Significance of macrophage infiltration in the prognosis of lung adenocarcinoma patients evaluated by scRNA and bulkRNA analysis

Huaiyang Zhu1†, Chunning Zheng2†, Hongtao Liu3†, Fanhua Kong4†, Shuai Kong2, Feng Chen5 and Yuan Tian6,7*

1Department of Thoracic Surgery, Shandong Second Provincial General Hospital, Jinan, China, 2Department of Gastrointestinal Surgery, Shandong Provincial Hospital affiliated to Shandong First Medical University, Jinan, China, 3Department of Pathology, The First Affiliated Hospital of Shandong First Medical University and Shandong Provincial Qianfoshan Hospital, Shandong Medicine and Health Key Laboratory of Clinical Pathology, Shandong Lung Cancer Institute, Shandong Institute of Nephrology, Jinan, China, 4Department of Thoracic Surgery, The Affiliated Taian City Central Hospital of Qingdao University, Taian, China, 5Department of Thoracic Surgery, Shandong Cancer Hospital and Institute, Shandong First Medical University and Shandong Academy of Medical Sciences, Jinan, China, 6Department of Otolaryngology-Head and Neck Surgery, Shandong Provincial ENT Hospital, Shandong University, Jinan, China, 7Radiotherapy Department, Shandong Second Provincial General Hospital, Shandong University, Jinan, China

Purpose: To investigate the significance of macrophage infiltration to the prognosis of lung adenocarcinoma.

Methods: R language bioinformatics analysis technology, was used to obtain macrophage infiltration-related module genes through WGCNA (Weighted Gene Co-Expression Network Analysis). Marker genes of macrophage subtypes were identified using single-cell sequencing of lung adenocarcinoma tissue. Risk score models were constructed and validated using external data cohorts and clinical samples.

Results: Analysis of cohorts TCGA-LUAD, GSE11969, GSE31210, GSE50081, GSE72094 and GSE8894, revealed a negative correlation between macrophage infiltration and survival. Immunohistochemical analyses of clinical samples were consistent with these data. Based on cell-cluster-markers and TAMs-related-genes, TOP8 genes were obtained (C1QTNF6, CCNB1, FSCN1, HMMR, KPNA2, PRC1, RRM2, and TK1) with a significant association to prognosis. Risk score models including 9 factors (C1QTNF6, FSCN1, KPNA2, GL2, TYMS, BIRC3, RBBP7, KRT8, GPR65) for prognosis were constructed. The efficacy, stability and generalizability of the risk score models were validated using multiple data cohorts (GSE19188, GSE26939, GSE31210, GSE50081, GSE42127, and GSE72094).

Conclusions: Macrophage infiltration negatively correlates with prognosis in patients with lung adenocarcinoma. Based on cell-cluster-markers and
1 Introduction

Lung cancer remains the most common malignancy worldwide and a leading cause of cancer-related death, despite advances in screening and treatment (1, 2). Whether it was for the non-small cell lung cancer (NSCLC) or small cell lung cancer (SCLC) patients, immunotherapy was the most shining one among many treatment methods, which had changed the landscape of anti-tumor therapy and brought anti-tumor therapy into a new era (3–7). However, there were still many details in the screening of immunotherapy benefit populations and related predictors needed to be further elucidated (8–16). Specific macrophage phenotypes can act as indicators of lung cancer prognosis and the efficacy of immunotherapy (17–24). Sequencing technologies and R language based bioinformatics, formerly reported (25–27), can be used for studies in this area (28–30). Based on our previous studies (28–30), we performed bioinformatics analysis and clinical sample validation to identify specific macrophage signatures that can act as indicators of therapeutic efficacy.

2 Methods

2.1 Data analysis

2.1.1 TCGA data

mRNA expression profiles, clinical information, copy number alterations and mutations of GDC TCGA Lung Adenocarcinoma (LUAD) samples were downloaded from https://xenabrowser.net/datapages/. Tumor samples were screened according to sample name. RNA-seq data for 513 tumor samples and 59 paracancerous samples were obtained.

2.1.2 GEO data

Expression data and sample survival information for GSE11969, GSE19188, GSE26939, GSE31210, GSE42127, GSE50081, GSE72094 and GSE8894 were downloaded from the GEO database (https://www.ncbi.nlm.nih.gov/geo). Survival information of the samples were summarized as follows: (Supplementary Table 1_train_clin.tsv; Supplementary Table 1_GSE11969_clin.txt; Supplementary Table 1_GSE19188_clin.txt; Supplementary Table 1_GSE26939_clin.txt; Supplementary Table 1_GSE31210_clin.txt; Supplementary Table 1_GSE42127_clin.txt; Supplementary Table 1_GSE50081_clin.txt; Supplementary Table 1_GSE72094_clin.txt; Supplementary Table 1_GSE8894_clin.txt). Single-cell sequencing data from GSE131907 were downloaded from the GEO database (https://www.ncbi.nlm.nih.gov/geo). A total of 42,995 cells and 29,634 genes were obtained.

2.2 Immune infiltration analysis

Immune infiltration for each sample was calculated using IObR of the R package for the training set TCGA expression matrix and GEO data, respectively (method = ‘cibersort’).

