Experimental Hematology & Oncology

RESEARCH

Open Access

Single-cell profiling reveals FAM117A as a key regulator linking macrophage–epithelial crosstalk with the progression of lung adenocarcinoma

Chao Wu[1*†], Zhipeng Ren[2†], Gang Che[3†], Hui Liu[4†], Haitao Tao[1], Dan Li[1], Jinliang Wang[1*], Xiaobin Hou[2*] and Yi Hu[1*]

Abstract

The tumor microenvironment (TME) has a profound influence on the progression of lung adenocarcinoma (LUAD) and its response to therapy. We identified a tumor-promoting communication network based on single-cell RNA sequencing. This shows an SPP1 + macrophage and basal epithelial cell communication module (CIM) that is enriched in invasive portions of tumors, and spatially associated with decreased cytotoxic T cell activity and poor prognosis. Transcriptomic modeling identified FAM117A as a major suppressor of the CIM network. FAM117A is a DYRK1A-interacting cell-cycle regulator. Loss of FAM117A resulted in longer G1/S transition time, increased cell proliferation, and an enhanced macrophage-epithelial feedback loop. Immunohistochemical studies showed decreased FAM117A expression in LUAD tissues, which correlated with SPP1 + macrophage density and poor outcome. Treatment with FAM117A or pharmacological inhibition of CDK4/6 reduced the in vitro and in vivo tumor growth. Thus, FAM117A links intrinsic cell cycle regulation with the extrinsic immune microenvironment, providing a rationale for combined therapies that address macrophage-tumor and cell cycle regulatory events in LUAD.

Keywords Lung adenocarcinoma, SPP1 + macrophage, FAM117A, Cell–cell communication, Tumor microenvironment

†Chao Wu, Zhipeng Ren, Gang Che and Hui Liu have contributed equally to this work.

*Correspondence:
Chao Wu
wuchao@medmail.com.cn
Jinliang Wang
wangjinliang301@163.com
Xiaobin Hou
drhouxb@163.com
Yi Hu
huyi301zlxb@sina.com
[1]Department of Medical Oncology, Senior Department of Oncology, Chinese PLA General Hospital, 28 Fuxing Road, Beijing 100853, China
[2]Department of Thoracic Surgery, The First Medical Center, Chinese PLA General Hospital, 28 Fuxing Road, Beijing 100853, China
[3]Department of Surgical Oncology, The First Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou 310016, Zhejiang, China
[4]Day Clinic, Hainan Hospital of Chinese PLA General Hospital, 80 Jianglin Road, Haitang District, Sanya, Hainan, China

Introduction

Lung adenocarcinoma (LUAD) is the most prevalent histological subtype of lung cancer and continues to be a leading cause of cancer-related mortality worldwide [1]. Despite advances in the understanding of this disease and the use of molecular-targeted therapies and immunotherapy, the prognosis of patients with advanced LUAD remains poor. This is in large part due to profound tumor heterogeneity and an immunosuppressive tumor microenvironment (TME). Recent advancements in technology, particularly single-cell and spatial multi-omics [2], have led to a more precise understanding of LUAD biology through the critical dissection of tumor-intrinsic and microenvironmental mechanisms [3, 4].

The TME represents a dynamic and complex ecosystem comprised of malignant epithelial cells, infiltrating immune cell populations, fibroblasts, and endothelial networks [5]. These components communicate with each other by means of soluble factors, ligand-receptor signaling, and extracellular matrix (ECM) remodeling to influence tumor behavior. Tumor-associated macrophages (TAMs) play a key role in this ecosystem, directing inflammation, angiogenesis, and immune evasion. Increasing evidence indicates that tumor behavior is affected by the polarization and spatial localization of macrophages, with TAMs emerging as a signature of aggressive LUAD [6]. These cells are present in high numbers in invasive tumor compartments and secrete cytokines and growth factors that promote epithelial-mesenchymal transition (EMT), inhibit cytotoxic T-cell infiltration, and promote matrix remodeling. Expression of the macrophage signature is associated with poor prognosis and resistance to immunotherapy, thereby providing a key therapeutic target within the LUAD TME.

