

Unveiling the prognostic and immunotherapeutic role of Tregs in lung cancer using integrated analysis of single-cell and bulk RNA-sequencing

HONGWEI HE^{1*}, ZHIMING XIAO^{1*}, JIJUN XUE², YU ZHANG³, JICHEN QU⁴, PENG WU¹ and QIANGJUN HE¹

¹Department of Cardiothoracic Surgery, Shanghai General Hospital Jiuquan Branch (Jiuquan People's Hospital), Jiuquan, Gansu 735000, P.R. China; ²Department of Thoracic Surgery, Gansu Cancer Hospital (Sun Yat-sen University Cancer Center), Lanzhou, Gansu 730000, P.R. China; ³Department of Thoracic Surgery, Gansu Provincial Maternity and Child Health Care Hospital (Gansu Provincial Central Hospital), Lanzhou, Gansu 730000, P.R. China; ⁴Department of Thoracic Surgery, Shanghai General Hospital, Shanghai 200001, P.R. China

Received April 21, 2025; Accepted February 16, 2026

DOI: 10.3892/ol.2026.15588

Abstract. The infiltration of regulatory T cells (Tregs) in lung adenocarcinoma (LUAD) is associated with a poor prognosis. The present study aimed to explore the potential function of Treg marker genes in prognosis and immunotherapy using the transcriptome profiles of five LUAD cohorts sourced from public databases. Among them, the single-cell dataset GSE131907 was employed to identify cell types in lung cancer tissues and to identify Treg markers. The prognostic Treg markers were screened using univariate Cox and Least Absolute Shrinkage and Selection Operator regression analyses. Subsequently, a prognostic model was constructed and assessed using Kaplan-Meier and receiver operating characteristic curves. Furthermore, the effect of prognostic Treg markers on clinical characteristics, the immune microenvironment and tumor mutation burden (TMB) were evaluated. In addition, the function of genes strongly correlated with the prognostic risk score were investigated using Spearman and functional enrichment analysis. Finally, the genes expression levels were assessed using reverse transcription-quantitative PCR (RT-qPCR) and western blotting. By analyzing the single-cell data, 13 Treg markers [centromere protein M (CENPM), pituitary tumor-transforming gene 1 protein, interleukin 1 receptor type 2, baculoviral IAP repeat containing 3, glucocorticoid induced 1 (GLCCI1), melanoma-associated

antigen H1 (MAGEH1), CD5, cytokine inducible SH2 containing protein, zinc finger protein 101 (ZNF101), Ikaros family zinc finger protein 4 (IKZF4), ankyrin repeat and SOCS box protein 2, zinc finger CCCH-type containing 12D and C-C motif chemokine receptor 6] were identified as prognostic features. The prognostic model constructed using these 13 genes revealed that the high-risk group had a poorer prognosis than the low-risk group. Moreover, it was demonstrated that the risk score could be an independent prognostic factor affecting the prognosis of patients with LUAD. Additionally, the high-risk group had a lower ESTIMATE score, higher TMB score and lower T cell receptor richness than the low-risk group. Finally, RT-qPCR and western blotting showed that the expression levels of CENPM, ZNF101, MAGEH1 and IKZF4 were significantly altered in cancer tissues compared with the adjacent normal tissues. In conclusion, a reliable prognostic model based on 13 Treg markers was developed. The comprehensive characterization of the Treg markers of LUAD may help to monitor the prognosis and provide new strategies for LUAD treatment.

Introduction

Lung cancer (LC) is one of the most widespread malignancies worldwide. The predominant histological subtype of LC, lung adenocarcinoma (LUAD), currently accounts for >40% of LC cases (1). In recent years, LC cases have gradually increased among young people and women (2). Moreover, the prognosis for patients with LUAD is poor, with 5-year survival rates of only 10-20% (3,4). Patients with early stage LUAD lack typical symptoms making diagnosis difficult (5), and most patients with advanced LUAD are at an increasing risk of metastasis and recurrence (6). Immunotherapy, specifically immune checkpoint inhibitors, such as anti-programmed cell death protein 1 and programmed death-ligand 1, have been used in the treatment of LUAD (7). However, more biomarkers are urgently needed for the diagnosis and prognosis assessment of LUAD; thus, the construction of a prognostic model may

Correspondence to: Dr Qiangjun He, Department of Cardiothoracic Surgery, Shanghai General Hospital Jiuquan Branch (Jiuquan People's Hospital), 9 Wind Power Generation Avenue, Suzhou, Jiuquan, Gansu 735000, P.R. China
E-mail: 1153584580@qq.com

*Contributed equally

Key words: lung adenocarcinoma, regulatory T cells, single-cell RNA sequencing, prognostic model, model genes

help stratify the risk of patients with LUAD and implement personalized treatment.

The tumor microenvironment (TME) constitutes an intricate and dynamically evolving mix of cancer cells, immune cells, extracellular matrix and biomolecules (8); it serves an important role in tumor progression and notably influences therapeutic response (9). Single-cell RNA sequencing (scRNA-seq) is a powerful technology that can be used to explore the TME and tumor heterogeneity of patients with LUAD at a single-cell level (10-12). A previous study has reported that tumors attract immunosuppressive immune cells such as regulatory T cells (Tregs) to evade the immune system and thus affect the effectiveness of immunotherapy (13). Tregs are an immunosuppressive subset of CD4⁺ T cells, which are used to maintain immunological self-tolerance and homeostasis (14). A study has reported that an increase in Treg infiltration in patients with LUAD after inhibition and knockout of discoidin domain receptor 1 promotes tumor growth in LUAD (13). Tregs also suppress antitumor immune responses; however, the mechanisms by which Tregs affect the prognosis of patients with LUAD remain unclear.

The present study aimed to build a prognostic model associated with Treg infiltration to analyze LUAD prognosis and provide support for risk stratification and personalized treatment. The scRNA-seq (GSE131907) and bulk-RNA sequencing datasets of patients with LUAD were obtained from the Gene Expression Omnibus (GEO) and The Cancer Genome Atlas (TCGA) databases. After pre-processing and cell type annotation, Treg markers in GSE131907 were selected. Using univariate Cox and Least Absolute Shrinkage and Selection Operator (LASSO) regression analyses, the prognostic genes in TCGA-LUAD were identified to construct a prognostic model. The prognostic model was then verified using independent datasets (GSE30219, GSE26939 and GSE72094). Moreover, the relationship between the prognostic risk score and survival, the TME and immunotherapeutic response were explored using different bioinformatic strategies. Finally, the expression levels of prognostic genes were examined at both the mRNA and protein level using clinical samples.

Materials and methods

Public data collection. A total of 4 LUAD transcriptome-related datasets, including the scRNA-seq dataset [GSE131907 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE131907>) (15)] and three bulk transcriptomic cohorts [GSE30219 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE30219>) (16), platform GPL570; GSE26939 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE26939>) (17), platform GPL9053; and GSE72094 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE72094>) (18), platform GPL15048], were obtained from the GEO (<https://www.ncbi.nlm.nih.gov/geo/>) database. Detailed sample information of these datasets is presented in Table SI. In addition, the gene expression profiles and corresponding clinical information of the TCGA-LUAD cohort (<https://portal.gdc.cancer.gov/projects/tcga-luad>) (19) were downloaded from the TCGA (<https://tcga-data.nci.nih.gov/tcga/>) database. TCGA-LUAD contains an expression matrix of 585 tumor samples and the

overall survival (OS) data of 572 tumor samples. Specifically, 513 patients with both gene expression data and OS information were used to construct the prognostic model, and the GSE30219, GSE26939 and GSE72094 datasets were used for external validation.

For bulk transcriptomic analyses, TCGA-LUAD and the GEO cohorts were included if they comprised histologically confirmed LUAD and provided genome-wide gene expression profiles together with OS data and basic clinicopathological information. Cohorts without survival data, with incomplete clinical annotations or with data quality issues were not considered. For scRNA-seq analysis, GSE131907 was selected as it profiled treatment-naïve patients with LUAD with well-annotated tumor, normal lung and lymph node samples, and provided sufficient T cell coverage for robust identification of Treg-related transcriptional programs.

In detail, TCGA-LUAD contains data for 276 female and 237 male patients with a median age of 66 years (range, 33-88 years). GSE30219 contains data for 43 female and 250 male patients with a median age of 63 years (range, 15-84 years). GSE26939 contains data for 62 female and 53 male patients with a median age of 65 years (range, 41-90 years). GSE72094 contains data for 222 female and 176 male patients with a median age of 70 years (range, 38-89 years). GSE131907 contains data for 17 female and 27 male patients with a median age of 63 years (range, 39-81 years).

Data processing. The ‘Seurat’ (v.4.1.0) R package (R Foundation for Statistical Computing, Vienna, Austria; <https://www.r-project.org/>) was employed for the data processing of GSE131907. Low-quality cells and genes were filtered out using $200 < n_{\text{Feature_RNA}} < 6,000$, $n_{\text{Count_RNA}} < 60,000$ and $\text{percent.mt} < 20\%$ as the criteria (15), otherwise the default was used for the other parameters. Subsequently, distinct cell clusters were generated using uniform manifold approximation and orjection clustering (resolution=1). These cell clusters were then annotated according to corresponding marker genes (15).

Prognostic model. Treg markers were selected in GSE131907 according to the following criteria: $\text{avg_llog2FoldChangel} \geq 1.5$ (Tregs vs. the other cells) and $\text{pct.m} > 10\%$ (the markers expressed in $\geq 10\%$ of Tregs). From these differentially expressed genes, the feature genes associated with prognosis were selected in TCGA-LUAD using univariate Cox analysis with $P < 0.05$ as the significance cut-off. Subsequently, LASSO regression was performed to select the prognostic genes using the ‘glmnet’ (4.1-4) R package. The risk score formula was follows: $\text{Risk score} = (\text{Expression}_1 \times \text{Coefficient}_1) + (\text{Expression}_2 \times \text{Coefficient}_2) + \dots + (\text{Expression}_n \times \text{Coefficient}_n)$, where n is the number of prognostic genes.

A total of 513 patients with LUAD were then divided into the high- and low-risk groups based on the median risk score. A Kaplan-Meier curve was generated to compare the OS difference between the high- and low-risk groups using the ‘survminer’ (v. 0.4.9) R package ($P < 0.05$). Meanwhile, the ‘survivalROC’ (v. 1.0.3.1) R package was used to construct the 1-, 2- and 3-year receiver operating characteristic (ROC) curves. The prognostic model was then verified using the GSE30219, GSE26939 and GSE72094 datasets using Kaplan-Meier and

ROC curves. In consideration that the external validation cohorts were profiled on different platforms, which introduce distributional and scale shifts in gene expression and in the resulting risk scores, using a single absolute cut-off across heterogenous cohorts can be poorly calibrated. Therefore, a cohort-specific optimal cut-off within each validation set was used, which is standard for platform-heterogenous validations. All bulk RNA-seq/microarray cohorts were analyzed independently without cross-cohort merging. Therefore, batch-effect correction across different datasets or platforms was not applicable.

