD cases with mRNA expression data, follow-up information, and clinical characteristics were used for further analysis. Additionally, another independent cohort, GSE68465, which contained 442 LUAD cases, was retrieved from the Gene Expression Omnibus database for external validation (<https://www.ncbi.nlm.nih.gov/geo/>). The genes involved in the myeloid cell differentiation—related pathway were downloaded from the Molecular Signatures Database (<https://www.gsea-msigdb.org/gsea/msigdb/human/search.jsp>)” (Supplementary Table 1).

### Construction and validation of myeloid cell differentiation—related prognostic signature

We applied the “edgeR” R package to screen out the differentially expressed genes (DEGs) between normal tissues and tumor tissues. Gene sets based on the criteria (False discovery rates (FDRs) < 0.05 and threshold of  $|\log_{2}FC| > 1$ ) were defined as DEGs. To elucidate the mechanism of DEGs, Gene Ontology (GO) analysis and the Kyoto Encyclopedia of Genes and Genomes (KEGG) annotation were performed using DAVID online tools.

Univariate regression analysis was used to analyze the association between DEGs and overall survival, respectively. Only genes with hazard ratio (HR) > 1 and  $\log_{2}FC > 1$  ( $P < 0.05$ ), or HR < 1 and  $\log_{2}FC < -1$  ( $P < 0.05$ ) will be included in the subsequent analysis. Further, the backward stepwise selection was applied using the likelihood ratio test with Akaike’s information criterion as the stopping rule to obtain the risk model, and the lowest value of AIC provided the sensitivity and specificity. Subsequently, six genes (F2RL1, PRKDC, TNFSF11, INHA, PLA2G3 and TUBB1) were selected. The risk score model was calculated as follows: risk score =  $\sum_{i=1}^n \text{Coeff}_i * X_i$ . In this formula,  $n$ ,  $X_i$ , and  $\text{Coeff}_i$  represented the number of hub genes, gene expression level, and coefficient of gene, respectively. In the TCGA cohort and GEO cohort, patients were divided into low- and high-risk groups according to the optimal cut-off value. The “surv\_cutpoint” function of “survminer” R package was used to measure optimal cut-off value of risk score. Then, K-M survival curves were carried out to compare the survival differences between low-risk and high-risk groups. Moreover, time-dependent ROC analysis was applied to assess the efficiency and accuracy of the established prognostic risk model of LUAD using the “survivalROC” R package.

### Assessment of immune microenvironment

The ESTIMATE algorithm was used to calculate the ratio of the immune and stromal components in the tumor microenvironment based on gene expression profiles<sup>9</sup>. Then, the tumor purity and TME scores which include ImmuneScore (indicating the level of immune cells infiltration), StromalScore (indicating the level of stromal cells infiltration) and ESTIMATEScore (reflecting the sum of both) for each patient were measured by the R package “estimate”<sup>23</sup>.

### Tumor immune cell infiltration analysis

For a more comprehensive estimation of immune cell infiltration, TIMER, XCELL, CIBERSORT abs, QUANTISEQ, MCPOUNTER, EPIC, CIBERSORT, and ssGSEA algorithms were applied to quantify the relative proportions of immune cells in the LUAD patients<sup>24</sup>. The relationship between risk score levels and immune infiltrating cells was explored using Spearman’s rank correlation analysis.

### Nomogram construction and verification

To provide clinicians with a quantitative method for predicting the prognosis of LUAD patients, the R packages “rms” and “survival” were applied to establish the nomogram, and the prognosis was predicted by combining the risk score and clinical characteristics (age, gender, smoking, T, and N) of LUAD patients. The accuracy of the model was evaluated by calibration curves. The calibration curve approaching to 45-degree diagonal line indicates perfect predictive capability.

### Drug sensitivity analysis

The “pRRophetic” R package was applied to predict the half-maximal inhibitory concentration (IC<sub>50</sub>) of each small molecular compounds for the low-risk and high-risk groups<sup>25</sup>.

## Cell lines and culture

The following human LUAD cell lines (A549 (RRID:CVCL\_0023), PC-9 (RRID:CVCL\_B260)) and human normal lung epithelium cell line BEAS-2B were purchased from the Chinese Academy of Sciences (Shanghai, China). A549 and PC-9 were cultured in RPMI-1640 (KeyGen Biotechnology, Nanjing, China), and BEAS-2B was cultured in DMEM (KeyGen Biotechnology, Nanjing, China), supplemented with 10% fetal bovine serum (FBS) at 37 °C with 5% CO<sub>2</sub> in a humidified atmosphere. All cell lines were authenticated by STR profiling.

## Quantitative real-time polymerase chain reaction (qPCR)

Total RNA were isolated from cells using TRIzol reagent (Invitrogen, Carlsbad, CA) according to the manufacturer's protocol, and then reversely transcribed into cDNA using 4 × EZscript RT Mix II (EZBioscience, USA) Kit. cDNA was amplified using SYBR Green PCR Kit (EZBioscience, USA) with a QuantStudio5 Sequence Detector. The primers for amplification were listed in Supplementary Table 2.

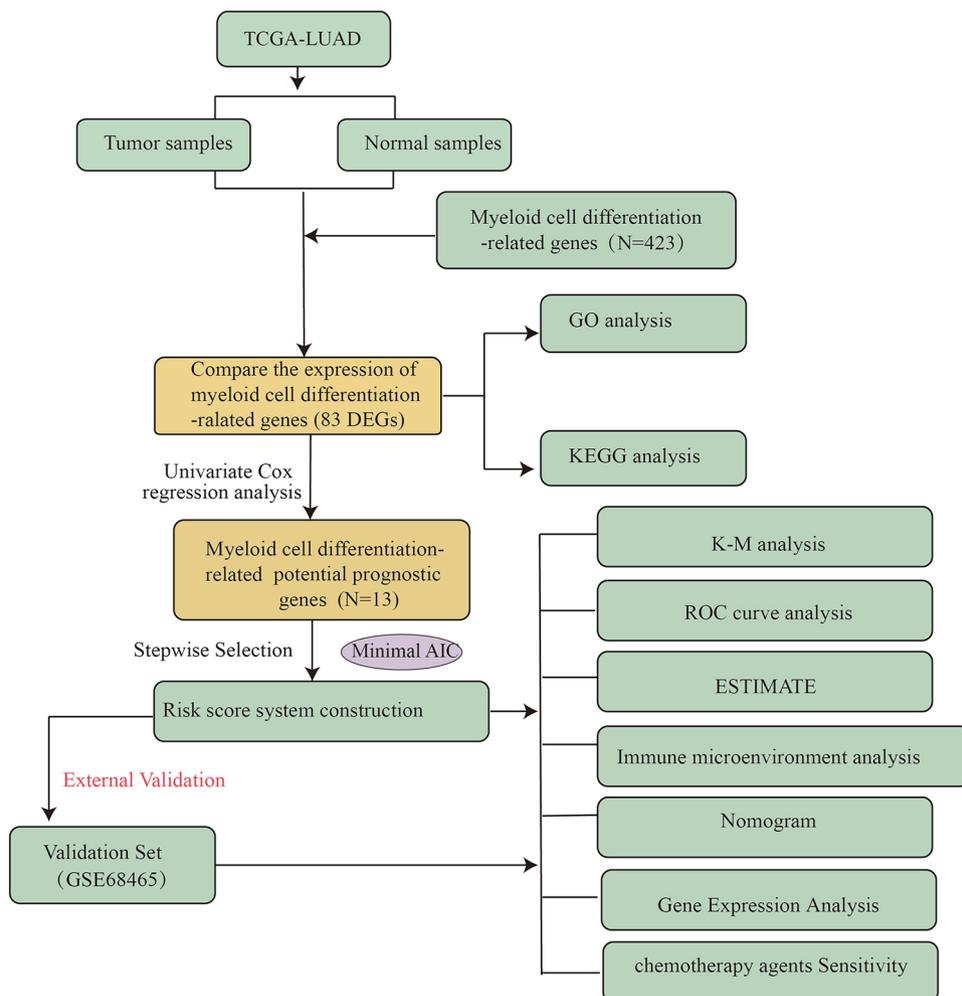
## Patient consent statement

TCGA and GEO belong to public databases. The patients involved in the databases provided their written informed consent to participate in this study. Users can download relevant data for free for research and publish relevant articles.