Concomitantly, tumor-intrinsic regulators of the cell cycle dictate malignant behavior and therapeutic sensitivity. Among those, Family with Sequence Similarity 117 A (FAM117A) was recently reported to be a novel modulator of the G1/S transition [7]. FAM117A interacts with the dual-specificity kinase DYRK1A to maintain cell-cycle arrest and genomic stabilityDysregulation of tumor-intrinsic cell-cycle–associated factors has been shown to activate oncogenic signaling pathways and promote recurrence and aggressive progression in lung cancer [8]. Targeting epigenetic and cell-cycle–related pathways has emerged as a promising strategy to enhance antitumor immunity and overcome resistance to immunotherapy in aggressive lung cancers [9]. Loss or downregulation of FAM117A occurs frequently in LUAD tissues and disrupts this checkpoint, thereby promoting uncontrolled proliferation and aggressive tumor growth [7]. Clinically, decreased expression of FAM117A is associated with advanced tumor stage and diminished overall survival, suggesting dual roles as a tumor suppressor and a prognostic biomarker [10, 11]. Building upon these insights, the aims of the present study were to delineate the crosstalk between macrophages (Mps) and epithelial tumor cells during the progression of LUAD, and to unveil the mechanistic role of FAM117A within this regulatory network. By combining single-cell transcriptomics and functional assays, we developed a CIM framework to map cell-type specific communication and identify prognostic gene signatures. Furthermore, we experimentally validated FAM117A as a key mediator of cell-cycle regulation and tumor growth. This integrative approach was designed to provide a comprehensive understanding of LUAD progression from both immune-microenvironmental and tumor-intrinsic perspectives. Our results should contribute to the development of targeted strategies that combine macrophage modulation with precision cell-cycle therapy.

Methods

Collection of clinical human samples

Tissue samples were collected from 76 patients who underwent surgical treatment in the Department of Thoracic Surgery, First Medical Center, PLA General Hospital. The lung lesions were pathologically confirmed as LUAD. A total of 57 patients with LUAD and complete tissue specimens, including cancer tissue and adjacent tissue, were screened (see Table S1 for details). After review of the project by the Ethics Committee, all patients signed informed consent.

Single-cell data processing

RPMI 1640 supplemented with 1 mM protease inhibitor was utilized for the collection and transportation of tissues. Subsequently, the tissues underwent digestion using a mixture of enzymes, including Dispase II, Type VIII Collagenase, and DNase I in PBS containing 5% fetal bovine serum. Following a 40-minute incubation at 37 °C, the cells were dissociated, collected at 20-minute intervals, and filtered through a 40 mm nylon cell strainer. Red blood cells were eliminated using a lysis buffer containing DNase I, followed by a wash in PBS containing 0.04% BSA. The cell concentration of the resulting single-cell suspension was determined and adjusted to 1000 cells/mL. Dead cells were removed using the Dead Cell Removal Kit before loading the cells as per the Chromium single-cell 30 kit standard protocol to capture 5,000–10,000 cells/chip position. Library construction and subsequent procedures adhered to the manufacturer's guidelines. Single-cell libraries were generated using Illumina HiSeq X Ten with 150 nt paired-end sequencing. The Cell Ranger software pipeline was employed to process the data, including demultiplexing cellular barcodes, mapping reads to the genome and transcriptome with the STAR aligner, and down-sampling reads as

necessary to generate normalized aggregate data across samples. This resulted in a gene count matrix versus cells.

The single-cell datasets GSE131907 and GSE189357 were downloaded from the GEO database. The matrix was processed using the Seurat R package. Only high-quality cells were fed for downstream analysis. These met the criteria of a minimum of 200 genes, a maximum of 6000 genes, and < 10% mitochondrial reads per cell. The raw counts were normalized and scaled. Batch effects were removed using the CCA algorithm, followed by principal component analysis. Clusters were identified using graph-based clustering and visualized by UMAP. The cell type was manually labeled using classical markers.

Analysis of single-cell trajectories was performed using the Monocle2 package with the features of each population as the input. Transcription factor (TF) analysis was performed with the SCENIC package using default parameters [12].

