International Journal of General Medicine

Dovepress

open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Integrative Analysis of Single-Cell and Bulk RNA Sequencing Reveals Prognostic Characteristics of Macrophage Polarization-Related Genes in Lung Adenocarcinoma

Ke Mi, Lizhong Zeng, Yang Chen, Shuanying Yang

Department of Respiratory and Critical Care Medicine, the Second Affiliated Hospital of Xi'an Jiaotong University, Xi'an, Shaanxi, People's Republic of China

Correspondence: Shuanying Yang, Department of Respiratory Medicine, the Second Affiliated Hospital of Xi'an Jiaotong University, Xi'an, Shaanxi, People's Republic of China, Email yangshuanying66@163.com

Background: Lung adenocarcinoma (LUAD) is a group of cancers with poor prognosis. The combination of single-cell RNA sequencing (scRNA-seq) and bulk RNA sequencing (RNA-seq) can identify important genes involved in cancer development and progression from a broader perspective.

Methods: The scRNA-seq data and bulk RNA-seq data of LUAD were downloaded from the Gene Expression Omnibus (GEO) database and the Cancer Genome Atlas (TCGA) database. Analyzing scRNA-seq for core cells in the GSE131907 dataset, and the uniform manifold approximation and projection (UMAP) was used for dimensionality reduction and cluster identification. Macrophage polarization-associated subtypes were acquired from the TCGA-LUAD dataset after analysis, followed by further identification of differentially expressed genes (DEGs) in the TCGA-LUAD dataset (normal/LUAD tissue samples, two subtypes). Venn diagrams were utilized to visualize differentially expressed and highly variable macrophage polarization-related genes. Subsequently, a prognostic risk model for LUAD patients was constructed by univariate Cox and Least Absolute Shrinkage and Selection Operator (LASSO), and the model was investigated for stability in the external data GSE72094. After analyzing the correlation between the trait genes and significantly mutated genes, the immune infiltration between the high/low-risk groups was then examined. The Monocle package was applied to analyze the pseudo-temporal trajectory analysis of different cell clusters in macrophage clusters. Subsequently, cell clusters of data macrophages were selected as key cell clusters to explore the role of characteristic genes in different cell populations and to identify transcription factors (TFs) that affect signature genes. Finally, qPCR were employed to validate the expression levels of prognosis signature genes in LUAD.

Results: 424 macrophage highly variable genes, 3920 DEGs, and 9561 DEGs were obtained from macrophage clusters, the macrophage polarization-related subtypes, and normal/LUAD tissue samples, respectively. Twenty-eight differentially expressed and highly mutated MPRGs were obtained. A prognostic risk model with 7 DE-MPRGs (RGS13, ADRB2, DDIT4, MS4A2, ALDH2, CTSH, and PKM) was constructed. This prognostic model still has a good prediction effect in the GSE72094 dataset. ZNF536 and DNAH9 were mutated in the low-risk group, while COL11A1 was mutated in the high-risk group, and they were highly correlated with the characteristic genes. A total of 11 immune cells were significantly different in the high/low-risk groups. Five cell types were again identified in the macrophage cluster, and then NK cells: CD56hiCD62L+ differentiated earlier and were present mainly on 2 branches. While macrophages were present on 2 branches and differentiated later. It was found that the expression levels of BCLAF1 and MAX were higher in cluster 1, which might be the TFs affecting the expression of the characteristic genes. Moreover, qPCR confirmed that the expression of the prognosis genes was generally consistent with the results of the bioinformatic analysis.

Conclusion: Seven MPRGs (RGS13, ADRB2, DDIT4, MS4A2, ALDH2, CTSH, and PKM) were identified as prognostic genes for LUAD and revealed the mechanisms of MPRGs at the single-cell level.

Keywords: lung adenocarcinoma, macrophage polarization-related genes, prognostic risk model, single-cell RNA sequencing, regulation mechanism

Received: 13 July 2023
Accepted: 30 October 2023
Published: 3 November 2023

Introduction

Lung cancer is a global health concern, being the most diagnosed cancer and the leading cause of cancer-related deaths worldwide. Among the various histological subtypes of lung cancer, lung adenocarcinoma (LUAD) is the most frequent.[1,2] According to the revised classification of lung cancer by the World Health Organization in 2021, LUAD is categorized into adenomatous atypical hyperplasia (AAH), adenocarcinoma in situ (AIS), minimally invasive adenocarcinoma (MIA), and invasive adenocarcinoma (AI).[3] Research has identified smoking as the major risk factor for LUAD, followed by air pollution, occupational exposure, and genetic predisposition.[4–7] While existing therapies such as surgical resection, chemotherapy, radiotherapy, and immunotherapy have been used to treat advanced LUAD, their efficacy in improving the prognosis of patients is limited.[8] Therefore, identifying novel characteristic genes and developing prognostic models may contribute to understanding the pathogenesis and molecular mechanisms of LUAD, as well as provide potential targets for clinical diagnosis and treatment.

Infiltrating immune and stromal cells are essential components for the function of tumor microenvironment (TME). Increasing data suggest that they have significant roles in tumor growth, recurrence, metastasis, and chemotherapy resistance. The significant role of infiltrating immune cells provides a novelty insight for revealing the tumor biology and will be an essential prognostic factor. Tumor-associated macrophages (TAMs) are a type of macrophage that infiltrates tumor tissues and constitute the major component of infiltrating leukocytes in the TME.[9,10] LUAD exhibits inter-patient and intra-tumoral heterogeneity in both tumor cells and TME.[11] The TME components vary markedly among different tumors and play crucial roles in tumor initiation, progression, and metastasis.[12,13] TAMs can be polarized into three main phenotypes: resting-state macrophages M0 (precursors of polarized macrophages), classically activated macrophages M1 (induced by Toll-like receptor and IFN-γ) and alternatively activated macrophages M2 (induced by IL-4 and IL-13).[14–16] In tumor progression, M1 macrophages typically play anti-tumor roles, while M2 macrophages exert pro-tumor functions.[17] M1 macrophages kill tumor cells through direct-mediated cytotoxicity and antibody-dependent cell-mediated cytotoxicity (ADCC).[18] In contrast, M2 macrophages can secrete tumor-stimulating factors, inhibit T cell-mediated immune response, and promote proliferation, invasion and angiogenesis.[19–22] Previous studies have shown that macrophages are involved in the development of lung adenocarcinoma (LUAD). For example, M2 macrophages enhance the expression of vascular endothelial growth factor (VEGF) family members A and C, which promote tumor angiogenesis and lymphangiogenesis. Additionally, M2 macrophages can promote LUAD cell epithelial-mesenchymal transition (EMT) by activating the ERK1/2/Fra-1/slug signaling pathway.[23,24] Zheng et al found that the lower density of M1 at the tumor center and higher proximity of tumor cells to M2 at the invasive margin predicted poor prognosis in lung cancer. It is worth mentioning that macrophages are highly plastic and heterogeneous.[16] Affected by the microenvironment changes, TAMs can be converted from one polarized phenotype to another. Inducing the phenotypic switch from M2 to M1 macrophages has been regarded as a novel therapeutic idea for malignant tumors. The study of macrophage polarization-related genes (MPRGs) helps to understand the underlying mechanism of macrophage polarization and shed new light on polarization control.

Integrating single-cell RNA sequencing (scRNA-seq) and bulk RNA sequencing has significant potential in identifying biomarkers. scRNA-seq provides a more detailed and accurate understanding of the heterogeneity of cell populations by analyzing gene expression at the single-cell level.[25,26] In contrast, bulk RNA sequencing is more cost-effective, has higher throughput, and provides greater coverage.[27,28] By integrating scRNA-seq data to identify specific cell types or subpopulations with bulk RNA sequencing data to identify differentially expressed genes or pathways, more accurate and reliable biomarkers can be identified.[29–32] These biomarkers can be used for diagnosis, prognosis, and personalized treatment of diseases. Therefore, integrating scRNA-seq and bulk RNA sequencing can provide a more comprehensive understanding of the molecular mechanisms underlying disease and improve patient outcomes.

In this study, the researchers utilized a combination of bulk RNA sequencing and single-cell RNA sequencing to investigate the role of macrophage polarization-related genes in LUAD. By integrating these techniques with bioinformatics methods, the team was able to identify novel biomarkers for malignant tumors and construct a new prognostic model for LUAD. This approach enabled a more comprehensive understanding of the heterogeneity of a particular cell

subset in the tumor microenvironment and provided valuable insights into the molecular mechanisms underlying LUAD progression.

Materials and Methods

Data Sources

Bulk RNA-seq data and clinical information of 59 normal tissue samples and 513 LUAD tissue samples with survival information in the TCGA-LUAD dataset were downloaded from the TCGA database (https://portal.gdc.cancer.gov/). The scRNA-seq data of 11 LUAD tissue samples in the GSE131907 dataset, the bulk RNA-seq data, and survival information of 398 LUAD tissue samples in the GSE72094 dataset were downloaded from the GEO database (https://www.ncbi.nlm.nih.gov/). Referring to the reported studies,[33] 35 macrophage polarization genes (MPGs) were obtained (Supplementary Table 1). The flow chart of this research was shown in Figure 1.

Analyzing of scRNA-Seq Data of LUAD

The "Seurat" package[34] was utilized to obtain the core cells, and several quality criteria were applied, including the exclusion of genes detected in only 3 or fewer cells, low-quality cells with less than 200 detected genes, and cells with mitochondrial expression genes ≥5%. The top 2000 highly variable genes were identified using the FindVariableFeatures and NormalizeData functions, and Principal Component Analysis (PCA) was performed on all samples, with the top 50 principal components selected for further analysis. UMAP algorithm[35] was used to perform overall dimensionality reduction on the samples based on the top 50 principal components. Marker genes were identified using the FindAllMarkers function in the Seurat package, and the 'SingleR' package[36] was used to annotate and visualize different cell types and extract high-margin genes in macrophage clusters. Furthermore, cellular communication between different cell types was analyzed using the 'cellchat' package.[37] These rigorous computational analyses helped to identify and characterize different cell types, extract key marker genes, and provide insights into cellular communication and molecular mechanisms underlying LUAD progression.[37]

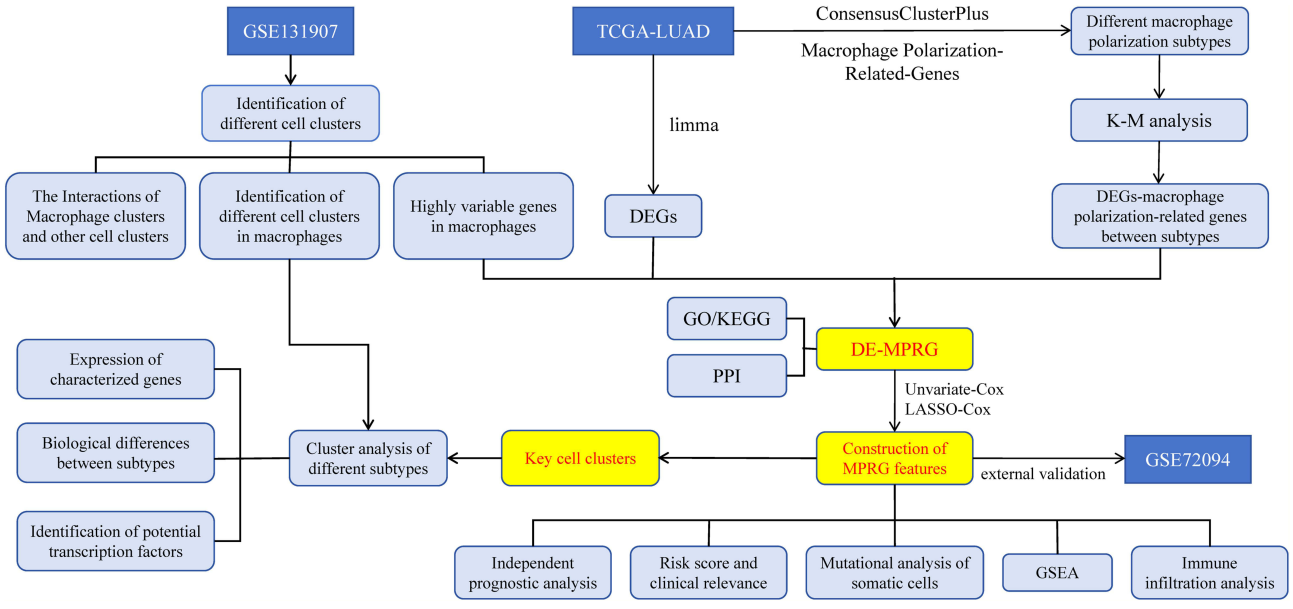


Figure 1 The flow chart of the research. Abbreviations are defined as follows.
Abbreviations: DEGs, Differential expressed genes; K-M, Kaplan-Meier; DE-MPRGs, Differentially expressed macrophage polarization-related genes; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; PPI, Protein-protein interaction (PPI) network; LASSO, Least absolute shrinkage and the selection operator.

