Articles

Dynamic Profiling of the Tumor Microenvironment in Lung Squamous Cell Carcinom by Multidimensional Transcriptomics

DOI:https://dx.doi.org/10.71373/WEQA7681
Submitted 24 December 2025
Accepted 25 December 2025
Pulished 31 December 2025

Dandan Sun[1], Zixuan Liu[1], Jiaqi Gu[1], Quanlei Wang[2*], Anxin Gu[1*]

Therapeutic options for lung squamous cell carcinoma (LSCC) remain limited and vary by TNM stage. Achieving precise, stage-dependent treatment requires an understanding of how the tumor microenvironment (TME) changes during disease progression. Therefore, we investigated stage-related TME changes and explored prognostic strategies based on multidimensional transcriptomic data. In this study, we comprehensively characterize the TME during different stages of LSCC progression by integrating single-cell RNA sequencing (scRNA-seq), spatial transcriptomics (ST), and bulk RNA sequencing (bulk RNA-seq). The TME of LSCC displayed marked heterogeneity across TNM stages. Notably, the proportion of SPP1⁺ macrophages increased progressively from early to late disease and was associated with heightened immunosuppressive activity. We constructed a prognostic model based on differentially expressed genes in SPP1⁺ macrophages between stage IV and stage I. We found SPP1⁺ macrophages and fibroblasts exhibited the most frequent ligand-receptor interactions at the tumor edge. Our study revealed stage-dependent remodeling of the LSCC TME and highlighted SPP1⁺ macrophage subgroup as key contributors to tumor progression. The SPP1⁺ macrophage–based prognostic model may facilitate risk stratification and inform future therapeutic strategies.

Introduction

Lung squamous cell carcinoma (LSCC) is a major histologic subtype of non-small cell lung cancer (NSCLC)[1]. Despite considerable advances in lung cancer therapy, effective LSCC-specific options remain limited and vary significantly across TNM stages. For metastatic LSCC, first-line chemo-immunotherapy achieves a median overall survival (OS) of only 17.1 months[2]. Therefore, further delineation of the cellular and molecular interaction characteristics within the immune niche of normal lung tissue and during the progression of LSCC is crucial for precisely clinical diagnosis and therapy options. Moreover, it is important to explore more effective treatments and clinically applicable prognostic models to achieve precise individualized treatment.

Previous studies showed that the tumor microenvironment (TME) of LSCC comprises epithelial cells, infiltrating immune cells, stromal elements, other resident cell types, and non-cellular components that collectively influence tumor initiation and progression[3]. However, the TME exhibits high heterogeneity within immune cell populations, including the presence of specific immune cell subsets that play stage-specific roles during tumor progression. Therefore, further dissection of the molecular features and spatiotemporal interaction heterogeneity of these specific immune subsets is critical for understanding the pathogenesis of LSCC and optimizing immunotherapy strategies.

Single-cell RNA sequencing (scRNA-seq) enables high-resolution profiling of thousands of cells per sample, providing critical insight into intratumoral heterogeneity and the cellular complexity of the TME[4]. Furthermore, spatial transcriptomics (ST) can provide the spatial distribution of various cell types[5]. Recent studies have applied scRNA-seq to LSCC: Zhang et al. identified distinct basal-cell subclusters within LSCC[6]; Sandra et al. examined basal-cell clonal dynamics and showed how fate shifts contribute to early LSCC development[7]; and Wu et al. delineated the TME landscape of advanced-stage LSCC[8]. However, these investigations are incomplete with respect to disease progression and do not comprehensively assess how the LSCC TME changes across pathological stages I−IV. To our knowledge, no study has systematically profiled stage-dependent TME remodeling throughout the full course of LSCC.

To elucidate the changes in LSCC progression, in the presented study, we integrated scRNA-seq datasets from LSCC stages I to IV to delineate the TME atlas. This stage-dependent atlas delineates the cellular heterogeneity and remodeling of the TME during disease progression. Within this atlas, we identified a tumor-associated macrophage (TAM) subset—SPP1⁺ TAMs—associated with immunosuppressive TME, and constructed a prognostic model based on their differential features. We further used publicly available ST datasets, which demonstrated the localization and interactions of SPP1⁺ macrophages within the TME. Collectively, these analyses provide a detailed molecular view of TME remodeling across LSCC progression and underscore the important role of SPP1⁺ macrophages.

Materials and Methods

Data Collection

We obtained scRNA-seq data from 53 tumor and adjacent normal tissue samples from 17 untreated male patients with LSCC through three databases: the Gene Expression Omnibus (GEO) database (GSE194070), the National Genomics Data Center (https://ngdc.cncb.ac.cn/cancerscem), and the lung cancer database of West China Hospital, Sichuan University (http://lungcancer.chenlulab.com). These datasets included tumor samples categorized by TNM stages I, II, III, and IV (5,

1Department of Radiation Oncology, Harbin Medical University Cancer Hospital, 150040, Harbin, China.
2Dongguan Institute of Gallbladder Disease Research, Dongguan Nancheng Hospital, Dongguan, 523000, China.
** Corresponding author:*
Anxin Gu, Department of Radiation Oncology, Harbin Medical University Cancer Hospital, Harbin, China. E-mail: guanxin@hrbmu.edu.cn
Quanlei Wang, Dongguan Institute of Gallbladder Disease Research, Dongguan Nancheng Hospital, Dongguan, China. E-mail: 408470802@qq.com

14, 6, and 10 tissues, respectively), along with 18 adjacent normal tissue samples. The ST data used in the analysis were obtained from BioStudies (https://www.ebi.ac.uk/biostudies/) with accession number E-MTAB-13530. Bulk RNA-seq data comes from the TCGA database (https://portal.gdc.cancer.gov/).

Single-cell RNA-seq Data Preprocessing

We processed scRNA-seq data in R (version 4.4.1) with Seurat[9] (version 5.1.0). Quality control removed cells with <200 detected genes or >20% mitochondrial expression, and genes present in <3 cells. Datasets were integrated and batch-corrected using Harmony[10] (version 1.2.1). After normalization, we identified 2,000 highly variable genes (HVGs), computed S and G2/M scores, and regressed cell-cycle effects (CellCycleScoring /ScaleData). PCA on scaled HVGs yielded 50 principal components (PCs) used to construct a shared nearest-neighbor (SNN) graph (FindNeighbors) and perform clustering (FindClusters). Uniform Manifold Approximation and Projection (UMAP) was used for dimensionality reduction and visualization.

Cell Type Annotation

Cell clusters were annotated based on canonical marker genes[11], including eight distinct cell types: T/NK cells (CD3D), B cells (CD79A), macrophages (CD14), dendritic cells (CD1C), neutrophils (CSF3R), mast cells (TPSAB1), epithelial cells (EPCAM), and fibroblasts (COL1A1).

Identification of Cancer Cells and Cell Stemness Analysis

To identify cancer cells, we used epithelial cells from normal control samples as reference cells and employed Inference of Copy Number Variations (inferCNV)[12] (version 1.20.0) to infer copy number variations (CNVs) in epithelial cells from tumor samples. CopyKAT[13] was used to automatically identify diploid cells as normal cells. For each non-diploid cell, breakpoints of CNVs were identified, and segments were determined using a Markov Chain Monte Carlo (MCMC) method. Normal and tumor cells were distinguished based on their distinct gene expression profiles. Finally, the results from inferCNV and CopyKAT were combined to define the tumor cell populations within the epithelial cells. Furthermore, cell stemness and differentiation potential was inferred using the CytoTRACE[14] (version 0.3.3) software package.

Construction and validation of prognostic risk model

We selected genes that were highly expressed in SPP1+ macrophages in stage IV compared to stage I. Using the "glmnet" package[15], we performed a least absolute shrinkage and selection operator (LASSO) Cox proportional hazards regression analysis on genes that overlapped with those in the TCGA-LUSC dataset. We compiled a list of genes with nonzero coefficients, and the resulting risk model was carefully constructed by linearly summing the products of the genes and their corresponding risk coefficients. Univariate Cox regression analysis was performed on these genes in conjunction with prognostic information from the TCGA-LUSC dataset to validate their prognostic association. Multivariate Cox regression analysis was performed using age, TNM stage, and risk score to demonstrate the independent predictive power of the model. Patients were stratified into low-risk or high-risk groups based on the median risk score threshold. To systematically validate the prognostic performance of the model, we calculated the area under the curve (AUC) using the "timeROC" package[16]. Survival analysis was also performed based on the Kaplan-Meier method. The predictive robustness of the model was rigorously validated using AUC calculations from other GEO datasets.