Selection of independent prognostic indicators. To assess the effect of clinical indicators on the prognosis of patients with LUAD, 5 clinical indicators [age, sex, tumor (T) stage, lymph node (N) stage and metastasis (M) stage] and risk score were evaluated. Subsequently, independent prognostic indicators were identified using univariate and multivariate Cox analyses ($P < 0.05$). Finally, the differences in risk score in different clinical subgroups were analyzed (Wilcoxon rank-sum test; $P < 0.05$).

TME analysis. To evaluate the TME of patients with LUAD, the CIBERSORT algorithm was utilized. The infiltration ratio differences of 22 immune cells were compared in TCGA-LUAD [adjusted (adj.) $P < 0.05$]. Subsequently, 7 immune-system-related metagene clusters [hemopoietic cell kinase, IgG, interferon, lymphocyte-specific kinase (LCK), major histocompatibility complex (MHC)-I, MCH-2 and STAT1] were obtained from the published literature (20) and the immune scores of each sample were calculated. In addition, the T-cell receptor (TCR) richness and TCR Shannon diversity were determined in the low- and high-risk groups. TCR richness refers to the number of distinct productive TCR clonotypes detected in a sample; it reflects how many different T cell clones are present (breadth of the repertoire). TCR Shannon diversity summarizes both the number of clonotypes and how evenly they are distributed. Higher values indicate that numerous different clones are present and none of them overwhelmingly dominates (greater breadth and evenness) (21).

Selection and functional enrichment analysis of risk genes. In TCGA-LUAD, genes that were moderately and highly correlated with risk scores were selected based on the Spearman's correlation ($|cor| > 0.5$; adj. $P < 0.05$). Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were performed to evaluate the function of these genes using the 'clusterProfiler' (v. 4.2.2) R package (adj. $P < 0.05$) (22).

Patient sample collection. The tumor and para-carcinoma tissues of 5 patients with LUAD (2 male and 3 female patients; median age, 60 years; age range, 52-65 years) were collected from the Shanghai General Hospital Jiuquan Branch (Jiuquan People's Hospital; Jiuquan, China) from March 2025 to April 2025. The inclusion criteria for patient selection were as follows: i) Histopathologically confirmed primary LUAD; ii) patients who underwent curative-intent surgical resection and could provide paired LUAD and adjacent non-tumorous lung tissues; iii) no neoadjuvant chemotherapy, radiotherapy, targeted therapy or immunotherapy before surgery; and

iv) basic clinicopathological information were available and the patient provided written informed consent. The exclusion criteria were as follows: i) History of other malignant tumors; ii) recurrent or metastatic LUAD at the time of surgery; iii) severe concomitant infectious, autoimmune or systemic inflammatory diseases that could markedly affect immune status; and iv) insufficient tissue quality or quantity for RNA and protein extraction.

Reverse transcription-quantitative PCR (RT-qPCR). Total RNA from all tissues was extracted using the E.Z.N.A.[®] Total RNA Kit I (cat. no. R6834-01; Omega Bio-tek, Inc.). The synthesis of cDNA was then performed using the 5X Evo M-MLV RT Master Mix (cat. no. AG11728; Hunan Accurate Bio-Medical Technology Co., Ltd.) according to the manufacturer's instructions. Subsequently, qPCR was performed using the 2X SYBR Green Premix Pro Taq HS qPCR Tracking Kit (cat. no. AG1170; Hunan Accurate Bio-Medical Technology Co., Ltd.) on the StepOne Plus[™] Real-Time PCR System (Thermo Fisher Scientific, Inc.) under the following thermocycling conditions: Initial denaturation at 95°C for 30 sec, followed by 40 cycles of denaturation at 95°C for 5 sec and annealing at 60°C for 30 sec. The primer sequences for all prognostic genes are listed in Table SII. The data was then exported and analyzed using the $2^{-\Delta\Delta Cq}$ method (23), with β -actin as the endogenous control.

Western blotting. Fresh-frozen LUAD tumor and adjacent normal tissues were homogenized on ice in RIPA buffer (MilliporeSigma). Lysates were centrifuged at 8,000 x g at 4°C for 15 min and supernatants were collected. Protein concentrations were determined using the BCA assay following the manufacturer's instructions (Thermo Fisher Scientific, Inc.) and quantified on a microplate reader at 562 nm. Aliquots containing equal amount of protein (20 μ g per lane) were mixed with 5X loading buffer (Beyotime Biotechnology), heated at 98°C for 5 min, rapidly cooled on ice and loaded onto 4-12% SDS-PAGE gels (Thermo Fisher Scientific, Inc.). Proteins were then transferred to PDVF membranes, followed by blocking with 5% BSA (MilliporeSigma) at room temperature for 1 h. Primary antibodies against centromere protein M (CENPM) (cat. no. 19840-1-AP), zinc finger protein 101 (ZNF101) (cat. no. 25599-1-AP), melanoma-associated antigen H1 (MAGEH1) (cat. no. 12424-1-AP), Ikaros family zinc finger protein 4 (IKZF4) (cat. no. 31094-1-AP) and β -actin (cat. no. 20536-1-AP) purchased from Proteintech Group, Inc., were diluted to 1:1,000 and incubated with the membranes overnight at 4°C. β -actin was used as the loading control. After washing with PBST, membranes were incubated with HRP-conjugated goat anti-rabbit recombinant secondary antibody (H+L) (cat. no. RGAR001; Proteintech) diluted to 1:5,000 at room temperature for 1 h. After washing with PBST, signals were developed using ECL chemiluminescence (Beyotime Biotechnology) and imaged on a chemiluminescence system. Band intensities were quantified in ImageJ (version 1.53a; National Institutes of Health) and target proteins were normalized to β -actin to obtain relative expression.

Statistical analysis. All statistical analyses were performed using R software and GraphPad Prism 9.0 (Dotmatics).

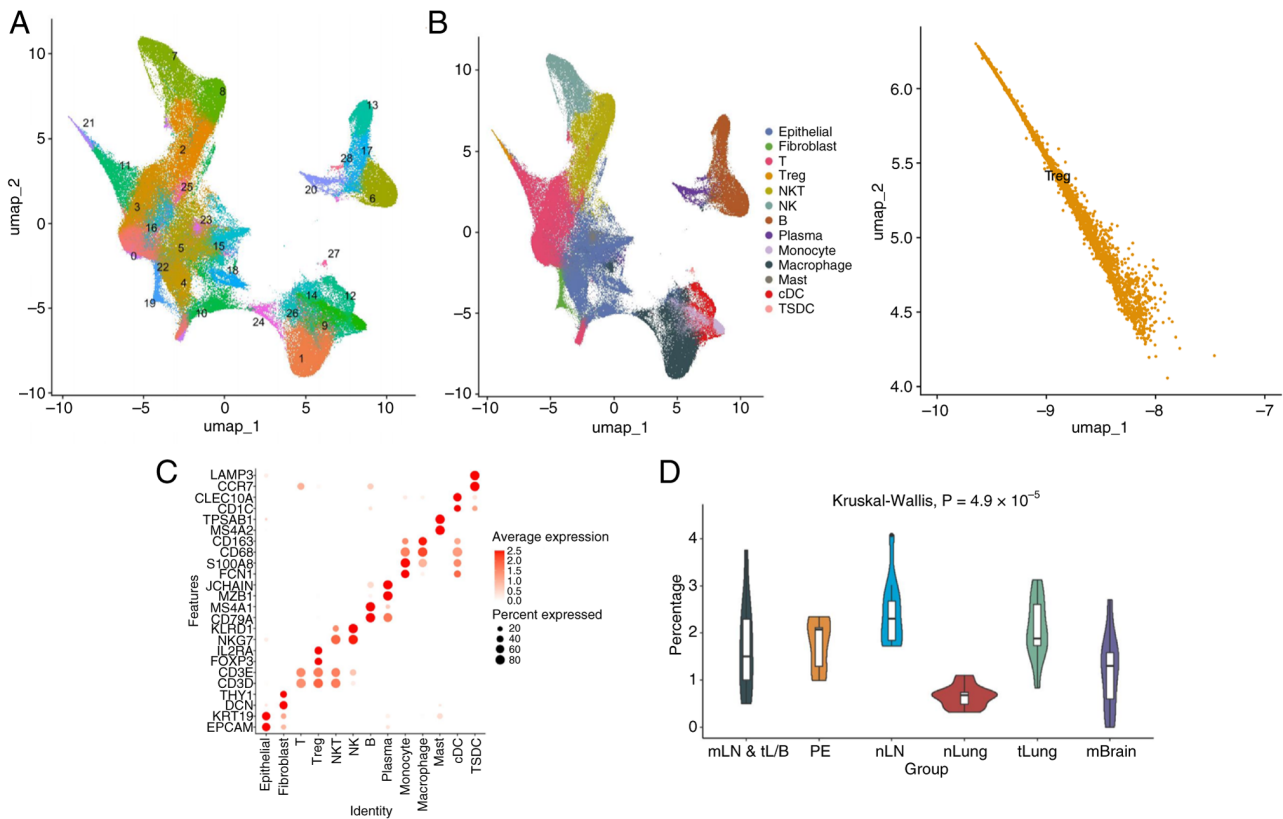


Figure 1. Identification of 13 cell types in the GSE131907 dataset. (A) UMAP plot of the distribution of 29 clusters. (B) UMAP plot of the distribution of 13 cell types (left) with an inset showing a magnified view of the Treg cluster (right). (C) Expression distribution of marker genes in each cell type. (D) Average proportion of Tregs in the mLN & tL/B, PE, nLN, nLung, tLung and mBrain groups. UMAP, uniform manifold approximation and projection; NKT, natural killer T cells; NK, natural killer cells; cDC, conventional dendritic cells; TSDC, tumor-specific dendritic cells; Tregs, regulatory T cells; LUAD, lung adenocarcinoma; mLN, metastatic lymph node; tL/B, tumor lung of advanced-stage LUAD; PE, pleural effusion; nLN, normal lymph node; nLung, normal lung; tLung, tumor lung of early-stage LUAD; mBrain, metastatic brain.

Continuous variables between two groups in the TCGA and GEO cohorts were compared using the Wilcoxon rank-sum test. For RT-qPCR and western blot experiments, expression levels of key genes are presented as mean \pm SD and differences between tumor and adjacent tissues were assessed using the paired Student's t-test ($n=5$ for each group). Survival curves were estimated by the Kaplan-Meier method and compared using the log-rank test. $P < 0.05$ was considered to indicate a statistically significant difference. For multiple testing in the differential expression and enrichment analyses, the Benjamini-Hochberg method was used to control the false discovery rate and adj. $P < 0.05$ was considered to indicate a statistically significant difference.