## Results

### Construction and evaluation of myeloid cell differentiation—related gene signature

To describe our study more clearly, a flow chart is provided in Fig. 1 to summarize the entire study. The RNA-seq data of 58 normal lung tissues and 501 LUAD tissues from the TCGA-LUAD cohort were merged and normalized to make further comparisons. The detailed basic characteristics of these patients are showed in Table 1. In the present study, 83 DEGs were identified, the Volcano plot and heatmap of DEGs are shown in Fig. 2A and Fig. 2B. Then, we performed GO functional annotation and KEGG analyses for DEGs using DAVID online analyses



**Figure 1.** Flow chart of data collection and analysis in the present study.

Characteristics	Groups	TCGA cohort (N = 501)	GSE68465 cohort (N = 442)
		No (%)	No (%)
Age	≤ 65	247 (49.3)	231 (52.3)
	> 65	254 (50.7)	211 (47.7)
Gender	Female	270 (53.9)	219 (49.5)
	Male	231 (46.1)	223 (50.5)
Pathologic Stage	I	277 (56.2)	–
	II	114 (23.1)	–
	III	77 (15.6)	–
	IV	25 (5.1)	–
T	T1	169 (33.9)	150 (34.1)
	T2	267 (53.6)	251 (57.0)
	T3	45 (9.0)	28 (6.4)
	T4	17 (3.4)	11 (2.5)
N	N0	328 (67.5)	299 (68.1)
	N1	89 (18.3)	87 (19.8)
	N2	67 (13.8)	53 (12.1)
	N3	2 (0.4)	–
M	M0	332 (93.3)	–
	M1	24 (6.7)	–
Smoking	No	71 (14.6)	49 (14.0)
	Yes	416 (85.4)	300 (86.0)
Survival status	Death	123 (24.6)	236 (53.4)
	Alive	378 (75.4)	206 (46.6)

**Table 1.** Basic characteristics of lung adenocarcinoma patients.

(Fig. 2C,D). Univariate cox regression analysis was conducted to estimate the prognostic significance of DEGs. Of the DEGs, 13 were correlated with OS in the univariate Cox regression analyses (all  $P < 0.05$ ). Then, based on the selection criteria, we carried out backward stepwise selection to determine the best performance efficacy predictive model with the minimum AIC value. Ultimately, a risk model including F2RL1, PRKDC, TNFSF11, INHA, PLA2G3 and TUBB1 was successfully constructed. Based on the formula ( $-0.102 \times \text{gene expression level of PLA2G3} + -0.129 \times \text{gene expression level of TUBB1} + 0.081 \times \text{gene expression level of TNFSF11} + 0.053 \times \text{gene expression level of INHA} + 0.199 \times \text{gene expression level of F2RL1} + 0.132 \times \text{gene expression level of PRKDC}$ ), each patient's risk score was calculated. The results obtained were used to divided patients into low-risk ( $n = 390$ ) and high-risk ( $n = 111$ ) subsets based on the optimal cut-off point, which was calculated by “survminer” package. The K-M curve showed that the high-risk group had a greater mortality risk compared with the low-risk group (Fig. 3A). Meanwhile, the risk heatmap clearly shows F2RL1, PRKDC, TNFSF11 and INHA were up-regulated as the risk score increases, while PLA2G3 and TUBB1 were down-regulated (Fig. 3B).

### Validation of the six-gene signature in the GSE68465 cohort

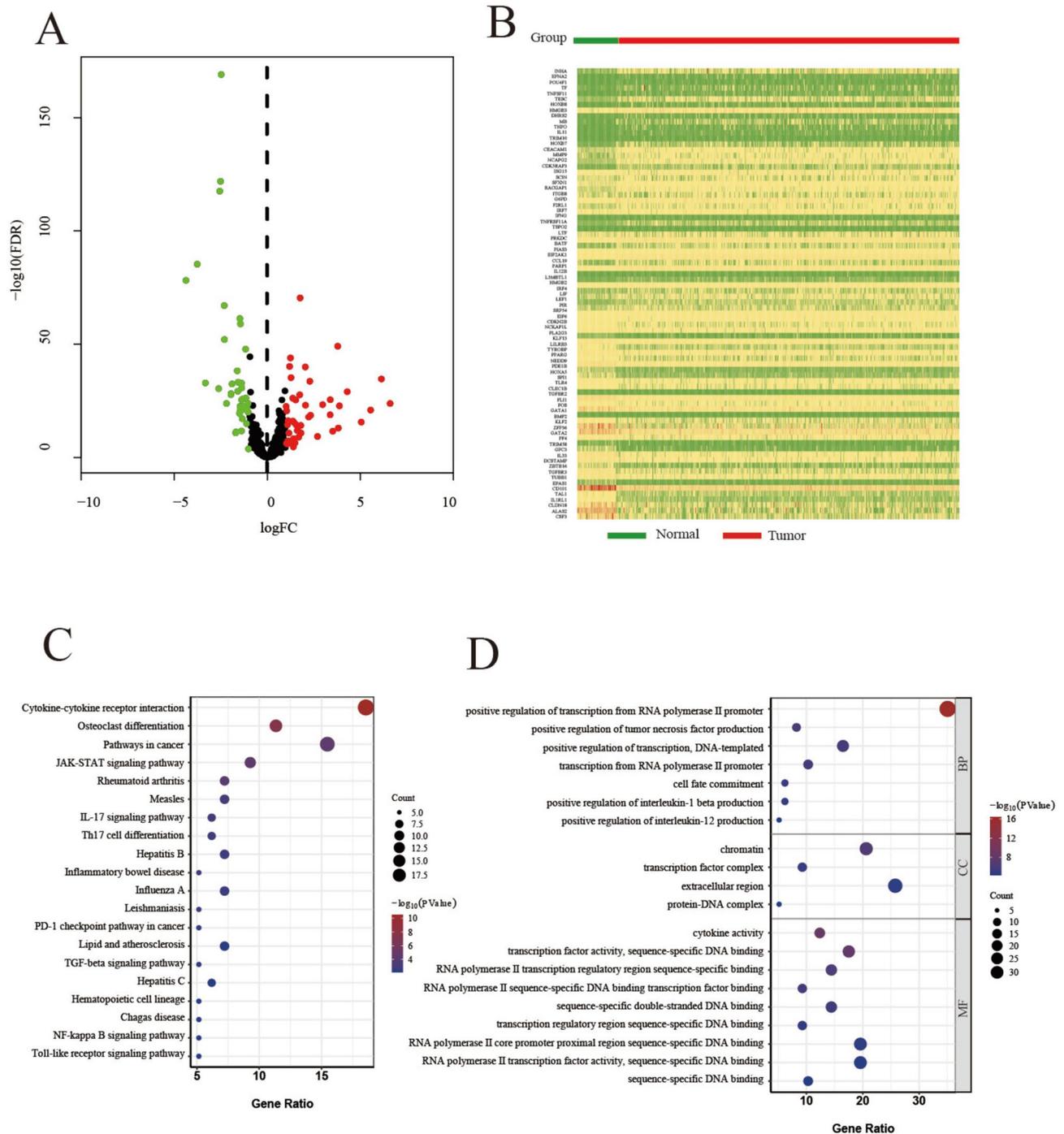
To further test the performance of the constructed model, the risk score of each sample ( $n = 442$ ) in GSE68465 was calculated using the same risk formula above. Ultimately, 49.32% of the LUAD patients ( $n = 218$ ) in the GSE68465 cohort were categorized into the low-risk subgroup, while the remaining patients ( $n = 224$ ) were categorized into the high-risk subgroup. Consistent with the results of the TCGA-LUAD dataset, the patients in the high-risk group also had lower survival rate than those in the low-risk group with  $p = 0.003$  (Fig. 3C and Fig. 3D).

### Correlation between risk score and immune landscape

To characterize the immune environment of LUAD patients, we assessed the immune infiltration of tumors using the TIMER, XCELL, CIBERSORT abs, QUANTISEQ, MCPOUNTER, EPIC, and CIBERSORT algorithms. Then, spearman's rank correlation analysis showed most immune and stromal cells were negatively correlated with risk scores (Figs. 4A and 5A). Furthermore, the results of ESTIMATE algorithm showed that the estimate score and immune score of the low-risk group were significantly higher than those of the high-risk group, and lower tumor purity was observed in the low-risk group (Figs. 4B and 5B). The relative proportion of 28 tumor infiltrating immune cells (TIICs) calculated by the ssGSEA algorithm in the both cohorts were presented in Fig. 4C and Fig. 5C. Compared with high-risk subset, the low-risk subset had higher percentages of Activated B cell, Activated CD4 T cell, T follicular helper cell, Eosinophil, Mast cell, Monocyte and Natural killer cell.

