Original Article

Integrated analysis reveals the microenvironment of non-small cell lung cancer and a macrophage-related prognostic model

Shenglong Xie[1,2#^], Guixiang Huang[3#^], Weiwei Qian[4^], Xuyang Wang[1^], Hanlu Zhang[1^], Zhiyang Li[1], Yu Liu[5^], Yun Wang[1^], Hongtao Yu[3^]

[1]Department of Thoracic Surgery, West China Hospital, Sichuan University, Chengdu, China; [2]Department of Thoracic Surgery, Sichuan Provincial People's Hospital, University of Electronic Science and Technology of China, Chengdu, China; [3]Department of Emergency Surgery, Sichuan Provincial People's Hospital, University of Electronic Science and Technology of China, Chengdu, China; [4]Department of Emergency, Shangjinnanfu Hospital, West China Hospital, Sichuan University, Chengdu, China; [5]Business School of Chengdu University, Chengdu, China

Contributions: (I) Conception and design: S Xie, H Yu; (II) Administrative support: W Qian; (III) Provision of study materials or patients: None; (IV) Collection and assembly of data: Y Wang; (V) Data analysis and interpretation: H Zhang, Z Li; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Hongtao Yu. Department of Emergency Surgery, Sichuan Provincial People's Hospital, University of Electronic Science and Technology of China, 32 West Section 2, 1st Ring Road, Chengdu 610072, China. Email: yhtflash@163.com; Yun Wang. Department of Thoracic Surgery, West China Hospital, Sichuan University, 37 Guoxue Alley, Chengdu 610041, China. Email: yunwwang@yeah.net.

Background: In the treatment of non-small cell lung cancer (NSCLC), recent advances in immunotherapy have heralded a new era. Despite the success of immune therapy, a subset of patients persistently fails to respond. Therefore, to better improve the efficacy of immunotherapy and achieve the purpose of precision therapy, the research and exploration of tumor immunotherapy biomarkers have received much attention.

Methods: Single-cell transcriptomic profiling was used to reveal tumor heterogeneity and the microenvironment in NSCLC. The Cell-type Identification by Estimating Relative Subsets of RNA Transcripts (CIBERSORT) algorithm was utilized to speculate the relative fractions of 22 infiltration immunocyte types in NSCLC. Univariate Cox and least absolute shrinkage and selection operator (LASSO) regression analyses were used for the construction of risk prognostic models and predictive nomograms of NSCLC. Spearman's correlation analysis was employed to explore the relationship between risk score and tumor mutation burden (TMB) and immune checkpoint inhibitors (ICIs). Screening of chemotherapeutic agents in the high- and low-risk groups was performed with the "pRRophetic" package in R. Intercellular communication analysis was conducted using the "CellChat" package.

Results: We found that most tumor-infiltrating immune cells were T cells and monocytes. We also found that there was a significant difference in the tumor-infiltrating immune cells and ICIs across different molecular subtypes. Further analysis showed that M0 and M1 mononuclear macrophages were significantly different in different molecular subtypes. The risk prediction model was shown to have to ability to accurately predict the prognosis, immune cell infiltration, and chemotherapy efficacy of patients in the high and low-risk groups. Finally, we found that the carcinogenic effect of migration inhibitory factor (MIF) is mediated by binding to CD74, CXCR4, and CD44 receptors involved in MIF cell signaling.

Conclusions: We have revealed the tumor microenvironment (TME) of NSCLC through single-cell data analysis and constructed a prognosis model of macrophage-related genes. These results could provide new therapeutic targets for NSCLC.

^ ORCID: Shenglong Xie, 0000-0003-4662-8832; Guixiang Huang, 0000-0003-4285-8999; Weiwei Qian, 0000-0002-2782-5550; Xuyang Wang, 0000-0002-5437-3430; Hanlu Zhang, 0000-0003-2964-0848; Yu Liu, 0000-0001-8996-7354; Yun Wang, 0000-0002-9685-7337; Hongtao Yu, 0000-0002-0630-3729.

Keywords: Non-small cell lung cancer (NSCLC); immune checkpoint inhibitors (ICIs); macrophages; migration inhibitory factor (MIF)

Submitted Nov 17, 2022. Accepted for publication Jan 20, 2023. Published online Feb 13, 2023.
doi: 10.21037/tlcr-22-866
View this article at: https://dx.doi.org/10.21037/tlcr-22-866

Introduction

Lung cancer is the third most common malignant tumor in humans and the leading cause of cancer-related mortality worldwide (1,2). Patients with non-small cell lung cancer (NSCLC) are more likely to die from the disease than those with small-cell lung carcinoma (3-5). NSCLC patients who are diagnosed during advanced stages typically have mild symptoms that are difficult to detect.

Tumor immunotherapy has gained increasing attention in recent years thanks to some remarkable discoveries (6). Currently, available immunotherapy modes include single-drug therapy, combined chemotherapy, chemotherapy plus immunotherapy, consolidation therapy after chemoradiotherapy, and combined immunotherapy (7). Although targeted therapies and immunotherapies have made significant progress in treating advanced NSCLC, the 5-year survival rate remains very low. For this reason, much work needs to be done to clarify the molecular mechanisms underlying NSCLC and to identify more specific biomarkers.

The availability of immune checkpoint inhibitors (ICIs) has improved the outlook for patients with advanced NSCLC. However, due to the heterogeneity of NSCLC, only 40% of those affected benefit from ICIs (8-13), which results in poor efficacy. Regardless of whether efficacy, prognosis, or toxic side effects or being assessed, the common challenge is the lack of perfect biomarkers. According to the current trend, integrating multiple biomarkers and establishing an efficacy prediction model is the future development direction.

Tumor microenvironment (TME) is composed of immune cells, fibroblasts and related cytokines, etc., which provides necessary conditions for the occurrence and development of tumors, in which macrophages play an important role (14). Tumor-associated macrophages (TAMs) are defined as the macrophages involved in the formation of TME, which affects the whole process of lung cancer development (15-17). TAMs are highly malleable, showing that they can polarize into different phenotypes and perform different functions under different microenvironmental stimuli (16). According to different functions and phenotypes, TAMs are mainly divided into classically activated M1 type macrophages and substitutivity activated M2 type macrophages (18). M1 and M2 macrophages have obvious functional differences, and can be transformed according to different environmental stimulus factors. Study has shown that M1-type macrophages are dominant in the early stage of NSCLC, while M2-type macrophages are dominant in the middle and late stages. With the progression of tumor, M1 gradually transforms to M2 phenotype (14). In this study, the heterogeneity of NSCLC was investigated by analyzing single-cell data. The M0 and M1 macrophages were significantly different in different clusters of NSCLCs, suggesting that macrophages play an important role in specific immunotherapy. At the same time, we analyzed the influence of macrophage-related genes (MRGs) on the prognosis and chemotherapeutic drug sensitivity of NSCLC to provide insights for personalized treatment. We further elucidated the critical role of the migration inhibitory factor (MIF) signaling pathway in the communication between NSCLC cells, which provides new guidance for patient immunotherapy. We present the

Highlight box

Key findings

- We have sufficiently revealed the tumor microenvironment of NSCLC through single-cell data analysis and constructed a prognosis model of macrophage-related genes.

What is known and what is new?

- We constructed a new macrophage prognostic model by means of effective univariate Cox and LASSO regression analyses model construction;
- We clarified the intercellular communication of NSCLC and found that the MIF signaling pathway plays the most significant role in the communication in epithelial cells-monocyte.

What is the implication, and what should change now?

- We need to verify the role played by MIF in cell signaling and cell communication through *in vitro* and *in vivo* experiments.

following article in accordance with the TRIPOD reporting checklist (available at https://tlcr.amegroups.com/article/view/10.21037/tlcr-22-866/rc).

Methods

Data acquisition and processing of expression data and clinical information

The small conditional RNA (scRNA)-seq data were obtained from the Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/) database (GSE148071) (19). In this study, scRNA-seq was analyzed primarily using the "Seurat" package in R (The R Foundation for Statistical Computing, Vienna, Austria) (20-22). PercentageFeatureSet from the "Seurat" package was used to calculate the mitochondrial percentages in each cell. Cells outside the threshold of 50 expressed genes and those with more than 5% mitochondrial content were excluded. LogNormalize was used to normalize the scRNA-seq data before 1,500 highly variable genes (HVGs) were identified based on the FindVariableGenes function. Moreover, the RNA-sequencing (RNA-seq) data of 108 normal and 1,041 NSCLC samples [The Cancer Genome Atlas (TCGA)-lung adenocarcinoma (LUAD) and TCGA-lung squamous cell carcinoma (LUSC)] were accessed from TCGA (http://cancergenome.nih.gov/) database, and additional NSCLC samples were obtained from the GEO database (GSE74706 and GSE103512) (23,24). The clinical data of TCGA-LUAD and TCGA-LUSC were extracted from the TCGA database; ambiguous survival status or unclear clinicopathological characteristics were excluded. NSCLC samples and corresponding adjacent normal tissues were collected from the West China Hospital, Sichuan University were isolated. After surgical removal, the samples were immediately frozen in liquid nitrogen. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The present study was approved by the Ethics Committee of West China Hospital (Ethics 2022 No. 74). Informed consent was obtained from the patients or their guardians. The quantitative polymerase chain reaction (qPCR) primer sequences are listed in Table S1.

Dimensionality reduction and cell visualization

Principal component analysis (PCA) was conducted for linear dimensional reduction of the scRNA-seq data, and the principal components (PCs) were then used for t-distributed stochastic neighbor embedding (tSNE) clustering. With the criteria of log2 fold change (FC) >1 and false discovery rate (FDR) <0.05, marker genes in each cluster were identified, and the top 10% of marker genes from clusters were laid out in the heatmap. Annotation of clusters was performed using the "SingleR" package (25), which compares the transcriptomes of single cells to reference datasets to determine cellular identity.

NSCLC differentiation-related genes (NDRGs)-based molecular subtypes

In this study, we analyzed NSCLC cell fate determination and pseudo time tracing using the "Monocle" package (21,26). Then, intracellular differentially expressed genes in cells with distinct differentiation states with log2 FC >1 and FDR <0.05 were designated as NDRGs. After the log2-scale transformation of NDRGs expression in the GSE103512 dataset, 75 NDRGs were retained for NSCLC molecular typing. Subtype identification was performed using the "ConsensusClusterPlus" package. For the best cluster number, the K-means algorithm and cumulative distribution function (CDF) curve were applied, and 50 iterations with max K=9 were conducted for stable subtypes. Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis was used for the degree of enrichment of NDRGs between 3 molecular subtypes in pathway terms.

TME scores in the three molecular subtypes

ESTIMATEScore, ImmuneScores, TumorPurity, and StromalScore were calculated using the Estimation of STromal and Immune cells in MAlignant Tumor tissues using Expression data (ESTIMATE) package. All 22 kinds of immune cells were calculated using the Cell-type Identification by Estimating Relative Subsets of RNA Transcripts (CIBERSORT) package. To analyze the association between molecular subtypes and immune cell infiltration, we adopted CIBERSORT to estimate the infiltration levels of 22 immune cell subtypes across molecular subtypes. Molecular subtype expression differences of common ICIs were analyzed using the expression profile of each molecular subtype.