Gene Enrichment Analysis

Gene Ontology (GO) term enrichment analysis was performed using the clusterProfiler R package[17] (version 4.12.6). For each cell cluster, marker genes with a fold change > 1.5 and adjusted P value < 0.01 were annotated based on the biological process category of GO terms. GO terms with an adjusted P value < 0.05 were considered statistically significant.

Survival Analysis

To evaluate the prognostic value of SPP1 expression, patients were stratified into high- and low-expression groups based on the optimal cutoff value determined by the surv_cutpoint() function from the survminer R package[18] (version 0.4.9). This function identifies the cutpoint that best separates patients according to overall survival using maximally selected rank statistics. Patients were then categorized using the surv_categorize() function, and survival differences between groups were assessed using Kaplan–Meier survival analysis and the log-rank test. Overall survival time was measured in months.

SCENIC Analysis

Single Cell Regulatory Network Inference and Clustering (SCENIC)[19] (version 1.3.1) analysis was performed in R to infer transcription factor (TF) regulatory networks, following the standard workflow. Input was the normalized gene expression matrix from Seurat. Regulon activity was quantified using AUC scores to reflect TF activity across cells.

Pathway Activity Scoring via GSVA

To assess hallmark pathway activity across macrophage subclusters, we performed GSVA with the GSVA R package[20] (version 1.52.3). Hallmark gene sets for Homo sapiens were obtained from MSigDB via msigdbr[21] (version 7.5.1) for Homo sapiens. Seurat's AverageExpression() generated subcluster-level expression matrices by averaging gene expression per gene, which served as GSVA input. Enrichment scores were visualized as heatmaps using pheatmap with row-wise Z-score normalization to highlight relative activity differences.

CellChat Analysis

Cell – cell communication was analyzed with CellChat[22] (version 1.6.1). A CellChat object was created from the merged Seurat object, using CellChatDB's Secreted Signaling and Cell–Cell Contact references. We computed communication probabilities with computeCommunProb, inferring and aggregating intercellular networks under default settings. Interaction strength was visualized to show aggregated networks and each cell type's outgoing signals.

Cell type deconvolution and annotation

We used the RCTD[23] (Deconvolution from Spatial Transcriptome Data) method to deconvolute and annotate cell types from ST data. RCTD is an innovative computational framework designed to leverage spatial transcriptome data and scRNA-seq datasets. In this analysis, we used scRNA-seq data from LSCC as a reference and inferred the cell abundance of each point in the ST data.

Spatial niche and cell distance analysis

To comprehensively understand the spatial niches and cell-to-cell interactions of cells within a sample, we used the MISTy (Multiview InterCellular Spatial Modeling Framework) approach. MISTy analyzes cell-to-cell interactions through different spatial environments, enabling the identification and understanding of the formation and changes of spatial niches. In this study, we used the MISTyR[24] (version 1.12.0) package to analyze spatial niches in ST data based on the spatial distances between cells.

Results

Single-cell landscape across differe- nt pathological stages of LSCC

The study workflow is shown in Figure 1A. After quality control, we retained 169,408 cells from 53 lung tissue samples and partitioned them into eight major lineages, based on canonical marker genes expression (Figure 1B, C): T/NK cells, B cells, macrophages, dendritic cells (DCs), neutrophils, mast cells, epithelial cells, and fibroblasts. Compared to adjacent normal lung, tumors exhibited lower infiltration of T cells and NK cells, enrichment of B cells, and reduced infiltration of macrophages and DCs (Supplementary Figure 1A). Interestingly, the abundances of inferred cell populations, for instance epithelial cells, T/NK cells, macrophages etc. in the TME of LSCC also varied markedly with advancing TNM stage (Figure 1D). Collectively, these shifts in cellular composition suggest stage-dependent remodeling of the LSCC tumor microenvir- onment that may contribute to disease progression.

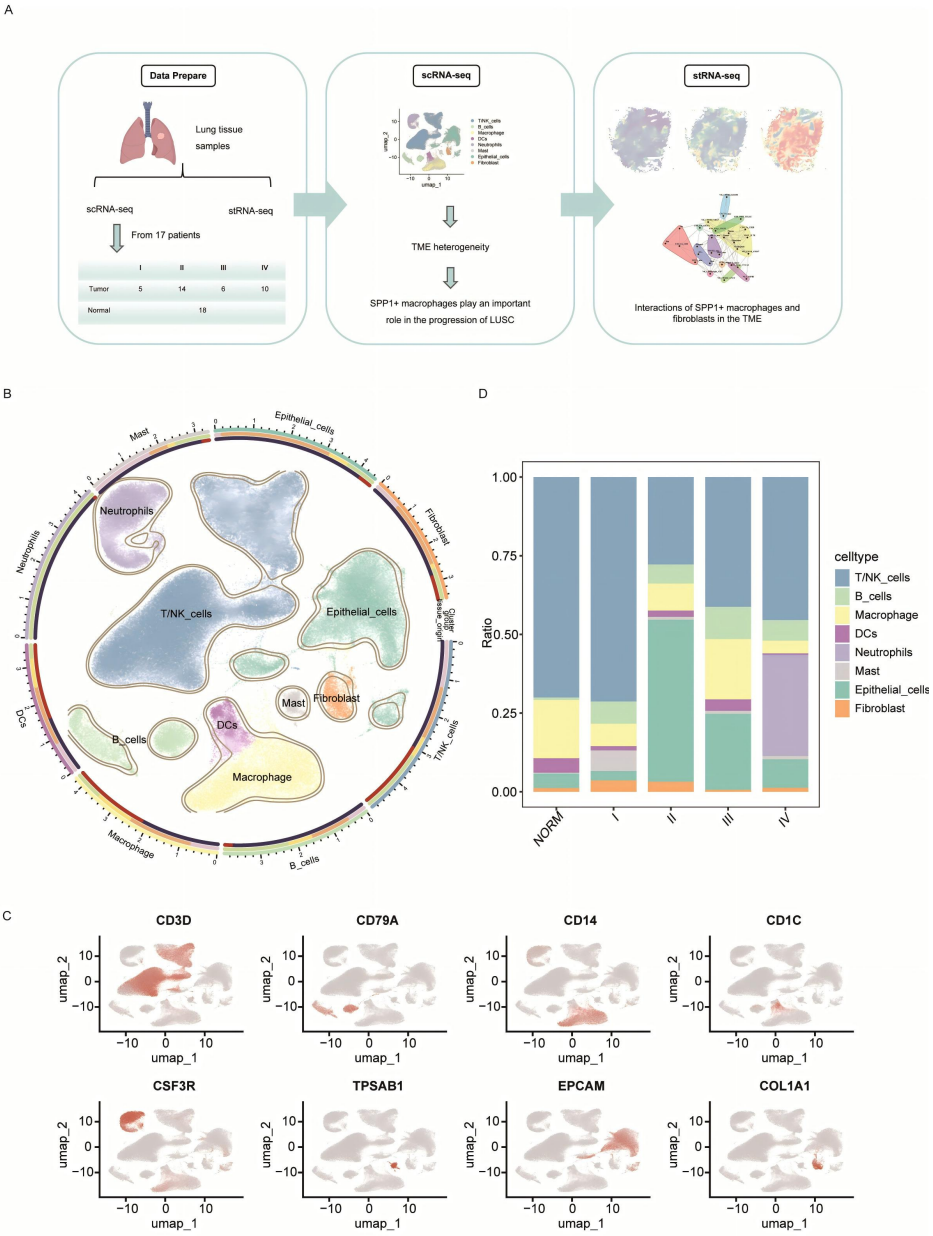

Figure 1. Workflow and single-cell RNA sequencing analysis reveal the TME of LSCC. A. Overall workflow diagram; B. UMAP plot showing cell clustering across 53 samples, with different colors representing different cell types; C. UMAP plot showing classic marker genes for each cell type; D. Histogram showing the proportions of various cell types across TNM stages and adjacent normal tissue.

Epithelial Cell Heterogeneity During LSCC Progression

To futher dissect the epithelial heterogeneity in LSCC, we applied inferCNV to 29,782 epithelial cells using matched normal epithelial cells as the reference. As expected, both the inferCNV heatmap (Figure 2A) and CNV scoring (Supplementary Figure 2A) revealed stage-dependent CNV patterns. Then, integration of CopyKAT results (Figure 2B) with inferCNV outputs enabled robust separation of malignant from normal epithelial cells. Futher, Reclustering of malignant cells classified into seven subclusters (Figure 2C). CytoTRACE analysis demonstrated stemness heterogeneity arcoss TNM stage (Figure 2D), with T_C1 exhibiting the highest stemness potential pattern (Figure 2E). Differential expression analysis highlighted up-regulated genes of T_C1, such as *IDO1*, *POLR2J3.1*, *FUT3*, *CLCA3P*, *TMPRSS2*, and *LRG1* (Figure 2F). Consistently, Hallmark GSVA showed T_C1 to be preferentially enriched for hypoxia, EMT, TGF-β, Hedgehog, and KRAS signaling relative to other epithelial clusters (Supplementary Figure 2B). Together, these findings strongly indicate that heterogeneity of tumor cells at different TNM stages and T_C1 represents a stem-like epithelial subpopulation within LSCC tumors niche.