### Therapeutic potential of the myeloid cell differentiation—related gene signature in LUAD

In order to elucidate the potential correlation between myeloid cell differentiation—related gene signature and immunotherapy, we compared the expression differences of 12 immune checkpoints (CD28, CD274, CTLA4,

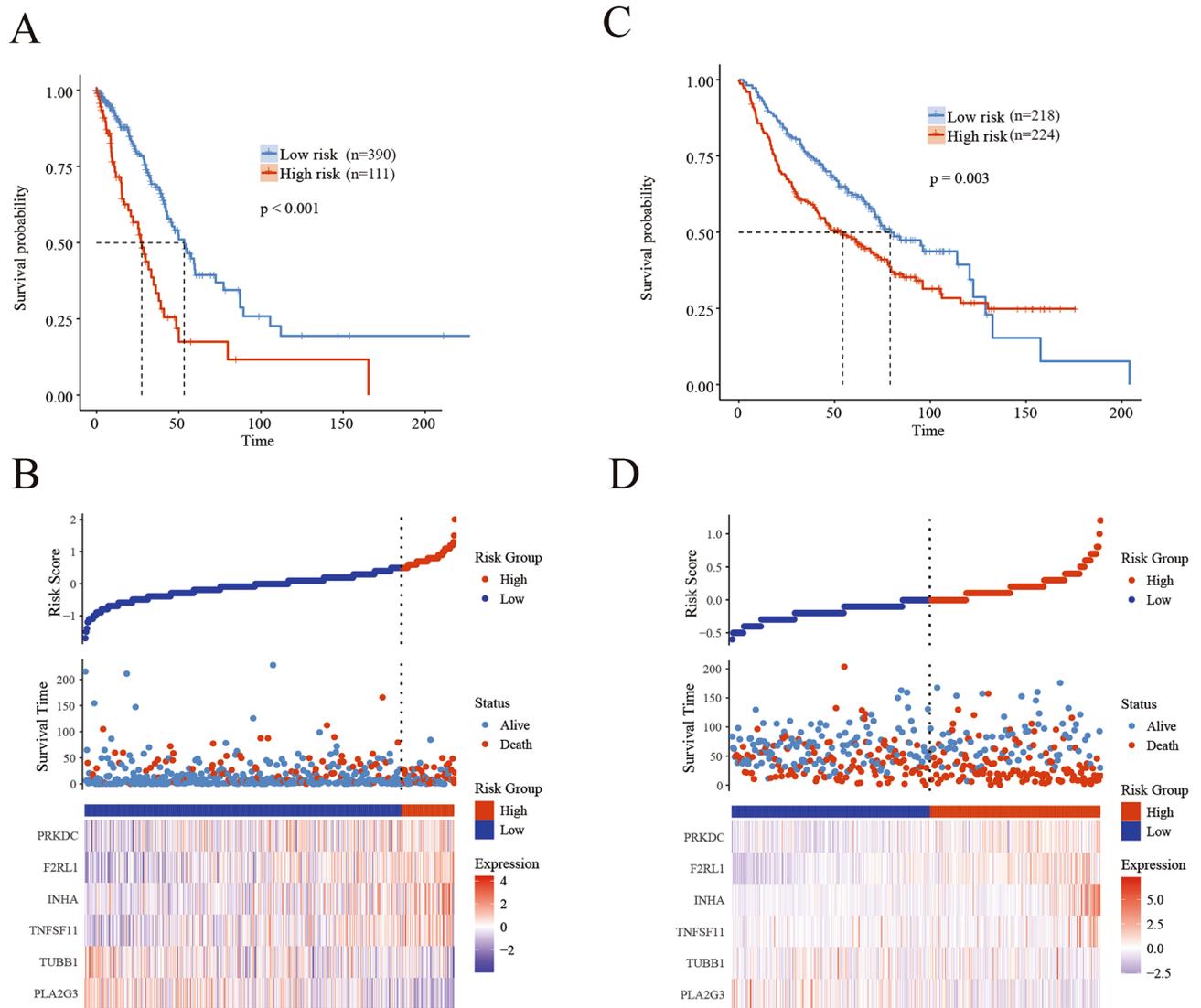


**Figure 2.** Identification of the myeloid cell differentiation -related DEGs in TCGA cohort. **(A, B)** Volcano plots and heatmap visualize the myeloid cell differentiation—related DEGs in TCGA-LUAD. The red dots represent up-regulated genes and the blue dots represent down-regulated genes ( $FDR < 0.05$ ,  $|FC| > 1$ ). **(C)** Top 20 terms of BP, CC, and MF in GO analysis for gene signature-related DEGs. **(D)** Top 20 terms of pathways in KEGG analysis for gene signature-related DEGs.

TNFSF4, BTLA, TGFBR1, TNFSF18, CD40LG, CD70, CD200, TNFSF9 and CD48) between the two subgroups and detected that five immunotherapy targets had high expression levels in the high-risk subgroup (Fig. 6A).

Increasing studies have focused on m6A, the most abundant mRNA modification and is involved in tumor occurrence and development<sup>26</sup>. Our results showed the expression of m6A-related genes (METTL14, YTHDC1, YTHDF1, YTHDF2, FMR1 and IGF2BP1) were remarkably higher among the high-risk group compared with low-risk patients (Fig. 6B).

As for chemotherapy sensitivity, we calculated six chemotherapeutic drugs via “pRRophetic” R package. The results showed JNK inhibitor VIII, cisplatin, paclitaxel, Cyclophosphamide, Rapamycin and Erlotinib displayed lower



**Figure 3.** The overall performance of the myeloid cell differentiation -related gene signature in TCGA and GEO cohorts. K-M survival analysis of myeloid cell differentiation—related gene signature in TCGA (A) and GEO (C) cohorts. (B, D) Distributions of risk score and patients' survival time in TCGA and GEO LUAD patients, and heatmaps of six gene expression profiles in both cohorts.

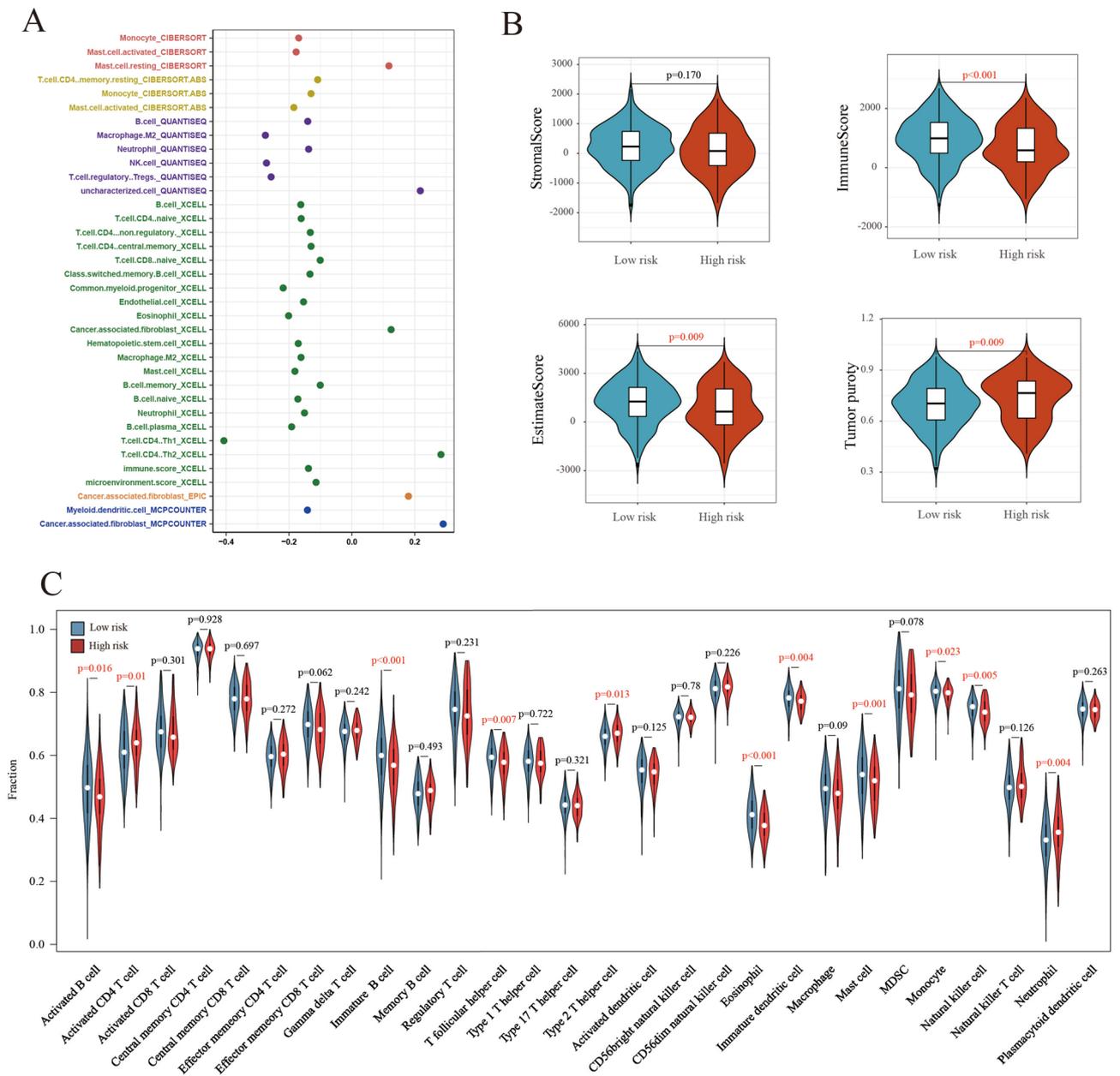
IC50 values in the high-risk group (Fig. 6C). Thus, high risk patients based on the myeloid cell differentiation -related gene signature will better benefit from chemotherapy.