2.3 Survival analysis

For survival assessments, R packages “survminer” and “survival” were analyzed and survival curves were constructed based on survival time and status. Differences in prognosis among the groups were assessed.

2.4 Screening of modules corresponding to macrophages using WGCNA

Hierarchical clustering analysis was performed on the TCGA expression matrix using the R package “hclust”, “method=average”. Phenotypic information was obtained
using the infiltration ratio of macrophages. A correlation between different modules and macrophages was obtained.

2.5 Clustering analysis of samples

The R package “ConsensusClusterPlus” was used to perform consensus clustering analysis. After clustering on the TCGA and GEO data, the optimal number of categories were determined according to the change of area under the CDF curve. The k value of the cluster category ranged from 2 to 6.

2.6 Analysis of single-cell data

Single-cell data were filtered using the R package "seurat" to remove cells with ≥ 20% mitochondrial expression. Data were analyzed using the "seurat" normalization pipeline. To identify tumor-associated macrophage (TAM) populations, marker genes from published studies were used to identify corresponding clusters. TAM populations were selected for standardization analysis using "Seurat".

2.7 Trajectory analysis of single-cell data

Trajectory analysis was performed on TAM subclasses using the R package "monocle" with default parameters. This resulted in differentiation trajectories and key genes determining these trajectories.

2.8 Gene set variation analysis

To investigate differences in the expression patterns of specific TAM isoforms in biological processes, GSVA enrichment analysis was performed using the R package "GSVA". GSVA is a nonparametric, unsupervised method primarily used to assess alterations in signaling pathways and biological processes in samples.

2.9 Construction of risk scoring model

Univariate cox regression analysis was performed on "cell-cluster-markers" and "TAMs-related-genes", and genes significantly associated with OS survival were screened at the p<0.05 level. According to the identified prognosis-related genes, the R package 'gimmet' was used to construct a prognosis model (or classifier model) with a 10-fold cross-validation fold using the cox method. Characteristic factors were then screened. Kaplan-Meier survival analysis and ROC curves were used to evaluate the predictive power of the prognostic model.

2.10 Clinical sample validation (sample collection and immunohistochemistry)

Lung Cancer samples were collected from the First Affiliated Hospital of Shandong First Medical University & Shandong Provincial Qianfoshan Hospital from June 2012 to February 2020. Written informed consent was provided by all participants. Tumor tissues were surgically resected, formalin fixed and paraffin embedded (FFPE) for histological evaluation. HE-stained and immunohistochemical (IHC) slides were examined by two independent and experienced pathologists according to the WHO criteria.

Samples were IHC stained with mouse anti-human CD68 monoclonal antibodies (MAB-0863, clone MX075) and mouse anti-human CD163 monoclonal antibodies (MAB-0869, clone MX081). CD68 was used as a general surface marker for macrophages, whilst CD163 was used as a marker for M2 macrophages (31). Double-labeled immunohistochemical staining was performed using alkaline phosphatase and horseradish peroxidase conjugated secondary antibodies. Substrates were fast red (AP-Red) and dianinobenzidine (DAB) (Roche Ltd) stained. Slides were processed using an automated Roche BenchMark XT staining system according to the manufacturer’s protocol.

3 Results

3.1 Proportion of immune infiltrating cells and the prognostic efficacy of macrophages

CIBERSORT was used to evaluate the levels of immune-infiltration from different lung adenocarcinoma datasets (TCGA-LUAD, GSE1969, GSE31210, GSE50081, GSE72094, and GSE8894). According to the median macrophage ratio, samples were divided into high- and low-levels of macrophage infiltration. Survival differences between high- and low-groups showed a significant correlation with macrophage infiltration (Figure 1; Supplementary Table 2_train_cibersort.txt; Supplementary Table 2_GSE11969_cibersort.txt; Supplementary Table 2_GSE31210_cibersort.txt; Supplementary Table 2_GSE50081_cibersort.txt; Supplementary Table 2_GSE72094_cibersort.txt; Supplementary Table 2_GSE8894_cibersort.txt).

3.2 Screening of modules corresponding to macrophages

To identify macrophage-related genes related to infiltration, WGCNA module analysis was performed on the training dataset (Supplementary Figure 1, Supplementary Table 3_gene_module.txt).
Genes corresponding to red modules were named “TAMs-related-genes” for subsequent analysis.

3.3 TAMs-related-gene-based clustering analysis, molecular typing and prognostic assessments

We analyzed the expression profiles of TAMs-related-genes in samples from different lung adenocarcinoma datasets (TCGA-LUAD, GSE13213, GSE31210, GSE72094, and GSE8894) to construct consistent clustering profiles. Based on cumulative distribution functions and incremental area maps, we selected stable clusters of TAMs-related-genes to obtain multiple subtypes (Supplementary Figures 2A–E, Supplementary Table 4_TCGA_consensusClass.csv; Supplementary Table 4_GSE13213_consensusClass.csv; Supplementary Table 4_GSE31210_consensusClass.csv; Supplementary Table 4_GSE72094_consensusClass.csv; Supplementary Table 4_GSE8894_consensusClass.csv). Survival analysis was performed on cluster subtypes from different datasets, revealing significant survival differences (Supplementary Figures 2F–I). Dimensionality reduction analysis was performed on each dataset, revealing significant differences in sample characteristics between different subtypes (Supplementary Figures 2K–O).