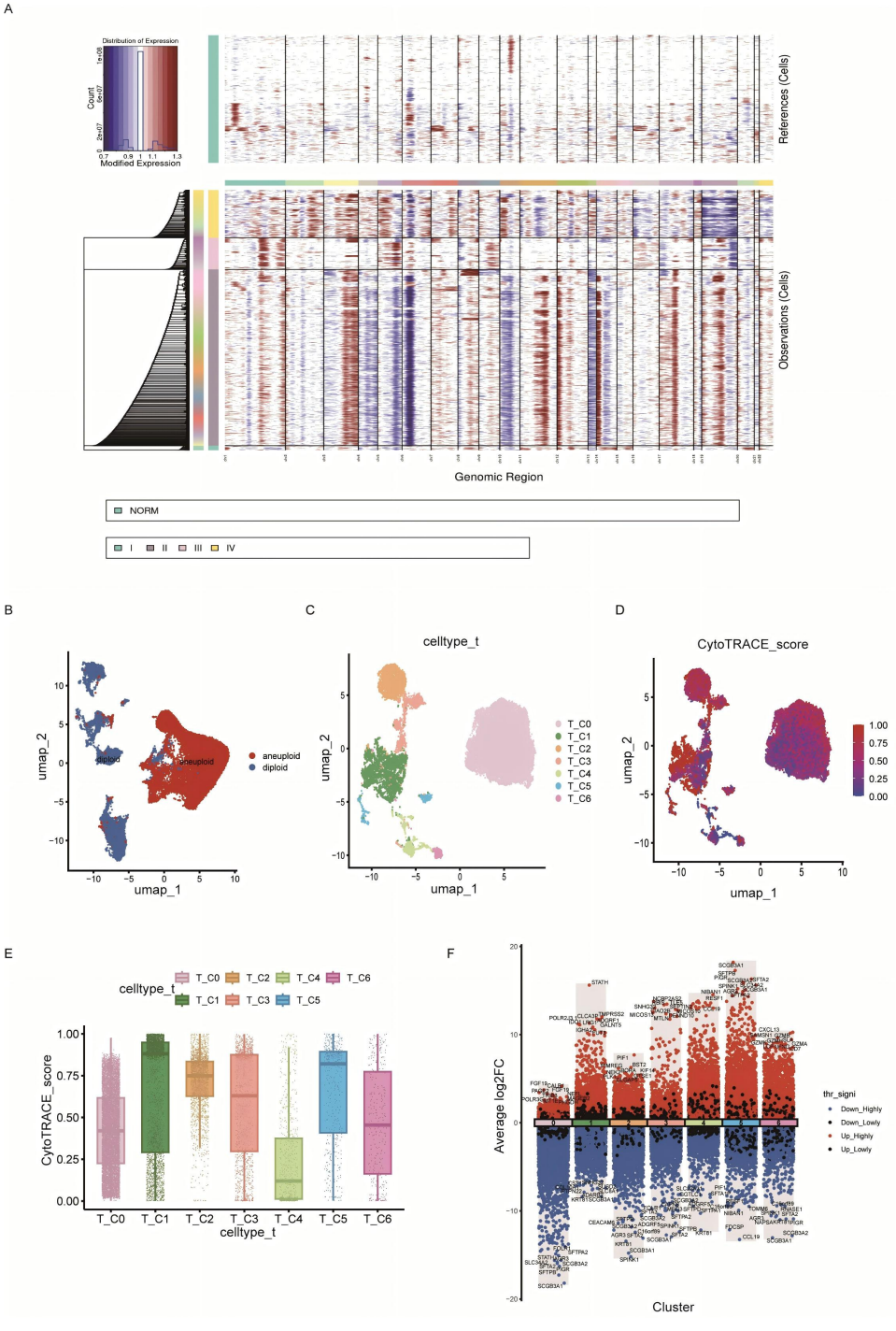

Figure 2. Expression patterns in epithelial cells of LSCC at different stages. A. Comparison of inferCNV analysis results between LSCC tumor cells at different stages and normal epithelial cells. Colors represent the log2(CNV) ratio, with red representing amplification and blue representing deletion; B. Results of the CopyKat algorithm for all

epithelial cells; C. The union of inferCNV and CopyKat results is identified as tumor cells. Seurat clusters are generated after dimensionality reduction and clustering of tumor cells, where each point corresponds to a cell and different colors represent different clusters; D. The distribution of tumor cell stemness levels inferred by CytoTRACE. Red represents cells with high stemness, and blue represents cells with low stemness; E. Box plots showing the stemness scores of seven tumor cell subtypes; F. Display the differentially expressed genes of each tumor cell subtype relative to other subtypes.

Characteristics of immune cell infiltration during the progression of LSCC

We firstly analyzed the changes in T/NK cells, B cells and macrophages at different stages of LSCC. We divided T/NK cells into fourteen subclusters based on the expression of canonical marker genes, including CD4_Tn_LTB/CD28 (naive cells), CD4_Treg_TIGIT/CTLA4 (regulatory T cells), CD8_EM_GZMA (effector memory T cells), CD8_TRM_ITGAE (tissue resident memory T cells), NK_CD16hi_GNLY/NKG7/KLRF1 (NK cells with high CD16 expression), NK_CD56bright_CD7 (NK cells with bright CD56 expression), NK_CD11d_CD247 (CD11d-positive NK cells), NKT_S1PR1 (Natural Killer T cells), gdT_KLRC2 (Gamma delta T cells), and MAIT_IL7R (Mucosal-Associated Invariant T cells) (Figure 3A-B). Compared with other stages, the proportion of CD4_Tn cells in stage IV LSCC decreased significantly, while the proportion of CD4_Treg cells increased significantly (Figure 3C). CD4_Treg cells are the main immunosuppressive cell population in tumors. In line with previous study, higher proportion of CD4_Treg cells often indicates a "cold" state of TME, and the dysfunction or excessive presence of CD4_Treg cells subgroup will significantly weaken the effect of immunotherapy[25]. Our results are consistent with previous studies.