### Construction of a nomogram

Time-dependent ROC curves were plotted by “survivalROC” R package, and the AUC was computed at different time points to estimate the predictive performance of the gene signature. In TCGA dataset, the AUC of 1-, 3- and 5-year OS were 0.682, 0.675 and 0.656, respectively (Fig. 7A). Furthermore, the risk score model was an independent factor in both univariate and multivariate Cox analysis ( $P < 0.001$ , HR: 2.718, 95% CI: 1.957–3.776 and  $P < 0.001$ , HR: 2.315, 95% CI: 1.513–3.541, respectively) (Table 2).

Besides, in GEO dataset, the AUC for OS were 0.658 at 1 year, 0.619 at 3 years, and 0.589 at 5 years (Fig. 7B). Also, the six-gene signature was an independent factor in both the univariate and multivariate Cox analyses ( $p < 0.001$ , HR: 3.553, and 95% CI: 1.783–7.080; and  $p = 0.010$ , HR: 2.651, and 95% CI: 1.264–5.560, respectively) (Table 2).

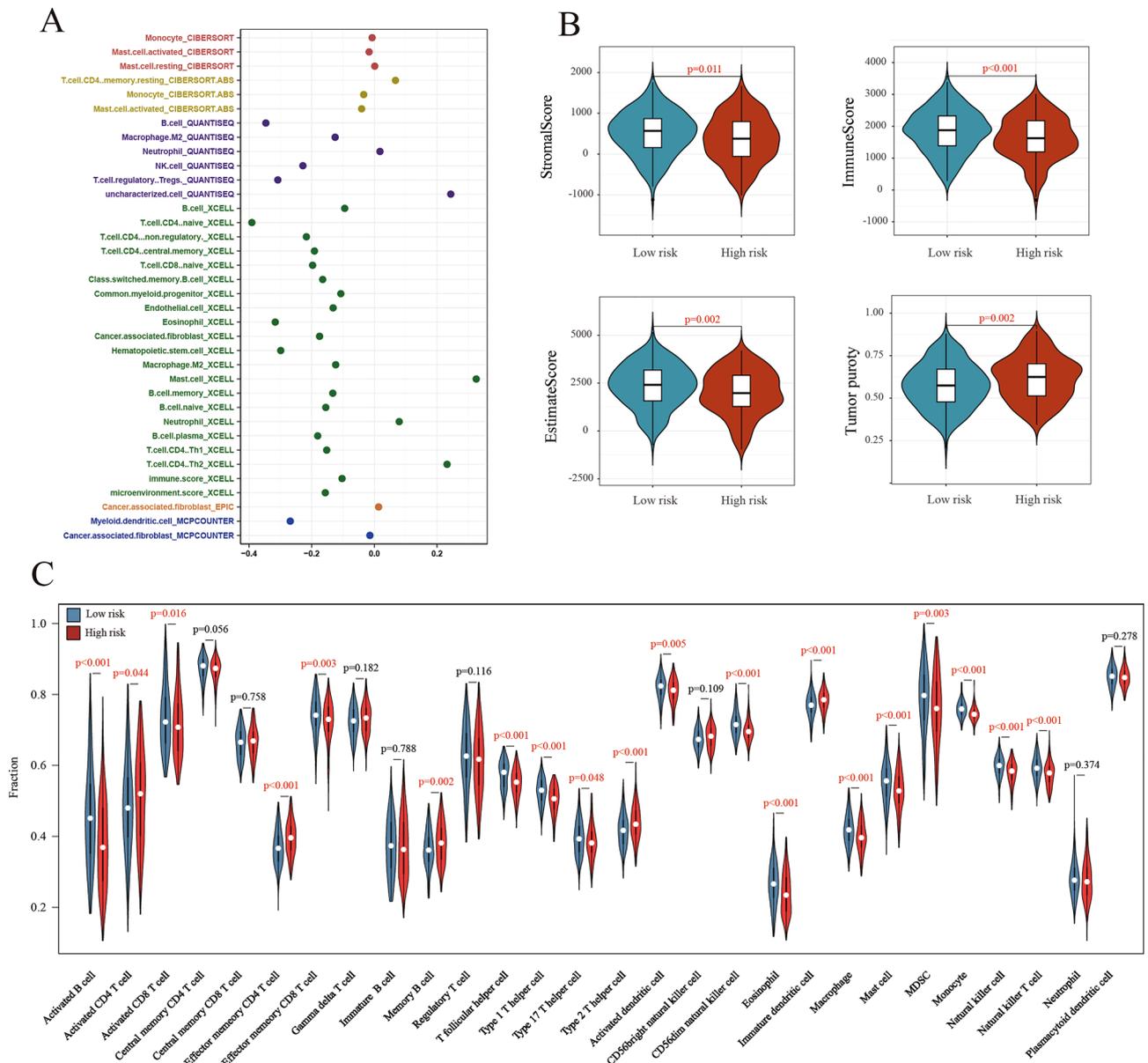
To further test the performance of the constructed model, we combined it with the clinical characteristics (included age, gender, T stage, N stage, and smoking) of the TCGA-LUAD cohort to construct a quantitative nomogram for predicting 1-year, 3-year, and 5-year OS of LUAD (Fig. 7C). Calibration curves were then plotted to compare the predicted probability of OS with the actual OS. The results showed that the predicted results of OS in terms of the 1-, 3-, and 5-year were consistent with the actual results in both TCGA-LUAD cohort (Fig. 7D–F) and the GSE68465 cohort (Fig. 7G–I).



**Figure 4.** Evaluation of the TME between the low-risk and high-risk groups in TCGA cohort. **(A)** Bubble chart revealed that most immune and stromal cells were critically negatively correlated with the risk score. **(B)** Violin plots revealed the difference of immune scores, stromal score, estimate score and tumor purity between high-risk and low-risk groups in TCGA-LUAD patients **(C)** Different infiltrating abundances of 28 TIICs estimated by ssGSEA between subgroups.

### Analysis of the six gene expression in silico and in cell lines

We analyzed the normalized mRNA expression of six genes in TCGA cohort and the result indicated F2RL1, INHA, PRKDC and TNFSF11 were significantly higher in LUAD tumor tissues than that in adjacent normal samples, while PLA2G3 and TUBB1 were opposite (Fig. 8A). qPCR was used to validate mRNA level of six genes in LUAD cell lines (PC-9 and A549) and human normal lung epithelial cells (BEAS-2B). The results revealed that the expression of F2RL1, INHA, PRKDC and TNFSF11 were significantly increased in PC-9 and A549 cell lines. In addition, PLA2G3 and TUBB1 were down-regulated significantly (Fig. 8B). These data was in accordance with our bioinformatics analysis results.