Differential-expression analysis and functional enrichment analysis

The "FindAllMarkers" function was used in Seurat to identify differentially expressed genes. The non-parametric Wilcoxon rank-sum test was used to obtain *p*-values for comparisons, as well as the adjusted *p*-values based on Bonferroni correction, for all genes in the dataset. The differential genes were enriched in biological process terms of Gene Ontology using the ClusterProfiler package, and the results were plotted with the ggplot2 package.

Survival analysis

To assess the gene expression signatures or pathways associated with the infiltration of subclusters, the top 20 features of subclusters were considered as gene sets for GSVA enrichment based on the TCGA LUAD dataset. For survival analyses, patients were stratified into high- and low-score groups based on the median value of the corresponding GSVA score or CIM score within the TCGA LUAD cohort. Survival analysis was performed by R package survival based on the GSVA score. Kaplan–Meier survival curves were modeled by the survfit function, with the two-sided log-rank test used to compare the curves.

Cell-cell interaction analysis

Cellchat was applied to analyze the strength and number of information interactions between various subpopulations [13]. NicheNet was used to infer the ligand and target gene pairs between SPP1 + macrophages and basal cells [14].

CIM scoring system

An innovative CIM scoring system was devised, drawing upon adaptations from existing research [14]. This system incorporated a distinctive CIM signature composed of the top 20 marker genes present in SPP1 + Mps, basal cells, and genes derived from interacting pairs. The CIM model was then meticulously crafted by employing the least absolute shrinkage and selection operator (LASSO). The CIM scoring system was formulated by integrating individual normalized gene expression values, each with assigned weights based on their respective LASSO Cox coefficients. Ten-fold cross-validation was performed to determine the optimal penalty parameter (λ) that minimized the partial likelihood deviance. Genes with non-zero coefficients at the optimal λ value were retained, resulting in a final set of 13 genes that comprised the CIM signature.

Multiplex immunohistochemistry (mIHC)

The Opal Polaris™ 7-color Manual IHC Kit (AKOYA Biosciences, NEL861001KT) was utilized for the multiplex immunohistochemistry (mIHC) experiment. Initially, 4 μm tissue sections were heated at 65 °C for 1 h. Subsequently, the sections underwent dewaxing in xylene and rehydration through a gradient of alcohol solutions. Post rehydration, the sections were fixed in 10% neutral buffered formalin (NBF) (G2161, Solarbio, China) for 20 min. Epitope retrieval was carried out using an appropriate antigen retrieval buffer in a microwave, starting at 100% power for one minute and then reducing to 20% power for 15 min. After cooling to room temperature, the sections were blocked with the blocking/antibody diluent for 10 min, followed by a 10-minute incubation with primary antibodies. Following exposure to the primary antibodies, the sections were washed three times with 1× TBST (T1082, Solarbio, China). The tissue sections were then treated with Opal Polymer HRP Ms + Rb for another 10 min. After three additional rinses with TBST, the sections were treated with opal reagent, repeating the process. A comprehensive application of Opal TSA-DIG and Opal Polaris 780 labeling ensued. Finally, the slides were washed with TBST and distilled water before staining the nuclei with Mowoil (81381, Sigma-Aldrich) supplemented with DAPI (C1002, Beyotime). The 7-color slides were visualized using the Vectra Polaris Quantitative Imaging System.

Hematoxylin and eosin (H&E) staining

The tissues were fixed in 4% paraformaldehyde for a duration of one day before being embedded in paraffin. Subsequent to this, tissue slices of 4 μm thickness underwent deparaffinization in xylene and gradual rehydration through a series of alcohol dilutions. Following three washes with PBS (P1020, Solarbio), the sections were

stained with hematoxylin (G1080, Solarbio) for a period of thirty minutes. Post-staining, the slides were rinsed in PBS and then immersed in 1% ammonia water (G1822, Solarbio) to impart a blue-violet hue to the nuclei. Thereafter, the sections were briefly immersed in 75% alcohol and stained with eosin (G1100, Solarbio) for 1 h to accentuate the cytoplasm. Following rinses with a series of alcohol solutions, xylene was employed to displace the dehydrated alcohol before mounting. The tissues were observed with a Leica light microscope and quantified with image processing software (Image-Pro Plus).