3.4 Preprocessing of single-cell data

To further investigate the role of macrophages in lung adenocarcinoma, published single-cell sequencing data of lung adenocarcinoma patients was analyzed (PMC7210975) (32). Gene distribution and mitochondrial gene expression were screened (Supplementary Figures 3A–C). Cells with mitochondrial expression ≥20% were identified as dead and removed.

3.5 Identification of TAMs in total cells

Markers were used to detect the presence of TAMs in the lung adenocarcinoma single-cell datasets (Supplementary Figures 3D–I). TAMs were then extracted and subtype analysis performed to obtain a TAMs subtype map (Supplementary Figure 3I).

3.6 Screening of differential expression genes among TAMs subsets

To identify marker genes amongst the different TAM subgroups, samples were screened in “Seurat”. Dot and violin plots revealed the top 5 marker genes for each TAM subtype (Supplementary Figure 4; Supplementary Table 5_TAM_marker_genes.txt).
3.7 Simulation of dynamic changes in macrophages

"Monocle" was used to identify dynamic changes of macrophages in the tumors and cell polarization (Supplementary Figures 5A–C). Cluster 0 could be divided into Cluster 1 and Cluster 2 amongst TAM subtypes. The identified genes were found to regulate differentiation (Supplementary Figure 5D). Gene enrichment analysis on the subtypes of TAM showed that Cluster 2 positively correlated with E2F TARGETS and G2M CHECKPOINT, whilst Cluster 4 negatively correlated with these pathways (Supplementary Figure 5E).

3.8 Screening of prognostic factors based on cell-cluster-markers and TAMS-related-genes using univariate cox regression analysis

Markers of each TAM subtype and TAMS-related-genes were used to identify genes related to the prognosis. Samples were divided into high- and low-expression groups according to the median of gene expression. Univariate Cox analysis was performed and survival curves of the top8 prognostic genes were displayed (Figures 2A–H; Supplementary Table 6_cox_significant.txt).

3.9 Construction of risk score models and evaluation of the prognostic efficacy

Based on the “GLMNET” of the R package, LASSO (Least Absolute Shrinkage and Selection Operator) regression analysis was used to construct a regression model for the expression matrix of prognosis related genes corresponding to “Cell-Cluster-Markers” and “Tams-Related-Genes”. By analysis, when the value of the freedom degree was 9, the model was accurate (Figures 2I–O; Supplementary Table 7_forest.univariate_cox.txt, Supplementary Table 7_Signature_Coeff.txt). The calculation formula of the risk score model are listed as follows:

Risk Score = 0.0354754835*C1QTNF6 (Expression Value) + 0.0023344103* FSCN1 (Expression Value) + 0.002229189*GLI2 (Expression Value) + 0.0001616254 * KPNA2 (Expression value) + 0.0005176419*TYMS (Expression Value) + 0.0037948174 *BIRC3 (Expression Value) + 0.003257017*RBBP7 (Expression Value) + 0.0002465129 *KRT8 (Expression Value) - 0.0263442444 *GPR65 (Expression Value). Kaplan-Meier survival curves indicated a significant difference in survival between high and low risk groups. The ROC curve indicated high performance of the risk score model.

3.10 Validation of risk score prognostic models in external datasets

To further verify the stability of the risk score model, external and independent data GSE19188, GSE26939, GSE31210, GSE50081, GSE42127 and GSE72094 were used to verify predictive efficacy. Through Kaplan-Meier survival analysis, the constructed risk score model performed well for all external data predictions (Figure 3).

3.11 Robust principal component analysis of risk scoring models in clinical factors

To confirm the stability of the risk score model according to clinical characteristics, differences in survival status between high- and low-risk groups in terms of age, gender, radiotherapy, clinical characteristics and Pathlogic M were explored. Significant differences in survival between high- and low-risk groups were observed in those aged ≥ 60 and ≤ 60 years (Figures 4A, B; Supplementary Table 8_clinical_inf.txt). Similar differences were observed between gender subgroups (Figures 4C, D). In the radiotherapy group, differences between high- and low-risk groups were more pronounced (Figures 4E, F). In Pathologic M (Figures 4G, I), significant differences between high- and low-risk groups were observed for M0, indicative of higher stability.

3.12 Differences in risk score models among cancer clinical factors

To investigate the relationship between the risk score model and clinical characteristics, specific features were selected for analysis. The risk score was found to be related to radiation therapy, pathologic T and tumor stage. No significant relationship to age or gender were observed (Supplementary Figure 6).

3.13 Evaluation of risk score models through univariate and multivariate cox regression analysis

To determine whether the risk score model could act as an independent prognostic factor for cancer, the "coxph()" function in the R package “survival” was adopted for univariate and multivariate regression analysis on training and test sets, respectively. We found that in all validation and test sets, the p value of the risk score was ≤ 0.05 (Figure 5; Supplementary Table 9_TCGA_clinical.multivariate_cox.txt; Supplementary Table 9_TCGA_clinical.univariate_cox.txt; Supplementary
Survival analysis of top 8 genes is significantly associated with prognosis. Abscissa axis: survival time. Ordinate axis: survival probability. Colors: differential gene expression. Survival analysis of (A) C1QTNF6, (B) CCNB1, (C), FSCN1, (D) HMMR, (E) KPNA2, (F) PRC1, (H) RRM2, and (I) TK1.