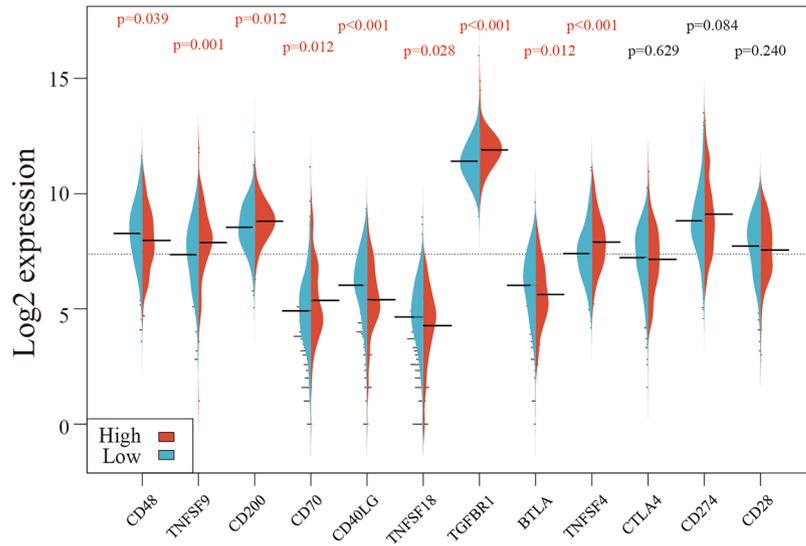
**Figure 5.** Evaluation of the TME between the low-risk and high-risk groups in GEO cohort. (A) Bubble chart revealed that most immune and stromal cells were critically negatively correlated with the risk score. (B) Violin plots revealed the difference of immune scores, stromal score, estimate score and tumor purity between high-risk and low-risk groups in GSE68465 patients (C) Different infiltrating abundances of 28 TIICs estimated by ssGSEA between subgroups.

### Discussion

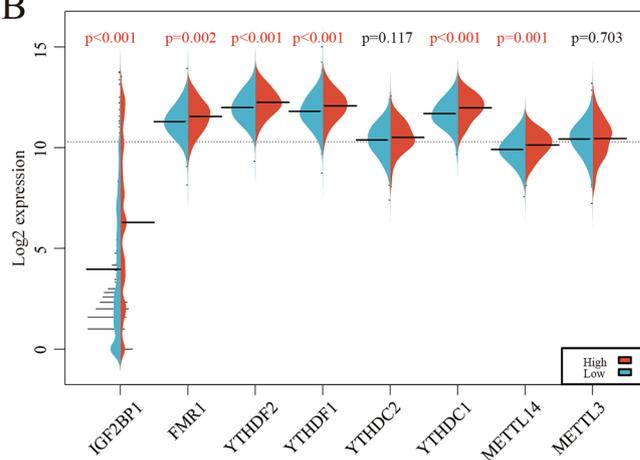
LUAD is the most common malignant tumor and a major cause of cancer-related deaths worldwide<sup>27</sup>. Due to sequencing technologies, molecular targeted therapy and immunotherapy, significant progresses have been achieved in treating LUAD, but LUAD prognosis remains poor<sup>28</sup>. Thus, identifying reliable prognostic biomarkers will help predict patient survival and guide precision medical intervention.

Myeloid cells constitute the primary component of the immunosuppressive tumor microenvironment<sup>29,30</sup>. Particularly, the existence of tumor-infiltrating myeloid cells in LUAD tumors has been confirmed to be associated with LUAD patient prognosis, implying that myeloid cell differentiation—related genes could be sever as prognostic indicators for LUAD patient survival<sup>31,32</sup>. Increasing studies indicated that targeting myeloid cells in TME might be an emerging and promising treatment for immunotherapy, reshaping the immunosuppressive TME and enhancing the efficacy of tumor immunotherapy<sup>33</sup>. In this research, we explored the prognostic value of risk model based on myeloid cell differentiation—related pathway genes in LUAD, in place of directly detecting the presence of myeloid cells in LUAD tissue. Given the crucial role that myeloid cells play in cellular immune surveillance, it is necessary to identify potential biomarkers based on myeloid cell differentiation—related pathway genes to stratify patients for the personalized therapy.

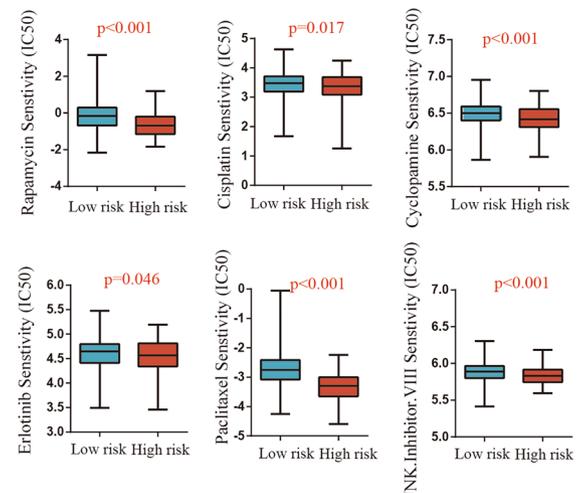
A



B



C

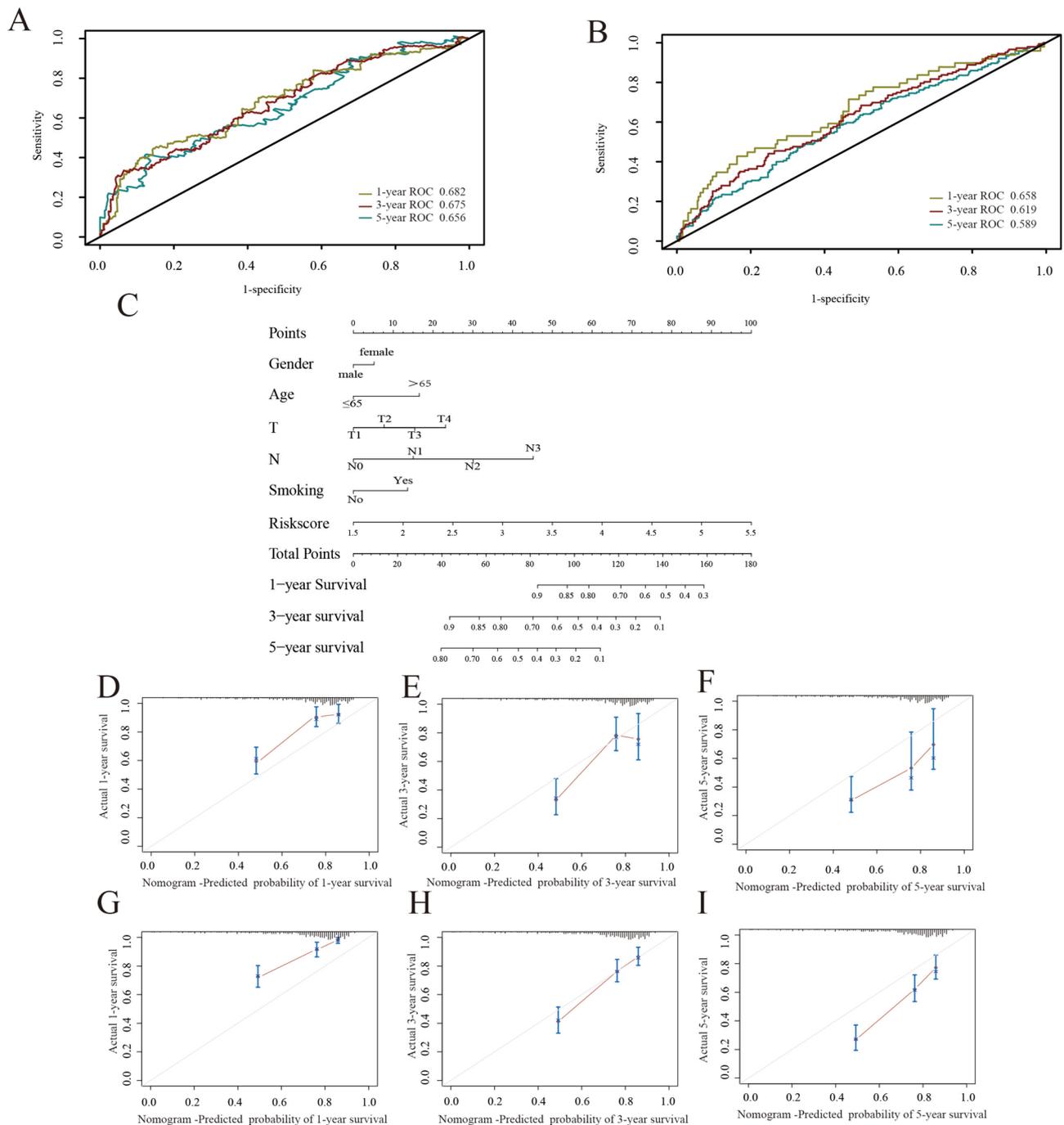


**Figure 6.** Expression of immune checkpoints, m6A genes and sensitivity to chemotherapeutic drugs between the high-risk and low-risk subgroups in TCGA cohort. **(A)** The expression level of key immune checkpoint genes and **(B)** m6A genes between the two groups divided by risk score level. **(C)** The IC50 of six chemotherapeutic drugs compared between high-risk and low-risk groups.