Western blotting

Cells were lysed with ice-cold radioimmunoprecipitation assay (RIPA) lysis buffer (R0020, Solarbio) supplemented with a protease inhibitor (P6731, Solarbio). The resulting protein concentration of the lysate was determined with a BCA Protein Assay Kit (P0010S, Beyotime). Subsequently, proteins were segregated via sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and later transferred onto nitrocellulose membranes (HATF00010, Merck Millipore, USA) for immunoblotting procedures. To avert non-specific interactions, the membranes were subjected to a 1 h blocking step at room temperature and then incubated overnight at 4 °C with specific primary antibodies. The next day, the membranes were washed three times with PBS containing 0.1% Tween® 20 (PBST) and then incubated with the secondary antibody for 1 h at room temperature. Additional PBST washes were carried out before visualizing on a LI-COR Odyssey imager. The images obtained were quantitatively assessed using Image Studio software (Version 4.0). All Western blot experiments were performed using three independent biological replicates.

Real-time PCR

Total RNA was extracted from individual cell suspensions utilizing either the RNeasy Mini Kit (QIAGEN, 74106) or RNA-Bee (Tel-Test). The poly(A)-tailed mRNA was subsequently reverse transcribed with SMART MMLV reverse transcriptase (Takara Bio, 639524). PCR amplification was carried out on a BioRad real-time PCR machine utilizing the SYBR green QPCR mix (Quantabio, 95056–02 K) and gene-specific primers that span introns. The expression levels of the target genes were standardized to GAPDH. The primer sequences utilized in this experiment are detailed in the key resource table. Quantitative PCR analyses were conducted on a minimum of three independent biological replicates, each measured in technical triplicate.

Mass spectrometry analysis

Trichloroacetic acid (TCA) was used to precipitate proteins, which were then resuspended in buffer containing 8 M urea and 100 mM Tris (pH 8.5). In preparation for tryptic digestion, the proteins were reduced and alkylated, and digested overnight at 37 °C with sequencing-grade trypsin (Promega). The peptides produced by this digestion were analyzed on an Easy-nLC 1000 UPLC system (Thermo Fisher Scientific) directly coupled online to a Q Exactive mass spectrometer (Thermo Fisher Scientific). Peptides were first loaded onto a trap column (75 μm inner diameter [ID] packed with 8 cm ODS-AQ 12 nm S-10 μm resin; YMC Co., Ltd) before separation on an analytical column (75 μm ID packed with 11 cm Luna 3 μm 100 Å C18 resin, Phenomenex). The instrument was set to allow acetonitrile separation with a linear gradient of 0–30% at a flow rate of 300 nL/min. The dynamic exclusion setting was set for 60 s duration. For the identification of peptides, tandem mass spectra (MS/MS) were screened against a customized database using the ProLuCID search engine. Search results were filtered using DTASelect 2.0, with 7 ppm precursor mass tolerance and 5% false discovery rate (FDR).

Xenograft tumor formation assay

LUAD cells were cultured in complete growth medium in sterile conditions. To assess the influence of FAM117A expression on LUAD cells, triplicate cultures were grown to approximately 70–80% confluence. The medium was then removed and replaced with fresh medium. After allowing the cells to reattach for an additional 3–4 h without antibiotics, the toxicity of the growth medium to trypan blue exclusion was replaced with phosphate-buffered saline (PBS). After washing with PBS, cells were treated with trypsin-EDTA for the harvesting process. The culture flasks were centrifuged for 2–5 min at ≤ 1,500 rpm, and the cell pellet was washed with PBS and re-suspended. The number of cells was determined with a hemocytometer, and their viability was assessed by trypan blue exclusion (0.8 mM trypan blue in PBS at a ratio of 1:1 with the cell suspension). Cell suspensions were used if their viability was > 95%. The suspensions were diluted so that the appropriate number of cells was present in 300 μL. Each injection consisted of 3.0×10^6 viable cells per animal. The female BALB/c-nu/nu mice (4–6 weeks old) used in these studies were obtained from Vital River Laboratories (Beijing, China) and were acclimatized for 3–5 days in the animal care facility prior to the experiments. The mice were assigned to experimental groups ($n = 5$ per group) in a random fashion to include two groups with FAM117A-overexpression and three groups with FAM117A-silenced. All animal use procedures were approved by the Ethics Committee of the PLA General Hospital, while animal care followed institutional biosafety and ethics guidelines.