The macrophage-related prognostic model for NSCLC

Based on the CIBERSORT algorithm, immune cell

distribution was estimated in TCGA-LUAD and TCGA-LUSC. A Pearson correlation analysis of the MRGs co-expression relationship was carried out, with cutoffs of R>0.4 and P<0.001 used to represent co-expression. Gene Ontology (GO) enrichment and KEGG analysis were used to reveal the MRGs-related biological functions in NSCLC.

Cox regression and least absolute shrinkage and selection operator (LASSO) regression analyses were used to screen for prognostic MRGs in TCGA. The risk score of the gene signature was calculated as follows: risk score = (Coefficientgene 1 × expression of gene 1) + (Coefficientgene 2 × expression of gene 2) + … + (Coefficientgene n × expression gene n). The diagnostic value of the risk score was assessed using receiver operating characteristic (ROC) curves to divide the patients with NSCLC into low-risk and high-risk groups.

Next, to determine the risk score of each patient and to analyze the overall survival (OS) time between risk groups, Kaplan-Meier survival analysis was performed. A multivariable Cox regression analysis of a risk-score model and clinical characteristic parameters was conducted to evaluate 1-, 3-, and 5-year OS probability in the TCGA cohort using the "rms" package.

Our analysis of enrichment terms in the entire TCGA cohort was carried out using version 4.1.0 of the gene set enrichment analysis (GSEA) software (http://www.gsea-msigdb.org/gsea/index.jsp).

The function of risk scores for chemotherapeutic drugs predicted

To evaluate the potential significance of the model in the clinical treatment of NSCLC, we calculated the half-maximal inhibitory concentration (IC50) values of chemotherapeutic drugs using the "pRRophetic" package (27). The differences in the IC50 values between the high- and low-risk groups were evaluated using the Wilcoxon signed-rank test.

Inference and analysis of cell-cell communication

Extraction was based on single-cell receptor and ligand expression levels used to infer intercellular communication. The intercellular communication networks were analyzed based on scRNA-seq data (GSE148071) using the "CellChat" package (28).

Statistical analysis

All statistical analyses were carried out in R and Perl, and P<0.05 was considered indicative of statistical significance.

Results

Single-cell profiling of gene expression in NSCLC cells

Analyzing single-cell RNA-seq data with the Seurat V4.1 R package was performed. In this study, 1,858 cells from 4 NSCLC samples were obtained from GSE148071. A total of 3,146 genes were analyzed, of which 1,646 had low intercellular variation and 1,500 had high variation (*Figure 1A*). For dimensionality reduction, the top 1,500 HVGs with the highest variance were selected for PCA. From this analysis, 11 PCs were selected to run the tSNE algorithm. A heat map showing the marker genes of each cluster is displayed in *Figure 1B*. We identified 7 types of cells, including T cells, monocytes, B cells, epithelial cells, tissue stem cells, endothelial cells, and natural killer (NK) cells (*Figure 1C*). In addition, we analyzed the expression of the 8 greatest variation genes in NSCLC tissues and normal tissues (*Figure 1D*). These results are consistent with the results obtained using gene expression profiling interactive analysis (GEPIA) (Figure S1).

Clustering analysis revealed the heterogeneity of the immune microenvironment in NSCLC

The study identified 209 NDRGs. NDRG-based consensus clustering analysis was completed in GSE148071, and 3 molecular subtypes that contained all the NSCLC samples were identified at a clustering threshold of max K=9 (*Figure 2A-2C*). Using KEGG function enrichment, the NGDRGs in subset 1 were involved in cytokine production and cell activation (*Figure 2D*), those in subset 2 were also associated with phagosome and lysosome (*Figure 2E*), and those in subset 3 were closely related to phagosome and tuberculosis (*Figure 2F*). Then, we compared the immune scores of each molecular subtype. The immune scores of the molecular subtypes were correlated using ESTIMATE. After that, we compared the ESTIMATEScore, ImmuneScores, TumorPurity, and StromalScore of the 3 molecular subtypes. The current study found that the differences in estimate/immune/stromal scores were highly significant in subtypes C1/2/3 (*Figure 3A-3D*). The results of this study indicate that NSCLC molecular subtypes have different effects on the efficacy of therapeutics. The CIBERSORT software was further used to generate gene expression configuration files for 22 immune cells, which were then combined

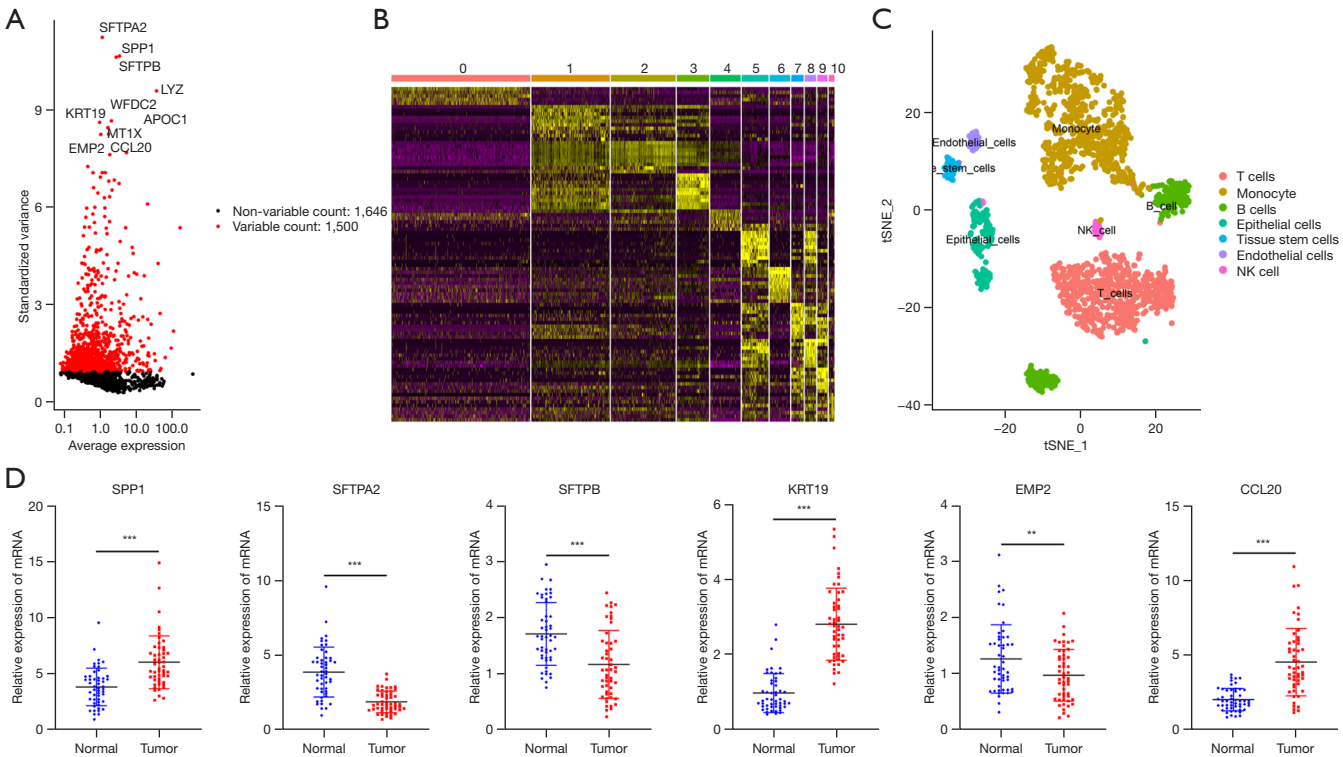

Figure 1 Analysis of NSCLC heterogeneity and microenvironment using single cells. (A) A total of 3,146 genes were analyzed, of which 1,500 had high variation. (B) 500 NSCLC cells were aggregated into 11 clusters (principal components 0–10) and the 1,500 genes in each cluster are displayed on the heat map. (C) Single-cell clustering analysis based on the full scRNA-seq data and the annotation of each cluster based on canonical marker analysis. (D) Expression of genes was examined by qRT-PCR. **, P<0.01; ***, P<0.001. tSNE, t-distributed stochastic neighbor embedding; NK, natural killer; NSCLC, non-small cell lung cancer; scRNA, small conditional RNA; qRT-PCR, quantitative real-time polymerase chain reaction.

with the gene expression matrix of lung cancer samples to generate 22 immune cell counts (*Figure 3E*). To investigate the potential for altered immune response caused by 3 molecular subtypes, we estimated the 22 immune cell infiltration levels by CIBERSORT. B cells memory, plasma cells, T cells CD4 memory resting, T cells regulatory, NK cells activated, monocytes, macrophages M0, macrophages M2, dendritic cells resting, and mast cells resting differed significantly between molecular glioma subtypes (*Figure 3F*). Further expression profiles among immune checkpoints in 3 subtypes of NSCLC were also confirmed (*Figure 3G*). These discoveries, based on molecular subtypes, potentiate future personalized and optimal treatments.

The expression of MRGs in NSCLC

TAMs are a critical component of the TME and are involved in various aspects of tumor behavior (16,29-31). According to the above results, we found that macrophage infiltration levels showed significant differences among the different molecular subtypes of NSCLC. We next aimed to analyze the role of macrophages in NSCLC more directly. First, the macrophage 0-, 1-, and 2-related genes were screened based on the Spearman correlation method with the absolute value of the correlation coefficient >0.4 and P value <0.001 as the threshold (*Figure 4A-4C*). To further elucidate the functions of the identified MRGs in NSCLC, GO and KEGG pathway enrichment analyses were employed, and MRGs were found to be enriched on immune-related pathways (*Figure 4D,4E*). These results reveal that macrophages infiltration plays a critical role in tumor progression.