Acquisition of Macrophage Polarization-Related Subtypes in the TCGA-LUAD Dataset

The R package "ConsensusClusterPlus"[38] was employed to identify macrophage polarization-related subtypes in the TCGA-LUAD dataset based on the expression of 35 MPGs. The overall survival (OS) among different subtypes was explored using the "Survival" package. Additionally, differential expression analysis was performed using "DESeq2"[39] to obtain the differentially expressed genes (DEGs) between the two subtypes, which were defined as MPRGs. The screening criteria for MPRGs were adj.P-value <0.05 and |log2FoldChange|>0.5. Similarly, the DEGs between normal and LUAD tissue samples in the TCGA-LUAD dataset were identified using the DESeq2 package. Volcano plots and heatmaps were generated using the R packages "ggplot2" and "pheatmap", respectively, to visualize DEGs and MPRGs.

Identification and Biological Functional Analysis of Differentially Expressed and Highly Mutated MPRGs in TCGA-LUAD

The DEGs, MPRGs, and hypervariable genes of macrophages were intersected to identify the differentially expressed and highly mutated MPRGs. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis of these intersecting genes were performed using the "clusterProfiler" package,[38] and the results were visualized through bubble plots created using the "ggplot2" package in R. Additionally, the protein-protein interaction (PPI) network of the intersecting genes was constructed using the STRING database (https://string-db.org).

Construction and Validation of the Prognostic Risk Model

The intersection genes were subjected to univariate Cox analysis in the TCGA-LUAD dataset to identify prognosis-related genes with a significance level of $P < 0.05$. Then, using the least absolute shrinkage and selection operator (LASSO)[40] analysis, the characteristic genes were identified and sequentially eliminated. The resulting genes were used to construct a prognostic risk model, and the LUAD patients were divided into two groups based on the median risk score: high-risk and low-risk groups. The Kaplan-Meier (KM) curves were used to display the difference in overall survival (OS) between the two groups. To assess the prognostic capability of the model, the ROC curves were displayed using the "survROC" package.[41] Finally, the external GSE72094 dataset was used to validate the stability of the model. The "rms" package was used to construct a nomogram that predicts survival probability based on the characteristic genes. The calibration curve was used to validate whether the nomogram can be used as an optimal model for clinical decision-making.

Assessment of the Prognostic Risk Model

The distribution and differences in clinical characteristics between the high and low-risk groups were analyzed using the Wilcoxon test after combining clinical characteristics such as age, sex, T-stage, N-stage, M-stage, and survival status with risk scores. To determine whether clinical characteristics and risk scores were independent predictive factors for LUAD patients, univariate and multifactorial Cox analyses were performed. The "rms" package was used to construct a nomogram that predicts survival probability based on independent prognostic criteria. The calibration curve and decision curve analysis (DCA) were both used to validate whether the nomogram can be used as an optimal model for clinical decision-making. The "maftools" package was utilized to analyze the mutation frequencies of genes between the two groups, while the Spearman correlation coefficient was used to analyze the correlation between characteristic genes and significantly mutated genes. Additionally, all genes in the high/low-risk groups were analyzed for correlation with characteristic genes, and the functional pathways of characteristic genes were analyzed using gene set enrichment analysis (GSEA).

Immune Infiltration Analysis in LUAD

The CIBERSORT algorithm[42] was used to assess the infiltration of immune cells in all samples between the high and low-risk groups. The differences in immune cell abundance between the two groups were examined using the Wilcoxon test. Spearman correlation analysis was performed to investigate the correlations between the 22 immune cell types.

Analysis of Characteristic Genes in Key Cell Clusters

To identify macrophage subtypes, we extracted the macrophage clusters in the GSE131907 dataset annotated in Analyzing of scRNA-Seq Data of LUAD and re-annotated them by dimensionality reduction clustering and subgroup annotation to subdivide the macrophage clusters. The resulting clusters were further analyzed using the Monocle package to perform trajectory analysis.[43] From these clusters, key cell clusters of macrophage types were selected for further analysis. The expression levels of characteristic genes were then analyzed among subtypes using consensus clustering and the Wilcoxon test. Additionally, GSEA was performed to understand the potential functions of the characteristic genes among subtypes.

Identification of Transcription Factors (TFs) Affecting the Expressions of Characteristic Genes

To explore the potential regulatory mechanisms affecting the characteristic genes in different subpopulations, we adopted the single-cell regulatory network inference and clustering (SCENIC) approach[44] to predict specific transcription factors (TFs). The Wilcoxon test was then utilized to analyze differences in the expression of TFs among subtypes.

Collection of Tissues and Analysis of Characteristic Genes Expression

In this study, we investigated the expression of 7-MPRGs in LUAD tissues. A total of 5 paired LUAD and adjacent normal lung tissue samples were collected from the Second Affiliated Hospital of Xi'an Jiaotong University. And this study was approved by the Ethics Committee of the Second Affiliated Hospital of Xi'an Jiaotong University. Written informed consent was obtained from all patients. Total RNA was extracted from tissue samples using the FAST1000 kit (Pioneer, China), and RNA reverse transcription was performed using the PrimeScript™ RT-PCR Kit (Takara, Japan). Quantitative real-time polymerase chain reaction (qRT-PCR) was performed using the TB Green® Premix Ex Taq™II (Takara, Japan) kit. All primers used in this study were synthesized by Sangon Biotech Co. Ltd. (Shanghai, China) (Supplementary Table 2).

Results

Integration and Clustering of scRNA-Seq Data

To screen for highly variable genes in LUAD scRNA-seq, we first performed quality control and screened a total of 29,747 core cells (Supplementary Figure 1). Following data normalization, 2000 highly variable genes were identified for subsequent analysis (Figure 2A). PCA analysis was performed on the single-cell samples, and the sample dispersion distribution was logical (Figure 2B). The top 50 principal components were selected for subsequent analysis (Figure 2C and D). The core cells were then divided into 23 independent cell clusters using the UMAP algorithm (Figure 2E), and the marker genes in each cluster are shown in Figure 2F. The "singleR" software package was used to annotate the different cell clusters, resulting in nine cell clusters (Figure 1G). Among them, the macrophage cluster had 424 highly variable genes (Supplementary Table 3). The number of ligand-receptor interactions is shown in Figure 1H, and macrophages could interact with other cells.

Identification of Macrophage Polarization-Related Subtypes

Based on the expression of 35 MPGs, we classified 513 LUAD patients into two subtypes (Figure 3A-C). It was observed that cluster2 had a worse prognosis than cluster1 (Figure 3D). This indicates that the degree of expression of MPGs had an impact on the prognosis of LUAD patients, adding rationality to the subsequent analysis. We found 3920 DEGs in the two subtypes, which were named macrophage polarization-related genes (MPRGs) (Figure 3E and F, Supplementary Table 4). In the TCGA-LUAD dataset, we observed a total of 9561 DEGs between normal/LUAD tissue samples (Figure 3G and H, Supplementary Table 5).

The Function Analysis of Differentially Expressed and Highly Mutated MPRGs

Firstly, considering that the degree of expression of MPGs had an impact on the prognosis of LUAD patients, we took the intersection of the two datasets of MPGs and DEGs. However, the intersection of MPGs and DEGs may not be expressed in macrophages. To ensure that the selected differential genes were expressed in macrophages, it was changed to take intersecting genes in DEGs, MPRGs, and highly variable genes, defined these 28 genes as differentially expressed

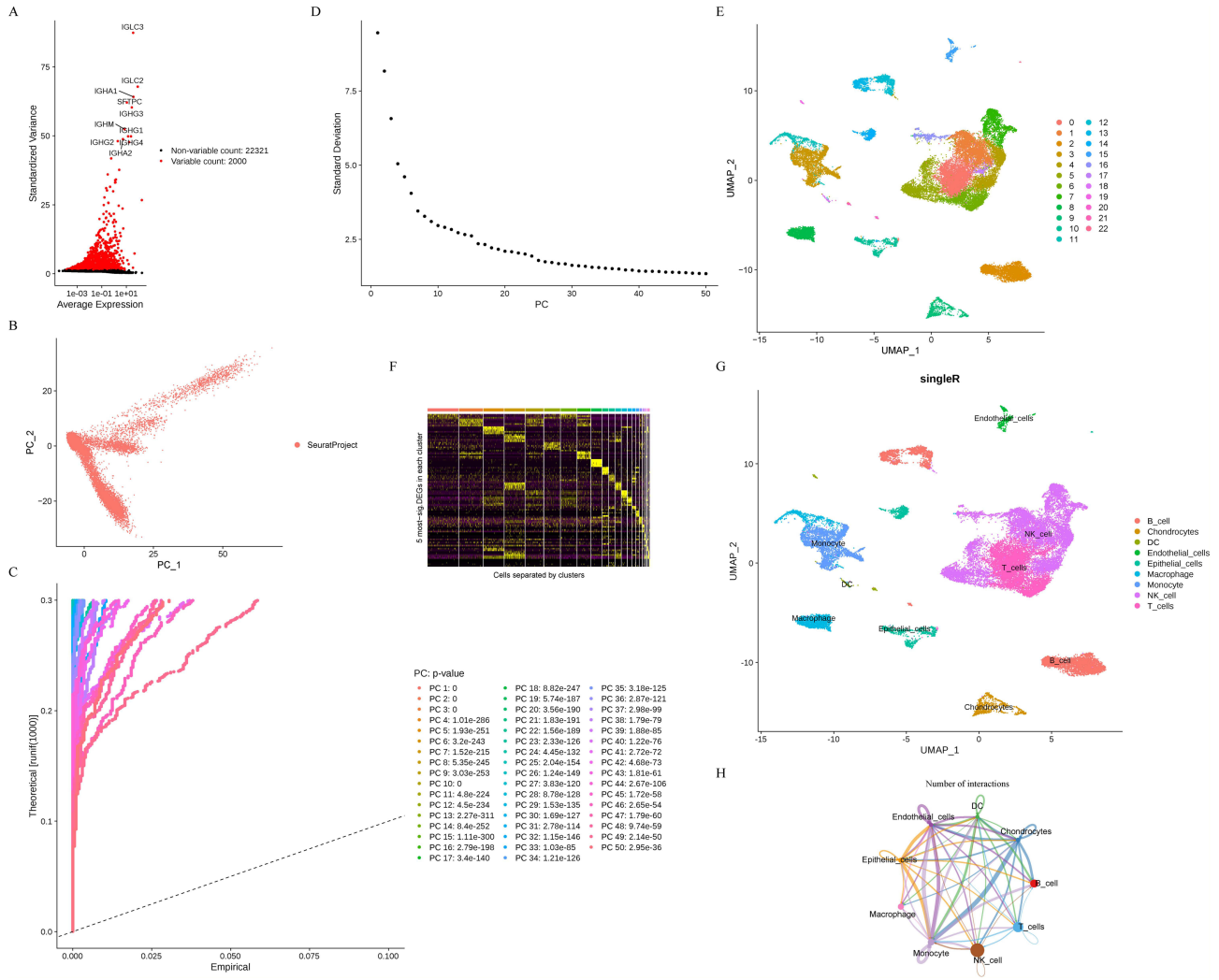


Figure 2 Highly variable gene screening and cell clustering of LUAD scRNA-seq data from GSE131907. (**A**) 2000 highly variable genes. (**B**) PCA results. (**C-D**) P-value of first 50 principal components (PCs). (**E**) Cluster diagram of the core cells. (**F**) Heatmap showing the top 5 marker genes in each cell cluster. (**G**) 23 cell clusters were annotated into 9 cell types. (**H**) CellChat diagram of each cell type.

MPRGs (DE-MPRG) (Figure 4A, Supplementary Table 6). These 28 genes are the genes whose expression in macrophages of LUAD tumor tissue is different from that in normal tissue. Biological process (BP) analysis revealed that these genes are involved in the activation of cysteine-type endopeptidase activity, apoptosis process, and leukocyte aggregation. In terms of cellular components (CC), these genes were associated with phagocytosis-related CC, including secretory granule lumen, cytoplasmic vesicle lumen, and vesicle lumen. Regarding molecular function (MF), these genes were enriched for immunoglobulin-binding functions (Figure 4B). The intersecting genes were also enriched in pathways such as IL-17 and Apoptosis (Figure 4C). The final PPI network of the intersecting genes containing 21 points with 26 interacting edges was obtained (Figure 4D). ALDH2 and PLB01 were found to only interact with each other in this network.