For tumor implantation, the lower flank was cleaned with a 70% ethanol and/or iodine solution prior to

injection. A suspension of 3.0×10^6 cells in 300 µL PBS/Matrigel mixture was loaded into a 1 mL syringe without a needle in order to reduce shear stress, and then injected subcutaneously through a 27- or 30-gauge needle. Tumor progression was monitored twice weekly, and tumor diameters were measured using digital calipers. Tumor volume (mm^3) was calculated using the formula: volume = (width² × length) / 2. Treatment was started when the tumors reached an average size of approximately 50–60 mm^3, which typically occurred 1–3 weeks post-inoculation.

Results

A single-cell atlas of the lung adenocarcinoma ecosystem

The research workflow is shown in Fig. 1A. Integration analysis was conducted on both single-cell RNA

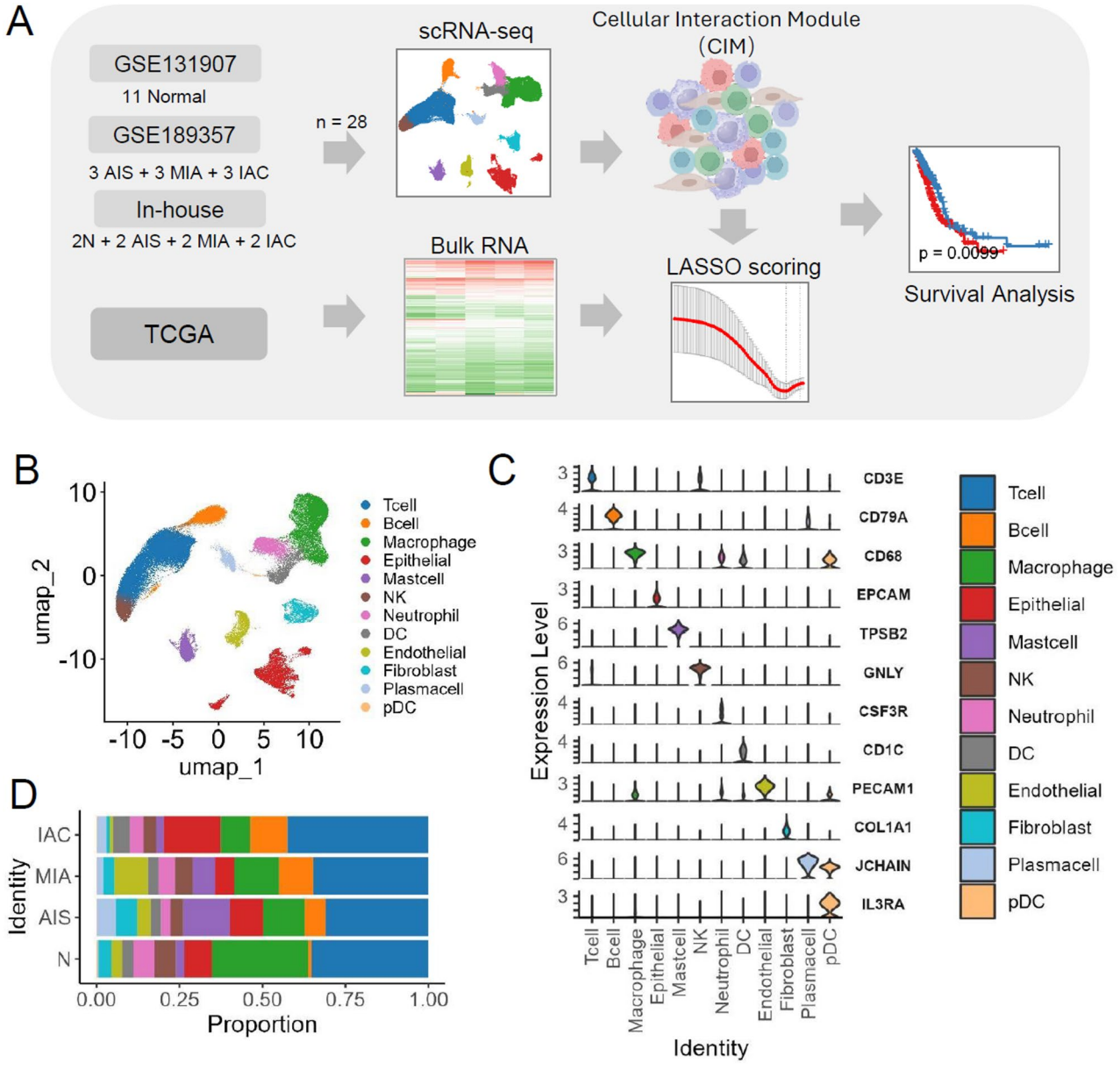


Fig. 1 Single-cell atlas of the lung adenocarcinoma ecosystem. **A** Overview of the study design integrating scRNA-seq and TCGA bulk RNA-seq datasets. scRNA-seq data were obtained from GSE131907 ($n = 11$ normal), GSE189357 ($n = 9$ sequential LUAD samples), and an in-house cohort ($n = 8$). This resulted in a total of 28 samples representing normal (N), adenocarcinoma in situ (AIS), minimally invasive adenocarcinoma (MIA), and invasive adenocarcinoma (IAC). **B** UMAP visualization of 152,686 high-quality cells from LUAD tissues, annotated into distinct clusters. **C** Expression of canonical marker genes used to identify major immune and non-immune populations, including T cells (CD3E), B cells (CD79A), NK cells (GNLY), plasma B cells (JCHAIN), macrophages (CD68), mast cells (TPSB2), neutrophils (CSF3R), dendritic cells (CD1C), plasmacytoid dendritic cells (IL3RA), epithelial cells (EPCAM), endothelial cells (PECAM1), and fibroblasts (COL1A1). **D** Distribution of 12 annotated cell types across normal, AIS, MIA, and IAC groups