Construction of an MRGs prognostic-predicting model

We next investigated the prognostic role of MRGs in

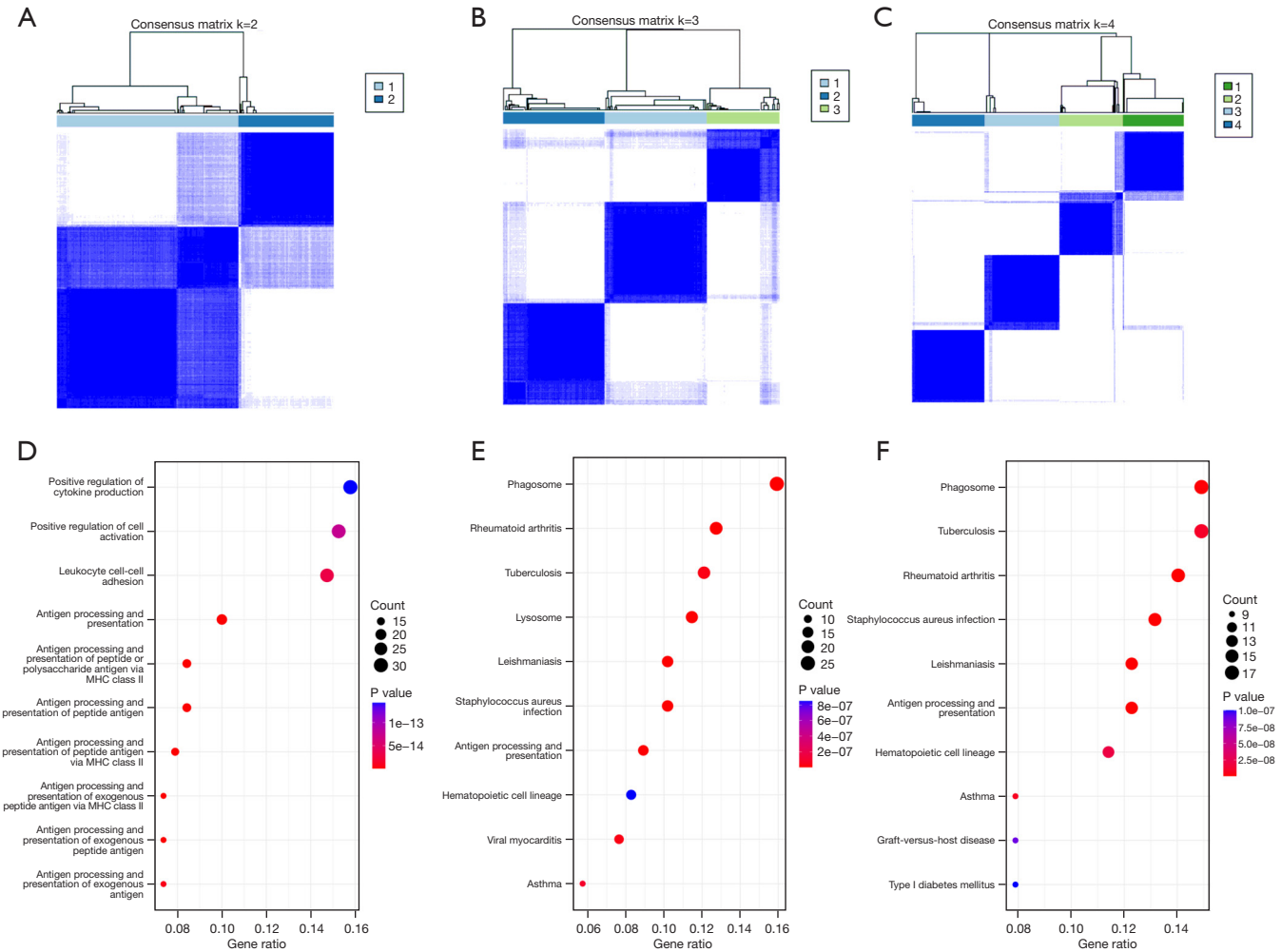

Figure 2 KEGG enrichment analysis of genes in different clusters. (A-C) Consensus clustering CDF for k=2 to k=4. (D) KEGG enrichment analysis of genes between clusters 1–2. (E) KEGG enrichment analysis of genes between clusters 1–3. (F) KEGG enrichment analysis of genes between clusters 2–3. MHC, major histocompatibility complex; KEGG, Kyoto Encyclopedia of Genes and Genomes; CDF, cumulative distribution function.

NSCLC patients. A univariate Cox regression analysis was conducted on 104 MRGs to identify prognosis-related genes. A total of 7 robust prognosis-related genes were identified (*Figure 5A*). Multivariate Cox regression analysis was then performed to determine the independent prognostic factors and to build prediction models (Figure S2A,S2B). NSCLC patients were divided into high- and low-risk groups using the median risk score. We analyzed whether risk score and other clinical traits were independent prognostic factors (Figure S2C,S2D). A Kaplan-Meier survival curve was applied to compare the survival rates between high- and low-risk scores (*Figure 5B,5C*). The ROC curves showed better efficacy to predict survival using the risk score (*Figure 5D*, Figure S2E). The ROC analyses for 1-year survival prediction indicated high areas under the curve (AUCs) (*Figure 5E*). Furthermore, patients with a high risk also showed a poor prognosis in terms of the OS time for age >65 years groups (*Figure 5F*), males (*Figure 5G*), stage III–IV (*Figure 5H*); patients with a low risk also showed a poor prognosis in terms of the OS time for age >65 years groups (*Figure 5F*), and stage III–IV (*Figure 5H*). Finally, a prognostic nomogram was established based on the TCGA (TCGA-LUAD and TCGA-LUSC) dataset, and the calibration curve indicated a high reliability of the nomogram (*Figure 5I,5J*). In summary, we built a risk

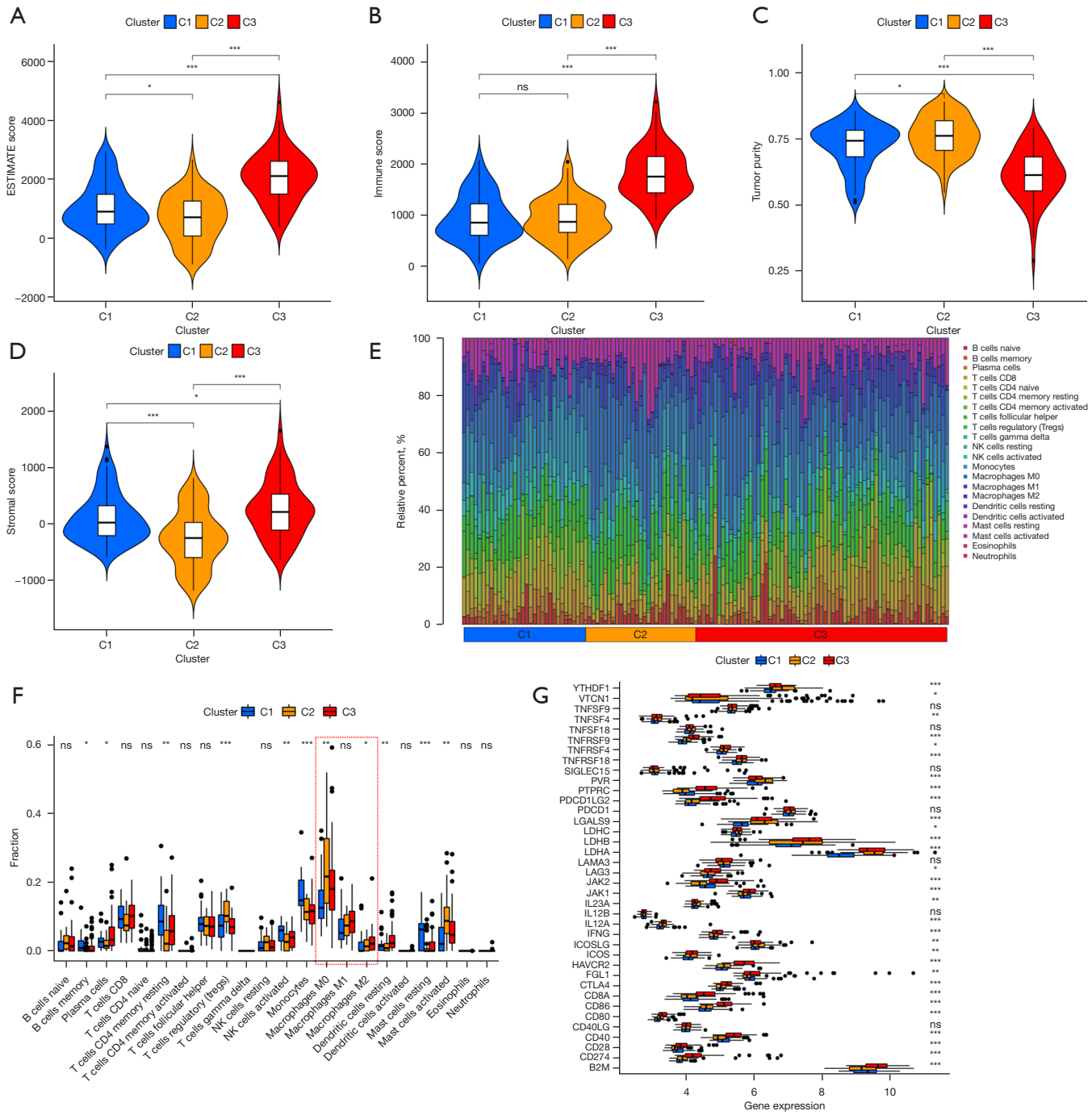

Figure 3 Analysis of overall tumor microenvironment scores and immune infiltrate in three molecular subtypes. (A-D) TME scores by three molecular subtypes. (E) The contents of 22 immune cells in each sample from the GSE103512 dataset. (F) The content difference analysis of 14 kinds of immune cells in 3 subtypes. (G) The difference expression analysis of 38 ICGs in 3 subtypes. *, P<0.05; **, P<0.01; ***, P<0.001. ESTIMATE, Estimation of STromal and Immune cells in MAlignant Tumor tissues using Expression data; NK, natural killer; ns, not significant; TME, tumor microenvironment; ICGs, immune checkpoint genes.

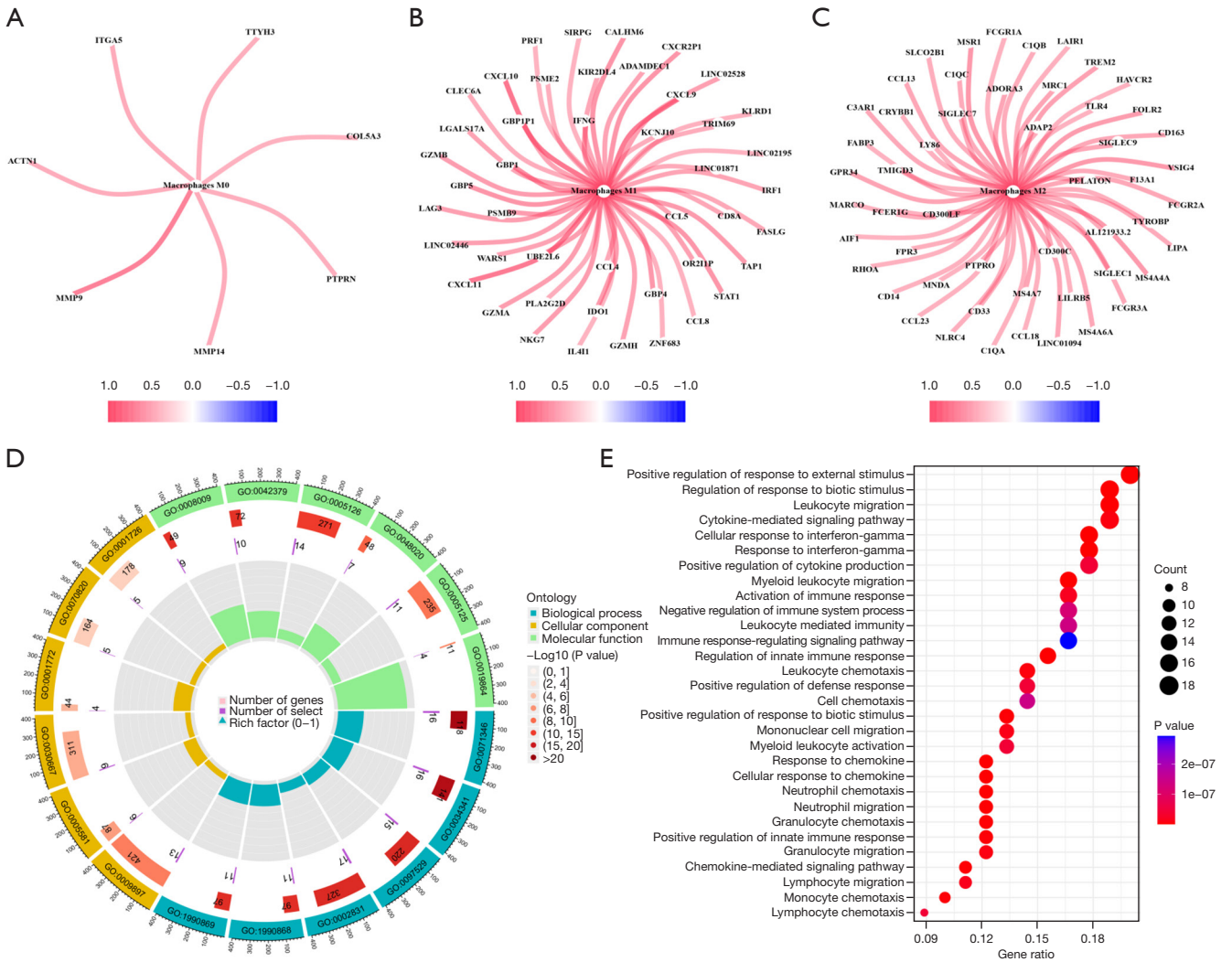

Figure 4 The macrophages associated genes and pathway enrichment (GO and KEGG) were carried out. (A) M0 macrophage-associated gene network. (B) M1 macrophage-associated gene network. (C) M2 macrophage-associated gene network. GO (D) and KEGG (E) enrichment analysis of macrophages-related genes. GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.

model based on MRGs.