Construction of the risk score model and evaluation of its prognostic efficacy. Forest plots of genes included in the risk score model. Right column: 9 genes included in the risk score model. Left column: corresponding forest plot. (K) Risk score plot for cancer samples (line graph).

Dynamic process diagram of variables screened by LASSO regression analysis and selection process diagram of the cross-validation parameter λ. (O) ROC curve of training datasets. Abscissa axis: survival time; ordinate axis: sensitivity. Colors represent different years.

**FIGURE 2**

Survival analysis of top 8 genes is significantly associated with prognosis. Abscissa axis: survival time. Ordinate axis: survival probability. Colors: differential gene expression. Survival analysis of (A) C1QTNF6, (B) CCNB1, (C), FSCN1, (D) HMMR, (E) KPNA2, (F) PRC1, (H) RRM2, and (I) TK1. (J) Construction of the risk score model and evaluation of its prognostic efficacy. Forest plots of genes included in the risk score model. Right column: 9 genes included in the risk score model. Left column: corresponding forest plot. (K) Risk score plot for cancer samples (line graph). (L) Risk score plot for cancer samples (scatter plot graph). (M) Dynamic process diagram of variables screened by LASSO regression analysis and selection process diagram of the cross-validation parameter λ. (N) Survival analysis of the training dataset. Abscissa axis: survival time; ordinate axis: survival probability. (O) ROC curve of training datasets. Abscissa axis: specificity, Ordinate axis: sensitivity. Colors represent different years.
Validation of the risk score model with external independent data. A1: Validation (ROC curve) of the risk score model using external independent data GSE19188; A2: survival analysis. B1: Validation (ROC curve) of the risk score model using GSE26939. B2: Survival analysis using GSE26939. C1: Validation (ROC curve) of the risk score model using GSE31210. C2: Validation (survival analysis) using GSE31210. D1: Validation results (ROC curve) of the risk score model using GSE50081. D2: (survival analysis) of the risk score model using GSE50081. E1: Validation results (ROC curve) of the risk score model using GSE42127. E2: survival analysis using GSE42127. F1: Validation (ROC curve) of the risk score model using GSE72094. F2: Survival analysis using GSE72094. Abscissa axis: survival time; Ordinate axis: survival probability. Colors: different risk groups.
This indicated that the risk score model was an accurate independent prognostic factor for cancer.

### 3.14 Construction of a nomogram model of risk scores and clinical factors to predict cancer progression

We next sought to apply the risk scoring model to the prediction of cancer progression in the clinic. The R package “rms” was adopted to construct a nomogram using a variety of clinical features. Calibration curves were used to calculate 1, 2, 3, and 5-year survival times (Supplementary Figure 6, Supplementary Table 10). All survival calibration curves were near the 45° slope, indicating high accuracy of the nomogram.
Univariate and multivariate Cox regression analysis of the risk score model in training and validation datasets. (A) Univariate Cox regression analysis. (B) Multivariate Cox regression analysis. (C) GSE19188 univariate Cox regression analysis. (D) GSE19188 multivariate Cox regression analysis. (E) GSE26939 univariate Cox regression analysis. (F) GSE26939 multivariate Cox regression analysis. (G) GSE42127 univariate Cox regression analysis. (H) GSE42127 multivariate Cox regression analysis. (I) GSE50081 univariate Cox regression analysis. (J) GSE50081 multivariate Cox regression analysis. (K) GSE72094 univariate Cox regression analysis. (L) GSE72094 multivariate Cox regression analysis.

FIGURE 5
3.15 Prediction of immunotherapy efficacy amongst subtypes

We next investigated whether the risk score model could predict the prognosis of immunotherapy. Data were calculated using the risk score model and the K-M survival status between high- and low-risk groups evaluated (Supplementary Figure 7A, Supplementary Table 11 _Immune_treatment.xlsx). Upon statistical analysis of the distribution of CR/PR and PD/SD, the proportion of treatment response rates significantly differed between high- and low-risk groups (Supplementary Figure 7B, chi-square test p= 0.004133). No significant differences in the risk scores between the different treatment response groups were observed (Supplementary Figure 7C).

3.16 Overall survival analyses of M1 and M2 macrophage subtypes in patients with lung cancer

A total of 32 patients with lung cancer were evaluated for M1 and M2 macrophage subtypes. Samples were stained using double-labeled IHC. The majority of patients were in pathologic Stage II (62.5%) and the dominant histopathological type was adenocarcinoma (68.8%). The clinicopathological characteristics of the lung cancer patients are shown in (Table 1).

To identify M1 and M2 macrophage subtypes, CD68 and CD163 antibodies were used for double-labeled IHC staining. CD163 (red) localized to the plasma membrane. M2 macrophages were identified through double staining for CD68 and CD163. M1 macrophages were identified through staining with CD68 alone. Representative IHC images are shown in (Figures 6A, B).