sequencing (scRNA-seq) and bulk RNA-seq datasets from the TCGA database. The scRNA-seq dataset comprised 11 normal samples from GSE131907, 9 samples representing the sequential progression of LUAD from GSE189357, and 8 samples from our in-house cohort. This comprised a total of 13 normal (N), 5 adenocarcinomas in situ (AIS), 5 microinvasive carcinomas (MIA), and 5 invasive adenocarcinomas (IAC). Although AIS, MIA, and IAC samples were analyzed in an order that reflects the established pathological stages of LUAD progression, it is important to note that this study was based on a cross-sectional sampling design rather than longitudinal follow-up of the same patients. Therefore, the inferred "temporal progression" represents a pseudo-temporal trajectory reconstructed across independent patient samples at different disease stages, rather than direct evidence of within-patient temporal evolution. This analytical framework allows the identification of stage-associated cellular and molecular patterns, but does not capture dynamic transitions at the individual level. Based on the integrative analysis of the scRNA-seq and bulk RNA-seq datasets, a cellular interaction module was identified as the CIM. Subsequently, by leveraging the established CIM, a LASSO score model was devised to stratify patients with poor prognosis using the TCGA dataset.

Our initial focus centered on investigating TME heterogeneity and the molecular signatures associated with the cellular components of LUAD by leveraging scRNA-seq techniques (Fig. 1B and C). Dimensionality reduction and cell clustering were facilitated through Uniform Manifold Approximation and Projection (UMAP), which was applied to genes exhibiting variable expression levels across all cells. By integrating data from the public datasets, we curated a total of 152,686 high-quality cells derived from 20 patients. These met the criteria of a minimum of 200 genes, a maximum of 6000 genes, and < 10% mitochondrial reads per cell. To further verify that CCA-based integration did not obscure dataset-specific biological variation, the integrated single-cell atlas was visualized using UMAP colored by the original data source (Figure S1). Cells derived from different datasets showed substantial overlap within shared cell-type clusters, rather than segregating by dataset origin, indicating effective mitigation of batch effects.