In the present study, we developed a six-gene myeloid cell differentiation—related gene signature that was comprised of F2RL1, PRKDC, TNFSF11, INHA, PLA2G3 and TUBB1 in TCGA-LUAD cohort via backward stepwise regression model, which could efficiently and independently predict the LUAD patients' prognosis and therapy response. Furthermore, we validated this signature in an independent LUAD cohort from the GEO database.

Based on the expression of the six genes, the LUAD patients were divided into high- and low-risk subsets. The prognosis of the low-risk subsets was significantly better than that of the high-risk subsets. Univariate and multivariate cox regression analysis confirmed that myeloid cell differentiation -related gene signature was an independent risk factor for LUAD. Moreover, when we further assessed the differences in subtypes of immune cells, we found that low-risk patients based on the gene signature was associated with more immune cells, namely Activated B cell, Activated CD4 T cell, T follicular helper cell, Eosinophil, Mast cell, Monocyte and Natural killer cell, showing benefits on tumors' prognosis<sup>34</sup>.

Subsequently, to further evaluate the performance of the constructed model, we constructed a nomogram by integrating clinical characteristics, providing clinicians with a personalized scoring system. The calibration curves confirmed that the nomogram predicted 1-, 3-, and 5-year survival probabilities were in accordance with observed survival results. Moreover, the analysis of immune genes and m6A genes of LUAD patients in the two risk subsets indicated that the risk score was significantly associated with immune checkpoint molecules and m6A genes.



**Figure 7.** Evaluation of the predictive ability of the myeloid cell differentiation—related gene signature. **(A, B)** Time-dependent ROC curve of risk model at 1-, 3-, 5- year in TCGA and GEO cohorts. **(C)** The nomogram for predicting the overall survival of TCGA-LUAD patients. **(D–F)** Calibration curves for evaluating the compatibility between the predicted and actual outcomes at 1-, 3-, 5- year in TCGA cohort. **(G–I)** Calibration curves for evaluating the compatibility between the predicted and actual outcomes at 1-, 3-, 5- year in GEO cohort.

We also found that JNK inhibitor VIII, cisplatin, paclitaxel, cyclophamide, Rapamycin and Erlotinib showed better responses in the treatment of the high-risk score LUAD patients, indicating that gene signature was a powerful predictive tool for immune therapy, drug sensitivity and could promote personalized precision medical treatment in the future.

Among the six genes in the risk signature, F2R-like Trypsin Receptor 1 (F2RL1), also known as PAR2, as a G-protein-coupled receptor, played an oncogenic role in a wide variety of cellular processes and can be activated by serine proteases<sup>35,36</sup>. In lung cancer, F2RL1 can promote tumor proliferation and angiogenesis by promoting EGFR phosphorylation and upregulated VEGFA expression<sup>37</sup>. Protein kinase, DNA-activated, catalytic polypeptide (PRKDC), is a component of the non-homologous end joining pathway involved in the DNA double-strand

Variable	Univariate cox analysis		Multivariate cox analysis	
	HR (95% CI)	Pvalue	HR (95% CI)	Pvalue
TCGA-LUAD (overall survival)				
Age	1.293 (0.902–1.851)	0.161	1.584 (1.017–2.467)	0.042
Gender	0.945 (0.662–1.350)	0.757	0.785 (0.505–1.219)	0.281
T	1.563 (1.252–1.951)	<0.001	1.149 (0.854–1.545)	0.358
N	1.755 (1.432–2.151)	<0.001	0.998 (0.636–1.566)	0.994
M	1.718 (0.916–3.223)	0.091	0.311 (0.078–1.226)	0.095
Pathologic stage	1.660 (1.410–1.954)	<0.001	1.989 (1.183–3.346)	0.009
Smoking	0.910 (0.510–1.625)	0.752	1.501 (0.713–3.159)	0.284
risk score	2.718 (1.957–3.776)	<0.001	2.315 (1.513–3.541)	<0.001
GSE68465 (overall survival)				
Age	1.546 (1.194–2.001)	<0.001	1.723 (1.270–2.338)	<0.001
Gender	1.390 (1.073–1.801)	0.012	1.066 (0.782–1.453)	0.685
T	1.672 (1.394–2.007)	<0.001	1.467 (1.149–1.874)	0.002
N	2.012 (1.712–2.365)	<0.001	2.037 (1.688–2.458)	<0.001
Smoking	1.246 (0.790–1.965)	0.344	1.475 (0.908–2.397)	0.116
Risk score	3.553 (1.783–7.080)	<0.001	2.651 (1.264–5.560)	0.010

**Table 2.** Univariate analysis and multivariate analysis of the correlation of risk score with outcomes among lung adenocarcinoma patients in two cohorts.