Construction and Validation of a Seven-Characteristic Gene-Based Prognostic Risk Model

In this study, univariate Cox regression analysis was performed using the TCGA-LUAD dataset, and 13 genes were found to be significantly associated with OS (Figure 5A). Next, a LASSO algorithm was used to screen for genes to construct a model (Figure 5B). The results showed that seven characteristic genes were screened at the lowest cross-

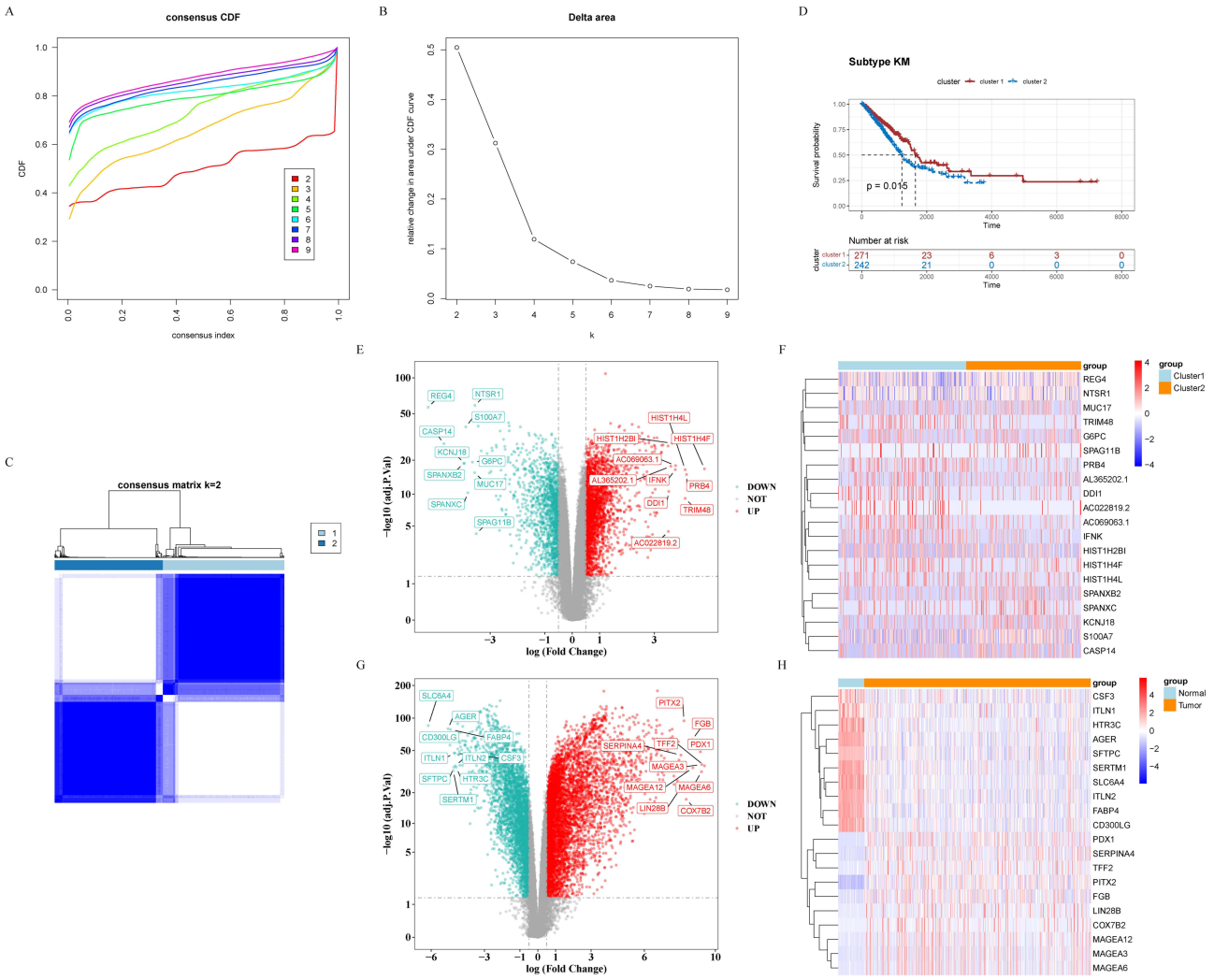


Figure 3 Identification of macrophage polarization-related subtypes of TCGA-LUAD samples. (**A**) Cumulative distribution function (CDF) curve of K=2-9. (**B**) The relative change in area under the CDF curve of K=2-9. (**C**) The consensus matrix heatmap of K=2. (**D**) Kaplan-Meier survival curves of the two subtypes. (**E-F**) Volcano plot and heat map representing DEGs (|log2FC|>0.5, adj.P-value<0.05) in the two subtypes. (**G-H**) Volcano plot and heat map of DEGs (|log2FC|>0.5, adj.P-value<0.05) between TCGA normal and LUAD samples.

validation error. The risk score was calculated using the following formula: Risk score = (−0.15120133 × RGS13) + (−0.10521700 × ADRB2) + (0.10214897 × DDIT4) + (−0.01788635 × MS4A2) + (−0.07556347 × ALDH2) + (−0.09135983 × CTSH) + (0.19599794 × PKM). Based on the median risk score (0.9332402), patients were divided into high-risk and low-risk groups. The K-M analysis revealed that the prognosis was better for low-risk patients, and the ROC curves indicated that the model was feasible (Figure 5C). Furthermore, the prognostic risk model demonstrated strong predictive power in the GSE72094 dataset (Figure 5D). To construct a reliable prognostic model, a nomogram was developed based on these 7 characteristic genes (Figure 5E). The nomogram indicated that the overall survival rate of patients decreased as the overall score increased. The calibration curve of the nomogram was close to the diagonal, indicating that the prediction of the model was accurate and reliable (Figure 4F).

Assessment of Prognostic Risk Models

Next, we will construct a nomogram model based on the seven characterized genes of LUAD. Risk scores differed between age, gender, T.stage, and N.stage (Supplementary Figure 2). Further analysis using univariate and multifactorial Cox regression revealed that the risk score, T.stage, N.stage, and M.stage were all independent prognostic factors for LUAD (Figure 6A and B). To construct a reliable prognostic model, a nomogram was developed based on these

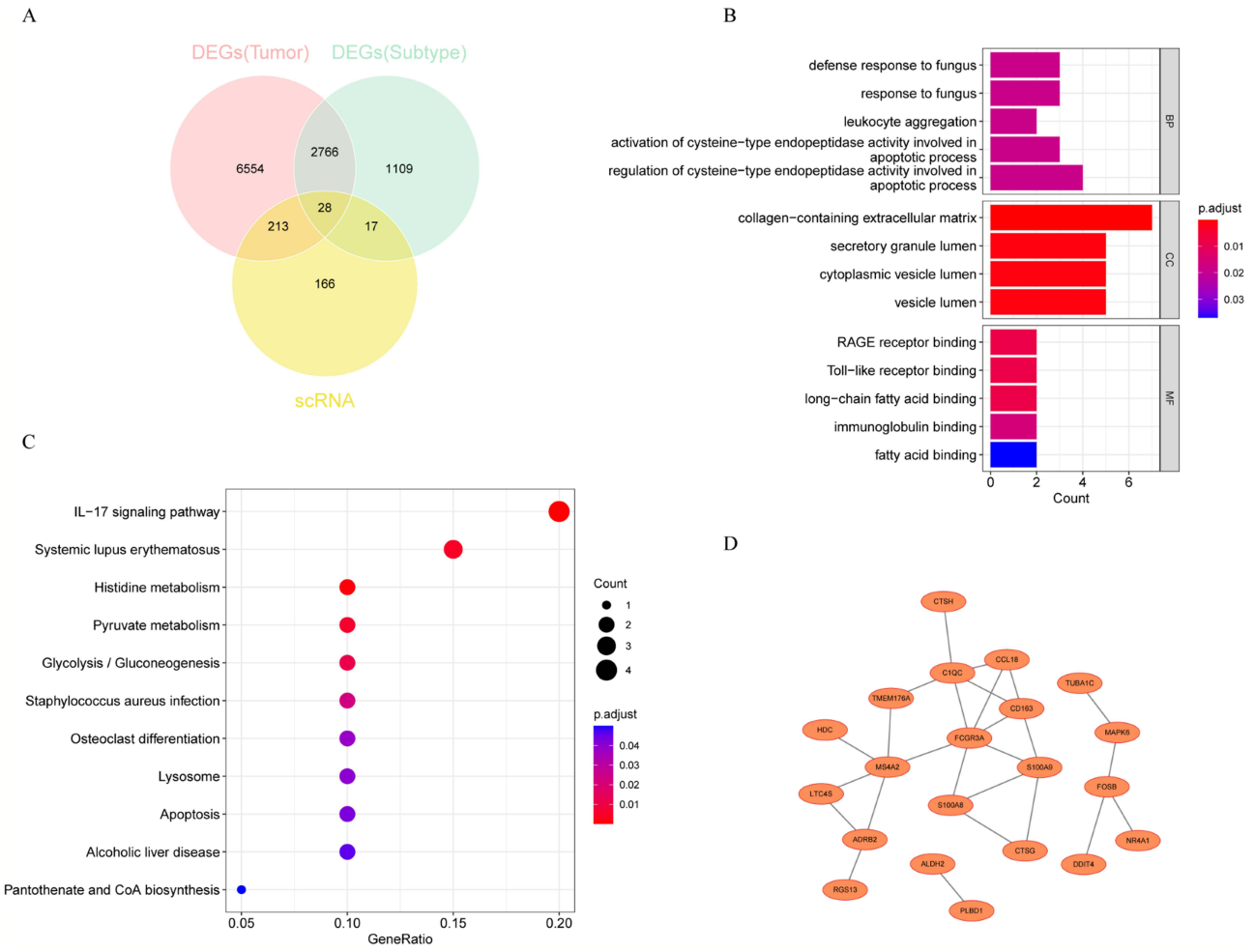


Figure 4 KEGG and GO analysis of differentially expressed and highly mutated MPRGs. (**A**) Venn diagram showing the 28 genes obtained by intersecting. (**B**) GO terms on BP, CC and MF levels enriched for the crossover genes. (**C**) KEGG pathways enriched for the 28 genes obtained. (**D**) PPI network of 26 intersecting genes.

independent prognostic factors (Figure 6C). The nomogram indicated that the overall survival rate of patients decreased as the overall score increased. The calibration curve of the nomogram was close to the diagonal, indicating that the prediction of the model was accurate and reliable (Figure 6D). Therefore, the nomogram was considered the optimal model for predicting the prognosis of LUAD patients (Figure 6E).

GSEA Analysis of the High/Low-Risk Groups

The top 5 mutated genes in the high/low-risk groups were TP53, TTN, CSMD3, MUC16, and RYR2 (Supplemental Figure 3). In addition, ZNF536 and DNAH9 were mutated in the low-risk group, while COL11A1 was mutated in the high-risk group, and these genes were highly correlated with the characteristic genes (Figure 7A). These findings suggested that these mutated genes may be involved in the regulation of characteristic genes, which in turn affects the prognosis of LUAD patients. Further gene set enrichment analysis (GSEA) showed that genes associated with characteristic genes were mainly enriched in the cell cycle pathway, which is closely related to tumor development and is an important cause of tumorigenesis and malignant growth due to impaired cycle regulation (Figure 7B).