Association between risk score and immune infiltration and tumor mutation burden (TMB)

To better understand the difference in immune cell infiltration between low- and high-risk groups, we carried out an infiltration analysis of immune cells. Among them, mast cells activated, macrophages M0, macrophages M2, neutrophils, NK cells resting, and T cells CD4 memory activated were positively correlated with a risk score, the remaining were negatively correlated (*Figure 6A*,*6B*). To verify the results of TCGA, a search was performed in the GEO (GSE74706 and GSE103512). The TAM infiltration was enriched in NSCLC (*Figure 6C*). Besides, the correlation between risk score and immune checkpoints was also investigated. The expression of ICIs was strongly associated with risk score (*Figure 6D*). The risk score may be used to predict NSCLC immunotherapy response based on its associations with immune checkpoint genes and tumor immune infiltration. Moreover, TMB was measured as a potential biomarker for immunotherapy. The TMB analysis showed that the risk scores of the high/low-risk TMB were significantly different (Figure S2F). The TMB

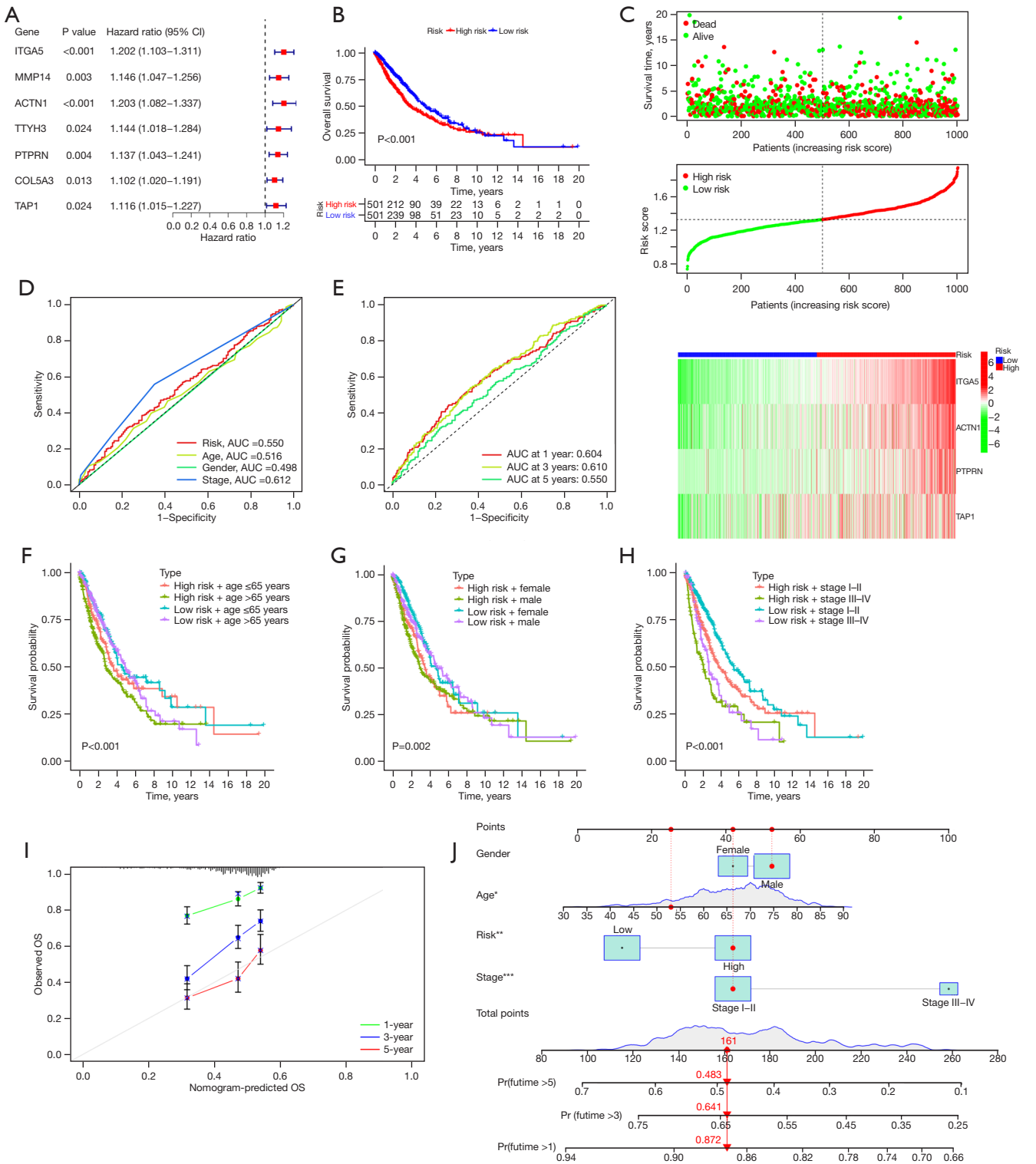

Figure 5 Construction of the prognostic model based on macrophage-related coding genes. (A) Prognosis analysis of macrophage-related gene using univariate analysis. (B) Kaplan-Meier curves for OS in risk groups according to the risk score. (C) Survival duration and status of NSCLC patient cases. macrophages-related gene risk score analysis of NSCLC patients. (D,E) ROC curve analysis according to the clinicopathological features and 1, 3, and 5-year survival of the area under the AUC value. (F-H) The prognosis of NSCLC patients with high/low-risk scores. (I,J) The nomogram to predict the 1-, 3-, and 5-year survival risk of NSCLC patients. *, P<0.05; **, P<0.01; ***, P<0.001. AUC, area under the curve; OS, overall survival; NSCLC, non-small cell lung cancer; ROC, receiver operating characteristic.

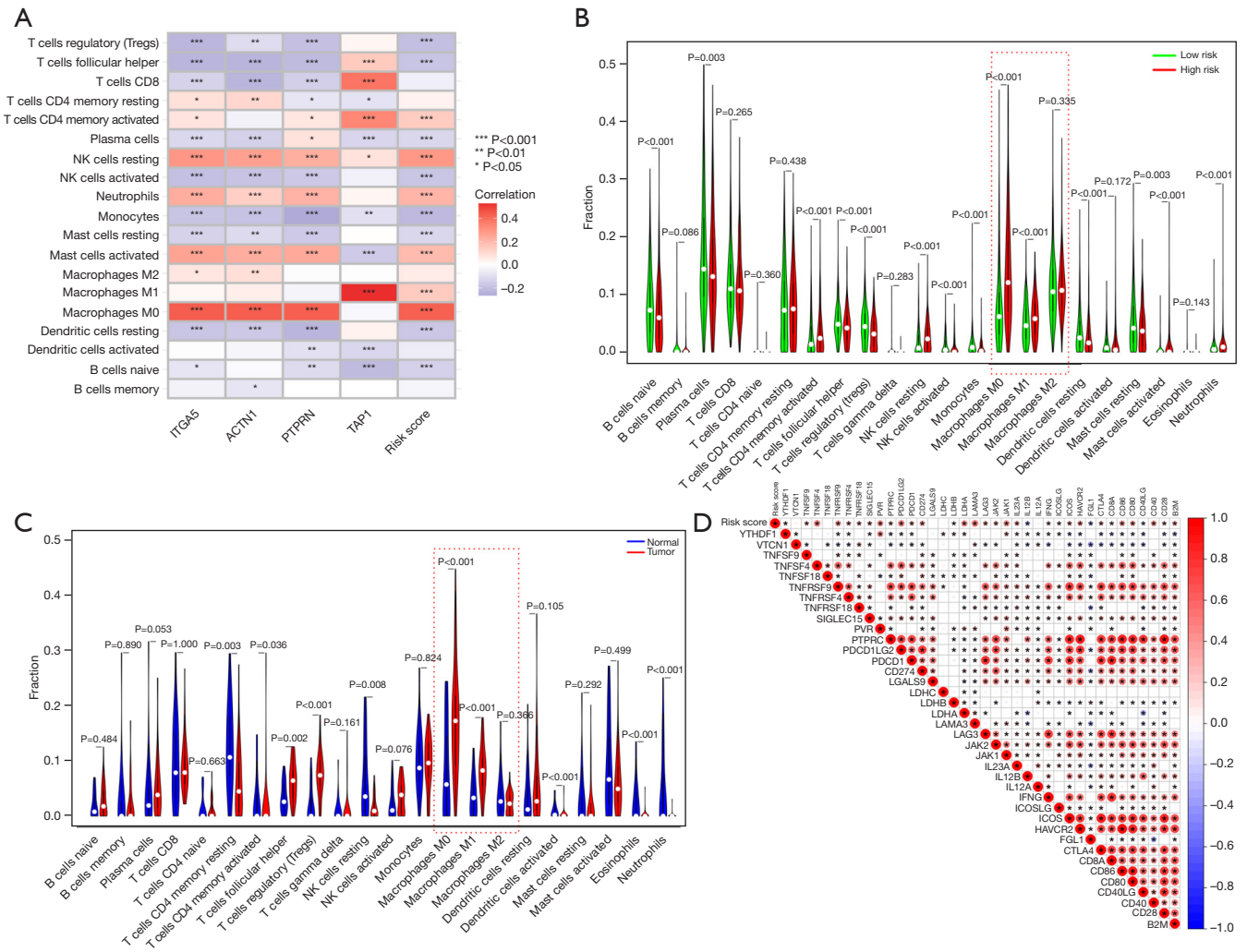

Figure 6 A correlation analysis of immune cell infiltration and risk score. (A) The heatmap shows the relationship between immune cells and risk score. (B) The infiltrating proportion of immune cells in high- and low-risk groups. (C) The infiltrating proportion of immune cells in normal and tumor groups. (D) The relationship between immune checkpoints and risk score. *, P<0.05; **, P<0.01; ***, P<0.001. NK, natural killer.

of each gene was significantly different between high- and low-risk groups (*Figure* 7). This also provides guidance for practical applications based on TMB.