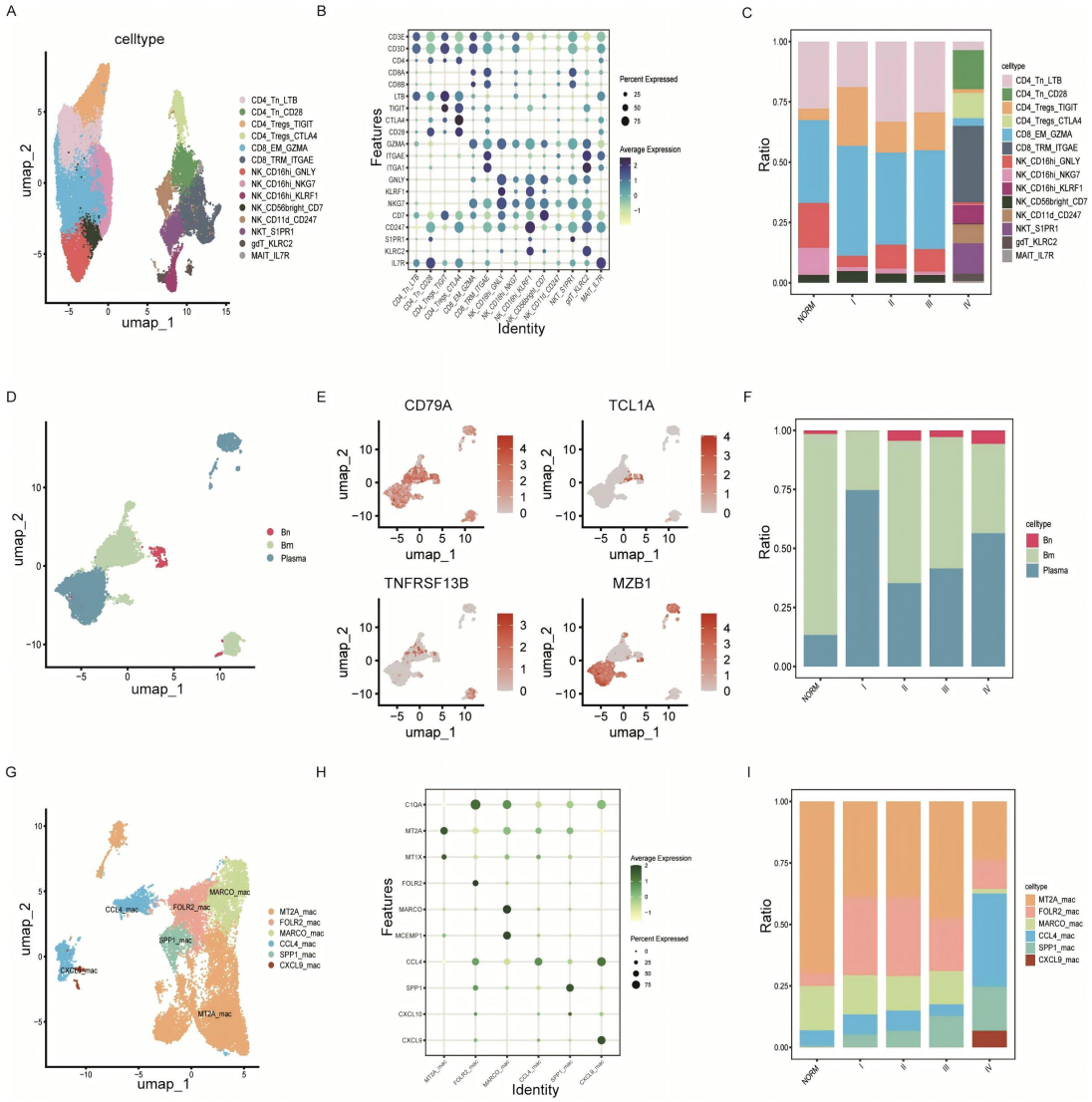

Figure 3. Distribution of T/NK cells, B cells and macrophages in different TNM stages. A. UMAP plots show T/NK cell clusters across all scRNA-seq samples from LSCC patients; B. Dot plots show the expression profiles of key surface markers or functional marker genes across all T/NK cell clusters. Dot color indicates mean expression level, and dot size indicates expression percentage; C. The proportion of various T/NK cell subtypes across different LSCC stages is shown; D. UMAP plots show B cell clusters across all scRNA-seq samples from LSCC patients; E. UMAP plots show the expression profiles of key surface markers or functional marker genes across all B cell clusters; F. The proportion of various B cell subtypes across different LSCC stages is shown; G. UMAP plots show macrophages clusters across all scRNA-seq samples from LSCC patients; H. Dot plots show the expression profiles of key surface markers or functional marker genes across all macrophage clusters. Dot color indicates mean expression level, and dot size

indicates expression percentage; I. The proportion of various macrophage subtypes across different LSCC stages is shown.

Previous study showed infiltration of B cells was associate with prognosis of lung cancer patients[26]. In the presented study, B cells are annotated into three subclusters, including Bn (naive B cells), Bm (memory B cells) and plasma cells (Figure 3D-E). Interestingly, the proportion of plasma cells was higher in stage I and IV tumor niche, while the proportion of memory B cells was higher in stage II and III tumor tissues (Figure 3F). The above results indicated B cell subtypes-mediated immunity pathways activation might be associated with LSCC prognosis.

We investigated the distribution patterns and cellular states of macrophage subtypes across different pathological stages of LSCC. In total, 17,889 macrophages were categorized into six subclusters, including MT2A$^+$ _mac, FOLR2$^+$ _mac, MARCO$^+$ _mac, CCL4$^+$ _mac, SPP1$^+$ _mac, and CXCL9$^+$ _mac (Figure 3G). The annotation of these macrophage subtypes was supported by the expression profiles of classical macrophage markers as well as subtype-specific signature genes (Figure 3H). We next compared the distribution of these macrophage subtypes across pathological stages I-IV of LSCC. Notably, the proportion of SPP1$^+$ mac progressively increased with advancing TNM stage (Figure 3I, Table 1), suggesting a potential association between SPP1$^+$ mac and LSCC progression. Previous studies have shown abundance of macrophages significantly increased in the mid - to - late - stage and higher levels of macrophage infiltration are associated with poor prognosis, which supports our hypothesis[27]. Taken together, these results indicated that SPP1+ mac subpopulation potentially play a central role in mediating cellular crosstalk network regulating LSCC prognosis within the immune microenvironment.

Table 1: Proportion of SPP1+ macrophages across TNM stages I – IV

TNM stage	The proportion of SPP1+ macrophages (%)
I	5.17
II	6.72
III	12.77
IV	17.96

Table 1: Proportion of SPP1+ macrophages across TNM stages I – IV. TNM stage: Pathological stage I – IV according to standard AJCC criteria. The proportion of SPP1+ macrophages (%): For each sample, the percentage of macrophages classified as SPP1-positive among all macrophages, calculated as 100 × (SPP1+ macrophages / total macrophages). Abbreviation: SPP1, secreted phosphoprotein 1 (osteopontin).

Prognostic model constructed based on SPP1+ macrophages

To test whether stage IV – specific alterations in SPP1$^+$ macrophages are linked to LSCC prognosis, we constructed a polygenic risk model, named SPP1$^+$ Mac Differential Gene Score (SPP1$^+$ Mac-DGS) based on the differentially expressed genes (DEGs) between stage IV and stage I SPP1$^+$ macrophages. Genes with avg_log$_2$ FC > 0 and $p < 0.05$ were considered up-regulated gene cocktails (Figure 4A). Using the male subset of TCGA-LUSC as the training cohort, 516 up-regulated genes as the input; 16 absent from TCGA were excluded, leaving 500 candidates for LASSO Cox modeling (Figure 4B, C). The final model comprised four genes, including *TBC1D1*, *GNG5*, *ID1*, and *MAX*; with the formula: SPP1+Mac-DGS =0.0473 × TBC1D1Exp + 0.0816 × GNG5Exp + 0.0065 × ID1Exp − 0.1092 × MAXExp. Univariate Cox analyses supported the prognostic relevance of these genes (Figure 4D), and multivariable Cox analysis confirmed SPP1$^+$ Mac-DGS as an independent predictor (Figure 4E).

Patients stratified by the median score showed significantly different overall survival (OS), with longer OS in the low-risk group (log-rank $p < 0.0001$; Figure 4F). In the training cohort, time-dependent ROC curves yielded AUCs of 0.602, 0.683, and 0.675 for 1-, 3-, and 5-year OS, respectively (Figure 4G); the external validation cohort (GSE30219) showed comparable performance with AUCs of 0.625, 0.655, and 0.682 (Figure 4H). We referenced the AUC values of prognostic models in previous literature. Although our model's AUC value was not as high, it had similar accuracy to those in previous literature[28]. Collectively, these results indicate that SPP1$^+$ Mac-DGS provides stable, modest prognostic discrimination in LSCC.