Correlation Analysis of Risk Scores and Immune Microenvironment

Supplemental Figure 4A shows the results of analyzing the percentage of 22 immune cell components in the TCGA-LUAD dataset after extracting features using data from scRNA-seq. Differential immune cells were identified and 11 of them were significantly different between the high- and low-risk groups (Figure 7C). Furthermore, the Spearman analysis

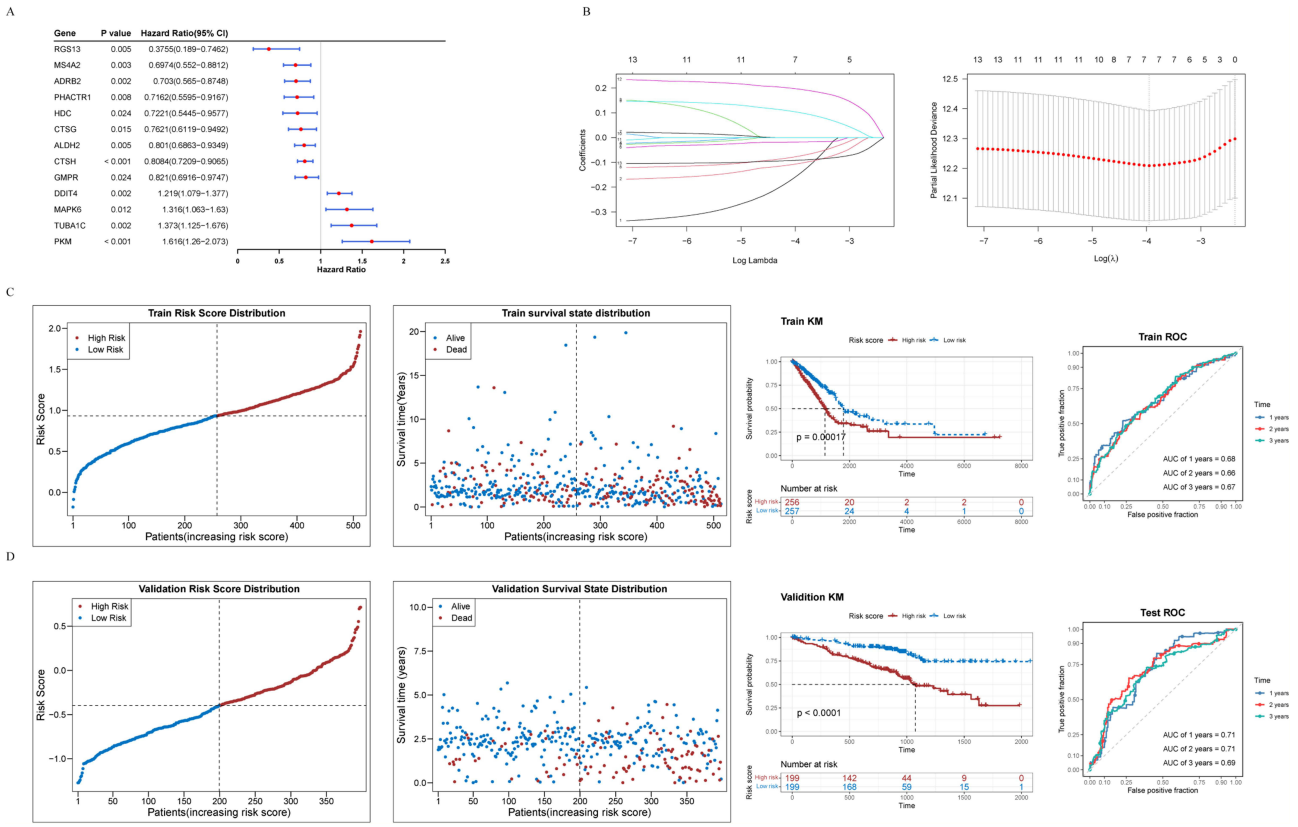


Figure 5 Construction and validation of differentially expressed and highly mutated MPRGs associated with LUAD. (**A**) Univariate Cox regression analysis of the TCGA-LUAD screened the differentially expressed and highly mutated MPRGs associated with LUAD. (**B**) 7 genes were further screened by LASSO regression analysis. The trajectory of each independent variable and the confidence interval under each lambda were showed. (**C**) Predictive value of 7 characteristic genes for LUAD in TCGA training set. (**D**) Predictive value of 7 characteristic genes for LUAD in GSE72094 validation set.

revealed the interactions between immune cells (Supplemental Figure 4B), and Monocyte, T cell CD8+, B cell plasma, and Mast cell activated were found to be highly correlated with other immune cells.

Role of Characteristic Genes in Macrophage Types Cell Clusters

To further subdivide macrophages, the macrophage clusters were further annotated and downscaled, resulting in the identification of five distinct classes of cell clusters (Figure 8A). Analysis of gene expression revealed that ALDH2, PKM, and CTSH were present in macrophages, while the remaining four characteristic genes were present in NK cells (Figure 8B). These findings suggested a potential association between macrophages and NK cells. The differentiation of macrophage clusters followed four directions (Figure 8C). Notably, NK cells: CD56hiCD62L+ differentiated earlier and were mainly present on wo branches, while macrophages were present on two branches and differentiated later (Figure 8D).

In examining the expression of genes in the various cell populations, three of the five cell populations annotated as subpopulations directly belonging to the macrophage type were selected as key cell populations for further follow-up analysis, given that our analysis was based on macrophages as a whole. The three cell clusters were Macrophage:Alveolar: B._anthacis_spore, Macrophage: Alveolar, Macrophage: monocyte-derived: M-CSF. Based on these three macrophage clusters, genes were extracted from the cells. The key cells were divided into two subpopulations based on the expression of characteristic genes (Figure 9A). Five characteristic genes, namely ADRB2, ALDH2, CTSH, DDIT4, and PKM, were found to be differentially expressed between the two subtypes. All of them were highly expressed in cluster2, except for PKM in cluster1 (Figure 9B). GSEA results between the two subpopulations suggested that these cells may be involved in the apoptotic signaling pathway and the ERBB signaling pathway, playing a role in LUAD (Figure 9C and D).

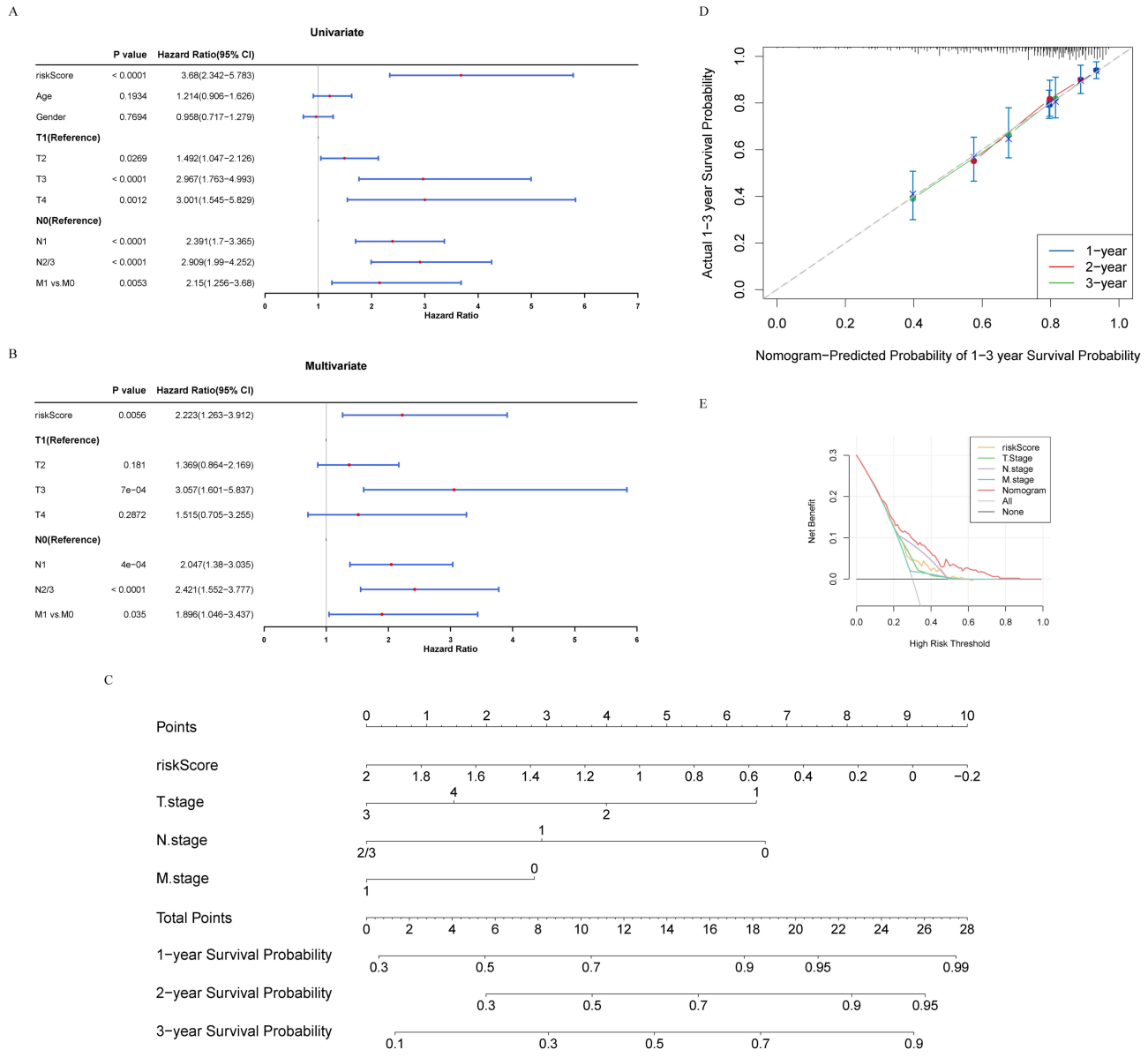


Figure 6 Construction of a nomogram model based on the 7 characteristic genes for LUAD. (**A-B**) Univariate and multifactorial Cox analysis of the TCGA-LUAD cohort. (**C**) Probabilistic nomogram for predicting 1-, 2-, and 3-year survival probability of patients with LUAD. (**D**) Calibration plot for conformance testing between 1–3 years survival predictions and actual outcomes. (**E**) Time-dependent ROC curve showing the risk score had significant prognostic value.

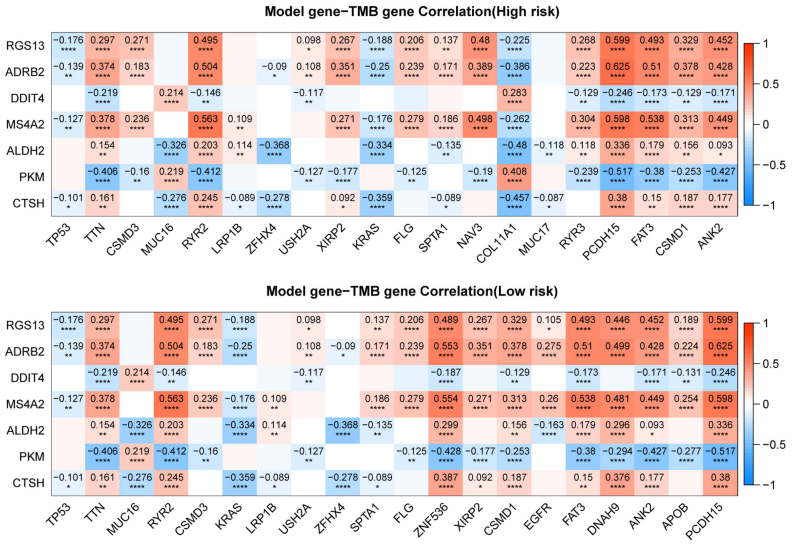
BCLAF1 and MAX May Be Transcription Factors (TFs) Affecting the Expression of Characteristic Genes

To identify transcription factors for characterized genes, the 10 transcription factors with the highest regulon specificity score (R-DSS) were shown as the subgroup-specific TFs in Figure 10A. To visualize the expression of the top 10 R-DSS TFs in different cell clusters for the two subgroups, UMAP plots were performed (Figure 10B). Among the top 10 R-DSS TFs, only BCLAF1 and MAX were found to be differentially expressed in the two subpopulations (Figure 10C). These results suggest that these two TFs may be associated with the expression of the characteristic genes.

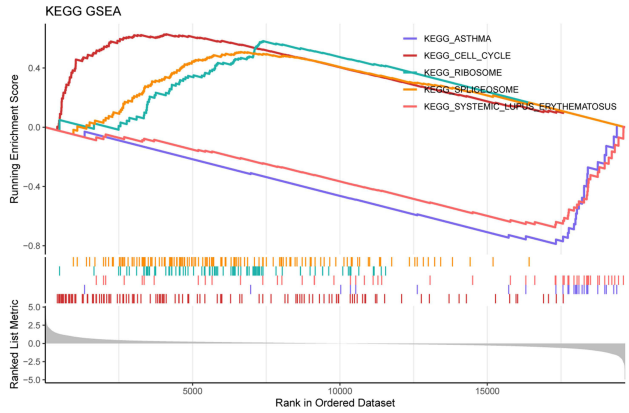
Expression of Seven Characteristic Genes in LUAD Tumor Tissues and TCGA-Database

We conducted a differential expression analysis on the data obtained from TCGA, which indicated that RGS13, ADRB2, MS4A2, ALDH2, and CTSH were significantly downregulated in LUAD tumor tissues, while DDIT4 and PKM were

A



B



C

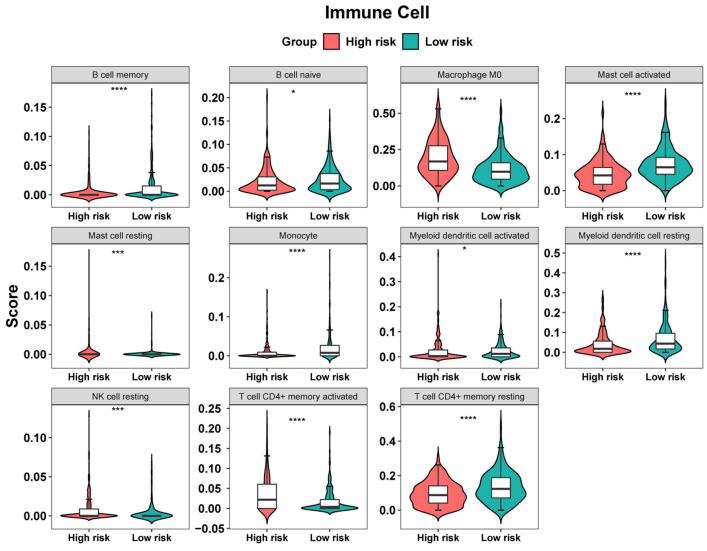


Figure 7 Mutated genes analysis and immune infiltration analysis in the high/low-risk groups. (**A**) Matrix heatmaps of the correlation analysis results. (**B**) GSEA enrichment analysis of high/low-risk groups. (**C**) Violin plots depicting differential analysis of immune cells in different risk groups. ****p<0.0001, ***p<0.001, *p<0.05.