Results

Normal lymph nodes (nLNs) contain the highest proportion of Tregs in patients with LUAD. First, quality control of all samples in GSE131907 was performed (Table SIII). Subsequently, the top 2,000 variable genes and the top 50 principal components were used for dimensionality reduction and all high-quality cells were grouped into 29 clusters (Fig. 1A). After annotation using canonical cell markers, 13 cell types, comprising epithelial cells, fibroblasts, T cells, Tregs, natural killer (NK)T cells, NK cells, B cells, plasma cells, monocytes, macrophages, MAST cells, classical dendritic cells (DCs) and

tumor specific DCs, were identified (Fig. 1B). The expression levels of the canonical marker genes in each cell type are displayed in a bubble plot shown in Fig. 1C. The average proportion of Tregs in the nLNs of patients with LUAD was 2.4346%, which was notably higher than that of the other five groups (Fig. 1D and Table SIV). In addition, the tumor lung (average proportion, 2.0831%) of patients with LUAD had markedly more Tregs than normal lung (average proportion, 0.6636%). By comparing gene expressions between Tregs and the other 12 cell types, a total of 326 Treg markers were obtained for the further analysis (Table SV).

Treg-related prognostic model composed of 13 prognostic genes in LUAD. A total of 17 Treg markers [C-C motif chemokine receptor 6 (CCR6), zinc finger CCCH-type containing 12D (ZC3H12D), ankyrin repeat and SOCS box protein 2 (ASB2), IKZF4, ZNF101, T cell receptor β constant 1, signaling lymphocytic activation molecule family member 1, CD28, cytokine inducible SH2 containing protein (CISH), CD5, MAGEH1, glucocorticoid induced 1 (GLCC1), lymphocyte transmembrane adaptor 1, baculoviral IAP repeat containing 3 (BIRC3), interleukin 1 receptor type 2 (IL1R2), pituitary tumor-transforming gene 1 protein (PTTG1) and CENPM] associated with prognosis were identified based on the univariate Cox analysis (Fig. 2A). To obtain more robust Treg markers, LASSO analysis was employed and 13 prognostic

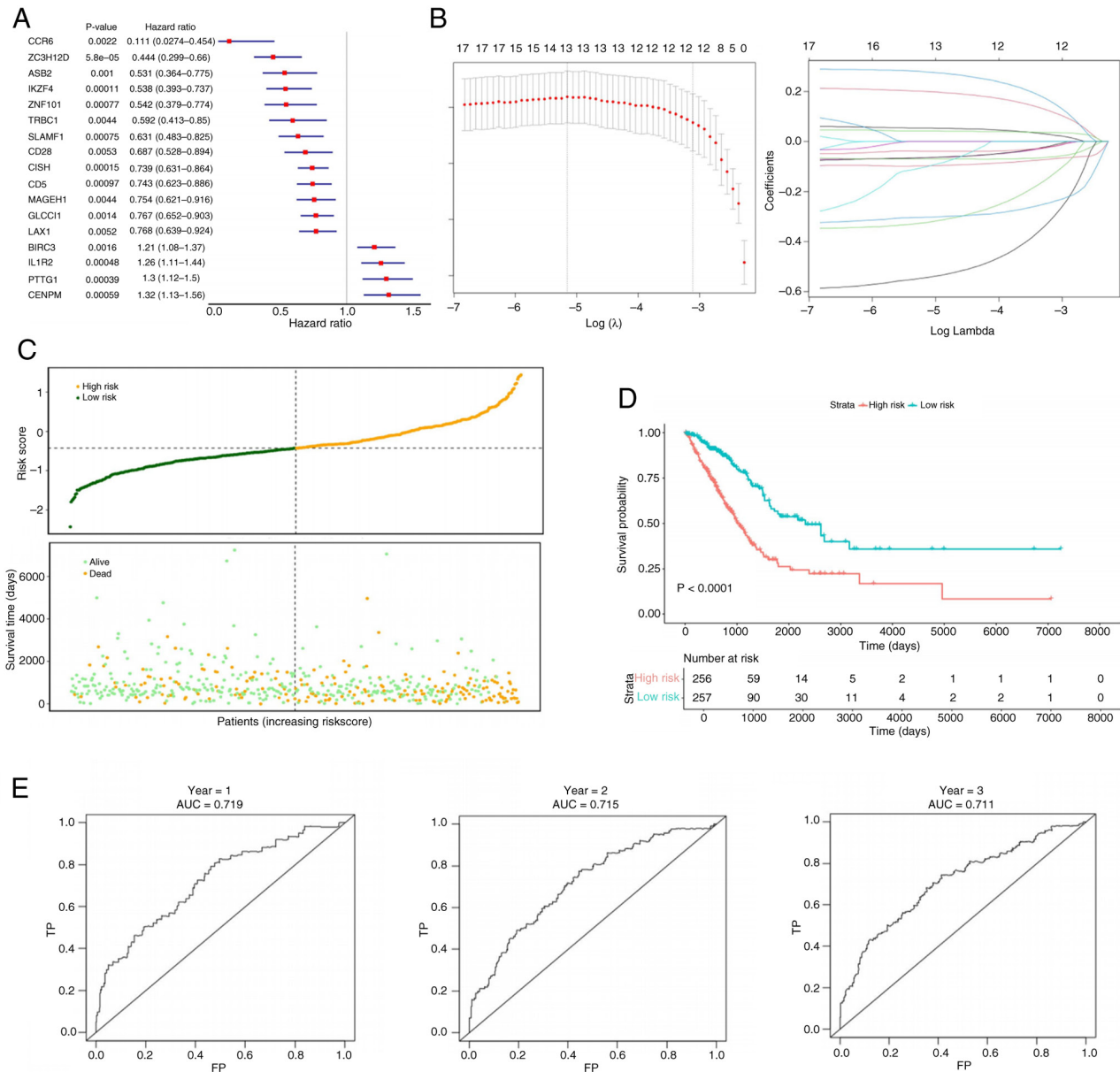


Figure 2. Selection of prognostic genes and construction of a prognostic model for LUAD. (A) Univariate Cox analysis using the TCGA-LUAD dataset. (B) Least Absolute Shrinkage and Selection Operator analysis using the TCGA-LUAD dataset. (C) Risk curves and survival state in the TCGA-LUAD dataset. (D) Kaplan-Meier curve for the TCGA-LUAD dataset. (E) Receiver operating characteristic curves for the TCGA-LUAD dataset. TCGA, The Cancer Genome Atlas; LUAD, lung adenocarcinoma; AUC, area under the curve; TP, true-positive; FP, false-positive.

Treg features with non-zero coefficients, including CENPM, PTTG1, IL1R2, GLCC11, BIRC3, MAGEH1, CD5, CISH, ZNF101, IKZF4, ASB2, ZC3H12D and CCR6, were selected ($\lambda_{\min}=0.005779793$) (Fig. 2B and Table SVI). The risk coefficients of CENPM, PTTG1, IL1R2 and BIRC3 were positive and were thus risk factors in the model. The risk coefficients of GLCC11, MAGEH1, CD5, CISH, ZNF101, IKZF4, ASB2, ZC3H12D and CCR6 were negative and were thus protective factors in the model. In addition, the expression levels of these 13 genes in the Treg cells of metastatic (m)LN & tumor lung of advanced-stage LUAD (tL/B), pleural effusion, nLN, nLung, tLung and mBrain groups were detected. It was found that, except CENPM, the other 12 genes were differentially expressed across different anatomical groups. For example, the expression levels of MAGEH1 and CCR6 were the highest

in the tLung group, but the abundances of IKZF4, BIRC3 and PTTG1 were the highest in the mLN & tL/B group (Fig. S1). Subsequently, based on the median risk score, 513 patients with LUAD were grouped into the high- (n=256) and low- (n=257) risk subgroups. As shown in Fig. 2C, the high-risk group had more deaths than the low-risk group. The survival probability of patients in the low-risk group was significantly higher than that in the high-risk group (Fig. 2D). Meanwhile, the area under ROC curves (AUC) for 1-, 2- and 3-years of the risk score model were 0.719, 0.715 and 0.711, respectively, demonstrating the predictive power of the prognostic model (Fig. 2E).

Reliability of the Treg-related prognostic model. To assess the universality of the prognostic model, three independent

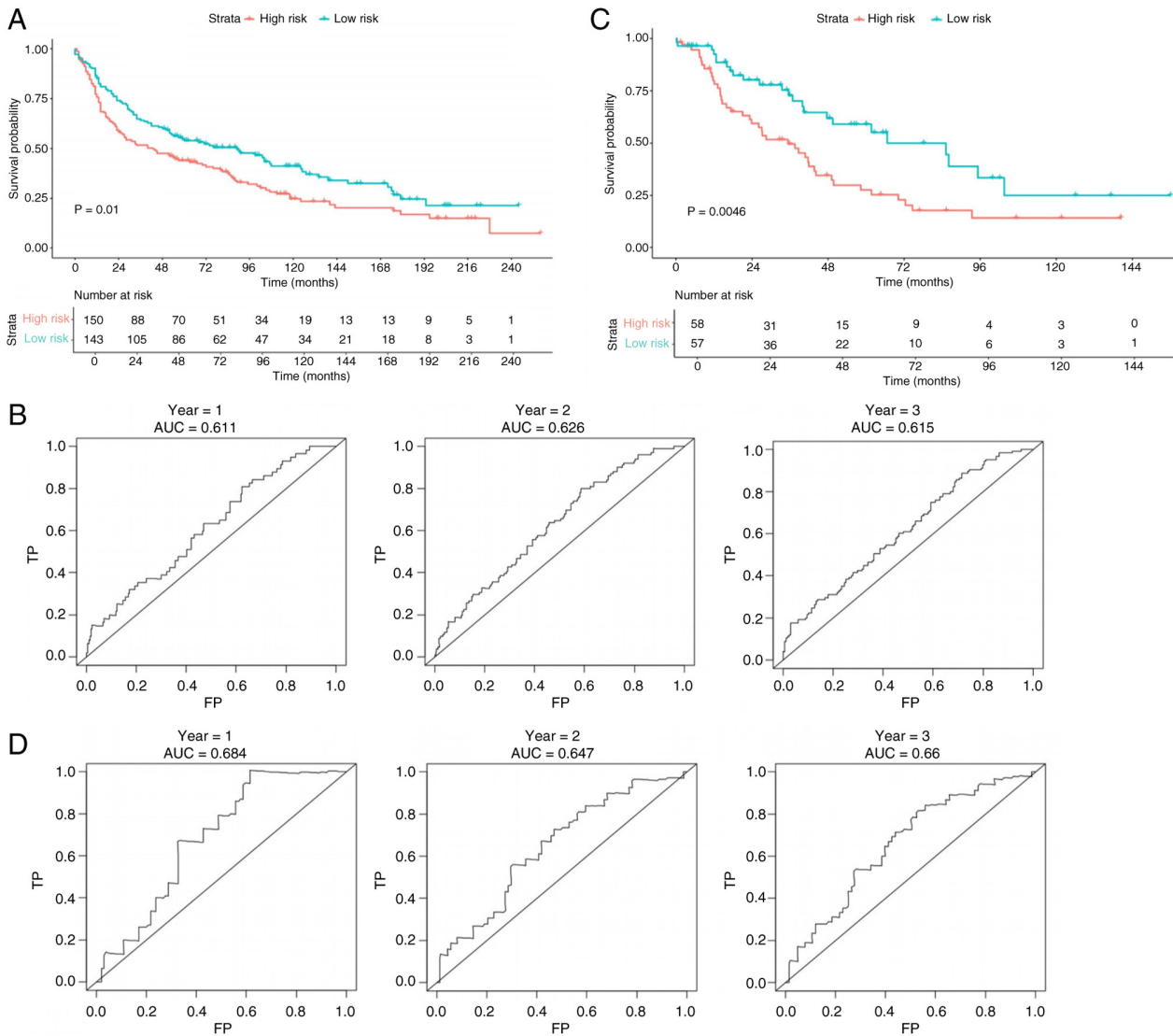


Figure 3. Validation of prognostic model. (A) KM and (B) ROC curves for the GSE30219 dataset. (C) KM and (D) ROC curves for the GSE26939 dataset. KM, Kaplan-Meier; ROC, receiver operating characteristic; AUC, area under the curve; TP, true-positive; FP, false-positive.