The expression patterns of canonical markers were examined in order to characterize each cell cluster (Fig. 1C). These canonical markers were utilized to manually identify cell clusters. Four lymphocyte populations—T cells, B cells, NK cells, and plasma B cells—were distinguished by their elevated expression levels of CD3E, CD79A, GNLY, and JCHAIN, respectively. Myeloid cells were characterized by their specific molecular signatures, including Mps (CD68), mast cells (TPSB2), neutrophils (CSF3R), dendritic cells (DC; CD1C), and plasmacytoid dendritic cells (pDC; IL3RA). Additionally, three non-immune cell populations were labeled: epithelial cells (EPCAM), endothelial cells (PECAM1), and fibroblasts (COL1A1). In total, 12 distinct cell types were identified, with their distribution across groups illustrated in Fig. 1D. The integrated scRNA-seq landscape comprehensively captures the primary cellular constituents of the LUAD TME, encompassing a diverse array of immune and non-immune cells from all stages of LUAD tissues. By resolving major immune and epithelial compartments and their stage-dependent distributions, this analysis provides the foundation for subsequent interrogation of macrophage heterogeneity and its functional relevance during LUAD progression.

Macrophage subpopulations and their immune activities in LUAD

We next investigated the macrophage populations in LUAD. These were subdivided into distinct subclusters: MARCO + Mps, FABP4 + Mps, SPP1 + Mps, APOE + Mps, S100A8 + monocytes, CD1C + dendritic cells (DC), and CLEC9A + DC, as shown in Fig. 2A. The specific markers for these populations are shown in Fig. 2B. Four macrophage subsets were identified based on the expression of specific macrophage markers—MACRO, FABP4, SPP1, and APOE. Additionally, one monocyte subset was characterized by expression of the classic monocyte marker S100A8. Two DC subsets were identified with the key lineage markers CD1C and CLEC9A.

The distribution of subsets among different groups is depicted in Fig. 2C. For a more granular examination of the subsets, the differential infiltration of macrophage subsets was compared across groups (Fig. 2D). Notably, we observed a consistent trend in SPP1 + Mps, with a gradual increase aligning with the progression of LUAD from non-invasive to invasive stages (N to IAC). This finding implies that SPP1 + Mps could play a crucial role in the development of LUAD.

Cluster-specific feature genes were computed (Fig. 2E), followed by Gene Ontology (GO) analysis to delineate the underlying biological processes of each subpopulation (Fig. 2F). CD1C + DC and CLEC9A + DC were notably enriched in pathways associated with immune responses, whereas S100A8 + monocytes exhibited enrichment in pathways related to cytokine and chemokine production. APOE + Mps displayed enrichment in pathways linked to carboxylic acid catabolic processes. On the other hand, MARCO + Mps were enriched in pathways associated with peptide secretion, while SPP1 + Mps were involved in processes related to metal ion transport.

To assess the impact of macrophage cell populations on the prognosis of LUAD patients, we integrated bulk transcriptome data with scRNA-seq data for in-depth