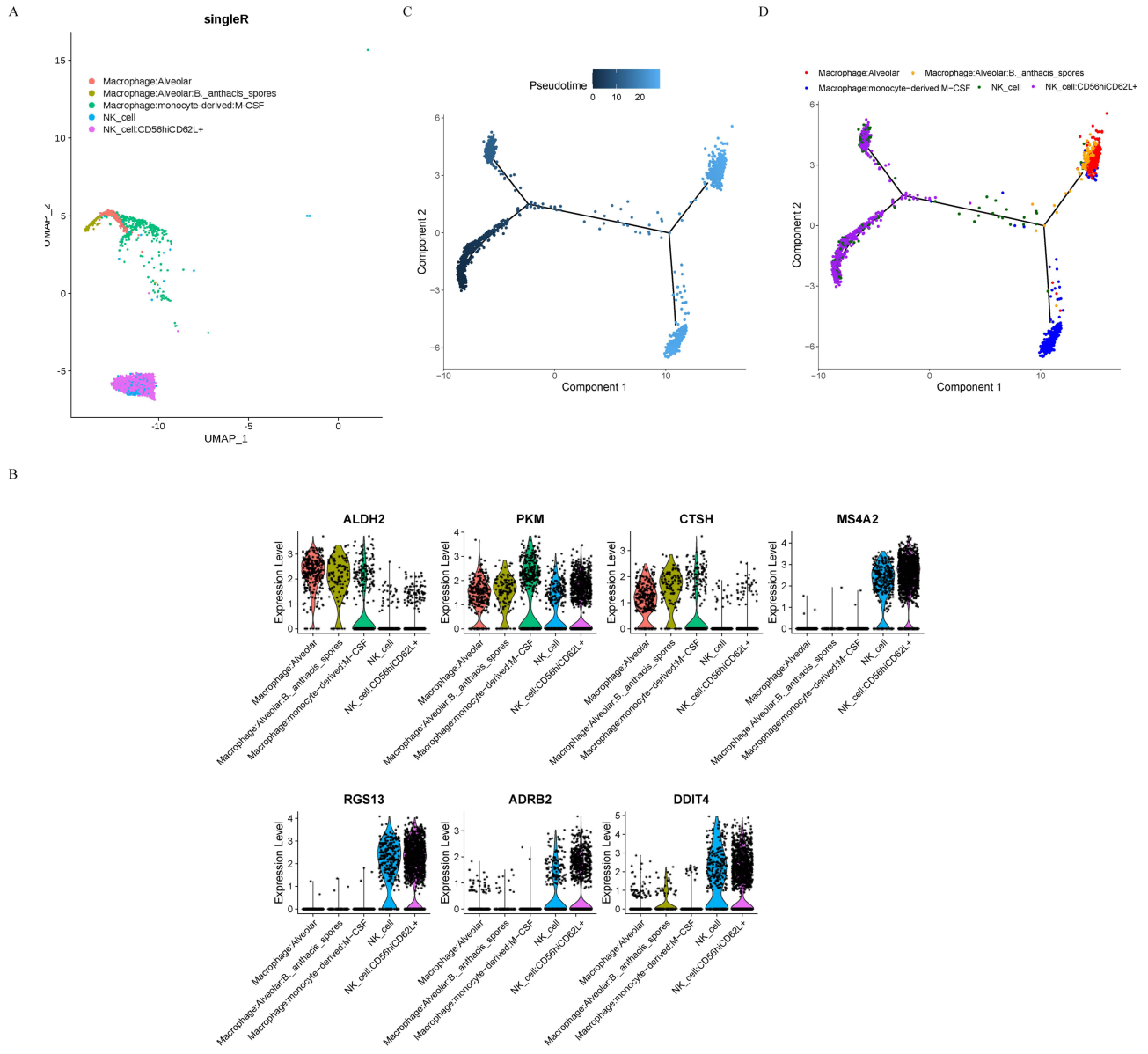


Figure 8 Reannotation and analysis of the macrophage clusters. (**A**) Five clusters were identified after further annotation. (**B**) Expression of characteristic genes in different cell clusters. (**C**) Pseudotime analysis of the macrophage clusters. (**D**) Cell annotation for the pseudotime analysis.

upregulated (Figure 11A). To validate our findings, we further analyzed the expression of the 7-MPRGs in 5 pairs of LUAD and adjacent non-tumor lung tissues using qRT-PCR. Our results showed that the expression levels of MS4A2 and CTSH were significantly decreased in the LUAD tumor tissues when compared to the adjacent non-tumor lung tissues (Figure 11B).

Discussion

In recent years, there have been significant improvements in the diagnosis and treatment of LUAD. However, the prognosis for many patients is still poor, especially those diagnosed at an advanced stage of the disease.[45] It is now widely recognized that the tumor microenvironment plays a critical role in tumor progression, and the interactions between immune cells and cancer cells are key factors in this process.[46,47]

In this study, we conducted a comprehensive analysis of gene expression profiles associated with macrophage polarization in LUAD tissues. By identifying and analyzing a panel of macrophage polarization-related genes, we were able to develop a risk score model that can be used to predict individual LUAD patient prognosis with high

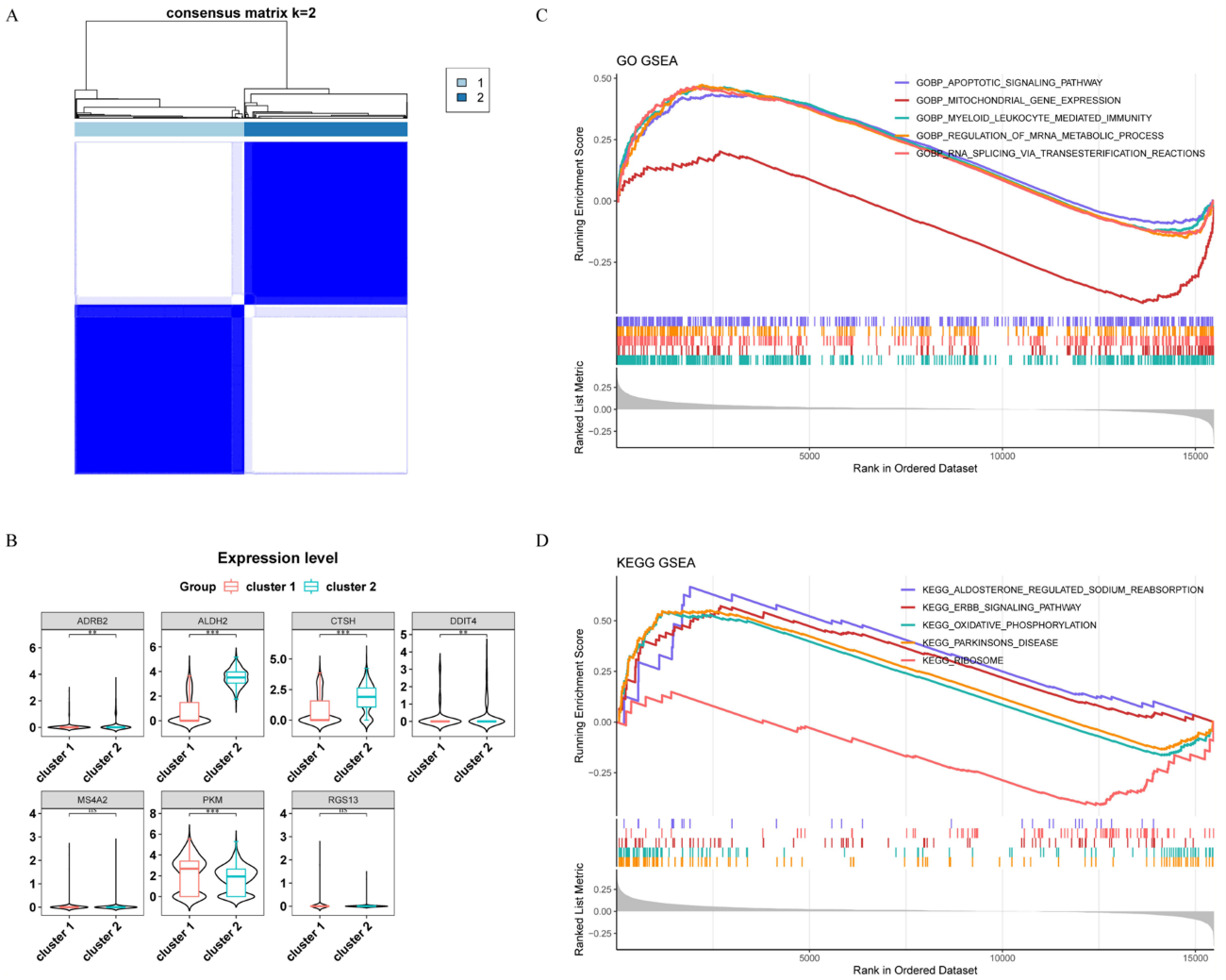


Figure 9 Expression and function of characteristic genes in macrophage types cell clusters. (**A**) Key cells from three macrophage-type-cell clusters were divided into two subpopulations. (**B**) Expression of characteristic genes in different cell subpopulations. (**C-D**) GSEA enrichment analysis of two subpopulations. ***p<0.001, **p<0.01, NS≥0.05.

accuracy. This study provides valuable insights into the underlying mechanisms of LUAD progression and highlights the importance of macrophage polarization in tumor microenvironment. It also underscores the potential of using prognostic gene expression profiles to develop personalized treatment strategies for LUAD patients.