GSEA was conducted between high- and low-risk groups

We performed a GSEA to determine which pathways were differentially enriched across risk groups. In terms of signaling pathway enrichment analysis, we found that the high-risk groups were mainly focused on the cell cycle, chemokine, cytokine-receptor, extracellular matrix (ECM)-receptor, and focal adhesion signaling pathway (*Figure 8A*). These findings may indicate that the above signaling is associated with the malignant prognosis of NSCLC. Meanwhile, GSEA revealed that drug metabolism by cytochrome p450, linoleic acid metabolism, maturity-onset diabetes of the young, metabolism of xenobiotics by cytochrome p45, and ribosome pathways were mainly enriched in the low-risk group (*Figure 8B*). The results of GSEA showed that significantly metabolism-related pathways were enriched in the low-risk group, which can provide critical clues for personalized treatment.

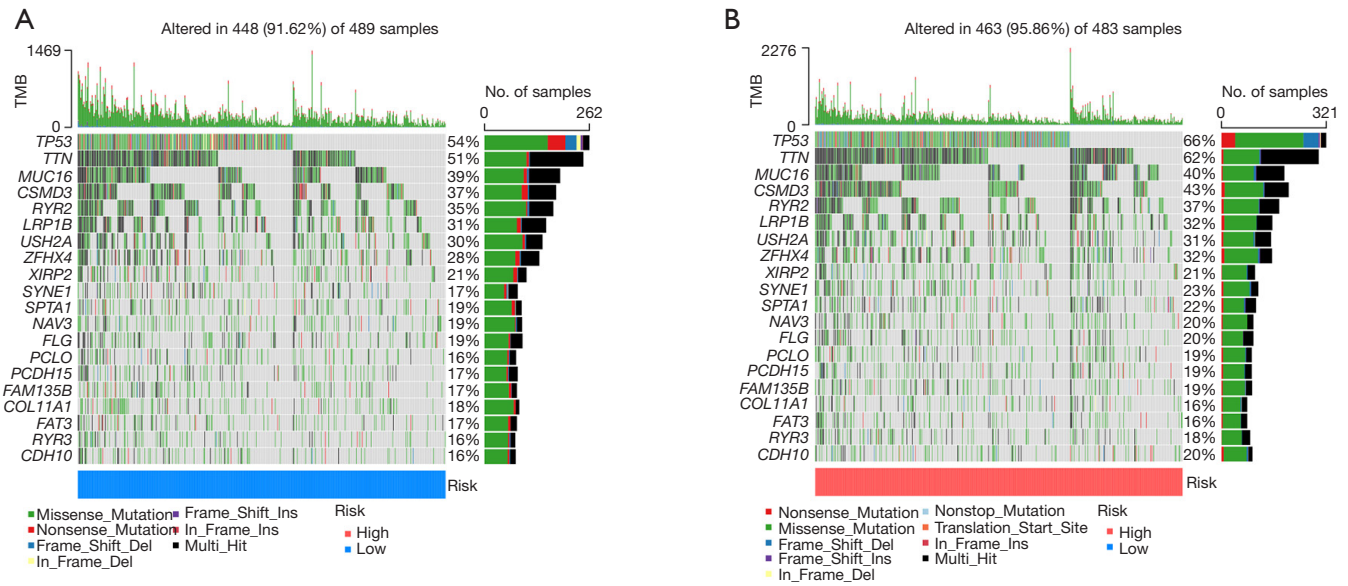

Figure 7 Correlation analysis between risk groups and total mutation count. The mutation in low-risk (A) and high-risk (B) groups. TMB, tumor mutation burden.

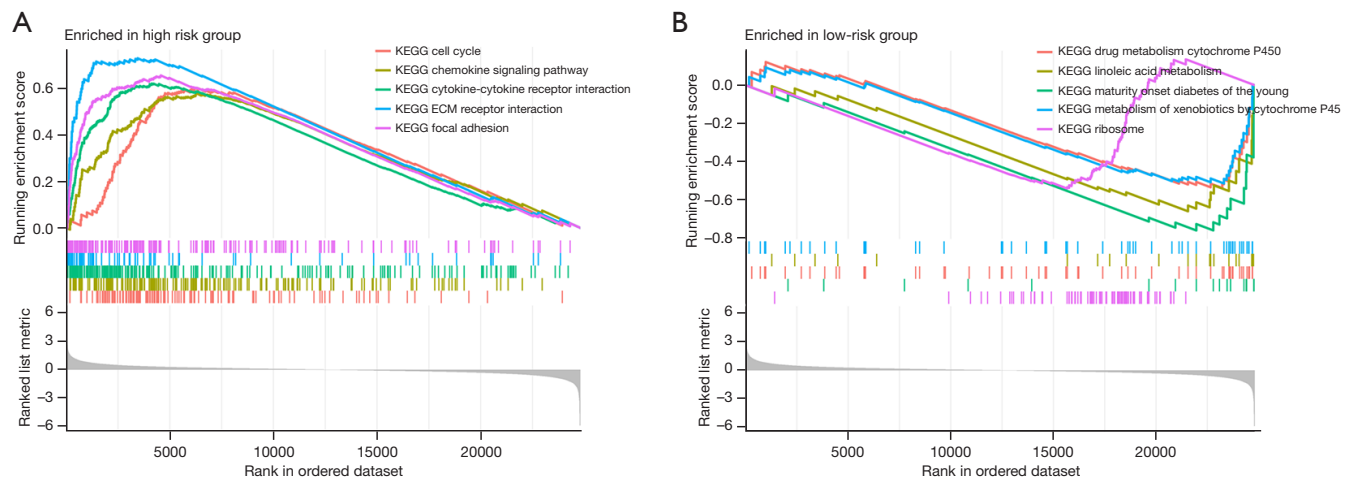

Figure 8 GSEA analysis. GSEA signal pathway enrichment analysis was used to analyze the high-risk group (A) and low-risk group (B). GSEA, gene set enrichment analysis; KEGG, Kyoto Encyclopedia of Genes and Genomes.

Drug sensitivity of NSCLC patients with high/low risk

It is necessary to improve outcomes through more accurate risk group stratification and effective personalized risk-adapted treatment. Patients scoring low risk received single-agent chemotherapy, whereas those scoring high risk received different chemotherapy regimens. Furthermore, we predicted which chemotherapeutic drugs should be prescribed based on the tumor heterogeneity among different risk groups. The data showed that a low-risk score was associated with lower IC50 values for drugs such as ribociclib (*Figure 9A*), PF-4708671 (*Figure 9B*), ABT737 (*Figure 9C*), OSI-027 (*Figure 9D*), doramapimod (*Figure 9E*), and AZD6482 (*Figure 9F*), a high-risk score was associated with lower IC50 values for drugs such as BMS-536924 (*Figure 9G*), alisertib (*Figure 9H*), entospletinib (*Figure 9I*), BPD-00008900 (*Figure 9J*), AZ960 (*Figure 9K*), and

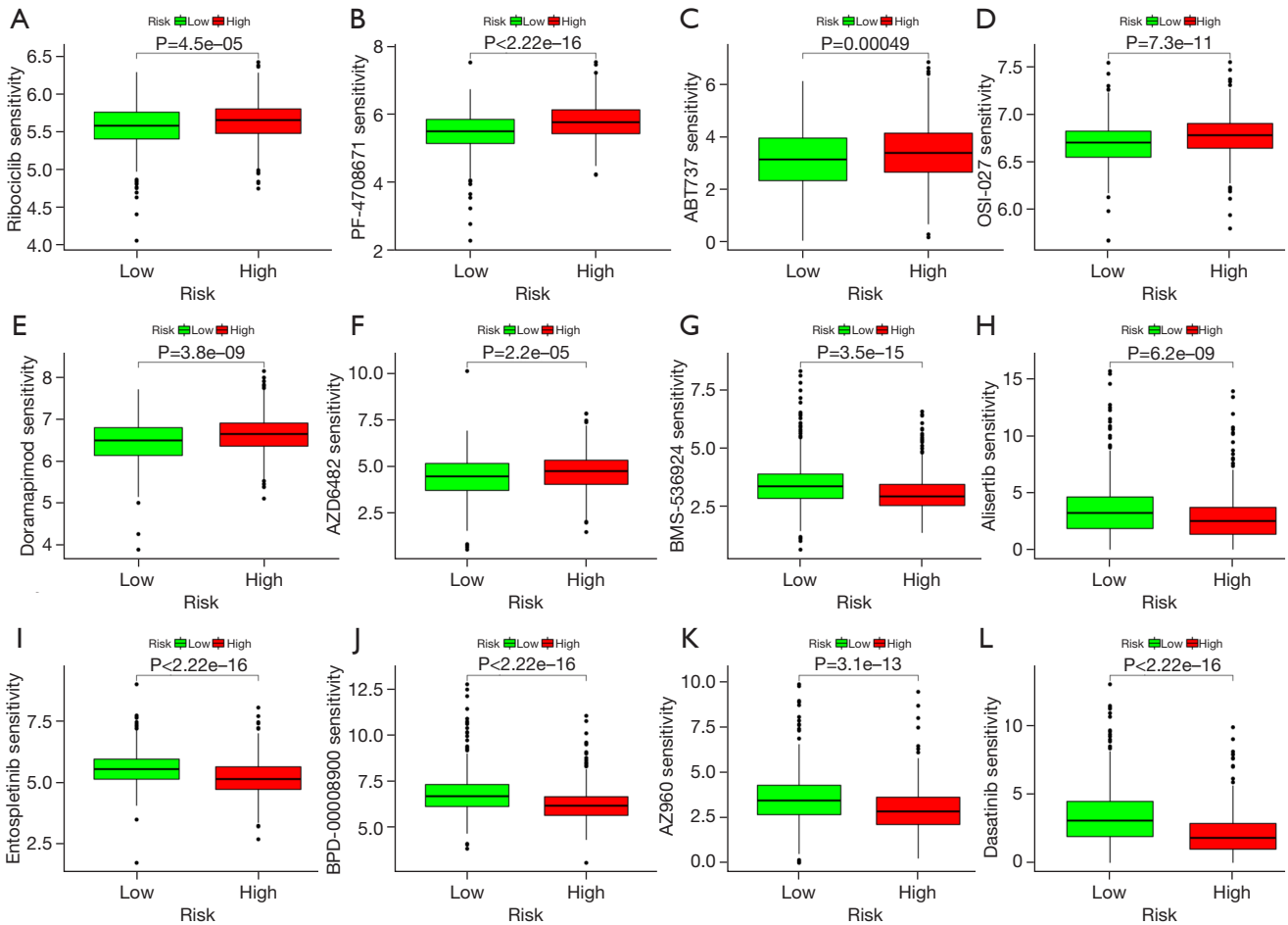

Figure 9 Drug sensitivity of NSCLC patients with high/low risk. (A-F) A low-risk score was associated with lower IC50 values for drugs. (G-L) A high-risk score was associated with lower IC50 values for drugs. NSCLC, non-small cell lung cancer.

dasatinib (*Figure 9L*). So, accurate prediction can potentially help to provide better treatment for patients.