The prognostic value of macrophage infiltration was next evaluated. Total macrophages, M2 to M1, and M2 macrophage infiltration were identified as detrimental to patient survival (Figures 6C, E, F), whilst M1 macrophage infiltration was beneficial to prognosis (Figure 6D). The infiltration of M1 macrophages in adenocarcinoma was significantly higher than that in squamous cell carcinoma of lung cancer. No significant differences in M2 nor total macrophage infiltration were observed between these two histological subtypes (Figure 6G).

4 Discussion

Macrophages with different phenotypes are frequently cited as indicators of the prognosis of lung cancer patients and the efficacy of immunotherapy (17–24). In our preliminary analyses, macrophage infiltration, rarely reported in lung cancer, had a significant detrimental effect on the prognosis of lung cancer patients (Figure 1). These data were consistent across cohorts (Figures 1A–F) and further verified in follow-up immunohistochemical analysis of clinical samples (Figure 6C). Collectively, these data highlight how macrophages not only act as innate immune cells to regulate immunological responses (23), but play an important role in the prognosis of lung cancer. This lays the foundation for subsequent module analysis based on macrophage infiltration (Supplementary Figure 1).

Based on the expression profiles of TAMs-related-genes, a consistent clustering profile was constructed (Supplementary Figures 2A–E). Significant differences in both survival analysis and PCA (Supplementary Figures 2F–O) were observed. These apparent differences were further identified in single-cell data (Supplementary Figure 3) confirming the importance of macrophages to the prognosis of lung cancer patients (17–24, 32). These data also highlight the need for further refinement of relevant factors to more favorably evaluate patient prognosis.

Given the advantages and progress of single-cell sequencing in lung cancer immunity (33–35), the single-cell data was further analyzed (Supplementary Figures 4, 5, and Figure 2) (32). Based on cell-cluster-markers and TAMs-related-genes, TOP 8 genes (CIQTNF6, CCNB1, FSCN1, HMMR, KPN2, PRC1, RRM2, and TK1) significantly associated with prognosis were obtained (Figure 2). These have obvious benefits to clinicians for the assessment of patient prognosis (36–49). The same data were used to construct a risk score model containing 9 factors (CIQTNF6, FSCN1, KPN2, GLI2, TYMS, BIRC3, RBBP7, KRT8, and GPR65) for prognostic evaluation (Figure 2) (50–55). The model was validated using external data cohorts (Figure 3) and identified as robust and accurate for prognostic evaluation (Figure 4). Significant differences in the risk scores
Validation analysis of clinical samples. A1: Immunohistochemical staining of M1 macrophages. A2: Enlargement of the boxed regions. B1: Immunohistochemical staining of M2 macrophages. B2: Enlargement of the boxed regions. Images were obtained at 40×10 magnification under a light microscope. (C) Kaplan-Meier survival analysis of macrophage infiltration. (D) M1 macrophage infiltration. (E) M2 macrophage infiltration. (F) M2 to M1 macrophage infiltration. (G) M1, M2 and total macrophage infiltration between adenocarcinoma and squamous cell carcinoma. Horizontal axis: survival time. Vertical axis: survival probability. Colors: macrophage infiltration.
were observed for clinical characteristics including radiation therapy, pathologic T, and Tumor stage (Supplementary Figures 6A–E). This further highlighted the efficiency of the risk score to predict therapeutic efficacy.

Through univariate and multivariate cox regression analysis, the risk score model held utility as an independent prognostic factor for cancer, further affirming its clinical benefits (Figure 5). Furthermore, cancer progression could be more accurately predicted using nomogram models constructed based on risk scores and clinical factors (Supplementary Figures 6F–J). For prognostic assessments of immunotherapy, the risk score model could also act as an accurate evaluation tool (Supplementary Figure 7). Upon immunohistochemical analysis of clinical tissue samples to verify the correlation between the macrophage phenotype and patient prognosis, similar conclusions were obtained (Table 1; Figures 1, 6). Macrophage infiltration, particularly for the M2 phenotype, were not conducive to the prognosis and survival of patients, consistent with previous studies (20–24, 56, 57).

We used WGCNA to identify macrophage infiltration-related module genes and single-cell sequencing of lung adenocarcinoma tissue to identify marker genes of macrophage subtypes. This permitted the construction of a risk assessment model with high prognostic efficacy. The model performed well on external and independent datasets. Immunohistochemistry analysis of clinical samples were consistent with our data. We therefore infer that the risk score has both high clinical practicability and application.

5 Conclusion

Macrophage infiltration was negatively correlated with prognosis for patients with lung adenocarcinoma. Based on cell-cluster-markers and TAMs-related-genes, both TOP8 genes (C1QTNF6, CCNB1, FSCN1, HMMR, KPNA2, PRC1, RM2, TK1) and the risk score model containing 9 risk factors (C1QTNF6, FSCN1, KPNA2, GLI2, TYMS, BIRC3, RBBP7, KRT8, GPR65) had a high efficacy for the prediction of prognosis.

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.

Ethics statement

The ethics involved in the clinical samples in this study have been approved by the Ethics Committee of Qianfoshan Hospital of Shandong Province (2022-S527). The patients/participants provided their written informed consent to participate in this study.