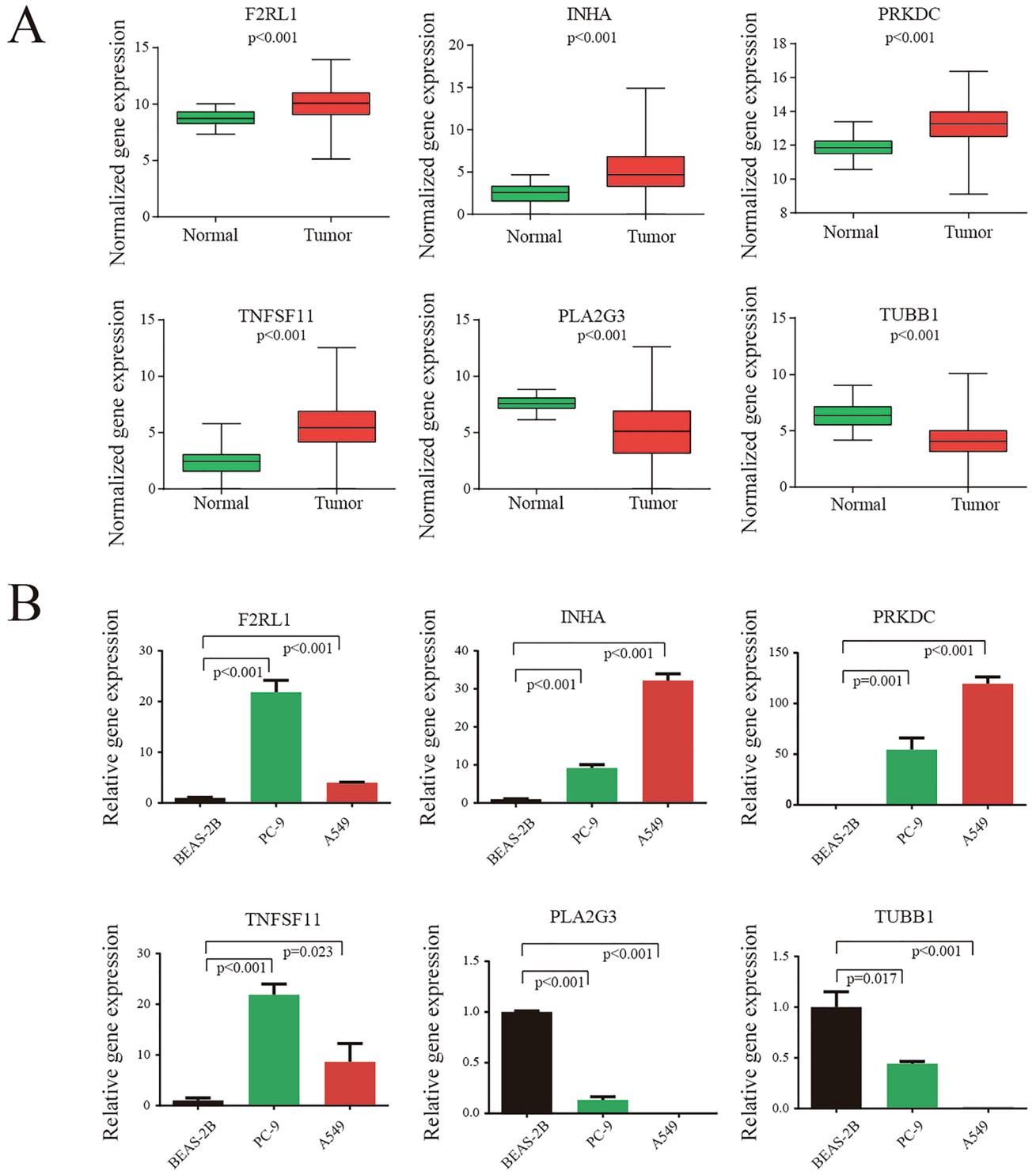
break (DSB) repair response<sup>38</sup>. Several studies have shown that PRKDC plays an important role in regulating the motility, migration, and development of lung cancer<sup>38,39</sup>. Also, Tan et al. demonstrated PRKDC mutations can be potential biomarkers to guide LUAD patient selection for checkpoint blockade immunotherapy<sup>40</sup>. TNFSF11, also known as RANKL, promotes osteoclastogenesis and facilitates bone metastasis of lung cancer cell<sup>41</sup>. The RANKL-mediated pathway is also an important immune mediator in lung cancer, and anti-RANKL therapy has been linked to improved responses to immunotherapy<sup>42</sup>. As for INHA, was found to be remarkably overexpressed in lung cancer tissue compared with normal tissue. Also, high INHA expression associated with shorter and a higher pathological stage<sup>43</sup>. Group III sPLA2 (PLA2G3), regulates antigen- and IgE-dependent mast cell degranulation and can induce mast cell maturation<sup>44</sup>. Also, PLA2G3 is a possible target to adjust neuronal damage induced by free radicals<sup>45</sup>. However, as far as I know, the regulation function of PLA2G3 in lung cancer has not been reported. The last TUBB1, encodes the tubulin b-1 chain, has been shown to be particularly expressed in platelets and megakaryocytes, and TUBB1 dysfunction will cause genome instability. Also, at the biological level in lung cancer progression is, as yet, unknown.

Unlike the previous risk model in LUAD, our research was the first myeloid cell differentiation—associated gene signature, which was thought to be highly associated with the immune cell infiltration in LUAD. It can predict which LUAD patients are more prone to chemotherapy agents and immune checkpoint inhibitors.

Of course, our study has several limitations. Firstly, the construction and validation of myeloid cell differentiation-related gene signature is based on data from public databases. Its robustness should be further enhanced and validated by multicenter clinical research in the future. Secondly, more experiments are needed to verify the individual or combined roles of the six genes involved in gene signature in LUAD. Additional investigations, both in vitro and in vivo, are required to explore these interactions fully.

## Conclusions

To sum up, myeloid cell differentiation—related gene signature is an efficient and independent prognostic model for assessing the prognosis of LUAD patients. We concluded the risk model was significantly associated with immune cells infiltration in LUAD, providing new sights of the myeloid cell differentiation-related genes in anti-tumor immunity and can be used to guide the therapy selection for LUAD patients in the foreseeable future.



**Figure 8.** differences of six-gene expression levels. (A) F2RL1, PRKDC, TNFSF11, INHA, PLA2G3 and TUBB1 mRNA expression levels were significantly different in the unpaired tumor-adjacent normal TCGA LUAD samples. (B) F2RL1, PRKDC, TNFSF11, INHA, PLA2G3 and TUBB1 mRNA expression levels were significantly different in LUAD cell lines and human normal lung epithelial cells.

### Data availability

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.