Integrated analysis reveals the interaction between the same or different cell types

To determine possible interactions between immune cells, we performed CellChat analysis on GEO data (GSE148071). The CellChat tool can identify and predict the putative functions of poorly understood signaling pathways in a scRNA-seq dataset. The cell communication includes 3 main parts: secreted signaling, ECM-receptor, and cell-cell contact mediated through heterodimer interactions, and information on ligand-receptor was sourced from the KEGG database (*Figure 10A*). The results showed that epithelial cells, endothelial cells, B cells, tissue stem cells, T cells, NK cells, monocytes, and macrophages interact closely (*Figure 10B,10C*). Monocyte-macrophage (*Figure 10D*), tissue stem cells-monocyte (*Figure 10E*), T cells-macrophage (*Figure 10F*), macrophage-monocyte (*Figure 10G*), endothelial cells-monocyte (*Figure 10H*), B cells-monocyte (*Figure 10I*), epithelial cells-monocyte (*Figure 10J*), and NK cells-macrophage (*Figure 10K*) were found to be significant interactions. The results provide a clear understanding of the cell communication of NSCLC.

The function of the MIF signaling network in TME

Although studies have consistently reported that the cells communicate by releasing and receiving secreted signaling molecules, the interacting effects of signaling pathways on cell-cell communication in NSCLC are poorly understood.

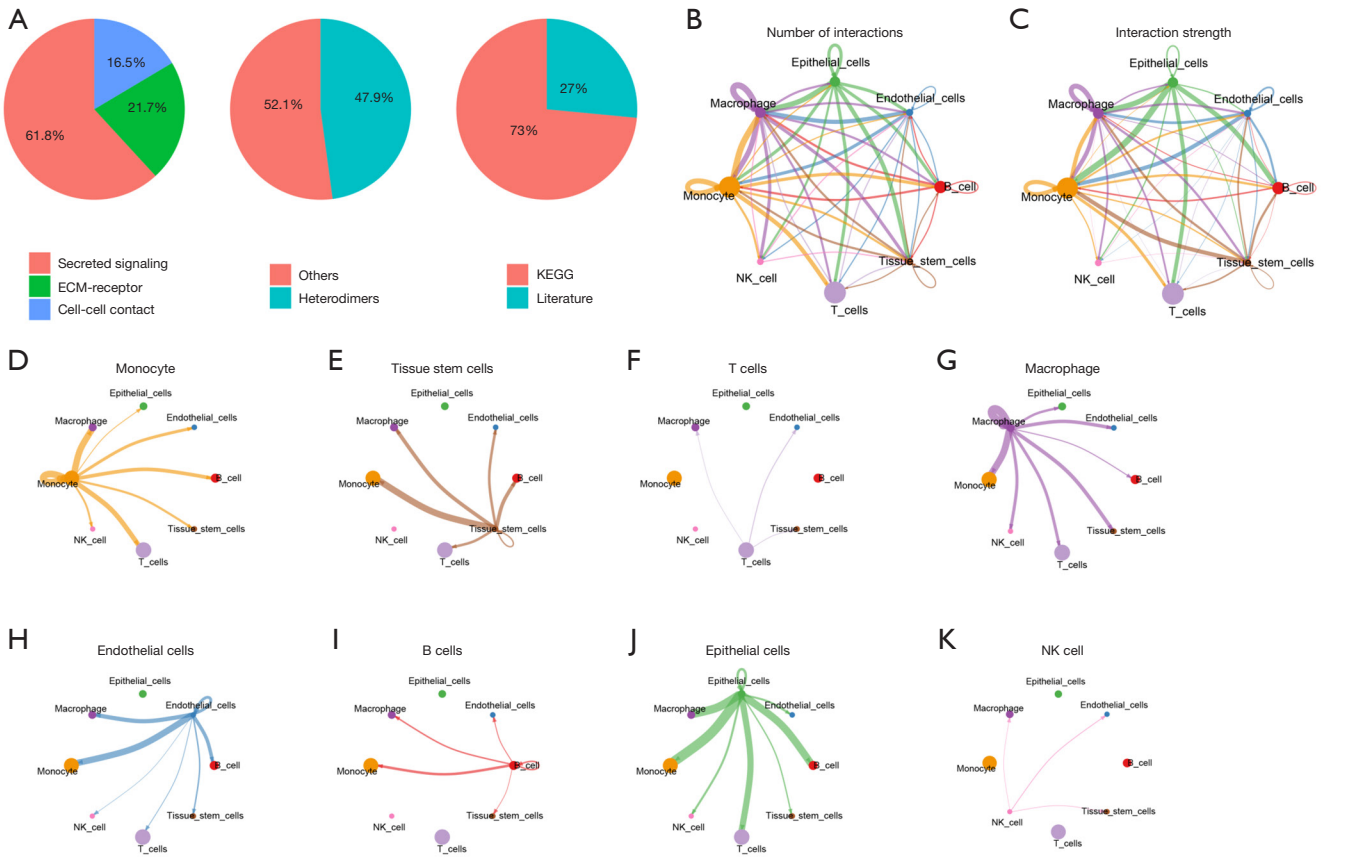

Figure 10 Network diagram of the immuno-infiltrating cells communication. (A) Types of receptor-ligand interactions; the way receptors and ligands interact; receptor-ligand source. (B) The total number of ligand-receptor interactions between cells of the same or different cell types. (C) The strength of ligand-receptor interactions between cells of the same or different cell types. (D-K) Receptor ligand pair interactions between immune cells. ECM, extracellular matrix; KEGG, Kyoto Encyclopedia of Genes and Genomes; NK, natural killer.

Next, we analyzed the effects on the cells' communication signaling pathways molecules. The SPP1, RETN, NAMPT, MIF, and LGALS9 signaling pathways might play a role in cell-cell communication (*Figure 11A*). Our analysis found that the MIF signaling pathway plays the most significant role in the communication in epithelial cells-monocyte (*Figure 11B*). The MIF signaling pathway might affect the B cell influencer, endothelial cells sender, epithelial cells sender, macrophage influencer, monocyte receiver or mediator, T cells receiver or influencer, and tissue stem cells influencer to exert its effect (*Figure 11C*). Among them, the oncogenic effects of MIF are mediated by binding to receptors CD74, CXCR4, and CD44, which are involved in MIF cell-signaling (*Figure 11D-11F*). Clearly, exosomes do not just play a role in cell-cell communication, and the mechanism of action of MIF needs more elucidation.

Discussion

In recent years, with the continuous improvement of medical research level, people have an increased understanding of immunosuppression and realize the value of the immune system in the treatment of advanced NSCLC (19,32,33), so more and more researchers have begun to discuss and analyze the effect of immunotherapy. It has been established that immune-related biomarkers are associated with a better prognosis in various types of cancer (34-37). In recent years, immunotherapy has brought epoch-making changes to the treatment of NSCLC. In particular, programmed cell death ligand 1 (PD-1)/programmed cell death ligand 1 (PD-L1), The ICIs of PD-L1 have been approved as first-line and second-line therapies in patients with metastatic NSCLC or some locally advanced NSCLC. However, only 15–30% of patients with advanced NSCLC

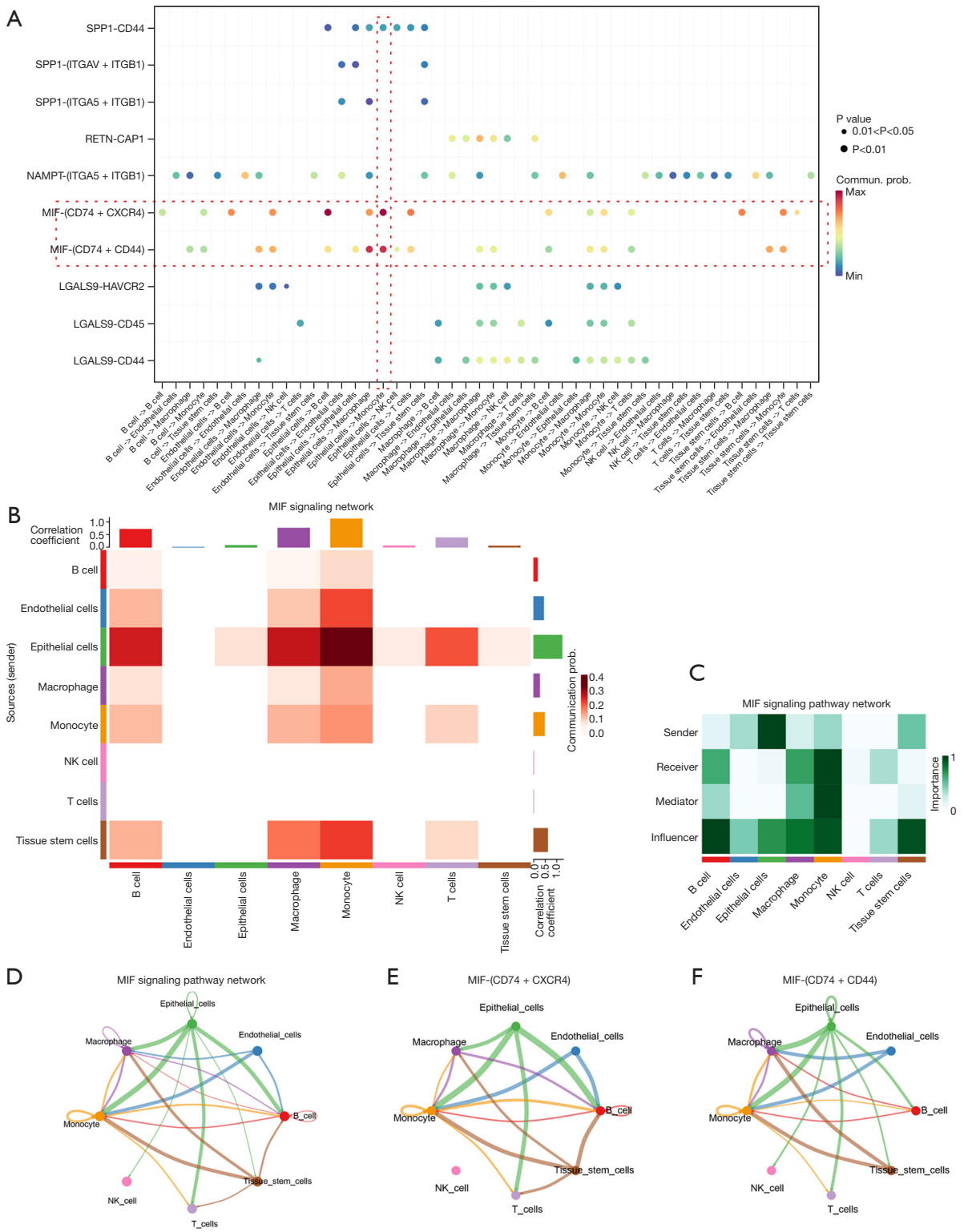

Figure 11 MIF signaling pathway networks play important roles in immuno-infiltrating cell communication. (A) Bubble diagram of cell communication signaling pathway. (B) The role of MIF signaling pathway in cell-cell communication. (C) Heat map of cell communication in the sender, receiver, mediator, and influencer. (D-F) MIF signaling pathways are mediated by binding to receptors CD74, CXCR4, and CD44. Commun. prob., communication probability; MIF, migration inhibitory factor; NK, natural killer.

can achieve sustained remission and long-term survival from immunotherapy (38). How to search for good biomarkers to effectively predict the efficacy of immunotherapy is one of the great challenges currently faced. Thus, our study aimed to develop promising immune-related biomarkers and prognostic risk models for early diagnosis and treatment of NSCLC patients.