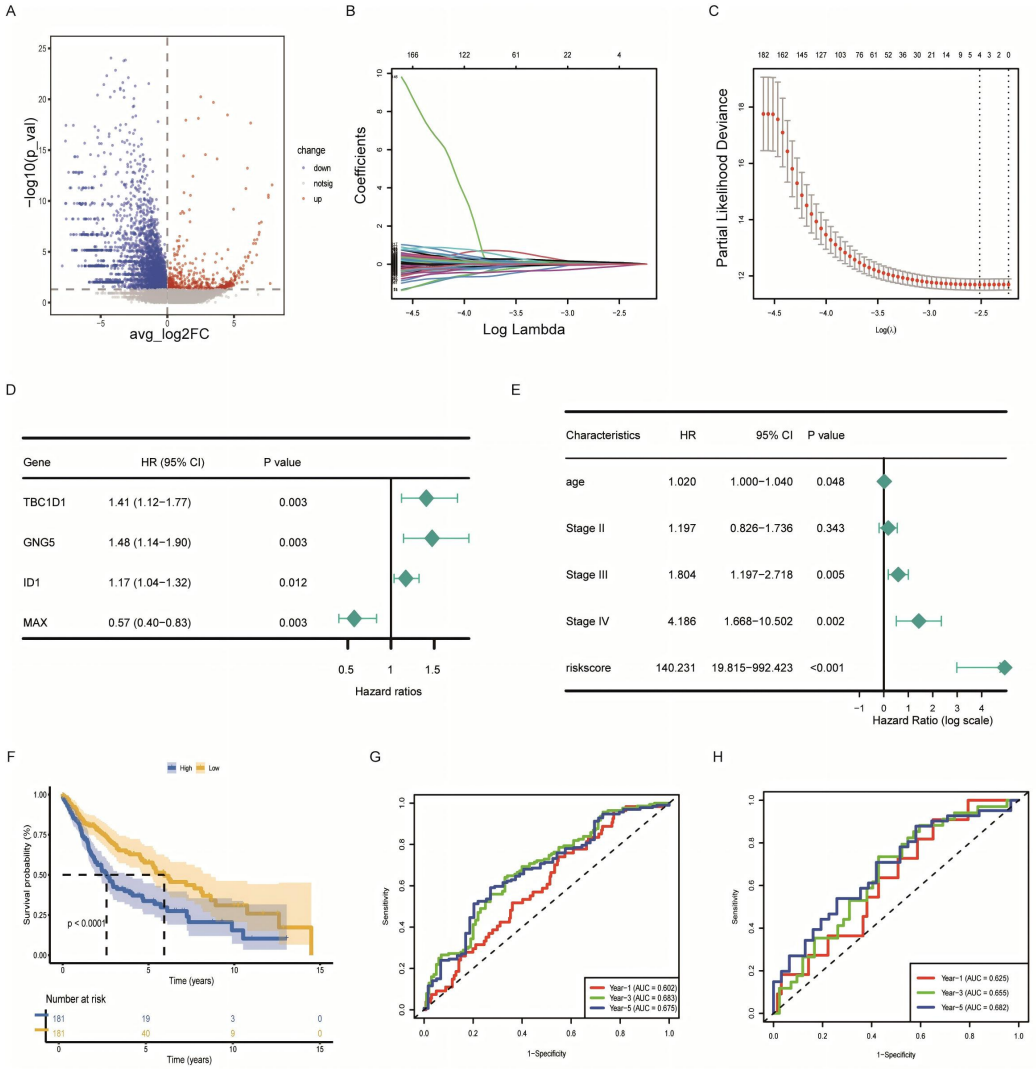

Figure 4. Construction and validation of the prognostic model SPP1+mac-DGS. A. Volcano plot showing significantly differentially expressed genes between stage IV and stage I SPP1+ macrophages; B-C. Lambda trace and distribution of each independent variable; D. Univariate COX analysis of key genes and prognostic information of TCGA-LUSC; E. Multivariate COX regression analysis of age, stage, risk score and prognostic information; F. Kaplan − Meier plot of prognostic survival of the male cohort in TCGA-LUSC; G-H. ROC curves for the TCGA (G) test set and validation set GSE30219 (H). Red indicates 1-year survival rate, green indicates 3-year survival rate, and blue indicates 5-year survival rate.

Functional Characterization of SPP1+ Macrophages

To delineate subtype-specific programs, we compared DEGs between SPP1$^+$ macrophages and other macrophage subsets (Figure 5A). Genes enriched in SPP1$^+$ macrophages, such as *S100A10*, *TPT1*, and *PTGES3* have been implicated in immunosuppression and tumor invasion/metastasis, whereas genes preferentially expressed in other subsets (e.g., *PTP4A3*, *SPDYA*, *HP*, *CCL17*) encompass cell-cycle control and T-cell activating functions and immunomodulatory activities. These patterns suggest that multiple macrophage lineages may influence tumor progression via distinct mechanisms. The GO enrichment analysis further highlighted biological differences between SPP1$^+$ macrophages and other subtypes (Figure 5B–C). SPP1$^+$ macrophages showed up-regulation of antiviral response and metabolic reprogramming pathways accompanied by down-regulation of T-cell co-stimulation, antigen-receptor signaling, and leukocyte adhesion. This transcriptomic signature indicates a deviation from classical pro-inflammatory macrophage states, consistent with reduced immunostimulatory capacity and an immunosuppressive phenotype. Such features align with the known tissue-remodeling and tumor-supportive roles of SPP1-high cells.

In the TCGA-LUSC cohort, SPP1 expression was significantly higher in tumors than in normal lung (Figure 5D), and high SPP1 expression level was associated with worse overall survival (Figure 5E), supporting the link between SPP1$^+$ macrophage programs and adverse prognosis.

Pathway activity profiling revealed subtype-specific signaling (Figure 5F). Notably, SPP1$^+$ macrophages exhibited the highest activity in MYC targets, angiogenesis, oxidative phosphorylation, DNA repair, and EMT, underscoring their pro-tumor features and suggesting contributions to progression through metabolic remodeling, immune suppression, and promotion of tumor cell plasticity. Finally, SCENIC analysis identified possible regulators of SPP1$^+$ macrophages, including *MAX*, *BCLAF1*, *HDAC2*, *ELF1*, and *ETV6* (Figure 5G).

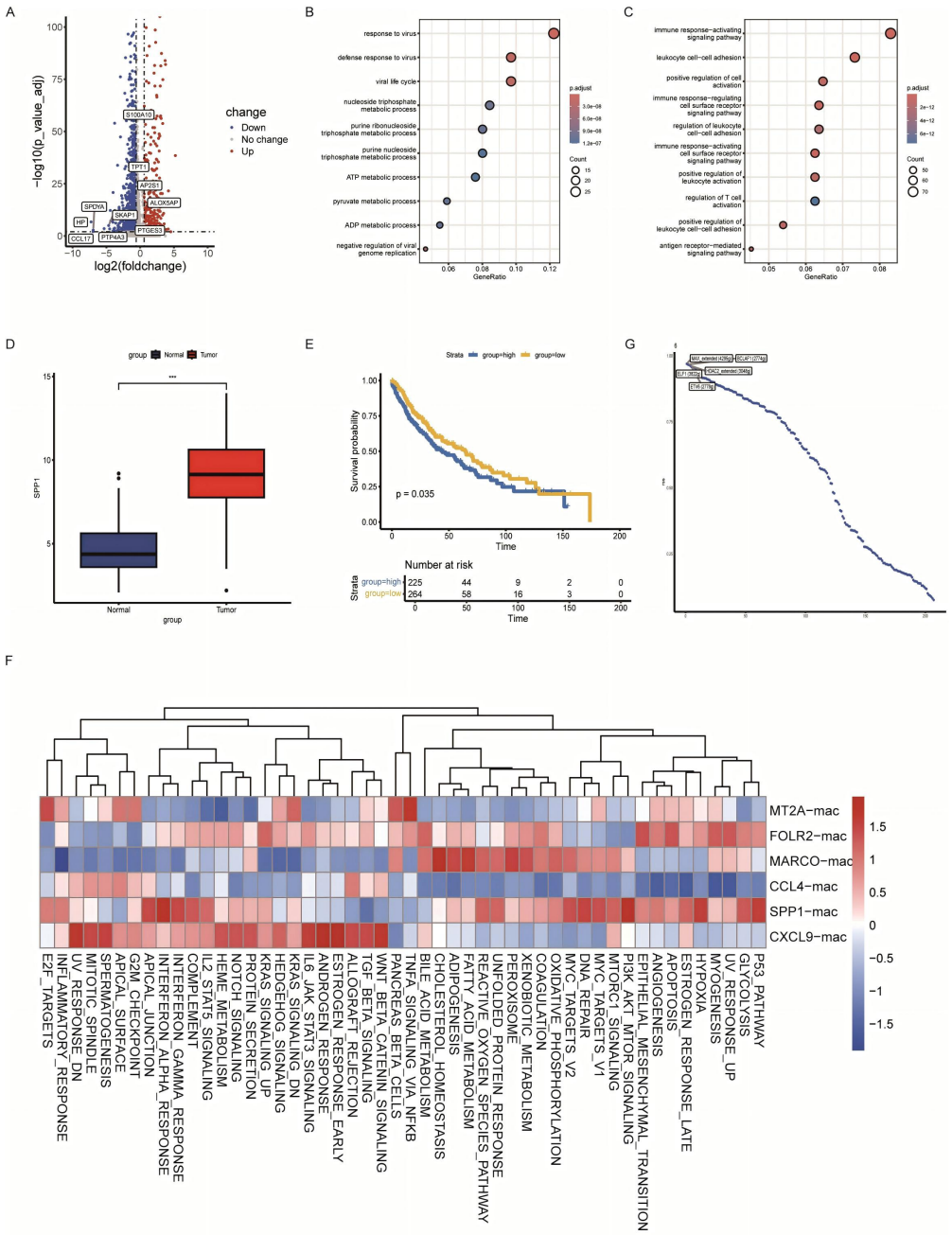

Figure 5. Functional characteristics of SPP1+ macrophages. A. Volcano plot showing genes significantly differentially expressed in SPP1+ macrophages compared to other macrophage subtypes; B-C. Dot plot showing GO enrichment analysis results of significantly upregulated and downregulated genes in SPP1+ macrophages; D. The boxplot shows the expression differences of SPP1 in tumor and normal tissues in the TCGA-LUSC dataset; E. Survival analysis of patients with high and low SPP1 expression; F. Heatmap of pathway activity scores in different macrophage subtypes; G. Major factors regulating SPP1+ macrophage subtypes.