datasets were examined, including GSE30219, GSE26939 and GSE72094. In GSE30219, 193 patients with LUAD were classified into the high- (n=150) and low- (n=143) risk subgroups based on the optimal threshold of -5.823151. The OS of the low-risk subgroup was significantly higher than that of the high-risk subgroup (P=0.01; Fig. 3A), and the 1-, 2- and 3-year AUC values of the risk score model were 0.611, 0.626 and 0.615, respectively (Fig. 3B). In GSE26939, 115 patients with LUAD were grouped into the high- (n=58) and low- (n=57) risk subgroups based on the optimal threshold of -0.932411. The patients with LUAD in the low-risk subgroup had a higher OS rate than those in the high-risk group (Fig. 3C) and the AUC values of the risk score model for 1-, 2- and 3-year were 0.684, 0.647 and 0.660, respectively (Fig. 3D). In GSE72094, 398 patients with LUAD were divided into the high- (n=200) and low- (n=198) risk subgroups based on the optimal threshold of -2.738428. The OS of the low-risk subgroup was significantly higher than that of the high-risk subgroup (P=0.00036; Fig. 4A) and the 1-, 2- and 3-year AUC values of the risk score model were 0.616, 0.657 and 0.663, respectively (Fig. 4B).

Treg-related risk score is an independent prognostic factor. In the TCGA-LUAD cohort, the risk score, T stage, N stage and M stage were determined to be associated with the prognosis of patients with LUAD using univariate Cox analysis (Fig. 5A). Moreover, the risk score, T stage and N stage were selected as independent prognostic indicators (Fig. 5B). In addition, an association was demonstrated between the risk score and clinical characteristics of patients with LUAD. Specifically, the risk score of patients with LUAD with stage T3-T4 and N1-N3 were markedly higher than those patients with stage T1-T2 and N1-N3, respectively (Fig. 5C). This indicates that the risk score may be relevant to LUAD development and progression. Moreover, male patients with LUAD had notably higher risk scores than female patients with LUAD (Fig. 5C); however, there was no significant difference in the risk score between the M0 and M1 stages (Fig. 5C).

Difference in the TME between two Treg-related risk groups. To assess differences in the TME of the low- and high-risk groups, the infiltration of 22 immune cells was first evaluated

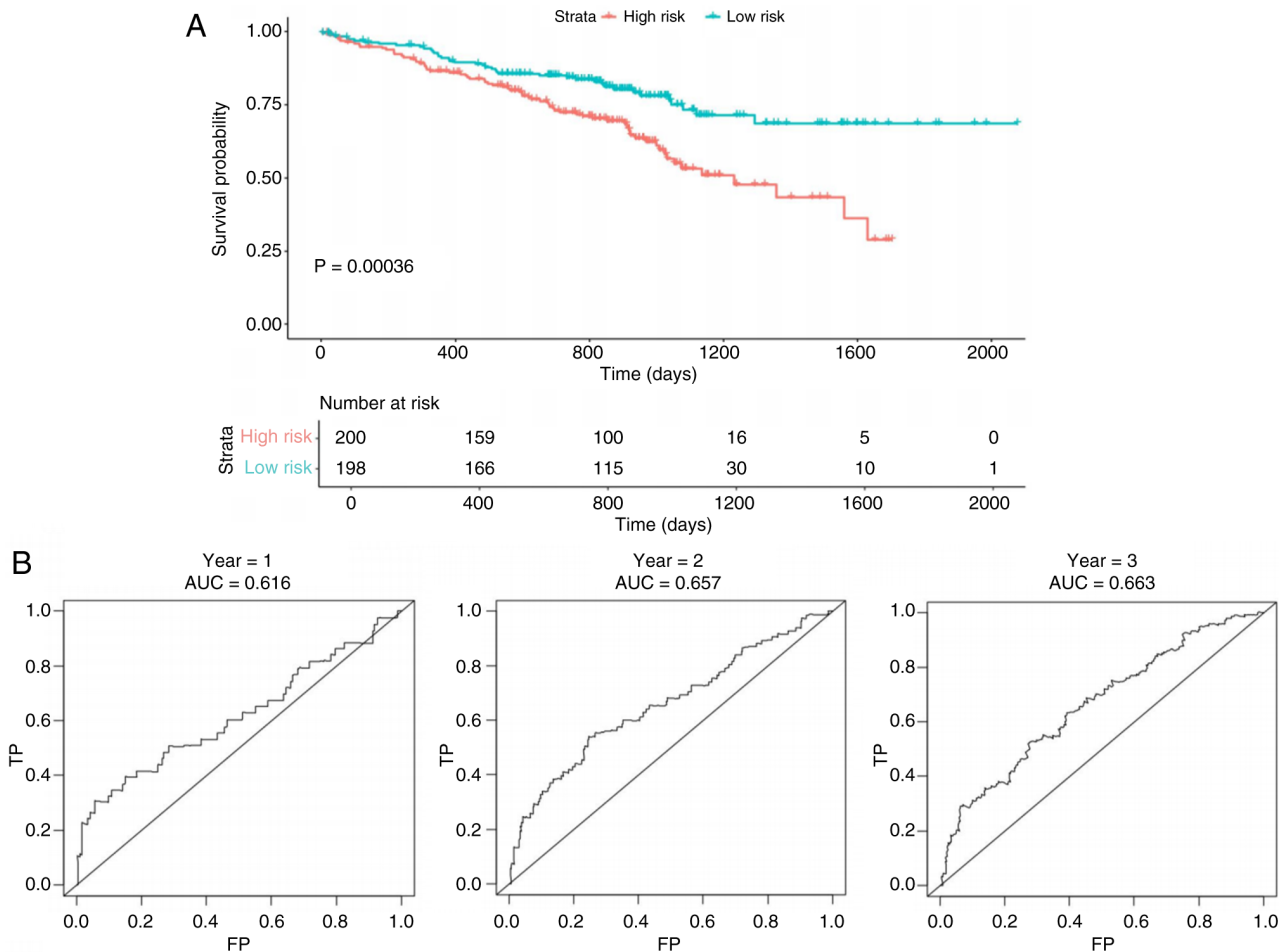


Figure 4. (A) Kaplan-Meier and (B) receiver operating characteristic curves for the GSE72094 dataset. AUC, area under the curve; TP, true-positive; FP, false-positive.

(Fig. 6A). The results revealed that the infiltration of M0 macrophages and activated mast cells in the high-risk group was notably higher, whilst the infiltration of plasma cells and Tregs was markedly lower, compared with the low-risk group (Fig. 6B). Moreover, the high-risk group had lower immune, stromal and ESTIMATE scores than the low-risk group (Fig. 6C), further supporting that the immune environment of the two groups was different.

To further assess the potential effect of the prognostic Treg markers on immunotherapeutic treatment, the correlations between risk score and inflammatory metagenes were then determined. The findings demonstrated that risk score had significant weak correlations with 4 metagene clusters, including IgG (cor=-0.22), interferon (cor=0.22), LCK (cor=-0.31) and MHC-II (cor=-0.17) (Fig. 7A). In addition, the TCR repertoire analysis revealed that the high-risk group had lower TCR Shannon diversity (P<0.0001; Fig. 7B) and lower TCR richness (P<0.0001; Fig. 7C) than the low-risk group, suggesting a reduced capacity to recognize neoantigens in the high-risk group. Furthermore, considering the association between tumor mutation and immune response, the tumor mutation burden (TMB) score was assessed and the results revealed that the TMB of the high-risk group was markedly higher than that of the low-risk group (Fig. 7D). These findings indicate that the prognostic Treg-related

markers may affect immune response and therapeutics in patients with LUAD.

Genes correlated with risk score are mainly involved in metabolism. To further evaluate the role of prognostic Treg-related genes, genes that were strongly correlated with the Treg-related risk score were identified. After correlation analysis, 33 genes were determined to be negatively correlated with the risk score and 4 genes were positively correlated (|cor| > 0.5; adj.P < 0.05; Fig. 8A and Table SVII). These genes were mainly associated with metabolic biological processes, such as ‘ADP catabolic process’, ‘pyridine nucleotide catabolic process’ and ‘purine nucleoside diphosphate catabolic process’ (Fig. 8B and Table SVIII) and KEGG pathways such as ‘Glycolysis/Gluconeogenesis’, ‘HIF-1 signaling pathway’, ‘Carbon metabolism’ and ‘Pyruvate metabolism’ (Fig. 8C and Table SIX). These results suggest the potential role of the prognostic Treg markers in metabolism.

Expression levels of the prognostic genes in LUAD tumors. The expression levels of these genes in the TCGA-LUAD cohort were examined and it was found that the CENPM, PTTG1, ZC3H12D and ZNF101 expression levels were significantly upregulated, while CISH, GLCC11, IKZF4 and MAGEH1 were significantly downregulated in the tumor tissues (Fig. 9A).