Author contributions

Corresponding authors (YT) conceived the idea and designed this study. TY was responsible for the final submission. YT, CZ and HZ were responsible for data collection, partial data analysis and drafting the manuscript. HL performed immunohistochemical staining of the clinical tissue and related data analysis. FK, SK, and FC were reviewed the manuscript and were responsible for data corrections. All authors contributed to the article and approved the submitted version.

Funding

The study was funded by Clinical Research Fund of Shandong Medical Association Qilu Special Project (YXH2022ZX02016; YT), and Jinan Clinical Medicine Technology Innovation Program (YT).

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.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2022.1028440/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Screening of modules corresponding to macrophages using WGCNA (Weighted Gene Co-Expression Network Analysis). (A) The cluster dendrogram of modular genes associated with macrophage infiltration.
**SUPPLEMENTARY FIGURE 2**

Sample clustering, molecular typing, and prognosis evaluation analyses based on TAMs-related-genes. (A) Subtypes of clustering profiles obtained through analysis of the expression profiles of TAMs-related-genes in TCGA-LUAD samples (consensus matrix k=3). (B) Subtypes of consistent clustering profiles obtained through analysis of the expression profiles of TAMs-related-genes in GSE5213 samples (consensus matrix k=3). (C) Subtypes of the consistent clustering profiles obtained from GSE31210 samples (consensus matrix k=2). (D) Clustering profiles from GSE72094 samples (consensus matrix k=2). (E) GSE8894 samples (consensus matrix k=5). (F) Survival analysis of different cluster subtypes in TCGA-LUAD samples. (G) Results of survival analysis of different cluster subtypes in GSE8894 samples; The horizontal axis represents survival time; the vertical axis represents survival probability; Curves with different colors represent different cluster subtypes. (H) Results of survival analysis of different cluster subtypes in GSE13213 samples; The horizontal axis represents survival time; the vertical axis represents survival probability; Curves with different colors represent different cluster subtypes. (I) Results of survival analysis of different cluster subtypes in GSE72094 samples. The horizontal axis represents survival time; the vertical axis represents survival probability; Curves with different colors represent different cluster subtypes. (J) Results of survival analysis of different cluster subtypes in GSE8894 samples. The horizontal axis represents survival time; the vertical axis represents survival probability; Curves with different colors represent different cluster subtypes.

**SUPPLEMENTARY FIGURE 3**

Single-cell data. (A) Number of genes expressed in cells. (B) Total counts. (C) Mitochondrial gene expression. (D–I) UMAP (Uniform Manifold Approximation and Projection) dimensionally reduction analysis results of TAM subgroups. (D) CD20. (E) CD16. (F) APOE. (G) CD133. (H) Analysis of TAM clusters. (I) Subtypes derived from re-clustering of TAM cell populations.

**SUPPLEMENTARY FIGURE 4**

Screening of differentially expressed genes amongst tumor macrophage subsets. (A) Dotplot of Top5 maker genes of each subtype. Abscissa axis: marker genes. Ordinate axis: top5 of TAM subtypes. Colors: mean expression per-group; Dot sizes represent the fraction of cells in each group. (B–G) Violin plots of the expression of the top5 marker genes. Abscissa axis: different TAM subtypes. Ordinate axis: gene expression.

**SUPPLEMENTARY FIGURE 5**

Pseudo-chronological analysis of tumor macrophages for simulation of the dynamic changes of macrophages. (A–C) Differential states according to monocle trajectory analysis, distribution of TAMs in trajectories, and pseudo-sequences of differentiation. (D) Genes influencing differentiation states in the clusters. Left column: different clusters. Right column: names of genes. (E) Pathway enrichment analysis of different TAM subtypes. Color: correlation; Red: positive correlation; Blue: negative correlation. Numerical values: correlation p-value.

**SUPPLEMENTARY FIGURE 6**

Comparison of risk scores corresponding to the clinical characteristics of the different groups. (A) Age, (B) Gender, (C) Radiation therapy. (E) Pathologic T cells. (F) Tumor stage. Abscissa axis: different groups. Ordinate axis: risk scores. (F) Nomogram model for risk scores and clinical factors according to the clinical characteristics of prognosis. (G) Calibration curve for 1-year survival. (H) 2-year survival. (I) 3-year survival. (J) 5-year survival. Abscissa axis: predicted probability of survival. Ordinate axis: actual survival.

**SUPPLEMENTARY FIGURE 7**

Assessment of immunotherapy prognosis according to the risk score. (A) Survival analysis of immunotherapy responses in the training set. Abscissa axis: survival time. Ordinate axis: survival probability. Colors represent different risk groups. (B) Comparative analysis of the proportion of treatment response states between high and low risk groups. (C) Comparative analysis of risk scores for different treatment response states.