In recent years, several studies have highlighted the importance of post-transcriptional modifications such as polyadenylation and methylation alterations in the development of lung cancer, including LUAD. Differentially expressed genes, miRNAs, methylation sites, and APA-related genes associated with NSCLC (LUAD and LUSC) were identified in large datasets for the prediction of metastasis in cancer tissues.[48,49] Based on cell cluster analysis and microarray data, seven key genes associated with LUAD prognosis were identified in a July paper.[50] Considering the effect of macrophages, our study identified seven MPRGs using scRNA-seq and bulk RNA-seq data (RGS13, ADRB2, DDIT4, MS4A2, ALDH2, CTSH and PKM) as characterized genes and modeled the prognostic risk of LUAD. This comprehensive model, which includes both clinical and gene factors, can provide individualized assessment of LUAD prognosis and has better prediction ability for overall survival (OS) than single factors. While RGS13 has been associated with immune-related diseases, such as asthma and B lymphoma, its role in solid tumors and LUAD remains unknown and needs further validation.[51–54] ADRB2, on the other hand, has been shown to promote tumorigenesis and development in several malignancies, but its expression was significantly decreased in LUAD, and low expression was associated with shorter OS and disease-specific survival.[55–59] DDIT4 was confirmed to be an independent predictor of OS for LUAD

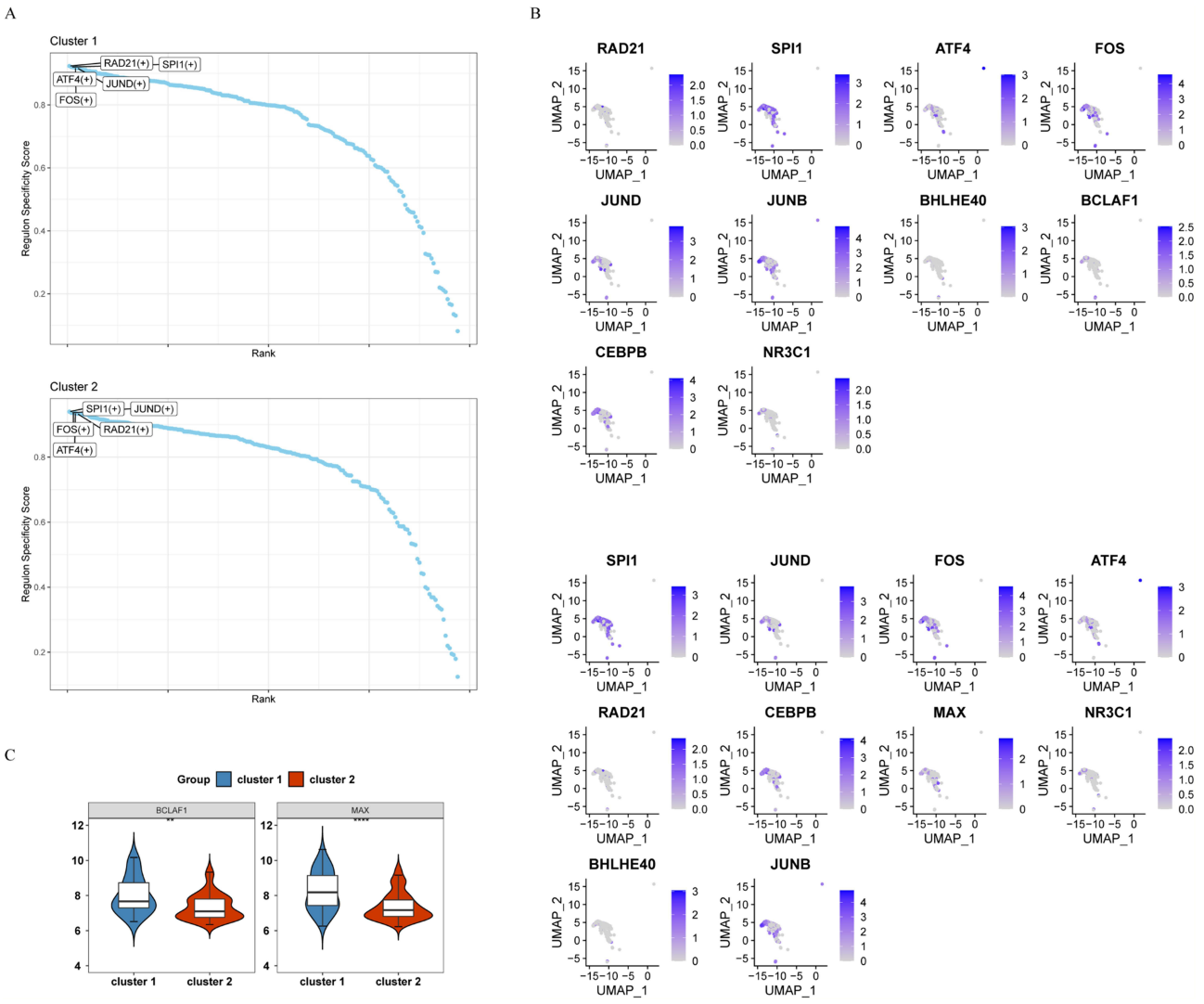


Figure 10 Identification of transcription factors of characteristic genes. (**A**) The top 10 TFs with the highest regulon specificity score in each subgroup. (**B**) Expression of the top 10 TFs in each cell cluster. (**C**) Expression of BCLAF1 and MAX in two subgroups. ****p<0.0001, **p<0.01.

patients and was shown to enhance the migration and invasion abilities of LUAD cells by activating the mTORC2/AKT pathway.[60,61] MS4A2, a favorable prognostic indicator, was identified as an independent prognostic marker in early-stage LUAD patients.[62] ALDH2, a major enzyme for detoxification of endogenous acetaldehyde, was shown to inhibit malignant features of LUAD cells when overexpressed.[63–65] CTSH, a lysosomal cysteine protease, had a controversial role in tumors, with high levels observed in gliomas and melanoma but lower levels associated with better prognosis in colorectal cancer and thyroid carcinoma.[66–69] Interestingly, its levels were decreased in tumor tissues but elevated in sera of lung cancer patients.[70] PKM, a well-known promoter of LUAD progression, was found to be consistent with our risk model.[71–74] In conclusion, the role played by the RGS13 and CTSH genes in the progression of the tumors is unclear. Elevated expression levels of the DDIT4 and PKM genes may indicate a worsening of the tumors, whereas elevated expression levels of the ADRB2, MS4A2, and ALDH2 genes may indicate an improvement in the tumor condition. Overall, our study provides new insights into the molecular mechanisms underlying LUAD and offers a potential tool for personalized clinical management.

GSEA revealed that several pathways were enriched in the high-risk group, including cell cycle, ribosome, and spliceosome pathways. The deregulation of cell cycle can lead to abnormal cell proliferation, which is a basic mechanism of cancer. NSCLCs, including LUAD, often have cell cycle abnormalities that have a negative effect on prognosis.[75]

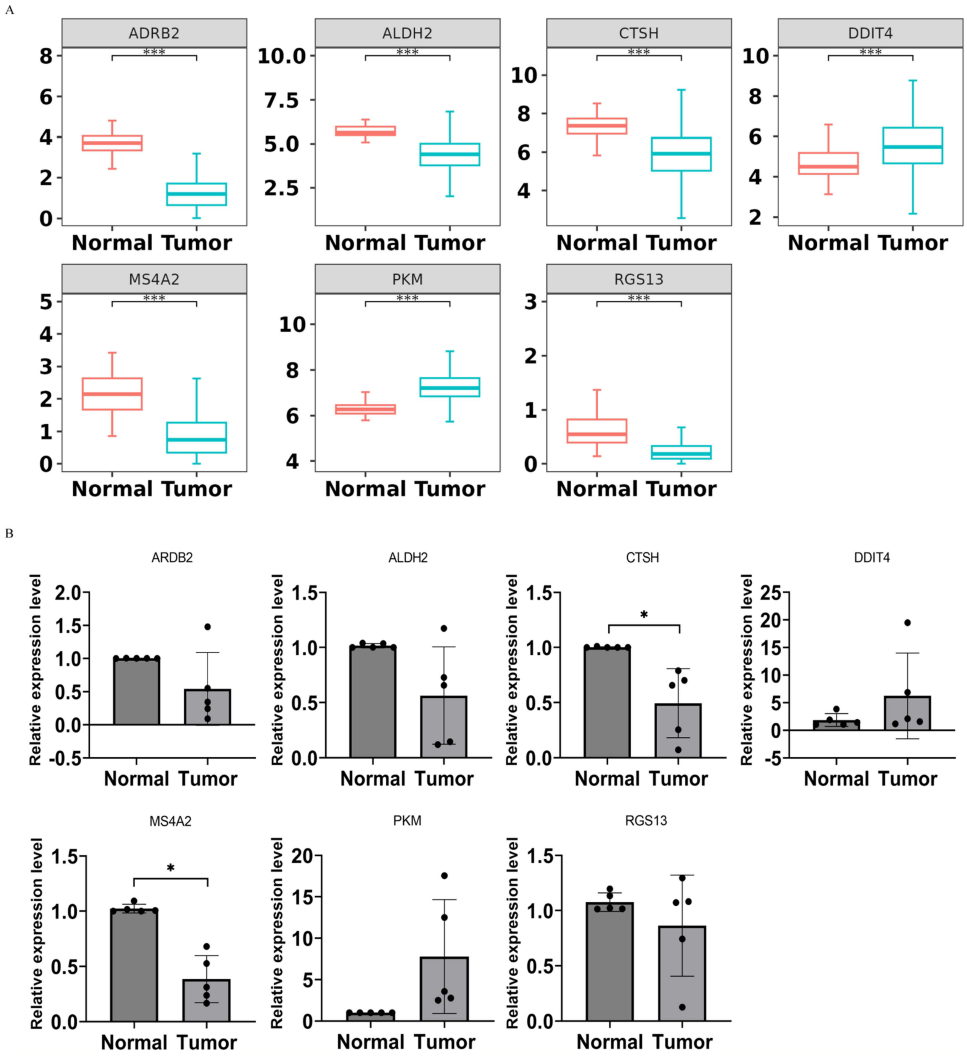


Figure 11 Relative mRNA expression levels of 7-MPRGs in TCGA-database and clinical samples. (**A**) Differential expression in TCGA-LUAD-database. (**B**) Differential expression in five paired LUAD tissues and adjacent non-tumor lung tissues. ***p<0.001, *p<0.05.

Ribosomes and spliceosomes are essential for protein synthesis, sustaining tumor cell growth and proliferation. Dysregulation of the splicing process and ribosomal modifications are both closely related to tumorigenesis.[76–80] These findings provide insight into the mechanisms underlying the high-risk group of LUAD patients and could potentially lead to the development of targeted therapies to improve prognosis for this group.

In our study, we found that TP53, TTN, CSMD3, MUC16, and RYR2 were among the top 5 genes that were mutated in both groups. These somatic mutations may impact LUAD by affecting the expression of characteristic genes. For instance, DDIT4 is a transcriptional target of p53 following DNA damage.[81] Additionally, a previous analysis involving 12 types of cancer found a strong association between TP53 mutations and PKM expression.[82] Moreover, TTN expression has been shown to be related to CTSH,[83] while Hu et al demonstrated that mice with ALDH2 mutations had higher levels of phosphorylated RYR2 protein than wild-type mice.[84]

To better understand how characteristic genes affect immune function, a CIBERSORT analysis was performed to assess the differences in immune infiltration between high- and low-risk groups. The results showed significant differences in 11 types of immune cell infiltration between the two groups, with certain immune cells being more abundant in the high-risk group. Specifically, macrophage M0, CD4+ memory activated T cells, and resting NK cells were found to be enriched in the high-risk group. M0 macrophages are thought to play a role in tumor initiation and progression by promoting angiogenesis, remodeling the extracellular matrix, and suppressing antitumor immune

responses.[85–87] M0 macrophages can promote tumor growth and invasion by secreting cytokines and growth factors such as IL-6, IL-10, VEGF, and MMPs.[88] In LUAD, M0 macrophages have been shown to play a role in tumor progression and metastasis. Studies have found that M0 macrophages are increased in both tumor tissues and peripheral blood of LUAD patients, especially those with lymph node metastasis.[89] CD4+ memory-activated T cells have a crucial role in the immune response against tumors, including lung adenocarcinoma. However, their role in cancer is complex and may depend on various factors. In some instances, CD4+ memory-activated T cells may promote cancer growth and progression by suppressing the immune system or promoting inflammation. Regulatory T cells, which are a type of CD4+ T cell that suppresses immune responses, have been found to be enriched in some cancers, including lung cancer, and can contribute to tumor immune evasion and progression.[90–92] Current evidence suggests that the role of resting NK cells in predicting the prognosis of lung cancer is not entirely clear. Some studies have found that the density of resting NK cells in lung adenocarcinoma tissues is associated with a worse prognosis, possibly due to their immunosuppressive effects. Resting NK cells have been shown to produce cytokines such as TGF-β and IL-10, which can inhibit the function of other immune cells such as T cells and dendritic cells. Additionally, resting NK cells can promote the accumulation of regulatory T cells (Tregs), which can further suppress the function of other immune cells.[93] These immunosuppressive effects can contribute to tumor immune evasion and progression, ultimately leading to a worse prognosis for the patient. When elevated numbers of macrophage M0, CD4+ memory activated T cells, and resting NK cells are found in vivo, it should be realized that this may indicate a worsening of the LUAD tumor situation. Adopting appropriate means to inhibit the function of these immune cells may be a new therapeutic idea to inhibit tumor progression. In addition to these types of cells, the level of monocyte, T-cell CD8+, and mast cell infiltration may also have a correlative link to the prognostic level of patients with LUAD, and appears useful in predicting survival in advanced LUAD.[94]

From the onset to the end of developmental differentiation, the major components of the cell clusters transition from NK cells to three types of macrophages, all of the three cell clusters are associated with the progression of LUAD. After dividing the three key clusters into two subgroups based on the expression levels of characteristic genes, GSEA analysis was performed, which revealed an enrichment of the ERBB signaling pathway. The ERBB tyrosine kinase family includes important members such as EGFR, HER2, HER3, and HER4. Oncogenic alterations of these genes have been known to activate tyrosine kinase aberrantly, thereby driving the development and progression of tumors and immune escape in LUAD.[95,96] These findings highlight the potential significance of ERBB signaling in the development of LUAD and could potentially lead to the development of new therapeutic targets for this disease.