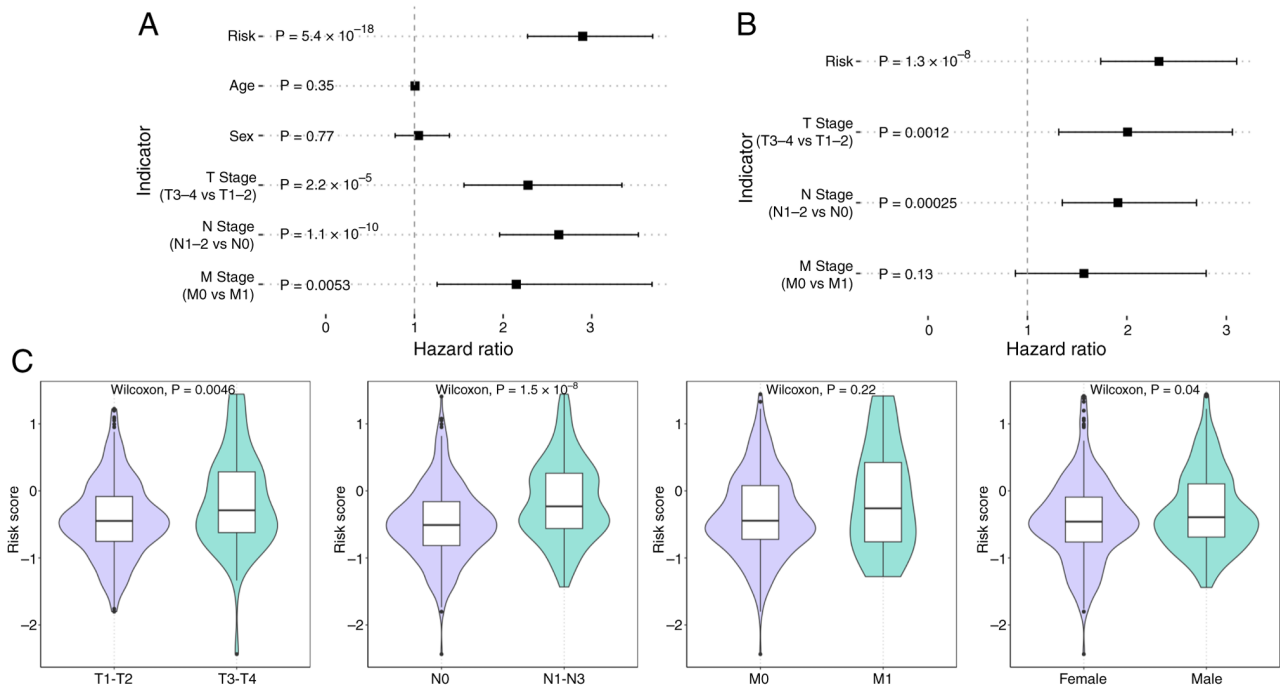


Figure 5. Clinical prognostic model and differences in risk scores between different clinical subgroups in the TCGA-LUAD dataset. (A) Univariate Cox analysis. (B) Multivariate Cox analysis. (C) Differences in the risk scores between different clinical subgroups. T stage, tumor stage; N stage, lymph node stage; M stage, metastasis stage.

Clinical samples were then collected to assess the expression levels of these 8 genes and it was found that the levels were altered consistently with the TCGA-LUAD results (Fig. 9B). The protein expression levels of CENPM, ZNF101, MAGEH1 and IKZF4 in the clinical samples were also examined. The results showed that CENPM and ZNF101 were significantly increased, while MAGEH1 and IKZF4 were reduced in the tumor tissues (Fig. 9C and D). These findings indicate that the expression levels of these genes were reproducible in the independent clinical LUAD samples.

Discussion

At present, T cell-targeted immunotherapy is employed in LUAD tumor therapy (11,14). However, an increase in the Treg infiltration of tumors and metastatic lymph nodes in patients with LUAD accelerates recurrence and metastasis (24,25). In previous LUAD studies, Treg-related genes were often selected from bulk RNA-seq data, either by weighted gene co-expression network analysis to identify modules correlated with estimated Treg infiltration (26) or by first screening survival-associated genes and then correlating them with the estimated Treg infiltration (27,28). The biomarkers from such bulk-first procedures may be Treg infiltration related genes, but not direct Treg markers. In the present study, intratumoral Treg clusters were first delineated at single-cell resolution, which identified 326 Treg-specific markers against other immune cell populations and then only these markers were projected onto bulk cohorts for survival modeling. This Treg-first, sc-RNA-anchored strategy improved cell-type specificity, resulting in more direct and accurate selection for Treg markers. Based on the LASSO-Cox regression analysis, 13 prognostic genes were selected in the TCGA-LUAD

cohort, including CENPM, PTTG1, IL1R2, BIRC3, GLCCI1, MAGEH1, CD5, CISH, ZNF101, IKZF4, ASB2, ZC3H12D and CCR6. Subsequently, a prognostic model was constructed and evaluated based on these 13 genes. The results revealed that the high-risk group had a poorer prognosis than the low-risk group. Furthermore, the low-risk group had notably higher immune, stromal and ESTIMATE scores than the high-risk group.

To the best of our knowledge, CD5, IKZF4, MAGEH1 and ZNF101 have not yet been reported as LUAD prognostic markers. By contrast, ASB2, BIRC3 and IL1R2 have been linked to LUAD prognosis mainly in retrospective transcriptomic analyses with limited direct experimental validation in LUAD. For CISH, CENPM, GLCCI1, CCR6, PTTG1 and ZC3H12D, previous studies have reported their associations with LUAD prognosis together with expression evidence and/or functional validation *in vitro* or *in vivo*. Specifically, among these 13 prognostic genes, the upregulation of CENPM, PTTG1, IL1R2 and BIRC3 was associated with an increase in the risk score of patients with LUAD. Liu *et al* (29) reported that high expression of CENPM is a risk factor to facilitate LUAD progression via the PI3K/AKT/mTOR pathway. Moreover, the expression of PTTG1 was reported to be upregulated in LUAD, which may regulate the p53 signaling pathway (30). Additionally, the expression of IL1R2 was reported to be notably higher in LUAD tumors and that IL1R2 was positively associated with the marker genes of CD8⁺ T cells (31). Moreover, BIRC3 was identified as a feature gene in prognostic models (32,33). In summary, each of these 4 prognostic risk genes have previously been reported to be highly expressed in LUAD tumors. For the other 9 prognostic genes, the upregulation of GLCCI1, MAGEH1, CD5, CISH, ZNF101, IKZF4, ASB2, ZC3H12D and CCR6 was associated

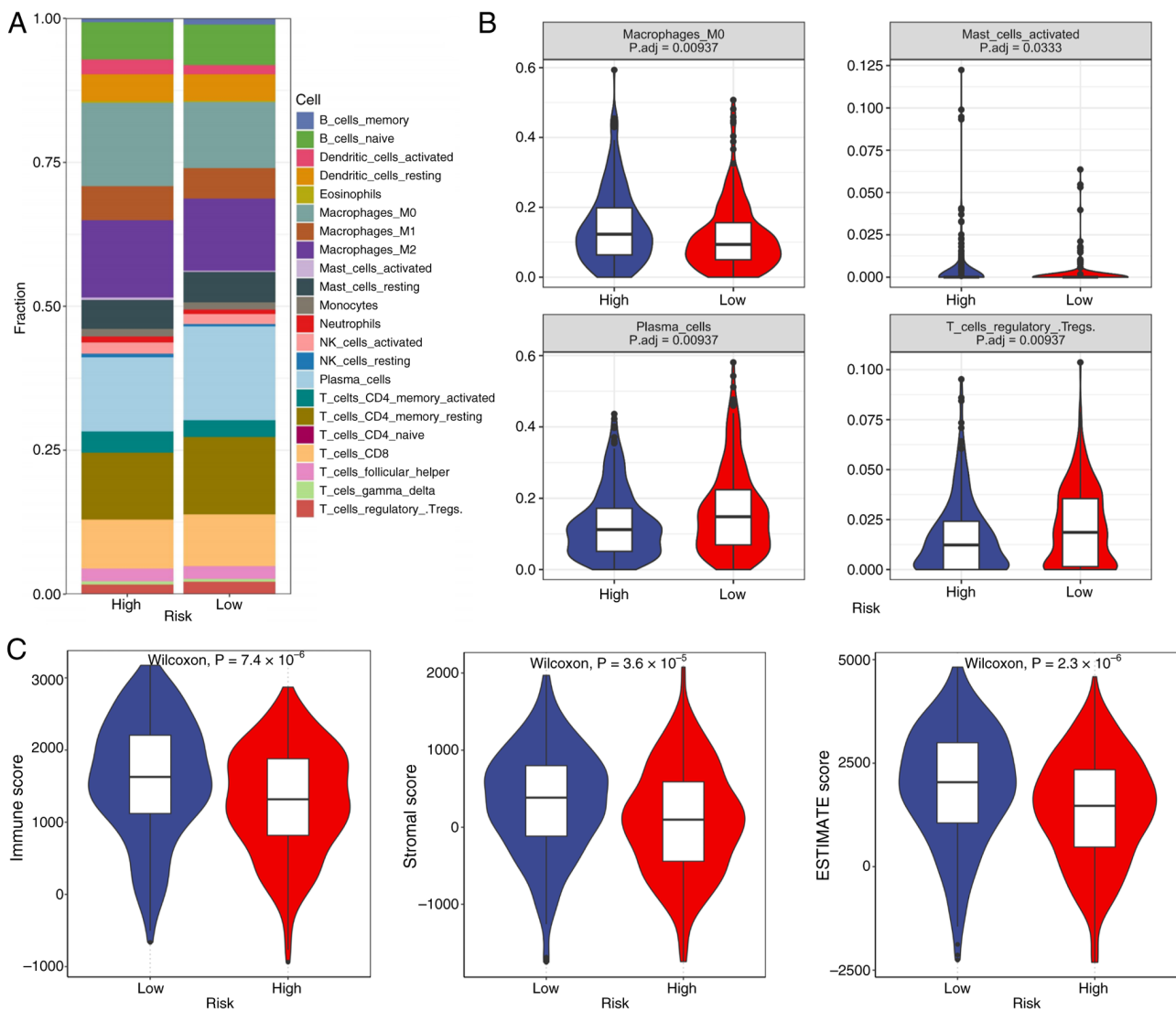


Figure 6. Tumor microenvironment differences between the high- and low-risk groups in the TCGA-LUAD dataset. (A) Infiltration of 22 immune cells. (B) Infiltration differences between the 2 risk groups. (C) Immune, stromal and ESTIMATE scores between the 2 risk groups.

with a decrease in the risk score of patients with LUAD. The expression of *GLCCII* was demonstrated to be downregulated in tumors using RT-qPCR and this suggests that this gene has potential as a prognostic marker for future studies. Moreover, *ZC3H12D* has been reported to be involved in the process of degrading inflammatory transcripts and attenuating macrophage response (34-37). Notably, the prognostic model constructed by Chen *et al* (38) showed that the *ZC3H12D* was significantly associated with a favorable prognosis in LUAD [hazard ratio (HR), 0.607; $P < 0.001$], which was consistent with the findings by Gong *et al* (39) (HR, 0.52; $P < 0.001$) and the present study (HR, 0.444; $P < 0.001$). However, Chen *et al* (38) reported that the expression of *ZC3H12D* demonstrated heterogeneity among different cell populations and that the expression of *ZC3H12D* in the Tregs of patients with LC was much lower. Consequently, we hypothesize that the high expression of *ZC3H12D* in Tregs and the high infiltration of Tregs in LUAD may worsen LUAD. Moreover, there was a negative correlation between *ZC3H12D* and the risk score in Spearman's correlation analysis ($cor = -0.5683$; $P < 0.001$). Finally, *CCR6* controls cell migration and immune

induction in certain inflammatory diseases, which interacts with C-C motif chemokine ligand 20 generating a chemokine receptor-ligand pair (40,41). Wei *et al* (42) reported that the apolipoprotein E⁺-cathepsin Z⁺ tumor-associated macrophage potentially interacts with Tregs via C-X-C motif chemokine ligand 16-CCR6 signals. However, the current evidence is still largely based on retrospective studies and lacks systematic experimental validation in LUAD, warranting further mechanistic and translational studies.