In the analysis of the immune infiltration landscape, we found that NSCLC had 7 immune infiltration subtypes, the most numerous were the population of monocytes with a classical phenotype. Monocytes are a type of white blood cell in the body's immune system which are produced in the bone marrow, and are classified monocytes inside blood vessels and macrophages outside blood vessels (39-43). Macrophages are highly heterogeneous and plastic immune cells belonging to the mononuclear macrophage system (44,45). Macrophages infiltrating tumor tissues are called TAMs, which are important types of immune cells in the TME. They are involved in regulating the growth, invasion, and distant metastasis of lung cancer by inhibiting tumor immunity and promoting angiogenesis (16,17). The infiltration level and polarization state of TAMs in lung cancer tissues are closely related to the prognosis of patients and may be a potential target for immunotherapy of lung cancer (46-49). Thus, in the current study, MRGs were analyzed in NSCLC.

A univariate Cox regression analysis was conducted on 104 MRGs to identify prognosis-related genes, and 7 robust prognosis-related genes were identified. MRGs were enriched on immune-related pathways and responded to interferon (IFN) regulation. The polarization process of macrophages is affected by a variety of regulatory modes, including gene transcription level, post-transcription level, and post-translational protein modification, and they cross each other to form regulatory networks (50). The main stimulating factors that induce the polarization of macrophage M1 are the Toll-like receptor (TLR) ligand and IFN-γ (51-53). These results reveal that IFN-γ plays a critical role in macrophage's anti-tumor progression. The specific role of macrophages in NSCLC is worthy of further study.

Multivariate Cox regression analysis was performed to determine the independent prognostic factors and to build a prediction model. NSCLC patients were divided into high- and low-risk groups using the median risk score. A poorer survival was found in high-risk patients. We also explored the expression of the ICIs, immune infiltration, and TMB between the high- and low-risk NSCLC patients. The risk score may be used to predict NSCLC immunotherapy response based on its associations with ICIs and tumor immune infiltration. Furthermore, we predicted which chemotherapeutic drugs should be prescribed based on the tumor heterogeneity among different risk groups. These results suggest that prognostic models can play a very important role in predicting patients' high and low risk of NSCLC and in personalizing treatment.

Finally, we also revealed the intercellular communication of NSCLC and found that the MIF signaling pathway plays the most significant role in the communication in epithelial cells-monocyte. MIF can regulate the polarity of TAM, and in tumors, macrophages usually differentiate into tumor-promoting M2 TAM, which accelerates the proliferation and metastasis of tumor cells and inhibits anti-tumor immune response (54-56). MIF is an immunomodulator in the TME, inducing the formation of a tumor inhibitory immune microenvironment (54). At present, studies on the function of MIF have revealed some clinical applications, such as a diagnostic marker for some diseases. However, the pathophysiological function, enzyme activity, signal transduction mechanism, and related biological functions of MIF have not been fully clarified and require further exploration.

Conclusions

As a result of our comprehensive integrated analysis of MRGs in NSCLC, we can better understand the molecular events related to the progression of NSCLC. It was also found that a risk score model can accurately predict chemotherapy response in patients and prognosis.

Acknowledgments

The authors gratefully acknowledge each editor and reviewer for their profound insight into this study.
Funding: None.

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at https://tlcr.amegroups.com/article/view/10.21037/tlcr-22-866/rc

Data Sharing Statement: Available at https://tlcr.amegroups.com/article/view/10.21037/tlcr-22-866/dss

Conflicts of Interest: All authors have completed the ICMJE

uniform disclosure form (available at https://tlcr.amegroups.com/article/view/10.21037/tlcr-22-866/coif). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The present study was approved by the Ethics Committee of West China Hospital (Ethics 2022 No. 74). Informed consent was obtained from the patients or their guardians.

References

1. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin 2021;71:209-49.
2. Jemal A, Bray F, Center MM, et al. Global cancer statistics. CA Cancer J Clin 2011;61:69-90.
3. Del Mastro L, Gennari A, Donati S. Chemotherapy of non-small-cell lung cancer: role of erythropoietin in the management of anemia. Ann Oncol 1999;10 Suppl 5:S91-4.
4. Dohopolski M, Iyengar P. Oligometastatic non-small cell lung cancer: a narrative review of stereotactic ablative radiotherapy. Ann Palliat Med 2021;10:5944-53.
5. Pérol M, Arpin D. Angiogenesis and lung cancer. Bull Cancer 2007;94 Spec No:S220-31.
6. Zhang Y, Zhang Z. The history and advances in cancer immunotherapy: understanding the characteristics of tumor-infiltrating immune cells and their therapeutic implications. Cell Mol Immunol 2020;17:807-21.
7. Ren X, Zhang L, Zhang Y, et al. Insights Gained from Single-Cell Analysis of Immune Cells in the Tumor Microenvironment. Annu Rev Immunol 2021;39:583-609.
8. Duma N, Santana-Davila R, Molina JR. Non-Small Cell Lung Cancer: Epidemiology, Screening, Diagnosis, and Treatment. Mayo Clin Proc 2019;94:1623-40.
9. Lurienne L, Cervesi J, Duhalde L, et al. NSCLC Immunotherapy Efficacy and Antibiotic Use: A Systematic Review and Meta-Analysis. J Thorac Oncol 2020;15:1147-59.
10. Alessi JV, Ricciuti B, Spurr LF, et al. SMARCA4 and Other SWItch/Sucrose NonFermentable Family Genomic Alterations in NSCLC: Clinicopathologic Characteristics and Outcomes to Immune Checkpoint Inhibition. J Thorac Oncol 2021;16:1176-87.
11. Doroshow DB, Sanmamed MF, Hastings K, et al. Immunotherapy in Non-Small Cell Lung Cancer: Facts and Hopes. Clin Cancer Res 2019;25:4592-602.
12. Iams WT, Porter J, Horn L. Immunotherapeutic approaches for small-cell lung cancer. Nat Rev Clin Oncol 2020;17:300-12.
13. Li W, Yu H. Separating or combining immune checkpoint inhibitors (ICIs) and radiotherapy in the treatment of NSCLC brain metastases. J Cancer Res Clin Oncol 2020;146:137-52.
14. Qian BZ, Pollard JW. Macrophage diversity enhances tumor progression and metastasis. Cell 2010;141:39-51.
15. Chen D, Zhang X, Li Z, et al. Metabolic regulatory crosstalk between tumor microenvironment and tumor-associated macrophages. Theranostics 2021;11:1016-30.
16. Pan Y, Yu Y, Wang X, et al. Tumor-Associated Macrophages in Tumor Immunity. Front Immunol 2020;11:583084.
17. Wang H, Tian T, Zhang J. Tumor-Associated Macrophages (TAMs) in Colorectal Cancer (CRC): From Mechanism to Therapy and Prognosis. Int J Mol Sci 2021;22:8470.
18. Yang L, Zhang Y. Tumor-associated macrophages: from basic research to clinical application. J Hematol Oncol 2017;10:58.
19. Wu F, Fan J, He Y, et al. Single-cell profiling of tumor heterogeneity and the microenvironment in advanced non-small cell lung cancer. Nat Commun 2021;12:2540.
20. Yip SH, Sham PC, Wang J. Evaluation of tools for highly variable gene discovery from single-cell RNA-seq data. Brief Bioinform 2019;20:1583-9.
21. Mangiola S, Doyle MA, Papenfuss AT. Interfacing Seurat with the R tidy universe. Bioinformatics 2021. [Epub ahead of print]. doi: 10.1093/bioinformatics/btab404.
22. Pereira WJ, Almeida FM, Conde D, et al. Asc-Seurat: analytical single-cell Seurat-based web application. BMC

Bioinformatics 2021;22:556.
23. Marwitz S, Depner S, Dvornikov D, et al. Downregulation of the TGFβ Pseudoreceptor BAMBI in Non-Small Cell Lung Cancer Enhances TGFβ Signaling and Invasion. Cancer Res 2016;76:3785-801.
24. Brouwer-Visser J, Cheng WY, Bauer-Mehren A, et al. Regulatory T-cell Genes Drive Altered Immune Microenvironment in Adult Solid Cancers and Allow for Immune Contextual Patient Subtyping. Cancer Epidemiol Biomarkers Prev 2018;27:103-12.
25. Huang Q, Liu Y, Du Y, et al. Evaluation of Cell Type Annotation R Packages on Single-cell RNA-seq Data. Genomics Proteomics Bioinformatics 2021;19:267-81.
26. Hillje R, Pelicci PG, Luzi L. Cerebro: interactive visualization of scRNA-seq data. Bioinformatics 2020;36:2311-3.
27. Geeleher P, Cox N, Huang RS. pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels. PLoS One 2014;9:e107468.
28. Liu Z, Sun D, Wang C. Evaluation of cell-cell interaction methods by integrating single-cell RNA sequencing data with spatial information. Genome Biol 2022;23:218.
29. Binnewies M, Pollack JL, Rudolph J, et al. Targeting TREM2 on tumor-associated macrophages enhances immunotherapy. Cell Rep 2021;37:109844.
30. Dai X, Lu L, Deng S, et al. USP7 targeting modulates anti-tumor immune response by reprogramming Tumor-associated Macrophages in Lung Cancer. Theranostics 2020;10:9332-47.
31. Xiang X, Wang J, Lu D, et al. Targeting tumor-associated macrophages to synergize tumor immunotherapy. Signal Transduct Target Ther 2021;6:75.
32. Naylor EC, Desani JK, Chung PK. Targeted Therapy and Immunotherapy for Lung Cancer. Surg Oncol Clin N Am 2016;25:601-9.
33. Stankovic B, Bjørhovde HAK, Skarshaug R, et al. Immune Cell Composition in Human Non-small Cell Lung Cancer. Front Immunol 2019;9:3101.
34. Li Y, Jiang T, Zhou W, et al. Pan-cancer characterization of immune-related lncRNAs identifies potential oncogenic biomarkers. Nat Commun 2020;11:1000.
35. Wu Y, Zhang L, He S, et al. Identification of immune-related LncRNA for predicting prognosis and immunotherapeutic response in bladder cancer. Aging (Albany NY) 2020;12:23306-25.
36. Zhang J, Tian Q, Zhang M, et al. Immune-related biomarkers in triple-negative breast cancer. Breast Cancer 2021;28:792-805.
37. Jing Y, Liu J, Ye Y, et al. Multi-omics prediction of immune-related adverse events during checkpoint immunotherapy. Nat Commun 2020;11:4946.
38. Huang C, Yang X. Advances in Biomarkers for Immunotherapy of Non-small Cell Lung Cancer. Zhongguo Fei Ai Za Zhi 2021;24:777-83.
39. Olingy CE, Dinh HQ, Hedrick CC. Monocyte heterogeneity and functions in cancer. J Leukoc Biol 2019;106:309-22.
40. Gren ST, Grip O. Role of Monocytes and Intestinal Macrophages in Crohn's Disease and Ulcerative Colitis. Inflamm Bowel Dis 2016;22:1992-8.
41. Mitchell AJ, Roediger B, Weninger W. Monocyte homeostasis and the plasticity of inflammatory monocytes. Cell Immunol 2014;291:22-31.
42. Påhlson C, Lu X, Ott M, et al. Characteristics of in vitro infection of human monocytes, by Rickettsia helvetica. Microbes Infect 2021;23:104776.
43. Yona S, Jung S. Monocytes: subsets, origins, fates and functions. Curr Opin Hematol 2010;17:53-9.
44. Hume DA, Irvine KM, Pridans C. The Mononuclear Phagocyte System: The Relationship between Monocytes and Macrophages. Trends Immunol 2019;40:98-112.
45. Mantovani A, Sica A, Sozzani S, et al. The chemokine system in diverse forms of macrophage activation and polarization. Trends Immunol 2004;25:677-86.
46. Arora S, Singh P, Ahmad S, et al. Comprehensive Integrative Analysis Reveals the Association of KLF4 with Macrophage Infiltration and Polarization in Lung Cancer Microenvironment. Cells 2021;10:2091.
47. Larionova I, Tuguzbaeva G, Ponomaryova A, et al. Tumor-Associated Macrophages in Human Breast, Colorectal, Lung, Ovarian and Prostate Cancers. Front Oncol 2020;10:566511.
48. Xu F, Cui WQ, Wei Y, et al. Astragaloside IV inhibits lung cancer progression and metastasis by modulating macrophage polarization through AMPK signaling. J Exp Clin Cancer Res 2018;37:207.
49. Zhou Q, Liang J, Yang T, et al. Carfilzomib modulates tumor microenvironment to potentiate immune checkpoint therapy for cancer. EMBO Mol Med 2022;14:e14502.
50. Funes SC, Rios M, Escobar-Vera J, et al. Implications of macrophage polarization in autoimmunity. Immunology 2018;154:186-95.
51. Müller E, Christopoulos PF, Halder S, et al. Toll-Like Receptor Ligands and Interferon-γ Synergize for Induction of Antitumor M1 Macrophages. Front Immunol 2017;8:1383.