The study identified two potential transcription factors, BCLAF1 and MAX, that may be associated with the differential expression of characteristic genes, but conclusive evidence for their relationship with the 7-MPRGs is currently lacking.

Finally, the researchers analyzed the differential expression of the seven model genes in LUAD RNA sequencing data from the TCGA database. The results were consistent with the risk model developed in the study. Furthermore, the reliability of most of the conclusions was verified using qRT-PCR. However, we acknowledged limitations such as an insufficient sample size and the potential impact of individual differences and other confounding factors. Therefore, it is important to consider these limitations when interpreting the results obtained from existing databases. Overall, these findings provide important insights into the potential roles of the identified genes in the development and progression of LUAD.

In summary, this study identified seven MPRGs (RGS13, ADRB2, DDIT4, MS4A2, ALDH2, CTSH, and PKM) as prognostic genes for LUAD through integrative analysis of scRNA-seq and bulk RNA-seq data. The developed model showed improved prediction compared to single clinical factors. However, the molecular mechanisms of these genes in LUAD and macrophage polarization remain unclear and require further experimental validation. Moreover, the application of the predictive model requires expansion of clinical samples and data support. The researchers will continue to focus on follow-up studies of these genes to gain more insights into their potential roles in LUAD.

Data Sharing Statement

The authors confirm that the data supporting the findings of this study are available within the article and its Supplementary Materials.

Ethical Approval

We hereby confirm that our study complies with the Declaration of Helsinki. All methods described in this manuscript were carried out in strict accordance with the relevant guidelines and regulations. Ethical approval was obtained from the appropriate institutional review board prior to the commencement of the study. All study participants provided informed consent before participating in the research. We have taken all necessary precautions to ensure the safety and well-being of our study participants and adhered to all relevant protocols throughout the study.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This study was supported by a grant from the free exploration and innovation project of Xi'an Jiaotong University (NO. xjj2018144).

Disclosure

The authors declare no conflicts of interest in this work.

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(3):209–249. doi:10.3322/caac.21660
2. Herbst RS, Morgensztern D, Boshoff C. The biology and management of non-small cell lung cancer. *Nature*. 2018;553(7689):446–454. doi:10.1038/nature25183
3. Nicholson AG, Scagliotti G, Tsao MS, Yatabe Y, Travis WD. 2021 WHO Classification of Lung Cancer: a Globally Applicable and Molecular Biomarker-Relevant Classification. *J Thorac Oncol*. 2022;17(9):e80–e83. doi:10.1016/j.jtho.2022.07.006
4. Munzel T, Hahad O, Kuntic M, Keaney JF, Deanfield JE, Daiber A. Effects of tobacco cigarettes, e-cigarettes, and waterpipe smoking on endothelial function and clinical outcomes. *Eur Heart J*. 2020;41(41):4057–4070. doi:10.1093/eurheartj/ehaa460
5. Molina JR, Yang P, Cassivi SD, Schild SE, Adjei AA. Non-small cell lung cancer: epidemiology, risk factors, treatment, and survivorship. *Mayo Clin Proc*. 2008;83(5):584–594. doi:10.1016/S0025-6196(11)60735-0
6. Field RW, Withers BL. Occupational and environmental causes of lung cancer. *Clin Chest Med*. 2012;33(4):681–703. doi:10.1016/j.ccm.2012.07.001
7. Shen X, Wang L, Zhu L. Spatial Analysis of Regional Factors and Lung Cancer Mortality in China, 1973-2013. *Cancer Epidemiol Biomarkers Prev*. 2017;26(4):569–577. doi:10.1158/1055-9965.EPI-16-0922
8. Miller M, Hanna N. Advances in systemic therapy for non-small cell lung cancer. *BMJ*. 2021;375:n2363. doi:10.1136/bmj.n2363
9. Pathria P, Louis TL, Varner JA. Targeting Tumor-Associated Macrophages in Cancer. *Trends Immunol*. 2019;40(4):310–327. doi:10.1016/j.it.2019.02.003
10. Balkwill F, Mantovani A. Inflammation and cancer: back to Virchow? *Lancet*. 2001;357(9255):539–545. doi:10.1016/S0140-6736(00)04046-0
11. Chittezhath M, Dhillon MK, Lim JY, et al. Molecular profiling reveals a tumor-promoting phenotype of monocytes and macrophages in human cancer progression. *Immunity*. 2014;41(5):815–829. doi:10.1016/j.immuni.2014.09.014
12. Quail DF, Joyce JA. Microenvironmental regulation of tumor progression and metastasis. *Nat Med*. 2013;19(11):1423–1437. doi:10.1038/nm.3394
13. Hanahan D, Coussens LM. Accessories to the crime: functions of cells recruited to the tumor microenvironment. *Cancer Cell*. 2012;21(3):309–322. doi:10.1016/j.ccr.2012.02.022
14. Mills CD, Kincaid K, Alt JM, Heilman MJ, Hill AM. Pillars Article: m-1/M-2 Macrophages and the Th1/Th2 Paradigm. *J Immunol*. 2017;199(7):2194–2201. doi:10.4049/jimmunol.1701141
15. Gordon S. Alternative activation of macrophages. *Nat Rev Immunol*. 2003;3(1):23–35. doi:10.1038/nri978
16. Mosser DM, Edwards JP. Exploring the full spectrum of macrophage activation. *Nat Rev Immunol*. 2008;8(12):958–969. doi:10.1038/nri2448
17. Pan Y, Yu Y, Wang X, Zhang T. Tumor-Associated Macrophages in Tumor Immunity. *Front Immunol*. 2020;11:583084. doi:10.3389/fimmu.2020.583084
18. Bruns H, Buttner M, Fabri M, et al. Vitamin D-dependent induction of cathelicidin in human macrophages results in cytotoxicity against high-grade B cell lymphoma. *Sci Transl Med*. 2015;7(282):282ra247. doi:10.1126/scitranslmed.aaa3230
19. Mantovani A, Marchesi F, Malesci A, Laghi L, Allavena P. Tumour-associated macrophages as treatment targets in oncology. *Nat Rev Clin Oncol*. 2017;14(7):399–416. doi:10.1038/nrclinonc.2016.217
20. Liu Q, Song J, Pan Y, et al. Wnt5a/CaMKII/ERK/CCL2 axis is required for tumor-associated macrophages to promote colorectal cancer progression. *Int J Biol Sci*. 2020;16(6):1023–1034. doi:10.7150/ijbs.40535
21. Wang S, Ma L, Wang Z, et al. Lactate Dehydrogenase-A (LDH-A) Preserves Cancer Stemness and Recruitment of Tumor-Associated Macrophages to Promote Breast Cancer Progression. *Front Oncol*. 2021;11:654452. doi:10.3389/fonc.2021.654452

22. Lee CH, Liu SY, Chou KC, et al. Tumor-associated macrophages promote oral cancer progression through activation of the Axl signaling pathway. *Ann Surg Oncol*. 2014;21(3):1031–1037. doi:10.1245/s10434-013-3400-0
23. Guo Z, Song J, Hao J, et al. M2 macrophages promote NSCLC metastasis by upregulating CRYAB. *Cell Death Dis*. 2019;10(6):377. doi:10.1038/s41419-019-1618-x
24. Hwang I, Kim JW, Ylaya K, et al. Tumor-associated macrophage, angiogenesis and lymphangiogenesis markers predict prognosis of non-small cell lung cancer patients. *J Transl Med*. 2020;18(1):443. doi:10.1186/s12967-020-02618-z
25. Potter SS. Single-cell RNA sequencing for the study of development, physiology and disease. *Nat Rev Nephrol*. 2018;14(8):479–492. doi:10.1038/s41581-018-0021-7
26. Baslan T, Hicks J. Unravelling biology and shifting paradigms in cancer with single-cell sequencing. *Nat Rev Cancer*. 2017;17(9):557–569. doi:10.1038/nrc.2017.58
27. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods*. 2008;5(7):621–628. doi:10.1038/nmeth.1226
28. Qi R, Ma A, Ma Q, Zou Q. Clustering and classification methods for single-cell RNA-sequencing data. *Brief Bioinform*. 2020;21(4):1196–1208. doi:10.1093/bib/bbz062
29. Yao W, Liu X, He Y, et al. ScRNA-seq and bulk RNA-seq reveal the characteristics of ferroptosis and establish a risk signature in cholangiocarcinoma. *Mol Ther Oncolytics*. 2022;27:48–60. doi:10.1016/j.omto.2022.09.008
30. Zhang D, Lu W, Cui S, Mei H, Wu X, Zhuo Z. Establishment of an ovarian cancer omentum metastasis-related prognostic model by integrated analysis of scRNA-seq and bulk RNA-seq. *J Ovarian Res*. 2022;15(1):123. doi:10.1186/s13048-022-01059-0
31. Liu X, Li J, Wang Q, et al. Analysis on heterogeneity of hepatocellular carcinoma immune cells and a molecular risk model by integration of scRNA-seq and bulk RNA-seq. *Front Immunol*. 2022;13:1012303. doi:10.3389/fimmu.2022.1012303
32. Zhao J, Wang X, Zhu H, et al. Integrative Analysis of Bulk RNA-Seq and Single-Cell RNA-Seq Unveils Novel Prognostic Biomarkers in Multiple Myeloma. *Biomolecules*. 2022;12(12):1855. doi:10.3390/biom12121855
33. Zhao Y, Li M, Yang Y, et al. Identification of Macrophage Polarization-Related Genes as Biomarkers of Chronic Obstructive Pulmonary Disease Based on Bioinformatics Analyses. *Biomed Res Int*. 2021;2021:9921012. doi:10.1155/2021/9921012
34. Gribov A, Sill M, Luck S, et al. SEURAT: visual analytics for the integrated analysis of microarray data. *BMC Med Genomics*. 2010;3(1):21. doi:10.1186/1755-8794-3-21
35. Becht E, McInnes L, Healy J, et al. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat Biotechnol*. 2018. doi:10.1038/nbt.4314
36. Aran D, Looney AP, Liu L, et al. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat Immunol*. 2019;20(2):163–172. doi:10.1038/s41590-018-0276-y
37. Jin S, Guerrero-Juarez CF, Zhang L, et al. Inference and analysis of cell-cell communication using CellChat. *Nat Commun*. 2021;12(1):1088. doi:10.1038/s41467-021-21246-9
38. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. 2012;16(5):284–287. doi:10.1089/omi.2011.0118
39. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550. doi:10.1186/s13059-014-0550-8
40. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J Stat Softw*. 2010;33(1):1–22. doi:10.18637/jss.v033.i01
41. Hebert PD, Cywinska A, Ball SL, deWaard JR. Biological identifications through DNA barcodes. *Proc Biol Sci*. 2003;270(1512):313–321. doi:10.1098/rspb.2002.2218
42. Zhou S, Lu H, Xiong M. Identifying Immune Cell Infiltration and Effective Diagnostic Biomarkers in Rheumatoid Arthritis by Bioinformatics Analysis. *Front Immunol*. 2021;12:726747. doi:10.3389/fimmu.2021.726747
43. Borcherding N, Voigt AP, Liu V, Link BK, Zhang W, Jabbari A. Single-Cell Profiling of Cutaneous T-Cell Lymphoma Reveals Underlying Heterogeneity Associated with Disease Progression. *Clin Cancer Res*. 2019;25(10):2996–3005. doi:10.1158/1078-0432.CCR-18-3309
44. Lin C, Xu JQ, Zhong GC, et al. Integrating RNA-seq and scRNA-seq to explore the biological significance of NAD + metabolism-related genes in the initial diagnosis and relapse of childhood B-cell acute lymphoblastic leukemia. *Front Immunol*. 2022;13:1043111. doi:10.3389/fimmu.2022.1043111
45. Denisenko TV, Budkevich IN, Zhivotovsky B. Cell death-based treatment of lung adenocarcinoma. *Cell Death Dis*. 2018;9(2):117. doi:10.1038/s41419-017-0063-y
46. Mao X, Xu J, Wang W, et al. Crosstalk between cancer-associated fibroblasts and immune cells in the tumor microenvironment: new findings and future perspectives. *Mol Cancer*. 2021;20(1):131. doi:10.1186/s12943-021-01428-1
47. Hinshaw DC, Shevde LA. The Tumor Microenvironment Innately Modulates Cancer Progression. *Cancer Res*. 2019;79(18):4557–4566. doi:10.1158/0008-5472.CAN-18-3962
48. Liu D, Yao L, Ding X, Zhou H. Multi-omics immune regulatory mechanisms in lung adenocarcinoma metastasis and survival time. *Comput Biol Med*. 2023;164:107333. doi:10.1016/j.compbiomed.2023.107333
49. Qiu A, Xu H, Mao L, et al. A Novel apaQTL-SNP for the Modification of Non-Small-Cell Lung Cancer Susceptibility across Histological Subtypes. *Cancers*. 2022;14(21):5309. doi:10.3390/cancers14215309
50. Shu J, Jiang J, Zhao G. Identification of novel gene signature for lung adenocarcinoma by machine learning to predict immunotherapy and prognosis. *Front Immunol*. 2023;14:1177847. doi:10.3389/fimmu.2023.1177847
51. Johnson EN, Druey KM. Functional characterization of the G protein regulator RGS13. *J Biol Chem*. 2002;277(19):16768–16774. doi:10.1074/jbc.M200751200
52. Sun X, Hou T, Cheung E, et al. Anti-inflammatory mechanisms of the novel cytokine interleukin-38 in allergic asthma. *Cell Mol Immunol*. 2020;17(6):631–646. doi:10.1038/s41423-019-0300-7
53. Han JI, Huang NN, Kim DU, Kehrl JH. RGS1 and RGS13 mRNA silencing in a human B lymphoma line enhances responsiveness to chemoattractants and impairs desensitization. *J Leukoc Biol*. 2006;79(6):1357–1368. doi:10.1189/jlb.1105693