The present study constructed a prognostic model with 13 prognostic genes which had high universality in LUAD and all the AUC values at 1, 2 and 3 years were > 0.6 in four bulk RNA-seq databases. It was further found that, compared with the low-risk patients, high-risk patients had significantly lower Treg infiltration, higher TMB, lower immune score, lower TCR and lower TCR Shannon diversity. It has been reported that a prominent ratio of Tregs compared with conventional T cells within the TME is linked to an unfavorable prognosis, which may be due to Treg-mediated immunosuppression (43). However, the tumor immune microenvironment is a complex system, consisting of various immune cell types, not only

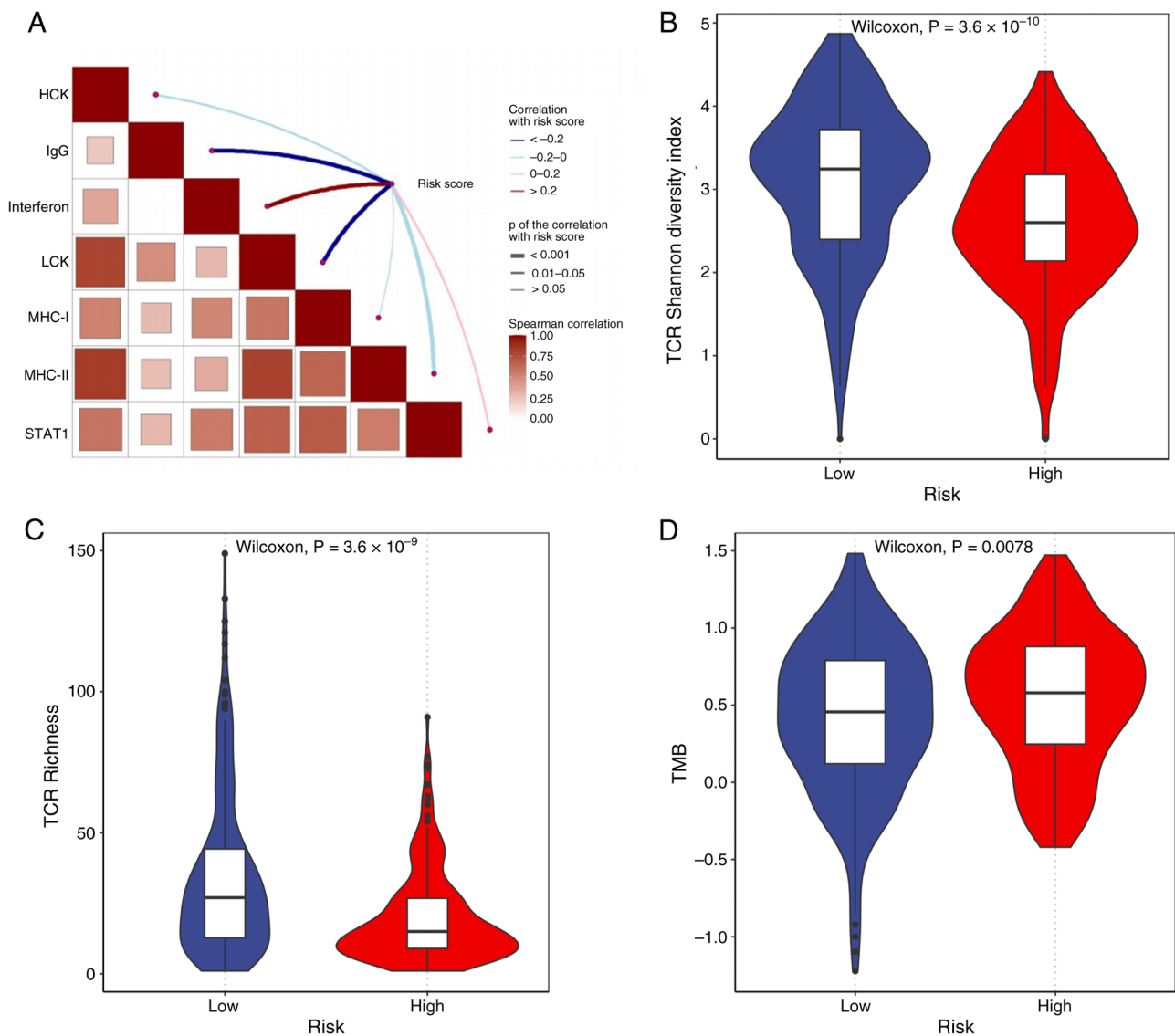


Figure 7. Association between risk score and immune-system-related metagene clusters, TCR richness and TMB in the TCGA-LUAD dataset. (A) Spearman's correlation analysis between 7 immune-system-related metagene clusters and risk score. Association between (B) TCR Shannon diversity, (C) TCR richness and (D) TMB scores between the two risk groups. TCR, T cell receptor; TMB, tumor mutational burden.

Tregs. In the present study, it was also observed that the high-risk group had a significantly lower global immune score, indicating that the immune function of high-risk patients was suppressed. Additionally, the low-risk group had more TCR and TCR Shannon diversity than the high-risk group, suggesting the decreased capacity of high-risk patient to capture and recognize neoantigens. Previous studies have shown that higher TMB levels are associated with an improved prognosis of patients with cancer (44-47). Theoretically, tumor cells with a higher TMB typically have more neoantigens, which could recruit immune cells (including T cells) into the TME. However, in the current study, the high-risk group had a higher TMB score. This may be explained by the fact that the calculated risk score was computed from the expression levels of 13 Treg-related genes using LASSO-derived coefficients and was not stratified by TMB. Therefore, in the present study, TMB, immune scores and survival outcomes should be interpreted as features associated with the risk score rather than as a causal sequence driven by TMB. Similar patterns, namely higher

TMB but a poorer prognosis together with lower immune infiltration scores in the high-risk group defined by immune and anoikis or immune and cell death related genes have been reported in LUAD cohorts (48,49). Notably, unlike these previously published signatures, the present model is derived from Treg-related transcriptional programs and integrates TCR repertoire metrics (richness and Shannon diversity), which provides an additional immunological layer to interpret why high-risk patients may still exhibit poorer outcomes despite a higher TMB. In summary, the heterogeneity between the high- and low-risk groups can further help determine targeted treatment courses for different populations.

From a translational perspective, although the model was derived from tumor tissue transcriptomes, it is based on only 13 genes and could therefore be measured in routine clinical specimens [such as formalin-fixed, paraffin-embedded (FFPE) core biopsies] using RT-qPCR. A practical next step will be to develop and analytically validate a standardized assay and scoring workflow, including normalization and predefined

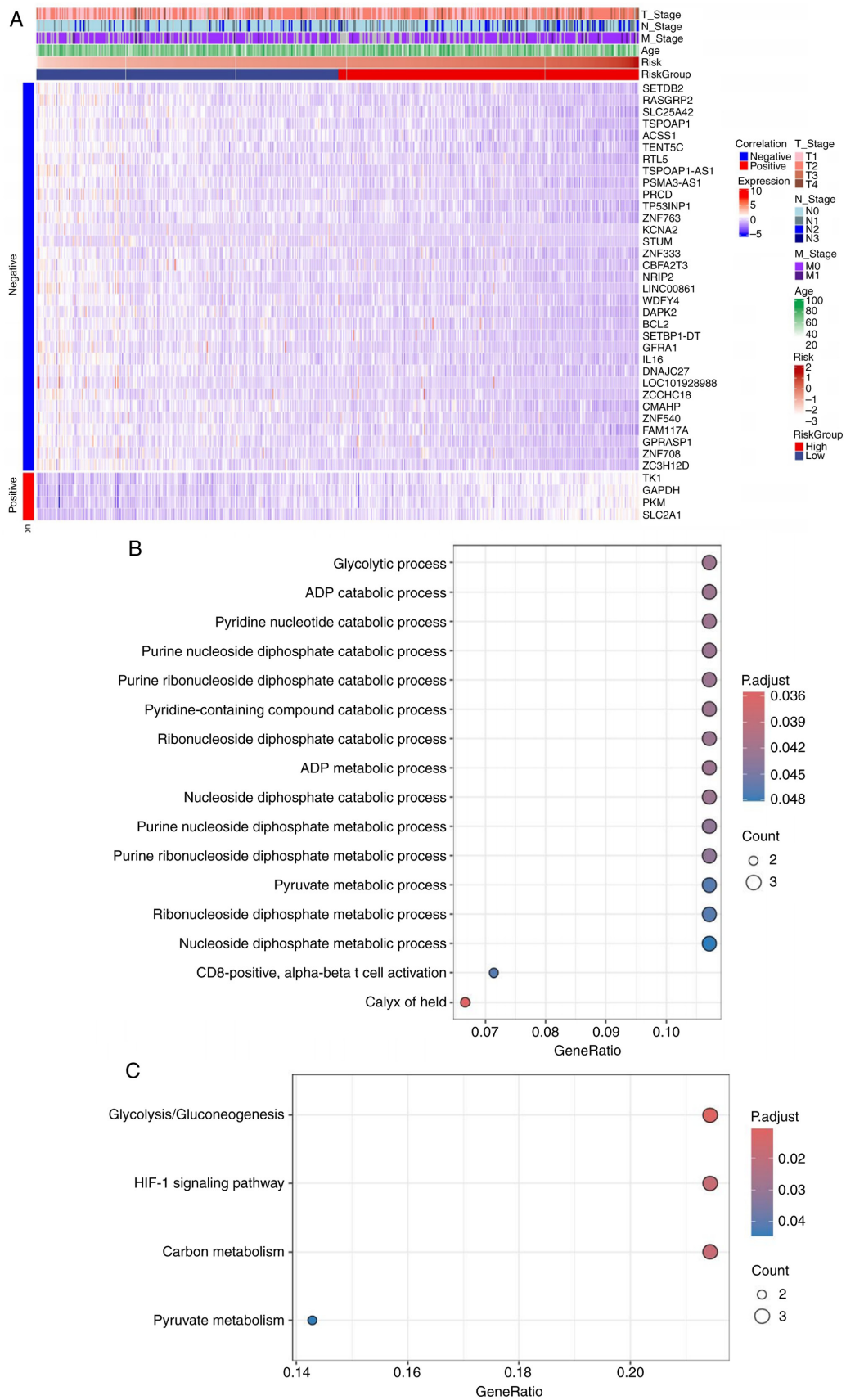


Figure 8. Identification and functional enrichment analysis of risk genes in the TCGA-LUAD dataset. (A) Spearman's correlation analysis between risk genes and risk score. (B) Gene Ontology enrichment analysis of risk genes. (C) Kyoto Encyclopedia of Genes and Genomes enrichment analysis of risk genes.