The SPP1$^+$ macrophages-fibroblasts interaction marks an edge-localized, immune-excluding niche in LSCC

To further decipher the signaling features of ligand-receptor interactions within the immune microenvironment at distinct TNM stages. CellChat revealed that patterns of incoming and outgoing interaction strengths varied across LSCC stages; notably, macrophages consistently exhibited higher outgoing strength than most other lineages (Figure 6A), indicating a dominant signal-sender role in the TME. Within the SPP1 signaling pathway, the macrophage-fibroblast axis was among the most pronounced interactions (Figure 6B). Across stages, CellChat predicted macrophage-derived SPP1 engaging integrin heterodimers on fibroblasts, including ITGA8+ITGB1 (Figure 6C). Violin plots of *SPP1*, *ITGA8*, and *ITGB1* expression across cell types further demonstrated the cellular sources and targets of these interactions (Figure 6D).

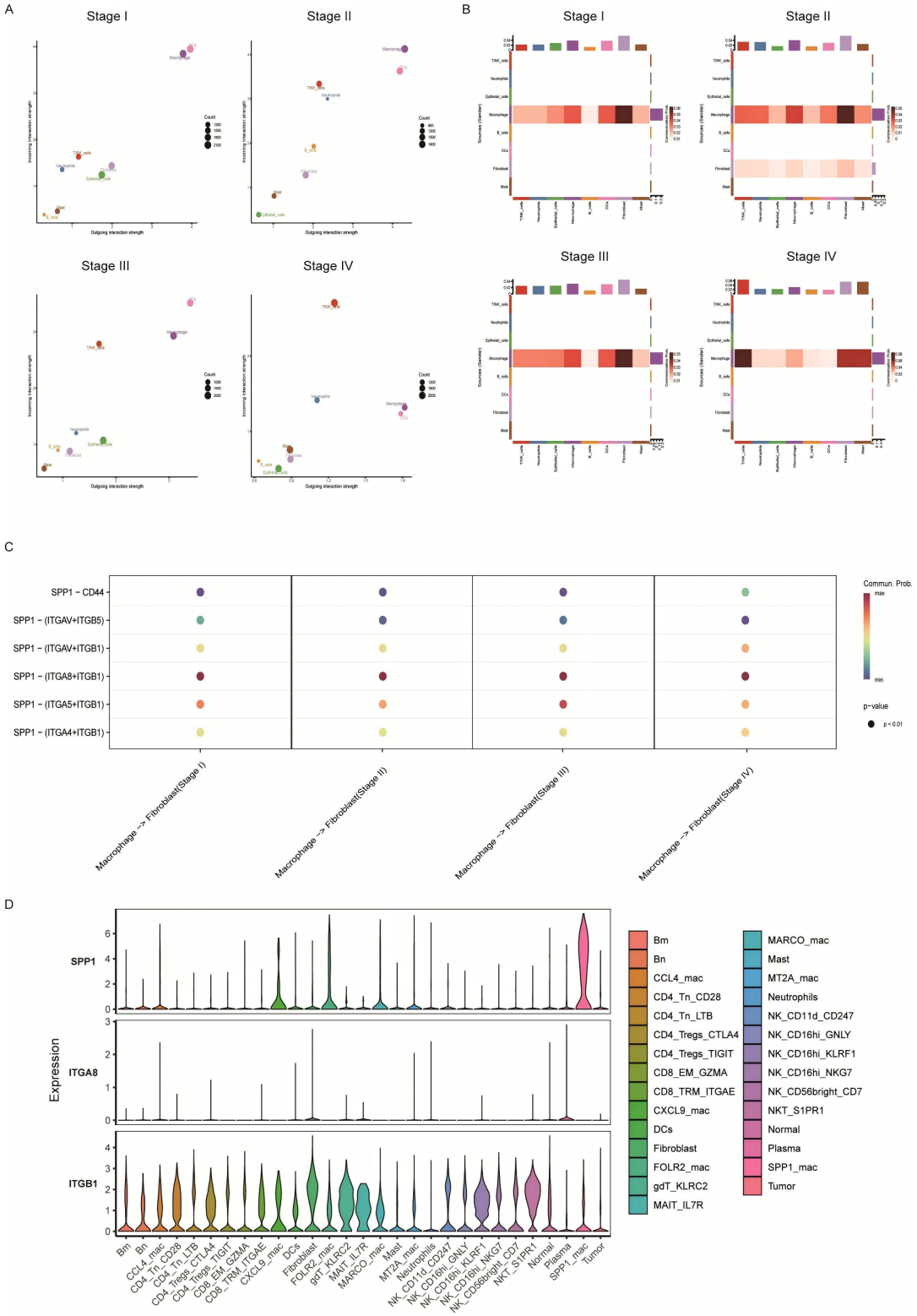

Figure 6. Interaction between SPP1+ macrophages and fibroblasts. A. Dot plots showing the strength of incoming and outgoing interactions between cell types at different disease stages; B. Heatmap showing the strength of interactions between cell types in the SPP1 signaling pathway at different stages. C. The dot plot shows the communication probability between SPP1+ macrophages and fibroblasts at different stages. D. Violin plots show the expression of key ligands (SPP1) and receptors (ITGA8, ITGB1) at different stages.

To futher dissect SPP1$^+$ macrophages-fibroblasts interactions spatial feature, we analyzed a stage IB LSCC spatial transcriptomic dataset. RCTD deconvolution mapped the locations and abundances of SPP1$^+$ macrophages, fibroblasts, and tumor cells (Figure 7A), revealing that SPP1$^+$ macrophages and fibroblasts were enriched at the tumor edge. MISTy further identified proximity-based associations (Figure 7B, C), and spatial distribution maps showed colocalization in the intra view, consistent with the CellChat-inferred communication. This organization mirrors reports in NSCLC of a marginalized, immune-exclusive TME in which SPP1$^+$ TAMs and stromal-active cancer-associated fibroblasts（CAFs） (e.g., *FAP*$^+$ */POSTN*$^+$) assemble at the tumor boundary and are associated with restricted T-cell entry[29].

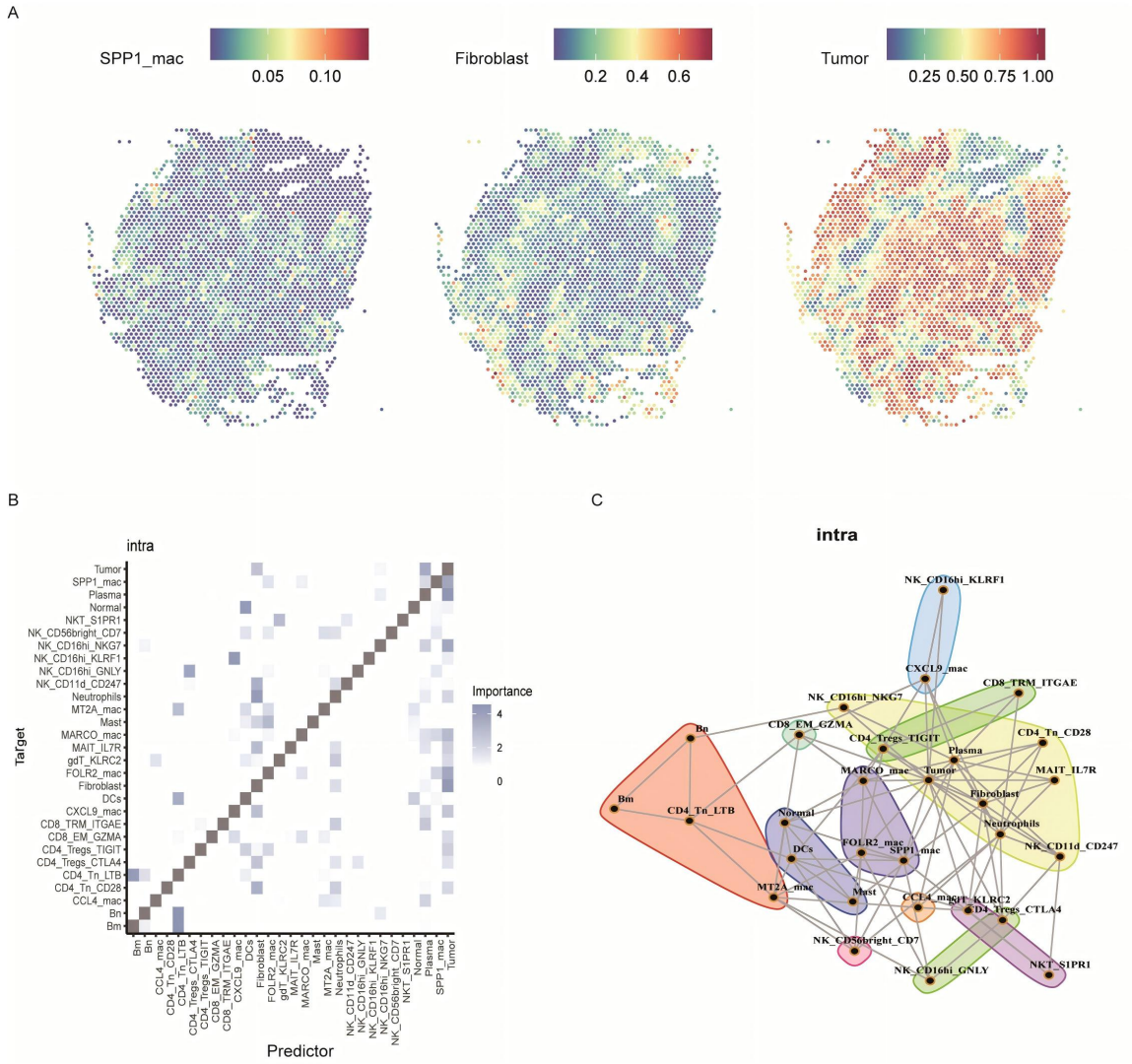

Figure 7. Spatial distribution of SPP1+ macrophages and fibroblasts. A. The spatial distribution of SPP1+ macrophages, fibroblasts, and tumor cells in LSCC tissue sections is visualized using the RCTD framework. B-C. Spatial microenvironmental and cell-cell interactions identified through MISTy analysis demonstrate that SPP1+ macrophages form a spatial microenvironment with fibroblasts and tumor cells.