54. Islam TC, Asplund AC, Lindvall JM, et al. High level of cannabinoid receptor 1, absence of regulator of G protein signalling 13 and differential expression of Cyclin D1 in mantle cell lymphoma. *Leukemia*. 2003;17(9):1880–1890. doi:10.1038/sj.leu.2403057
55. Kulik G. ADRB2-Targeting Therapies for Prostate Cancer. *Cancers*. 2019;11(3):358. doi:10.3390/cancers11030358
56. Zhang X, Zhang Y, He Z, et al. Chronic stress promotes gastric cancer progression and metastasis: an essential role for ADRB2. *Cell Death Dis*. 2019;10(11):788. doi:10.1038/s41419-019-2030-2
57. Renz BW, Takahashi R, Tanaka T, et al. beta2 Adrenergic-Neurotrophin Feedforward Loop Promotes Pancreatic Cancer. *Cancer Cell*. 2018;33(1):75–90 e77. doi:10.1016/j.ccell.2017.11.007
58. Wu FQ, Fang T, Yu LX, et al. ADRB2 signaling promotes HCC progression and sorafenib resistance by inhibiting autophagic degradation of HIF1alpha. *J Hepatol*. 2016;65(2):314–324. doi:10.1016/j.jhep.2016.04.019
59. Ji L, Xu F, Zhang J, et al. ADRB2 expression predicts the clinical outcomes and is associated with immune cells infiltration in lung adenocarcinoma. *Sci Rep*. 2022;12(1):15994. doi:10.1038/s41598-022-19991-y
60. Song L, Chen Z, Zhang M, et al. DDIT4 overexpression associates with poor prognosis in lung adenocarcinoma. *J Cancer*. 2021;12(21):6422–6428. doi:10.7150/jca.60118
61. Song X, Liu B, Zhao G, et al. Streptococcus pneumoniae promotes migration and invasion of A549 cells in vitro by activating mTORC2/AKT through up-regulation of DDIT4 expression. *Front Microbiol*. 2022;13:1046226. doi:10.3389/fmicb.2022.1046226
62. Ly D, Zhu CQ, Cabanero M, Tsao MS, Zhang L. Role for High-Affinity IgE Receptor in Prognosis of Lung Adenocarcinoma Patients. *Cancer Immunol Res*. 2017;5(9):821–829. doi:10.1158/2326-6066.CIR-16-0392
63. Chang JS, Hsiao JR, Chen CH. ALDH2 polymorphism and alcohol-related cancers in Asians: a public health perspective. *J Biomed Sci*. 2017;24(1):19. doi:10.1186/s12929-017-0327-y
64. Li K, Guo W, Li Z, et al. ALDH2 Repression Promotes Lung Tumor Progression via Accumulated Acetaldehyde and DNA Damage. *Neoplasia*. 2019;21(6):602–614. doi:10.1016/j.neo.2019.03.008
65. Li Y, Jiang L, Zhu Z, Fu B, Sun X, Jiao Y. Long Noncoding RNA SNHG16 Regulates the Growth of Human Lung Cancer Cells by Modulating the Expression of Aldehyde Dehydrogenase 2 (ALDH2). *J Oncol*. 2022;2022:2411642. doi:10.1155/2022/2411642
66. Sivaparvathi M, Sawaya R, Gokaslan ZL, Chintala SK, Rao JS. Expression and the role of cathepsin H in human glioma progression and invasion. *Cancer Lett*. 1996;104(1):121–126. doi:10.1016/0304-3835(96)04242-5
67. Kos J, Stabuc B, Schweiger A, et al. H, and L and their inhibitors stefin A and cystatin C in sera of melanoma patients. *Clin Cancer Res*. 1997;3(10):1815–1822.
68. Schweiger A, Christensen IJ, Nielsen HJ, Sorensen S, Brunner N, Kos J. Serum cathepsin H as a potential prognostic marker in patients with colorectal cancer. *Int J Biol Markers*. 2004;19(4):289–294. doi:10.1177/172460080401900406
69. Peng P, Chen JY, Zheng K, Hu CH, Han YT. Favorable Prognostic Impact of Cathepsin H (CTSH) High Expression in Thyroid Carcinoma. *Int J Gen Med*. 2021;14:5287–5299. doi:10.2147/IJGM.S327689
70. Schweiger A, Staib A, Werle B, et al. Cysteine proteinase cathepsin H in tumours and sera of lung cancer patients: relation to prognosis and cigarette smoking. *Br J Cancer*. 2000;82(4):782–788. doi:10.1054/bjoc.1999.0999
71. Hu Y, Mu H, Deng Z. The transcription factor TEAD4 enhances lung adenocarcinoma progression through enhancing PKM2 mediated glycolysis. *Cell Biol Int*. 2021;45(10):2063–2073. doi:10.1002/cbin.11654
72. Cao F, Yang D, Tang F, et al. Girdin Promotes Tumorigenesis and Chemoresistance in Lung Adenocarcinoma by Interacting with PKM2. *Cancers*. 2022;14(22):5688. doi:10.3390/cancers14225688
73. Chen W, Li X, Du B, Cui Y, Ma Y, Li Y. The long noncoding RNA HOXA11-AS promotes lung adenocarcinoma proliferation and glycolysis via the microRNA-148b-3p/PKM2 axis. *Cancer Med*. 2022.
74. Sun H, Zhu A, Zhang L, Zhang J, Zhong Z, Wang F. Knockdown of PKM2 Suppresses Tumor Growth and Invasion in Lung Adenocarcinoma. *Int J Mol Sci*. 2015;16(10):24574–24587. doi:10.3390/ijms161024574
75. Liu J, Peng Y, Wei W. Cell cycle on the crossroad of tumorigenesis and cancer therapy. *Trends Cell Biol*. 2022;32(1):30–44. doi:10.1016/j.tcb.2021.07.001
76. Sterlacci W, Fiegl M, Tzankov A. Prognostic and predictive value of cell cycle deregulation in non-small-cell lung cancer. *Pathobiology*. 2012;79(4):175–194. doi:10.1159/000336462
77. Gahete MD, Herman-Sanchez N, Fuentes-Fayos AC, Lopez-Canovas JL, Luque RM. Dysregulation of splicing variants and spliceosome components in breast cancer. *Endocr Relat Cancer*. 2022;29(9):R123–R142. doi:10.1530/ERC-22-0019
78. Sette C, Paronetto MP. Somatic Mutations in Core Spliceosome Components Promote Tumorigenesis and Generate an Exploitable Vulnerability in Human Cancer. *Cancers*. 2022;14(7):1827. doi:10.3390/cancers14071827
79. Wang E, Aifantis I. RNA Splicing and Cancer. *Trends Cancer*. 2020;6(8):631–644. doi:10.1016/j.trecan.2020.04.011
80. Pelletier J, Thomas G, Volarević S. Volarevic S: ribosome biogenesis in cancer: new players and therapeutic avenues. *Nat Rev Cancer*. 2018;18(1):51–63. doi:10.1038/nrc.2017.104
81. Ellisen LW, Ramsayer KD, Johannessen CM, et al. REDD1, a developmentally regulated transcriptional target of p63 and p53, links p63 to regulation of reactive oxygen species. *Mol Cell*. 2002;10(5):995–1005. doi:10.1016/S1097-2765(02)00706-2
82. Motlagh AV, Mahdevar M, Mirzaei S, et al. Introduction of mutant TP53 related genes in metabolic pathways and evaluation their correlation with immune cells, drug resistance and sensitivity. *Life Sci*. 2022;303:120650. doi:10.1016/j.lfs.2022.120650
83. Floyel T, Mirza AH, Kaur S, et al. The Rac2 GTPase contributes to cathepsin H-mediated protection against cytokine-induced apoptosis in insulin-secreting cells. *Mol Cell Endocrinol*. 2020;518:110993. doi:10.1016/j.mce.2020.110993
84. Hu Q, Chen H, Shen C, et al. Impact and potential mechanism of effects of chronic moderate alcohol consumption on cardiac function in aldehyde dehydrogenase 2 gene heterozygous mice. *Alcohol Clin Exp Res*. 2022;46(5):707–723. doi:10.1111/acer.14811
85. Huang L, Wang Z, Chang Y, et al. EFEMP2 indicates assembly of M0 macrophage and more malignant phenotypes of glioma. *Aging*. 2020;12(9):8397–8412. doi:10.18632/aging.103147
86. Zhang Y, Zou J, Chen R. An M0 macrophage-related prognostic model for hepatocellular carcinoma. *BMC Cancer*. 2022;22(1):791. doi:10.1186/s12885-022-09872-y
87. Xie Z, Li X, He Y, et al. Immune Cell Confrontation in the Papillary Thyroid Carcinoma Microenvironment. *Front Endocrinol (Lausanne)*. 2020;11:570604. doi:10.3389/fendo.2020.570604

88. Dong B, Wu C, Huang L, Qi Y. Macrophage-Related SPP1 as a Potential Biomarker for Early Lymph Node Metastasis in Lung Adenocarcinoma. *Front Cell Dev Biol*. 2021;9:739358. doi:10.3389/fcell.2021.739358
89. Liu X, Wu S, Yang Y, Zhao M, Zhu G, Hou Z. The prognostic landscape of tumor-infiltrating immune cell and immunomodulators in lung cancer. *Biomed Pharmacother*. 2017;95:55–61. doi:10.1016/j.biopha.2017.08.003
90. Li M, Zhao J, Yang R, et al. CENPF as an independent prognostic and metastasis biomarker corresponding to CD4+ memory T cells in cutaneous melanoma. *Cancer Sci*. 2022;113(4):1220–1234. doi:10.1111/cas.15303
91. Landskron J, Helland O, Torgersen KM, et al. Activated regulatory and memory T-cells accumulate in malignant ascites from ovarian carcinoma patients. *Cancer Immunol Immunother*. 2015;64(3):337–347. doi:10.1007/s00262-014-1636-6
92. Zhong R, Chen D, Cao S, Li J, Han B, Zhong H. Immune cell infiltration features and related marker genes in lung cancer based on single-cell RNA-seq. *Clin Transl Oncol*. 2021;23(2):405–417. doi:10.1007/s12094-020-02435-2
93. Bonnema JD, Rivlin KA, Ting AT, Schoon RA, Abraham RT, Leibson PJ. Cytokine-enhanced NK cell-mediated cytotoxicity. Positive modulatory effects of IL-2 and IL-12 on stimulus-dependent granule exocytosis. *J Immunol*. 1994;152(5):2098–2104. doi:10.4049/jimmunol.152.5.2098
94. Xue Q, Wang Y, Zheng Q, et al. Prognostic value of tumor immune microenvironment factors in patients with stage I lung adenocarcinoma. *Am J Cancer Res*. 2023;13(3):950–963.
95. Wang S, Shi J, Ye Z, et al. Predicting EGFR mutation status in lung adenocarcinoma on computed tomography image using deep learning. *Eur Respir J*. 2019;53(3):1800986. doi:10.1183/13993003.00986-2018
96. Wang Y, Yang N, Zhang Y, et al. Effective Treatment of Lung Adenocarcinoma Harboring EGFR-Activating Mutation, T790M, and cis-C797S Triple Mutations by Brigatinib and Cetuximab Combination Therapy. *J Thorac Oncol*. 2020;15(8):1369–1375. doi:10.1016/j.jtho.2020.04.014