52. Tiwari RK, Singh S, Gupta CL, et al. Repolarization of glioblastoma macrophage cells using non-agonistic Dectin-1 ligand encapsulating TLR-9 agonist: plausible role in regenerative medicine against brain tumor. Int J Neurosci 2021;131:591-8.
53. Ubil E, Caskey L, Holtzhausen A, et al. Tumor-secreted Pros1 inhibits macrophage M1 polarization to reduce antitumor immune response. J Clin Invest 2018;128:2356-69.
54. Calandra T, Bucala R. Macrophage Migration Inhibitory Factor (MIF): A Glucocorticoid Counter-Regulator within the Immune System. Crit Rev Immunol 2017;37:359-70.
55. Calandra T, Roger T. Macrophage migration inhibitory factor: a regulator of innate immunity. Nat Rev Immunol 2003;3:791-800.
56. Tilstam PV, Schulte W, Holowka T, et al. MIF but not MIF-2 recruits inflammatory macrophages in an experimental polymicrobial sepsis model. J Clin Invest 2021;131:e127171.

Cite this article as: Xie S, Huang G, Qian W, Wang X, Zhang H, Li Z, Liu Y, Wang Y, Yu H. Integrated analysis reveals the microenvironment of non-small cell lung cancer and a macrophage-related prognostic model. Transl Lung Cancer Res 2023;12(2):277-294. doi: 10.21037/tlcr-22-866

Supplementary

Table S1 Primers for quantitative real-time PCR

Genes	Forward primers (5'-3')	Reverse primers (5'-3')
SFTPA2	ACTTGGAGGCAGAGACCCAA	GGGCTTCCAACACAAACGTC
SPP1	CTCCATTGACTCGAACGACTC	CAGGTCTGCGAAACTTCTTAGAT
SFTPB	TGGAGCAAGCATTGCAGTG	ACTCTTGGCATAGGTCATCGG
KRT19	AACGGCGAGCTAGAGGTGA	GGATGGTCGTGTAGTAGTGGC
EMP2	GTGCTTCTTGCTTTCATCATCG	TGCAATTCGTGTTGTTGGTACA
CCL20	CCAAGAGTTTGCTCCTGGCT	GGATTTGCGCACACAGACAA

PCR, polymerase chain reaction.

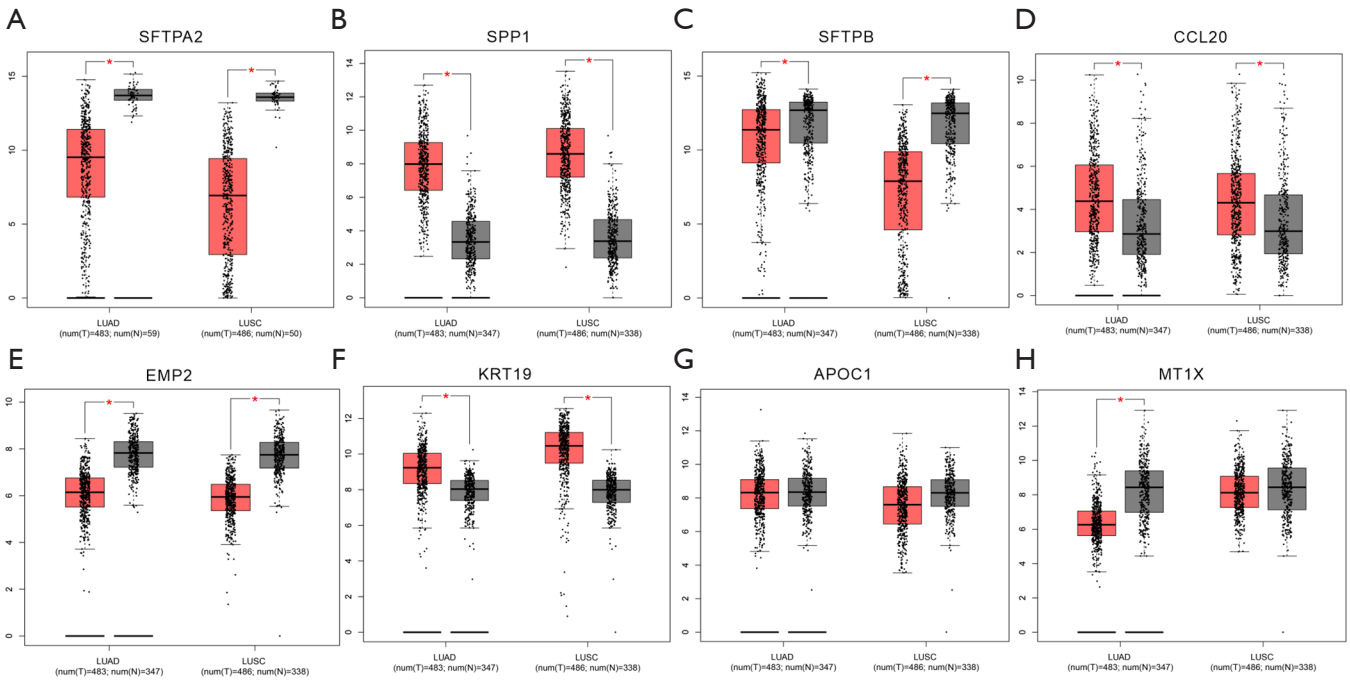

Figure S1 Analysis of GEPIA database data. (A-H) The GEPIA website was used to verify these 8 genes. The Y-axis represents gene expression. *, P<0.05. LUSC, lung squamous cell carcinoma; LUAD, lung squamous cell carcinoma; GEPIA, gene expression profiling interactive analysis.

https://dx.doi.org/10.21037/tlcr-22-866

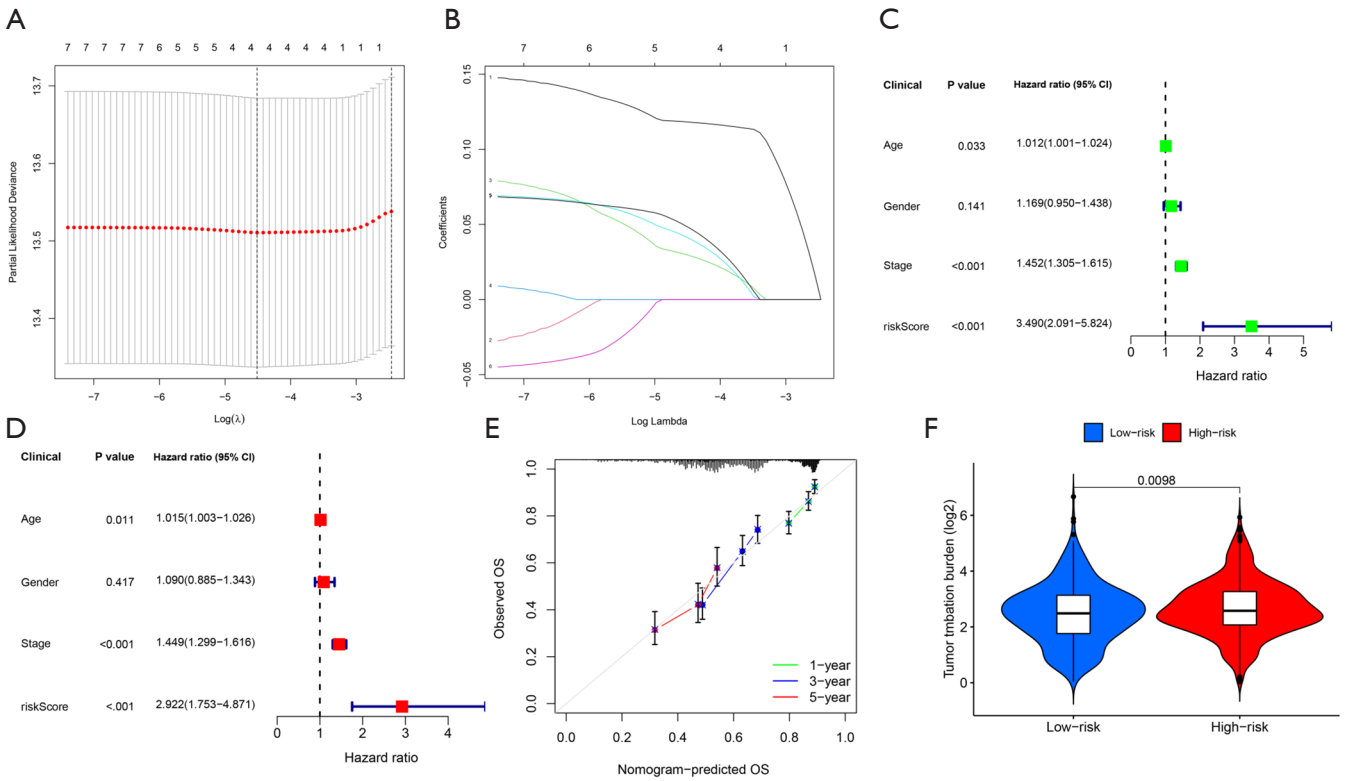

Figure S2 Additional information. (A) Cross-validation for tuning parameter (lambda, screening in the LASSO regression model. (B) LASSO coefficient profiles of 7 prognostic MRGs. (C,D) Univariate and multivariate Cox regression analyses. (E) The calibration curve of the 1, 3, and 5-year survival. (F) TMB of NSCLC patients with high/low risk. CI, confidence interval; LASSO, least absolute shrinkage and selection operator; NSCLC, non-small cell lung cancer; MRGs, metabolism-related genes; TMB, tumor mutation burden; OS, overall survival.

https://dx.doi.org/10.21037/tlcr-22-866