cut-offs that are compatible with FFPE-derived RNA. In addition, there is the possibility of adapting the signature

to minimally invasive blood-based assays using peripheral blood mononuclear cells and/or plasma cDNA. However, as

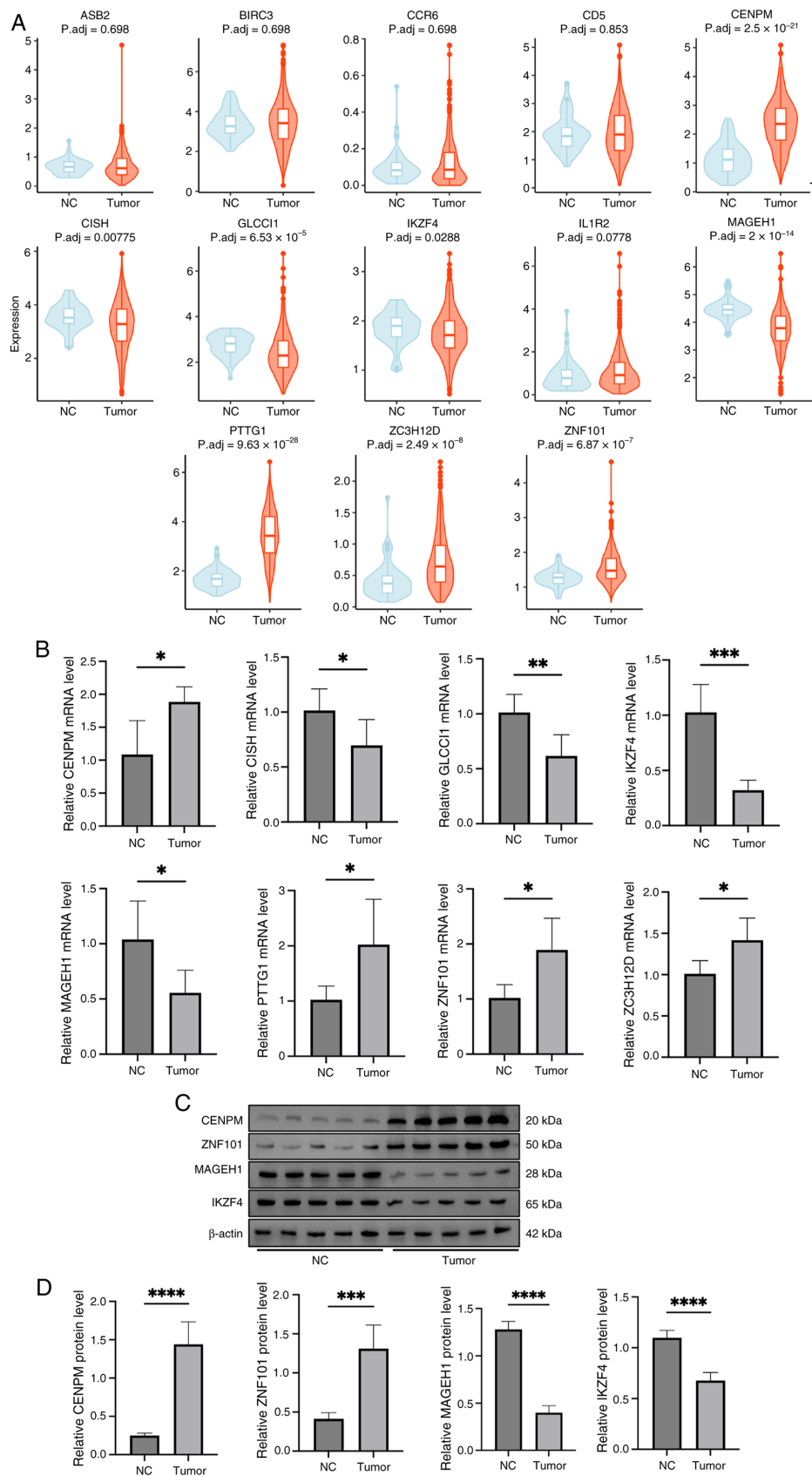


Figure 9. Expression levels of prognostic genes in LUAD tumor and adjacent tissues. (A) The expression levels of 13 prognostic genes in the TCGA-LUAD cohort. (B) The mRNA expression levels of 8 prognostic genes in clinic LUAD tumor and adjacent tissues were examined using quantitative PCR. (C) The protein expression levels of 4 prognostic genes in clinic LUAD tumor and adjacent tissues were examined using western blotting. (D) Semi-quantification of the western blotting results. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$. LUAD, lung adenocarcinoma; TCGA, The Cancer Genome Atlas; NC, normal control; ASB2, ankyrin repeat and SOCS box protein 2; BIRC3, baculoviral IAP repeat containing 3; CCR6, C-C motif chemokine receptor 6; CENPM, centromere protein M; CISH, cytokine inducible SH2 containing protein; GLCC1, glucocorticoid induced 1; IKZF4, Ikaros family zinc finger protein 4; MAGEH1, melanoma-associated antigen H1; PTTG1, pituitary tumor-transforming gene 1 protein; ZNF101, zinc finger protein 101; ZC3H12D, zinc finger CCCH-type containing 12D.

circulating signals may not fully mirror intratumoral immune programs, a bridging study is required. Therefore, in future work, prospective studies collecting paired tumor tissues and blood samples from patients with LUAD should be performed to evaluate concordance between circulating and tissue-based signals and to validate the prognostic utility of a blood-based test.

Nevertheless, there are some limitations in the present study. First, since the findings were derived from bioinformatic retrospective analyses, further validation using different LUAD cohorts in the real-world are needed to accurately examine the prognostic value of the model or any single prognostic gene within the model. Although the mRNA expression of 8 signature genes in paired LUAD and adjacent tissues was confirmed and western blotting for CENPM, ZNF101, MAGEH1 and IKZF4 was performed, protein-level validation was not extended to all remaining signature genes, including PTTG1, ZC3H12D, CISH and GLCCI1. In particular, protein-level validation of all 8 signature genes in larger, independent LUAD cohorts is required to further strengthen the translational robustness of the signature. Second, the high- and low-risk groups have a different TMB and immune microenvironment (such as immune infiltration, immune score and TCR repertoire features). These patterns suggest that the two groups may exhibit different immunotherapy responses. In addition, although the risk score showed statistically significant associations with several immune metagenes, the effect sizes were modest and these findings do not imply biologically meaningful correlations or causality. Therefore, the observed associations between risk score and immune metagenes should be validated in future mechanistic studies and higher-resolution immune profiling. Third, the present study did not include LUAD cohorts treated with immune checkpoint inhibitors (ICI), so this remains hypothesis-generating. To evaluate clinical utility, we plan to validate the prognostic model in independent LUAD ICI cohorts with adequate sample size/power. The analysis will: (i) Stratify patients by the risk score, (ii) compare objective response rate and survival endpoints [progression-free survival/OS] between risk groups, and (iii) use multivariable models that adjust for established covariates (such as programmed cell death 1 ligand 1, TMB and treatment line). Discrimination of the prognostic model will be assessed using the AUC values under the ROC curves and calibration will be evaluated using calibration plots. If pre-specified differences in ICI outcomes are confirmed, the model could then be considered as a predictive biomarker for immunotherapy response in LUAD. Forth, the detailed molecular mechanisms of the identified prognostic genes in regulating the development, progression and therapeutic response of LUAD and their functions in Tregs need to be explored by *in vitro* and *in vivo* experiments. In conclusion, the present study constructed a prognostic model with 13 prognostic genes (CENPM, PTTG1, IL1R2, BIRC3, GLCCI1, MAGEH1, CD5, CISH, ZNF101, IKZF4, ASB2, ZC3H12D and CCR6) by integrating scRNA-seq and bulk RNA-seq technology. Based on the TCGA-LUAD, GSE30219, GSE26939 and GSE72094 datasets, the results demonstrated good predictive performance across multiple datasets and has potential for clinical application.

Acknowledgements

Not applicable.

Funding

No funding was received.

Availability of data and materials

The data generated in the present study may be requested from the corresponding author.

Authors' contributions

HWH and ZMX designed the study, analyzed the data and drafted the manuscript. JJX and YZ performed the western blot experiments and critically revised the manuscript. JQ and PW acquired the qPCR data and critically revised the manuscript. QJH interpreted the results, critically revised the manuscript and supervised the study. HWH and QJH confirm the authenticity of all the raw data. All authors read and approved the final version of the manuscript.

Ethics approval and consent to participate

The experiments involving human participants were reviewed and approved by the Ethics Committee of Shanghai General Hospital Jiuquan Branch (Shanghai, China; approval no. 2025-C-013). The participants provided their written informed consent to participate in the present study.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

- Ren J, Zhang H, Wang J, Xu Y, Zhao L and Yuan Q: Transcriptome analysis of adipocytokines and their-related lncRNAs in lung adenocarcinoma revealing the association with prognosis, immune infiltration, and metabolic characteristics. *Adipocyte* 11: 250-265, 2022.
- Kleczyk EK, Kwak JW, Schenk EL and Nemenoff RA: Targeting the complement pathway as a therapeutic strategy in lung cancer. *Front Immunol* 10: 954, 2019.
- Herbst RS, Morgensztern D and Boshoff C: The biology and management of non-small cell lung cancer. *Nature* 553: 446-454, 2018.
- Schrader M and Laberke HG: Differential diagnosis of verrucous carcinoma in the oral cavity and larynx. *J Laryngol Otol* 102: 700-703, 1988.
- Cheng TY, Cramb SM, Baade PD, Youlden DR, Nwogu C and Reid ME: The international epidemiology of lung cancer: Latest trends, disparities, and tumor characteristics. *J Thorac Oncol* 11: 1653-1671, 2016.
- Kris MG, Gaspar LE, Chaft JE, Kennedy EB, Azzoli CG, Ellis PM, Lin SH, Pass HI, Seth R, Shepherd FA, *et al*: Adjuvant systemic therapy and adjuvant radiation therapy for stage I to IIIA completely resected non-small-cell lung cancers: American society of clinical oncology/cancer care ontario clinical practice guideline update. *J Clin Oncol* 35: 2960-2974, 2017.