In sum, these multi-modal data support an SPP1-centered axis at the tumor – stroma interface in LSCC: macrophages act as dominant senders, signaling through fibroblast integrin receptors to establish an immune-exclusive marginal niche that may impede lymphocyte infiltration.

Discussion

To our knowledge, this is the first study to construct a stage dynamic single-cell transcriptomic atlas of LSCC spanning TNM stages I – IV. We delineated dynamic feature in the TME, identified a progressive increase in SPP1$^+$ macrophages with advancing stage and their association with immunosuppressive programs, and constructed a prognostic model grounded in SPP1$^+$ macrophage – related transcriptional differences.

Patients with LSCC, particularly those with advanced disease, have limited therapeutic options and poor survival, underscoring the need to chart TME evolution and to develop improved treatment and prognostic strategies. Our research demonstrated marked TME heterogeneity across stages I – IV and showed that the proportion of SPP1$^+$ macrophages rose with increasing TNM stage, implicating stage-dependent recruitment or expansion of this subset in disease progression. SPP1 (secreted phosphoprotein 1; osteopontin) is a secreted, non-collagenous glycoprotein commonly upregulated in tumor and stromal compartments, and elevated levels in blood and tumor tissues have been linked to adverse outcomes across multiple cancers[30]. Consistent with these observations, SPP1$^+$ macrophages have been implicated in immunosuppressive TMEs in hepatocellular carcinoma, glioblastoma, and colorectal cancer[31-33], supporting our findings and the prognostic relevance of SPP1$^+$ macrophage – associated signatures in LSCC.

Based on genes up-regulated in stage IV relative to stage I within SPP1$^+$ macrophages, we developed a four-gene prognostic model comprising *TBC1D1*, *GNG5*, *ID1*, and *MAX*. *TBC1D1* encodes a Rab GTPase – activating protein that regulates vesicle trafficking and glucose uptake[34]. *GNG5* encodes the γ subunit of heterotrimeric G proteins; together with Gβ it forms the Gβγ complex that modulates chemotaxis, migration, and polarization, and has been linked to tumor-promoting phenotypes across multiple cancers[35,36]. *ID1*, a helix – loop – helix transcriptional regulator, is overexpressed in diverse malignancies and promotes tumor growth, metastasis, angiogenesis, and therapy resistance; in tumor-associated macrophages it contributes to an immunosuppressive milieu, limiting CD8$^+$ T-cell infiltration and supporting cancer stemness[37]. *MAX* is a context-dependent hub

of the MYC transcriptional network: as MYC's essential dimerization partner it enables oncogenic transcription and proliferation, whereas its inactivation functions as a tumor-suppressive event in specific tumor types[38]. Collectively, these mechanistic links support the biological plausibility of the model and its potential utility for prognostic risk stratification.

We used SCENIC to find five candidate regulators of SPP1⁺ macrophages — *MAX*, *BCLAF1*, *HDAC2*, *ELF1*, and *ETV6*. Among these, *HDAC2* (a class I histone deacetylase) emerged as a promising therapeutic entry point. In lung cancer, M2-like TAMs with high *HDAC2* expression are associated with inferior survival. Genetic or pharmacologic inhibition/silencing of *HDAC2* reprograms TAMs toward an M1-like state and alleviates immunosuppression[39], and *HDAC* inhibitors can enhance macrophage anti-tumor activity[40]. Together with our SCENIC results, these data show *HDAC2* as a potential therapeutic target for LSCC.

Our integrative analyses centered on macrophage–fibroblast crosstalk in LSCC. Consistent with CellChat inferences, macrophages acted as dominant "senders" that signal to fibroblasts through the SPP1 pathway, and spatial modeling placed these populations in close proximity at the tumor–stroma interface. This edge-localized pairing accords with reports that SPP1⁺ TAMs and stromal-active CAFs (e.g., FAP⁺/POSTN⁺) assemble at tumor margins to form a barrier that excludes T cells and modulates therapeutic response[41].

Our study for the first time revealed that the proportion of SPP1⁺ macrophages in LSCC progressively increased with higher TNM stages. We constructed a prognostic model based on the differences in SPP1+ macrophages between stage IV and stage I LSCC. These findings provide a rationale for developing stage-dependent treatment strategies in LSCC. However, inferences about the interaction between SPP1+ macrophages and fibroblasts are based on published association studies rather than direct experimental demonstration. Future studies should combine mechanistic validation with prospective cohort testing.

In summary, we chart TME remodeling across LSCC stages, highlight the progressive enrichment and immunosuppressive role of SPP1⁺ macrophages, and demonstrate the prognostic utility of the SPP1⁺ Mac-DGS model for risk stratification. These findings offer insights into immune cell infiltration and therapeutic responsiveness.

Limitations

Retrospective, public-dataset design. All analyses were performed on previously published scRNA-seq, spatial transcriptomics, and bulk RNA-seq datasets. As such, sampling frames, inclusion criteria, pre-analytical variables, and clinical covariates (e.g., prior therapy, smoking exposure, comorbidities) were heterogeneous and incompletely harmonized, which may introduce residual confounding.

Stage annotation and cohort comparability. TNM staging was derived from study metadata rather than uniform prospective assessment, and stage distributions differed across datasets. Therefore, some stage-dependent effects may reflect cohort composition or technical factors rather than biology alone.

Cross-platform integration. Integrating scRNA-seq, ST, and bulk data is sensitive to batch effects, platform chemistry, and normalization choices. Although integration reduces obvious batch structure, it cannot fully exclude alignment artifacts or attenuation of genuine biological variation.

Interaction inference is correlational. Ligand–receptor communication between SPP1⁺ macrophages and fibroblasts was inferred from transcript abundance and prior knowledge bases rather than direct protein-level or functional assays. Post-transcriptional regulation, ligand processing, receptor activation, and spatial proximity constraints were not experimentally verified.

No mechanistic validation. We did not perform perturbation experiments (e.g., macrophage depletion, SPP1 blockade) or orthogonal assays (protein, IHC/IF, multiplexed imaging) to demonstrate causal roles for SPP1⁺ macrophages or their crosstalk with fibroblasts.

Future work should incorporate multi-center prospective cohorts with standardized staging and treatment annotation, orthogonal protein/spatial validation, and mechanistic studies to test whether modulating SPP1⁺ macrophage–fibroblast axes alter LSCC progression and patient outcomes.

Reference

1. Barta JA, Powell CA, Wisnivesky JP: Global Epidemiology of Lung Cancer. Ann Glob Health 2019, 85(1).
2. Paz-Ares L, Vicente D, Tafreshi A et al: A Randomized, Placebo-Controlled Trial of Pembrolizumab Plus Chemotherapy in Patients With Metastatic Squamous NSCLC: Protocol-Specified Final Analysis of KEYNOTE-407. J Thorac Oncol 2020, 15(10):1657-1669.
3. Ren YF, Ma Q, Zeng X et al: Single-cell RNA sequencing reveals immune microenvironment niche transitions during the invasive and metastatic processes of ground-glass nodules and part-solid nodules in lung adenocarcinoma. Mol Cancer 2024, 23(1):263.
4. Papalexi E, Satija R: Single-cell RNA sequencing to explore immune cell heterogeneity. Nat Rev Immunol 2018, 18(1):35-45.
5. Rao A, Barkley D, França GS, Yanai I: Exploring tissue architecture using spatial transcriptomics. Nature 2021, 596(7871):211-220.
6. Zhang L, Zhang Y, Wang C et al: Integrated single-cell RNA sequencing analysis reveals distinct cellular and transcriptional modules associated with survival in lung cancer. Signal Transduct Target Ther 2022, 7(1):9.
7. Gómez-López S, Alhendi ASN, Przybilla MJ et al: Aberrant basal cell clonal dynamics shape early lung carcinogenesis. Science 2025, 388(6752):eads9145.
8. Wu F, Fan J, He Y et al: Single-cell profiling of tumor heterogeneity and the microenvironment in advanced non-small cell lung cancer. Nat Commun 2021, 12(1):2540.
9. Hao Y, Hao S, Andersen-Nissen E et al: Integrated analysis of multimodal single-cell data. Cell 2021, 184(13): 3573-3587.e3529.
10. Korsunsky I, Millard N, Fan J et al: Fast, sensitive and accurate integration of single-cell data with Harmony. Nat Methods 2019, 16(12):1289-1296.