**References**

1. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. CA Cancer J Clin (2022) 72(1):7–33. doi: 10.3322/caac.21708

2. Thai AA, Solomon BJ, Sequist LV, Gainor JF, Heist RS. Lung cancer. Lancet (2021) 398(10299):535–54. doi: 10.1016/S0140-6736(21)00312-3

3. Paz-Ares L, Dvorkin M, Chen Y, Reinmuth N, Hotta K, Trudkin D, et al. Durvalumab plus platinum-etoposide versus platinum-etoposide in first-line treatment of extensive-stage small-cell lung cancer (CASPIAN); a randomised, controlled, open-label, phase 3 trial. Lancet (2019) 394(10212):1929–39. doi: 10.1016/S0140-6736(19)32222-6

4. Kennedy LB, Salama AKS. A Rev Cancer immunotherapy toxicity. CA Cancer J Clin (2020) 70(2):86–104. doi: 10.3332/caac.21596

5. Beck M, Remon J, Hellmann MD. First-line immunotherapy for non-Small-Cell lung cancer. J Clin Oncol (2022) 40(6):586–97. doi: 10.1200/JCO.21.01497

6. Liu SV, Giacone G. Lung cancer: First-line immunotherapy in lung cancer - taking the first step. Nat Rev Clin Oncol (2016) 13(10):595–6. doi: 10.1038/nrclinonc.2016.148

7. Riuvi NA, Hellmann MD, Brahmer JR, Juergens RA, Borghia H, Gettinner S, et al. Nivolumab in combination with platinum-based doublet chemotherapy for first-line treatment of advanced non-Small-Cell lung cancer. J Clin Oncol (2016) 34(25):2969–78. doi: 10.1200/JCO.2016.66.9861

8. The Lancet Respiratory Medicine. Lung cancer immunotherapy biomarkers: refine not reject. Lancet Respir Med (2018) 6(6):403. doi: 10.1016/S2213-2600(18)30180-2

9. Rashdan S, Minna JD, Gerber DE. Diagnosis and management of pulmonary toxicity associated with cancer immunotherapy. Lancet Respir Med (2018) 6(6):472–8. doi: 10.1016/S2213-2600(18)30172-3

10. Duruisseaux M, Martinez-Cardas A, Calleja-Cervantes ME, Moran S, Castro de Moura M, Davalos V, et al. Epigenetic prediction of response to anti-PD-1 treatment in non-small-cell lung cancer: a multicentre, retrospective analysis. Lancet Respir Med (2018) 6(10):771–81. doi: 10.1016/S2213-2600(18)30284-4

11. Petersen I. Predictive pathology of lung cancer immunotherapy. Lancet Respir Med (2018) 6(10):731–3. doi: 10.1016/S2213-2600(18)30333-3

12. Camidge DR, Schenk EL. Blood-based biomarkers for predicting immunotherapy benefit in lung cancer. Cell (2020) 183(2):203–4. doi: 10.1016/j.cell.2020.09.052

13. Hendriks L, Besse B. New windows open for immunotherapy in lung cancer. Nature (2018) 558(7710):376–7. doi: 10.1038/s41586-018-05312-9

14. Hellmann MD, Ciuleanu TE, Pluzanski A, Lee JS, Otterson GA, Audigier-Valette C, et al. Nivolumab plus ipilimumab in lung cancer with a high tumor
mutational burden. *N Engl J Med* (2018) 378(22):2093–104. doi: 10.1056/NEJMoa1801180

11. Alessi JV, Awad MM. Immunotherapy in lung cancer: effective for patients with poor performance status? *Lancet Respir Med* (2020) 8(9):838–9. doi: 10.1016/s2213-2600(20)30107-7

12. Middelton G, Brock K, Savage J, Mant R, Summers Y, Connheimer J, et al. Pembrolizumab in patients with non-small-cell lung cancer of performance status 2 (P-PS2): a single arm, phase 2 trial. *Lancet Respir Med* (2020) 8(9):895–904. doi: 10.1016/s2213-2600(20)30033-3

13. Jeremy JW, Lin YC, Yao PL, Yuan A, Chen HY, Shun CT, et al. Tumor-associated macrophages: the double-edged sword in lung cancer. *J Clin Oncol* (2005) 23(5):953–64. doi: 10.1200/jco.2005.12.172

14. Aalipour A, Chauan HY, Murty S, D’Souza AL, Park SM, Galati GS, et al. Engineered immune cells as highly sensitive cancer diagnostics. *Nat Biotechnol* (2019) 37(5):531–9. doi: 10.1038/s41581-019-0064-8

15. Gordon SR, Maute RL, Dulken BW, Hutter G, George BM, McCracken MN, et al. PD-1 expression by tumour-associated macrophages inhibits phagocytosis and tumour immunity. *Nature* (2017) 545(7665):495–9. doi: 10.1038/nature22366

16. Ohi R, Shikuta A, Waller D, Bradding P. Macrophage phenotypes and their expression in non-small cell lung cancer. *J Clin Oncol* (2008) 26(15). doi: 10.1200/jco.2008.26.15_suppl.22132

17. Hsiao YL, Su KY, Chen HY, Yu SL. Opposite effects of M1 and M2 macrophage subtypes on lung cancer progression. *J Clin Oncol* (2015) 24(5):14273. doi: 10.1200/jco.2015.33.15_suppl.e19148

18. Lu C, Shiu AL, Su BH, Hsu TS, Wang CT, Su YC, et al. Oct4 promotes M2 macrophage polarization through upregulation of macrophage colony-stimulating factor in lung cancer. *J Hematol Oncol* (2020) 13(1):6. doi: 10.1186/s12004-020-00887-1