7. Reck M, Remon J and Hellmann MD: First-line immunotherapy for non-small-cell lung cancer. *J Clin Oncol* 40: 586-597, 2022.
8. Quail DF and Joyce JA: The microenvironmental landscape of brain tumors. *Cancer Cell* 31: 326-341, 2017.
9. Luo W, Zeng Z, Jin Y, Yang L, Fan T, Wang Z, Pan Y, Yang Y, Yao M, Li Y, *et al*: Distinct immune microenvironment of lung adenocarcinoma in never-smokers from smokers. *Cell Rep Med* 4: 101078, 2023.
10. Lambrechts D, Wauters E, Boeckx B, Aibar S, Nittner D, Burton O, Bassez A, Decaluwé H, Pircher A, Van den Eynde K, *et al*: Phenotype molding of stromal cells in the lung tumor microenvironment. *Nat Med* 24: 1277-1289, 2018.
11. Guo X, Zhang Y, Zheng L, Zheng C, Song J, Zhang Q, Kang B, Liu Z, Jin L, Xing R, *et al*: Global characterization of T cells in non-small-cell lung cancer by single-cell sequencing. *Nat Med* 24: 978-985, 2018.
12. Wu F, Fan J, He Y, Xiong A, Yu J, Li Y, Zhang Y, Zhao W, Zhou F, Li W, *et al*: Single-cell profiling of tumor heterogeneity and the microenvironment in advanced non-small cell lung cancer. *Nat Commun* 12: 2540, 2021.
13. Maitz K, Valadez-Cosmes P, Raftopoulou S, Kindler O, Kienzl M, Bolouri H, Houghton AM, Schicho R, Heinemann A and Kargl J: Altered treg infiltration after discoidin domain receptor 1 (DDR1) inhibition and knockout promotes tumor growth in lung adenocarcinoma. *Cancers (Basel)* 15: 5767, 2023.
14. Liang J, Bi G, Shan G, Jin X, Bian Y and Wang Q: Tumor-associated regulatory T cells in non-small-cell lung cancer: Current advances and future perspectives. *J Immunol Res* 2022: 4355386, 2022.
15. Kim N, Kim HK, Lee K, Hong Y, Cho JH, Choi JW, Lee JI, Suh YL, Ku BM, Eum HH, *et al*: Single-cell RNA sequencing demonstrates the molecular and cellular reprogramming of metastatic lung adenocarcinoma. *Nat Commun* 11: 2285, 2020.
16. Rousseaux S, Debernardi A, Jacquiau B, Vitte AL, Vesin A, Nagy-Mignotte H, Moro-Sibilot D, Brichon PY, Lantuejoul S, Hainaut P, *et al*: Ectopic activation of germline and placental genes identifies aggressive metastasis-prone lung cancers. *Sci Transl Med* 5: 186ra66, 2013.
17. Wilkerson MD, Yin X, Walter V, Zhao N, Cabanski CR, Hayward MC, Miller CR, Socinski MA, Parsons AM, Thorne LB, *et al*: Differential pathogenesis of lung adenocarcinoma subtypes involving sequence mutations, copy number, chromosomal instability, and methylation. *PLoS One* 7: e36530, 2012.
18. Schabath MB, Welsh EA, Fulp WJ, Chen L, Teer JK, Thompson ZJ, Engel BE, Xie M, Berglund AE, Creelan BC, *et al*: Differential association of STK11 and TP53 with KRAS mutation-associated gene expression, proliferation and immune surveillance in lung adenocarcinoma. *Oncogene* 35: 3209-3216, 2016.
19. Cancer Genome Atlas Research Network: Comprehensive molecular profiling of lung adenocarcinoma. *Nature* 511: 543-550, 2014.
20. Rody A, Holtrich U, Pusztai L, Liedtke C, Gaetje R, Ruckhaeberle E, Solbach C, Hanker L, Ahr A, Metzler D, *et al*: T-cell metagene predicts a favorable prognosis in estrogen receptor-negative and HER2-positive breast cancers. *Breast Cancer Res* 11: R15, 2009.
21. Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH, Porta-Pardo E, Gao GF, Plaisier CL, Eddy JA, *et al*: The immune landscape of cancer. *Immunity* 51: 411-412, 2019.
22. Yu G, Wang LG, Han Y and He QY: clusterProfiler: An R package for comparing biological themes among gene clusters. *OMICS* 16: 284-287, 2012.
23. Livak KJ and Schmittgen TD: Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25: 402-408, 2001.
24. Erfani N, Mehrabadi SM, Ghayumi MA, Haghshenas MR, Mojtahedi Z, Ghaderi A and Amami D: Increase of regulatory T cells in metastatic stage and CTLA-4 over expression in lymphocytes of patients with non-small cell lung cancer (NSCLC). *Lung Cancer* 77: 306-311, 2012.
25. Schneider T, Kimpfler S, Warth A, Schnabel PA, Dienemann H, Schadendorf D, Hoffmann H and Umansky V: Foxp3(+) regulatory T cells and natural killer cells distinctly infiltrate primary tumors and draining lymph nodes in pulmonary adenocarcinoma. *J Thorac Oncol* 6: 432-438, 2011.
26. Wang X, Xiao Z, Gong J, Liu Z, Zhang M and Zhang Z: A prognostic nomogram for lung adenocarcinoma based on immune-infiltrating Treg-related genes: From bench to bedside. *Transl Lung Cancer Res* 10: 167-182, 2021.
27. Sun S, Wang K, Guo D, Zheng H, Liu Y, Shen H and Du J: Identification of the key DNA damage response genes for predicting immunotherapy and chemotherapy efficacy in lung adenocarcinoma based on bulk, single-cell RNA sequencing, and spatial transcriptomics. *Comput Biol Med* 171: 108078, 2024.
28. Fan Z, Wu S, Sang H, Li Q, Cheng S and Zhu H: Identification of GPD1L as a potential prognosis biomarker and associated with immune infiltrates in lung adenocarcinoma. *Mediators Inflamm* 2023: 9162249, 2023.
29. Liu C, Wang Y, Dao Y, Wang S, Hou F, Yang Z, Liu P, Lv J, Lv L, Li G, *et al*: Upregulation of CENPM facilitates lung adenocarcinoma progression via PI3K/AKT/mTOR signaling pathway. *Acta Biochim Biophys Sin (Shanghai)* 54: 99-112, 2022.
30. Bai L, Li LH, Liang J and Li EX: Prognostic significance of PTTG1 and its methylation in lung adenocarcinoma. *J Oncol* 2022: 3507436, 2022.
31. Zhang Y, Ma D, Gong Y, Wang F, Wu J and Wu C: IL1R2 is a novel prognostic biomarker for lung adenocarcinoma. *Curr Mol Med* 24: 620-629, 2024.
32. Wu S, Pan J, Pan Q, Zeng L, Liang R and Li Y: Multi-omics profiling and experimental verification of tertiary lymphoid structure-related genes: Molecular subgroups, immune infiltration, and prognostic implications in lung adenocarcinoma. *Front Immunol* 15: 1453220, 2024.
33. Lei K, Tan B, Liang R, Lyu Y, Wang K, Wang W, Wang K, Hu X, Wu D, Lin H and Wang M: Development and clinical validation of a necroptosis-related gene signature for prediction of prognosis and tumor immunity in lung adenocarcinoma. *Am J Cancer Res* 12: 5160-5182, 2022.
34. Liang J, Wang J, Azfer A, Song W, Tromp G, Kolattukudy PE and Fu M: A novel CCCH-zinc finger protein family regulates proinflammatory activation of macrophages. *J Biol Chem* 283: 6337-6346, 2008.
35. Wang M, Vikis HG, Wang Y, Jia D, Wang D, Bierut LJ, Bailey-Wilson JE, Amos CI, Pinney SM, Petersen GM, *et al*: Identification of a novel tumor suppressor gene p34 on human chromosome 6q25.1. *Cancer Res* 67: 93-99, 2007.
36. Wawro M, Kochan J, Krzanik S, Jura J and Kasza A: Intact NYN/PIN-Like domain is crucial for the degradation of inflammation-related transcripts by ZC3H12D. *J Cell Biochem* 118: 487-498, 2017.
37. Yang B, Ji LL, Xu HL, Li XP, Zhou HG, Xiao T, Li XH, Gao ZY, Li JZ, Zhang WD, *et al*: Zc3h12d, a novel of hypomethylated and immune-related for prognostic marker of lung adenocarcinoma. *J Inflamm Res* 14: 2389-2401, 2021.
38. Chen W, Guo Z, Wu J, Lin G, Chen S, Lin Q, Yang J, Xu Y and Zeng Y: Identification of a ZC3H12D-regulated competing endogenous RNA network for prognosis of lung adenocarcinoma at single-cell level. *BMC Cancer* 22: 115, 2022.
39. Gong W, Dai W, Wei H, Chen Y and Zheng Z: ZC3H12D is a prognostic biomarker associated with immune cell infiltration in lung adenocarcinoma. *Transl Cancer Res* 9: 6128-6142, 2020.
40. Ranasinghe R and Eri R: pleiotropic immune functions of chemokine receptor 6 in health and disease. *Medicines (Basel)* 5: 69, 2018.
41. Zhang CY, Qi Y, Li XN, Yang Y, Liu DL, Zhao J, Zhu DY, Wu K, Zhou XD and Zhao S: The role of CCL20/CCR6 axis in recruiting Treg cells to tumor sites of NSCLC patients. *Biomed Pharmacother* 69: 242-248, 2015.
42. Wei J, Yu W, Chen J, Huang G, Zhang L, Chen Z, Hu M, Gong X and Du H: Single-cell and spatial analyses reveal the association between gene expression of glutamine synthetase with the immunosuppressive phenotype of APOE+CTS+TAM in cancers. *Mol Oncol* 17: 611-628, 2023.
43. Jiang J, Chen H, Zhao C, Li T, Zhang C, Ma L, Su H, Ma L, Duan Z, Si Q, *et al*: PRTN3 promotes IL33/Treg-mediated tumor immunosuppression by enhancing the M2 polarization of tumor-associated macrophages in lung adenocarcinoma. *Cancer Lett* 616: 217584, 2025.
44. Wang H, Ma J, Lu J, Wang Y, Zhang B, Zhang H and Peng H: TMB is associated with the prognosis of egfr-mutated non-small cell lung cancer in Xuanwei, China. *Biomark Med* 18: 1123-1133, 2024.
45. Park SE, Park K, Lee E, Kim JY, Ahn JS, Im YH, Lee C, Jung H, Cho SY, Park WY, *et al*: Clinical implication of tumor mutational burden in patients with HER2-positive refractory metastatic breast cancer. *Oncoimmunology* 7: e1466768, 2018.

46. Goodman AM, Kato S, Bazhenova L, Patel SP, Frampton GM, Miller V, Stephens PJ, Daniels GA and Kurzrock R: Tumor mutational burden as an independent predictor of response to immunotherapy in diverse cancers. *Mol Cancer Ther* 16: 2598-2608, 2017.
47. Zhang C, Shen L, Qi F, Wang J and Luo J: Multi-omics analysis of tumor mutation burden combined with immune infiltrates in bladder urothelial carcinoma. *J Cell Physiol* 235: 3849-3863, 2020.
48. Zhang JL, Dong YX, Di SY, Fan BS and Gong TQ: Identification and experimental verification of an anoikis and immune related signature in prognosis for lung adenocarcinoma. *Transl Cancer Res* 12: 887-903, 2023.
49. Xie Y, Chen H, Tian M, Wang Z, Wang L, Zhang J, Wang X and Lian C: Integrating multi-omics and machine learning survival frameworks to build a prognostic model based on immune function and cell death patterns in a lung adenocarcinoma cohort. *Front Immunol* 15: 1460547, 2024.



Copyright © 2026 He et al. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) License.