11. Madissoon E, Oliver AJ, Kleshchevnikov V et al: A spatially resolved atlas of the human lung characterizes a gland-associated immune niche. Nat Genet 2023, 55(1):66-77.
12. Patel AP, Tirosh I, Trombetta JJ et al: Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. Science 2014, 344(6190):1396-1401.
13. Gao R, Bai S, Henderson YC et al: Delineating copy number and clonal substructure in human tumors from single-cell transcriptomes. Nat Biotechnol 2021, 39(5):599-608.
14. Gulati GS, Sikandar SS, Wesche DJ et al: Single-cell transcriptional diversity is a hallmark of developmental potential. Science 2020, 367(6476):405-411.
15. Liang L, Yu J, Li J et al: Integration of scRNA-Seq and Bulk RNA-Seq to Analyse the Heterogeneity of Ovarian Cancer Immune Cells and Establish a Molecular Risk Model. Front Oncol 2021, 11:711020.
16. Wang P, Zhang T, Wang X et al: Aberrant human ClpP activation disturbs mitochondrial proteome homeostasis to suppress pancreatic ductal adenocarcinoma. Cell Chem Biol 2022, 29(9):1396-1408.e1398.
17. Lu Y, Yang A, Quan C et al: A single-cell atlas of the multicellular ecosystem of primary and metastatic hepatocellular carcinoma. Nat Commun 2022, 13(1):4594.
18. Wang L, Wang D, Yang L et al: Cuproptosis related genes associated with Jab1 shapes tumor microenvironment and pharmacological profile in nasopharyngeal carcinoma. Front Immunol 2022, 13:989286.
19. Aibar S, González-Blas CB, Moerman T et al: SCENIC: single-cell regulatory network inference and clustering. Nat Methods 2017, 14(11):1083-1086.
20. Hänzelmann S, Castelo R, Guinney J: GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics 2013, 14:7.
21. Liberzon A, Birger C, Thorvaldsdóttir H et al: The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst 2015, 1(6):417-425.
22. Jin S, Guerrero-Juarez CF, Zhang L et al: Inference and analysis of cell-cell communication using CellChat. Nat Commun 2021, 12(1):1088.
23. Cable DM, Murray E, Zou LS et al: Robust decomposition of cell type mixtures in spatial transcriptomics. Nat Biotechnol 2022, 40(4):517-526.
24. Tanevski J, Flores ROR, Gabor A, Schapiro D, Saez-Rodriguez J: Explainable multiview framework for dissecting spatial relationships from highly multiplexed data. Genome Biol 2022, 23(1):97.
25. Zhou Z, Xu J, Liu S et al: Infiltrating treg reprogramming in the tumor immune microenvironment and its optimization for immunotherapy. Biomark Res 2024, 12(1):97.
26. Siliņa K, Rulle U, Kalniņa Z, Linē A: Manipulation of tumour-infiltrating B cells and tertiary lymphoid structures: a novel anti-cancer treatment avenue? Cancer Immunol Immunother 2014, 63(7):643-662.
27. Song L, Yu X, Wu Y et al: Integrin β8 Facilitates Macrophage Infiltration and Polarization by Regulating CCL5 to Promote LUAD Progression. Adv Sci (Weinh) 2025, 12(2):e2406865.
28. Peng H, Wu X, Liu S et al: Cellular dynamics in tumour microenvironment along with lung cancer progression underscore spatial and evolutionary heterogeneity of neutrophil. Clin Transl Med 2023, 13(7):e1340.
29. Chen C, Guo Q, Liu Y et al: Single-cell and spatial transcriptomics reveal POSTN(+) cancer-associated fibroblasts correlated with immune suppression and tumour progression in non-small cell lung cancer. Clin Transl Med 2023, 13(12):e1515.
30. Lamort AS, Giopanou I, Psallidas I, Stathopoulos GT: Osteopontin as a Link between Inflammation and Cancer: The Thorax in the Spotlight. Cells 2019, 8(8).
31. Wei J, Marisetty A, Schrand B et al: Osteopontin mediates glioblastoma-associated macrophage infiltration and is a potential therapeutic target. J Clin Invest 2019, 129(1):137-149.
32. Qi J, Sun H, Zhang Y et al: Single-cell and spatial analysis reveal interaction of FAP(+) fibroblasts and SPP1(+) macrophages in colorectal cancer. Nat Commun 2022, 13(1):1742.
33. Liu Y, Xun Z, Ma K et al: Identification of a tumour immune barrier in the HCC microenvironment that determines the efficacy of immunotherapy. J Hepatol 2023, 78(4):770-782.
34. Cartee GD: Roles of TBC1D1 and TBC1D4 in insulin- and exercise-stimulated glucose transport of skeletal muscle. Diabetologia 2015, 58(1):19-30.
35. Senarath K, Payton JL, Kankanamge D et al: Gγ identity dictates efficacy of Gβγ signaling and macrophage migration. J Biol Chem 2018, 293(8):2974-2989.
36. Zhang W, Liu Z, Liu B et al: GNG5 is a novel oncogene associated with cell migration, proliferation, and poor prognosis in glioma. Cancer Cell Int 2021, 21(1):297.
37. Shang S, Yang C, Chen F et al: ID1 expressing macrophages support cancer cell stemness and limit CD8(+) T cell infiltration in colorectal cancer. Nat Commun 2023, 14(1):7661.
38. Augert A, Mathsyaraja H, Ibrahim AH et al: MAX Functions as a Tumor Suppressor and Rewires Metabolism in Small Cell Lung Cancer. Cancer Cell 2020, 38(1):97-114.e117.
39. Zheng X, Sarode P, Weigert A et al: The HDAC2-SP1 Axis Orchestrates Protumor Macrophage Polarization. Cancer Res 2023, 83(14):2345-2357.

40. Li X, Su X, Liu R et al: HDAC inhibition potentiates anti-tumor activity of macrophages and enhances anti-PD-L1-mediated tumor suppression. Oncogene 2021, 40(10):1836-1850.
41. Xiao M, Deng Y, Guo H et al: Single-cell and spatial transcriptomics profile the interaction of SPP1(+) macrophages and FAP(+) fibroblasts in non-small cell lung cancer. Transl Lung Cancer Res 2025, 14(7):2646-2669.

Funding Information

This work was supported by the National Natural Science Foundation of China [grant numbers 62376077] and Heilongjiang Province key research and development plan [grant numbers 2024ZX12C24].

Author contributions

DS: Conceptualization (lead), Formal Analysis (lead), Writing—Original Draft Preparation (lead), Data Curation (supporting); ZL and JG: Writing—Review & Editing (equal); QW and AG: Methodology (equal), Project Administration (equal), Writing—Review & Editing (equal). All authors read and approved the final manuscript.

Disclosure

Approval of the research protocol by an Institutional Reviewer Board: N/A.

Informed Consent: N/A.

Registry and the Registration No. of the study/trial: N/A.

Animal Studies: N/A.

Ethics Statement:

This study does not involve research with human participants or animals, as it is entirely based on publicly available datasets. Therefore, ethical approval and informed consent are not applicable.

Conflict of Interest:

The authors declare no conflicts of interest.

Data Availability:

The datasets utilized for the bioinformatics analysis in our study are available from the TCGA database (https://www.cancer.gov/ccg/research/genome-sequencing/tcga, project ID: TCGA-LUSC), GEO (https://www.ncbi.nlm.nih.gov/geo/, accession number: GSE194070), the National Genomics Data Center (https://ngdc.cncb.ac.cn/cancerscem, accession number: PRJNA976462), the lung cancer database of West China Hospital, Sichuan University (http://lungcancer.chenlulab.com, raw_data.csv.gz) and BioStudies (https://www.ebi.ac.uk/biostudies/, accession number: E-MTAB-13530).

Code availability

All code generated for analysis is available from the github repository: https://github.com/sdd0320/LUSC_sc2.