19. Conway EM, Pikor LA, Kung SH, Hamilton MJ, Lam S, Lam WL, et al. Macrophages, inflammation, and lung cancer. *Ann J Respir Crit Care Med* (2019) 199(2):116–30. doi: 10.1164/rccm.201805-0545sc

20. Chen J, Sun W, Zhang H, Ma J, Xu P, Yu Y, et al. Macrophages reprogrammed by lung cancer microtropics promote tumor development via release of IL-1β. *Cell Mol Immunol* (2017) 10(12):2333–44. doi: 10.1038/cmi.2017.161

21. Zhang Y, Zhang L, Li R, Chang DW, Ye Y, Minna JD, et al. Genetic variations in cancer-related significance mutated genes and lung cancer susceptibility. *Ann Oncol* (2017) 28(7):1625–30. doi: 10.1093/annonc/mdx161

22. Xu Y, Shen Y, Wu D, Zhang J, Lin C, Wang L, et al. CircularCGRAC3 accelerates esophageal cancer tumorigenesis and metastasis via sponging miR-27a-3p. *Mol Cancer* (2022) 21(1):145. doi: 10.1186/s12943-022-01615-8

23. Yang B, Zhang L, Cao Y, Chen S, Cao J, Wu D, et al. Overexpression of IncRNA IGFBP4-1 reprograms energy metabolism to promote lung cancer adenocarcinoma through GLI2 suppression and ROS production. *Cancer Lett* (2018) 452:14273. doi: 10.1016/j.canlet.2020.02.007

24. Tian Y, Liu H, Zhang C, Ma W, Huang A, Tian M, et al. A novel classification method for NSCLC based on the background interaction network and the edge-perturbation matrix. *Aging (Albany NY).* (2020) 12(7):1235–7. doi: 10.1164/ageing.2020.04044

25. Tian Y, Liu H, Zhang C, Liu W, Wu T, Yang X, et al. Comprehensive analyses of ferroptosis-related alterations and their prognostic significance in glioblastoma. *Front Mol Biosci* (2022) 9:940098. doi: 10.3389/fmolb.2022.940098

26. Tano Y, Ito H, Shiga T, Higashida N, Nakamura F, Nakamura H, et al. Accumulation of KPNA2 impacts radioresistance through positive regulation of the NF-kB signaling pathway in lung adenocarcinoma. *Cell Biosci* (2021) 11(1):74. doi: 10.3389/fcimb.2021.108872. doi: 10.3389/cas.2021.53176

27. Chen J, Sun W, Zhang H, Ma J, Xu P, Yu Y, et al. Macrophages reprogrammed by lung cancer microtropics promote tumor development via release of IL-1β. *Cell Mol Immunol* (2017) 10(12):2333–44. doi: 10.1038/cmi.2017.161

28. Zhang Y, Zhang L, Li R, Chang DW, Ye Y, Minna JD, et al. Genetic variations in cancer-related significance mutated genes and lung cancer susceptibility. *Ann Oncol* (2017) 28(7):1625–30. doi: 10.1093/annonc/mdx161

29. Xu Y, Shen Y, Wu D, Zhang J, Lin C, Wang L, et al. CircularCGRAC3 accelerates esophageal cancer tumorigenesis and metastasis via sponging miR-27a-3p. *Mol Cancer* (2022) 21(1):145. doi: 10.1186/s12943-022-01615-8

30. Tano Y, Ito H, Shiga T, Higashida N, Nakamura F, Nakamura H, et al. Accumulation of KPNA2 impacts radioresistance through positive regulation of the NF-kB signaling pathway in lung adenocarcinoma. *Cell Biosci* (2021) 11(1):74. doi: 10.3389/fcimb.2021.108872. doi: 10.3389/cas.2021.53176

31. Seidl C, Panzitt K, Bertsch A, Brcic L, Schein S, Mack M, et al. MicroRNA-let-7i/sponging miR-27a-3p. *Mol Cancer* (2022) 21(1):145. doi: 10.1186/s12943-022-01615-8

32. Tano Y, Ito H, Shiga T, Higashida N, Nakamura F, Nakamura H, et al. Accumulation of KPNA2 impacts radioresistance through positive regulation of the NF-kB signaling pathway in lung adenocarcinoma. *Cell Biosci* (2021) 11(1):74. doi: 10.3389/fcimb.2021.108872. doi: 10.3389/cas.2021.53176

33. Frazzi R. BIRC3 and BIRC5: multi-faceted inhibitors in cancer. *Drug Discov Today* (2022) 17:273868. doi: 10.1016/j.drudis.2022.875146

34. Tano Y, Ito H, Shiga T, Higashida N, Nakamura F, Nakamura H, et al. Accumulation of KPNA2 impacts radioresistance through positive regulation of the NF-kB signaling pathway in lung adenocarcinoma. *Cell Biosci* (2021) 11(1):74. doi: 10.3389/fcimb.2021.53176

35. Frazzi R. BIRC3 and BIRC5: multi-faceted inhibitors in cancer. *Drug Discov Today* (2022) 17:273868. doi: 10.1016/j.drudis.2022.875146