Received: 7 May 2024; Accepted: 19 July 2024

Published online: 29 July 2024

## References

- Siegel, R. L., Miller, K. D., Fuchs, H. E. & Jemal, A. Cancer Statistics, 2021. *CA: a cancer journal for clinicians* **71**, 7–33. <https://doi.org/10.3322/caac.21654> (2021).
- Lortet-Tieulent, J. *et al.* International trends in lung cancer incidence by histological subtype: adenocarcinoma stabilizing in men but still increasing in women. *Lung Cancer* **84**, 13–22. <https://doi.org/10.1016/j.lungcan.2014.01.009> (2014).
- Song, Y., Zhang, J., Fang, L. & Liu, W. Prognostic necroptosis-related gene signature aids immunotherapy in lung adenocarcinoma. *Front. Genet.* **13**, 1027741. <https://doi.org/10.3389/fgene.2022.1027741> (2022).
- Rizzo, A. *et al.* Impact of Proton Pump Inhibitors and Histamine-2-Receptor Antagonists on Non-Small Cell Lung Cancer Immunotherapy: A Systematic Review and Meta-Analysis. *Cancers (Basel)* <https://doi.org/10.3390/cancers14061404> (2022).
- Rizzo, A. Identifying optimal first-line treatment for advanced non-small cell lung carcinoma with high PD-L1 expression: a matter of debate. *Br. J. Cancer* **127**, 1381–1382. <https://doi.org/10.1038/s41416-022-01929-w> (2022).
- Dall’Olio, F. G. *et al.* Immortal time bias in the association between toxicity and response for immune checkpoint inhibitors: a meta-analysis. *Immunotherapy* **13**, 257–270. <https://doi.org/10.2217/imt-2020-0179> (2021).
- Rizzo, A. *et al.* Hypertransaminasemia in cancer patients receiving immunotherapy and immune-based combinations: the MOUSEION-05 study. *Cancer Immunol. Immunother.* **72**, 1381–1394. <https://doi.org/10.1007/s00262-023-03366-x> (2023).
- Guven, D. C. *et al.* The association between albumin levels and survival in patients treated with immune checkpoint inhibitors: A systematic review and meta-analysis. *Front. Mol. Biosci.* **9**, 1039121. <https://doi.org/10.3389/fmolb.2022.1039121> (2022).
- Lei, K. *et al.* Development and clinical validation of a necroptosis-related gene signature for prediction of prognosis and tumor immunity in lung adenocarcinoma. *Am. J. Cancer Res.* **12**, 5160–5182 (2022).
- Lamberti, G. *et al.* Beyond EGFR, ALK and ROS1: Current evidence and future perspectives on newly targetable oncogenic drivers in lung adenocarcinoma. *Crit. Rev. Oncol. Hematol.* <https://doi.org/10.1016/j.critrevonc.2020.103119> (2020).
- Mollica, V. *et al.* The impact of ECOG performance status on efficacy of immunotherapy and immune-based combinations in cancer patients: The MOUSEION-06 study. *Clin. Exp. Med.* **23**, 5039–5049. <https://doi.org/10.1007/s10238-023-01159-1> (2023).
- Santoni, M. *et al.* The impact of gender on The efficacy of immune checkpoint inhibitors in cancer patients: The MOUSEION-01 study. *Crit Rev Oncol Hematol* <https://doi.org/10.1016/j.critrevonc.2022.103596> (2022).
- Ries, A. *et al.* Mesothelioma-associated fibroblasts enhance proliferation and migration of pleural mesothelioma cells via c-Met/PI3K and WNT signaling but do not protect against cisplatin. *J. Exp. Clin. Cancer Res.* **42**, 27. <https://doi.org/10.1186/s13046-022-02582-0> (2023).
- Mittal, V. Epithelial Mesenchymal Transition in Tumor Metastasis. *Ann. Rev. Pathol.* **13**, 395–412. <https://doi.org/10.1146/annurev-pathol-020117-043854> (2018).
- Hanahan, D. & Coussens, L. M. Accessories to the crime: functions of cells recruited to the tumor microenvironment. *Cancer Cell* **21**, 309–322. <https://doi.org/10.1016/j.ccr.2012.02.022> (2012).
- Sato, R. *et al.* TGFbeta Signaling Activated by Cancer-Associated Fibroblasts Determines the Histological Signature of Lung Adenocarcinoma. *Cancer Res.* **81**, 4751–4765. <https://doi.org/10.1158/0008-5472.CAN-20-3941> (2021).
- Wu, J. *et al.* CXCL12 derived from CD248-expressing cancer-associated fibroblasts mediates M2-polarized macrophages to promote nonsmall cell lung cancer progression *Biochimica et biophysica acta. Mol. Basis Dis.* **1868**, 166521. <https://doi.org/10.1016/j.bbadis.2022.166521> (2022).
- Salazar, Y. *et al.* Microenvironmental Th9 and Th17 lymphocytes induce metastatic spreading in lung cancer. *J. Clin. Investing.* **130**, 3560–3575. <https://doi.org/10.1172/JCI124037> (2020).
- Gabrilovich, D. I., Ostrand-Rosenberg, S. & Bronte, V. Coordinated regulation of myeloid cells by tumours. *Nat. Rev. Immunol.* **12**, 253–268. <https://doi.org/10.1038/nri3175> (2012).
- Hirai, H. *et al.* CCR1-mediated accumulation of myeloid cells in the liver microenvironment promoting mouse colon cancer metastasis. *Clin. Exp. Metastasis* **31**, 977–989. <https://doi.org/10.1007/s10585-014-9684-z> (2014).
- Stefanescu, C., Van Gogh, M., Roblek, M., Heikenwalder, M. & Borsig, L. TGFbeta Signaling in Myeloid Cells Promotes Lung and Liver Metastasis Through Different Mechanisms. *Front. Oncol.* <https://doi.org/10.3389/fonc.2021.765151> (2021).
- Yan, H. H. *et al.* Gr-1+CD11b+ myeloid cells tip the balance of immune protection to tumor promotion in the premetastatic lung. *Cancer Res.* **70**, 6139–6149. <https://doi.org/10.1158/0008-5472.CAN-10-0706> (2010).
- Zeng, Z. *et al.* Identification and Application of a Novel Immune-Related lncRNA Signature on the Prognosis and Immunotherapy for Lung Adenocarcinoma. *Diagnostics* <https://doi.org/10.3390/diagnostics12112891> (2022).
- Qin, J. *et al.* Prognostic value of LECT2 and relevance to immune infiltration in hepatocellular carcinoma. *Front. Genet.* <https://doi.org/10.3389/fgene.2022.951077> (2022).
- Chang, W. *et al.* Identification of an amino acid metabolism-related gene signature for predicting prognosis in lung adenocarcinoma. *Genes (Basel)* <https://doi.org/10.3390/genes13122295> (2022).
- Gu, Y. *et al.* DMDRMR-Mediated Regulation of m(6)A-Modified CDK4 by m(6)A Reader IGF2BP3 Drives ccRCC Progression. *Cancer Res.* **81**, 923–934. <https://doi.org/10.1158/0008-5472.CAN-20-1619> (2021).
- Shi, Y., Zhao, Y. & Wang, Y. An Inflammatory Response-Related Gene Signature Can Predict the Prognosis and Impact the Immune Status of Lung Adenocarcinoma. *Cancers* <https://doi.org/10.3390/cancers14235744> (2022).
- Xie, T. *et al.* Analysis on methylation and expression of PSMB8 and its correlation with immunity and immunotherapy in lung adenocarcinoma. *Epigenomics* **14**, 1427–1448. <https://doi.org/10.2217/epi-2022-0282> (2022).
- Goswami, S., Anandhan, S., Raychaudhuri, D. & Sharma, P. Myeloid cell-targeted therapies for solid tumours. *Nat. Rev. Immunol.* **23**, 106–120. <https://doi.org/10.1038/s41577-022-00737-w> (2023).
- Barry, S. T., Gabrilovich, D. I., Sansom, O. J., Campbell, A. D. & Morton, J. P. Therapeutic targeting of tumour myeloid cells. *Nat. Rev. Cancer* **23**, 216–237. <https://doi.org/10.1038/s41568-022-00546-2> (2023).
- Wu, H., Qin, J., Zhao, Q., Lu, L. & Li, C. Microdissection of the Bulk Transcriptome at Single-Cell Resolution Reveals Clinical Significance and Myeloid Cells Heterogeneity in Lung Adenocarcinoma. *Front. Immunol.* <https://doi.org/10.3389/fimmu.2021.723908> (2021).
- De Cicco, P., Ercolano, G. & Ianaro, A. The New Era of Cancer Immunotherapy: Targeting Myeloid-Derived Suppressor Cells to Overcome Immune Evasion. *Front. Immunol.* **11**, 1680. <https://doi.org/10.3389/fimmu.2020.01680> (2020).
- Wu, Y., Yi, M., Niu, M., Mei, Q. & Wu, K. Myeloid-derived suppressor cells: An emerging target for anticancer immunotherapy. *Mol. Cancer* **21**, 184. <https://doi.org/10.1186/s12943-022-01657-y> (2022).
- Wang, M., Zhu, J., Zhao, F. & Xiao, J. Transcriptome Analyses Identify a Metabolic Gene Signature Indicative of Antitumor Immunosuppression of EGFR Wild Type Lung Cancers With Low PD-L1 Expression. *Front. Oncol.* <https://doi.org/10.3389/fonc.2021.643503> (2021).
- Yang, J. *et al.* Leveraging Tumor Microenvironment Infiltration in Pancreatic Cancer to Identify Gene Signatures Related to Prognosis and Immunotherapy Response. *Cancers* <https://doi.org/10.3390/cancers15051442> (2023).
- Mussbach, F. *et al.* Proteinase-activated receptor 2 (PAR2) in hepatic stellate cells - evidence for a role in hepatocellular carcinoma growth in vivo. *Mol. Cancer* **15**, 54. <https://doi.org/10.1186/s12943-016-0538-y> (2016).
- Li, Y. *et al.* PAR2 promotes tumor-associated angiogenesis in lung adenocarcinoma through activating EGFR pathway. *Tissue Cell* <https://doi.org/10.1016/j.tice.2022.101918> (2022).

38. Chen, X., Xu, Y., Jiang, L. & Tan, Q. miRNA-218-5p increases cell sensitivity by inhibiting PRKDC activity in radiation-resistant lung carcinoma cells. *Thoracic Cancer* **12**, 1549–1557. <https://doi.org/10.1111/1759-7714.13939> (2021).
39. Zhou, Z. *et al.* Identification of synthetic lethality of PRKDC in MYC-dependent human cancers by pooled shRNA screening. *BMC Cancer* **14**, 944. <https://doi.org/10.1186/1471-2407-14-944> (2014).
40. Tan, K. T. *et al.* Prkdc new biomarker and drug target for checkpoint blockade immunotherapy. *J. Immunother. Cancer* <https://doi.org/10.1136/jitc-2019-000485> (2020).
41. Chen, L. M. *et al.* RANKL increases migration of human lung cancer cells through intercellular adhesion molecule-1 up-regulation. *J. Cell. Biochem.* **112**, 933–941. <https://doi.org/10.1002/jcb.23009> (2011).
42. Casimiro, S., Vilhais, G., Gomes, I. & Costa, L. The Roadmap of RANKL/RANK Pathway in Cancer. *Cells* **10**, <https://doi.org/10.3390/cells10081978> (2021).
43. Zhang, X. *et al.* INHA acts as a novel and potential biomarker in lung adenocarcinoma and shapes the immune-suppressive tumor microenvironment. *Transl. Oncol.* <https://doi.org/10.1016/j.tranon.2023.101679> (2023).
44. Starkl, P., Marichal, T. & Galli, S. J. PLA2G3 promotes mast cell maturation and function. *Nat. Immunol.* **14**, 527–529. <https://doi.org/10.1038/ni.2612> (2013).
45. Martinez-Garcia, A. *et al.* PLA2G3, a gene involved in oxidative stress induced death, is associated with Alzheimer's disease. *J. Alzheimers Dis.* **22**, 1181–1187. <https://doi.org/10.3233/JAD-2010-101348> (2010).

### Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication. XT was responsible for the study design and paper writing. DW, YL performed the experiments and analyzed the data. JL was responsible for part of data analysis. DW, YL, LM were responsible for reviewed and revised the manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-68111-5>.

**Correspondence** and requests for materials should be addressed to X.T.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024