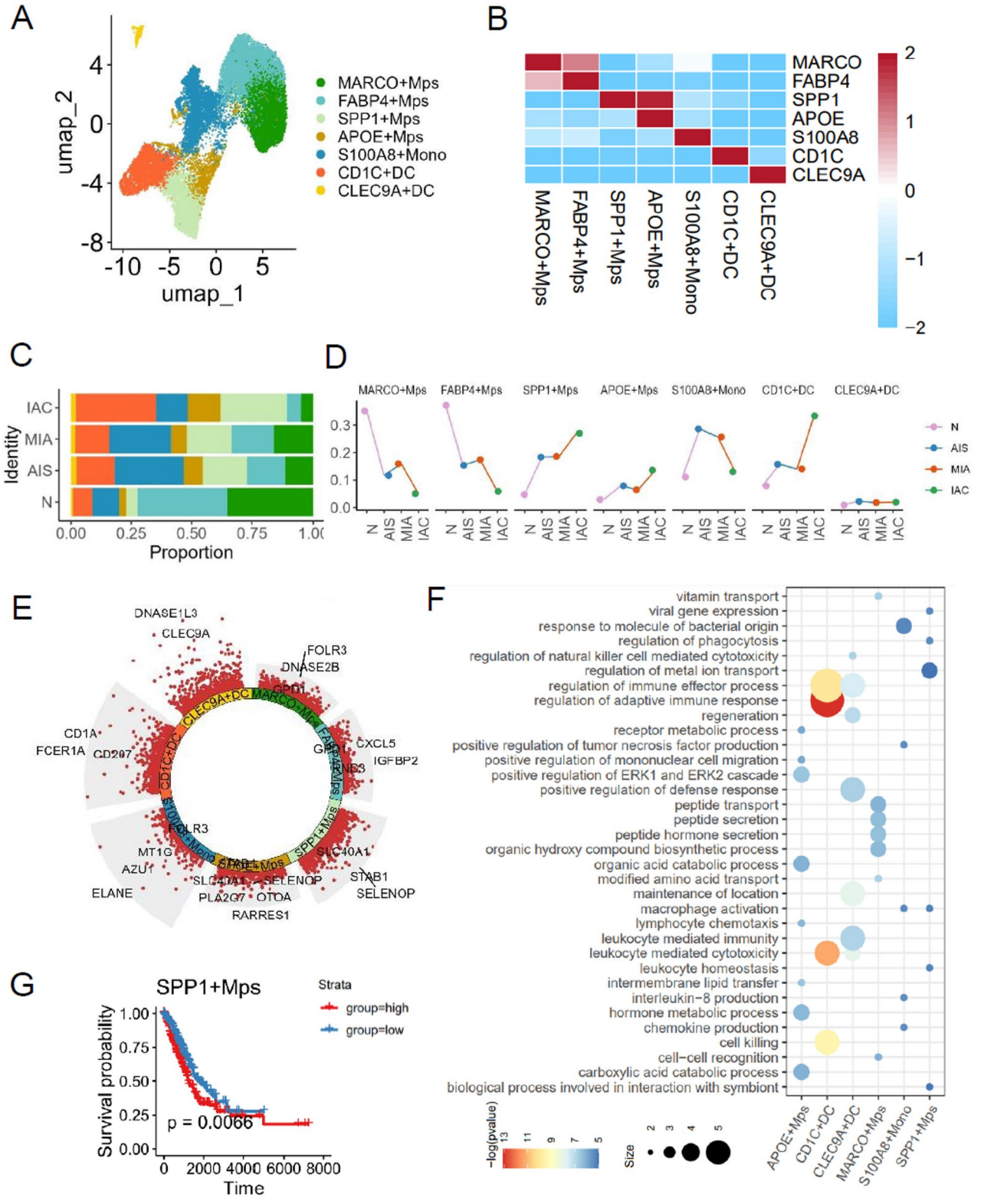


Fig. 2 Macrophage heterogeneity in LUAD. **A** UMAP visualization of macrophage and monocyte subsets. **B** Expression of representative markers defining MARCO + Mps, FABP4 + Mps, SPP1 + Mps, APOE + Mps, S100A8 + monocytes, CD1C + DC, and CLEC9A + DC. **C** Distribution of subsets across disease stages (N, AIS, MIA, IAC). **D** Progressive increase of SPP1 + Mps during LUAD progression. **E** Subset-specific feature genes. **F** GO enrichment analysis of macrophage and monocyte subsets. **G** Kaplan–Meier survival analysis of LUAD patients stratified by SPP1 + macrophage GSVA scores using TCGA data. Patients enriched in SPP1 + Mps showed significantly worse survival ($p < 0.05$, log-rank test)

analysis. By utilizing the top 20 feature gene signatures of each macrophage subpopulation derived from scRNA-seq, we investigated the relationship between macrophage populations and the overall survival (OS) of CRC patients. Patients exhibiting enrichment of SPP1 + Mps demonstrated a poorer prognosis (Fig. 2G), highlighting the potential of SPP1 + Mps as a prognostic biomarker.

Collectively, the above results demonstrate marked heterogeneity among macrophage populations in LUAD, and identify SPP1 + Mps as a stage-enriched subset associated with adverse prognosis. These observations prompted further investigation into the functional states, differentiation trajectories, and regulatory programs underlying the SPP1 + macrophage population.

The diverse functions within macrophage populations

To further analyze the functionality difference of the identified macrophage populations, we performed pseudotime analysis using the monocyte-macrophage lineage. For this analysis, single cells are ordered along a trajectory or lineage, and each cell receives a specific pseudotime value, representing its position along that trajectory. The location of the S100A8 + monocyte subsets was predicted to be mainly at the beginning of the differentiation site, followed by APOE + Mps and FABP4 + Mps. The

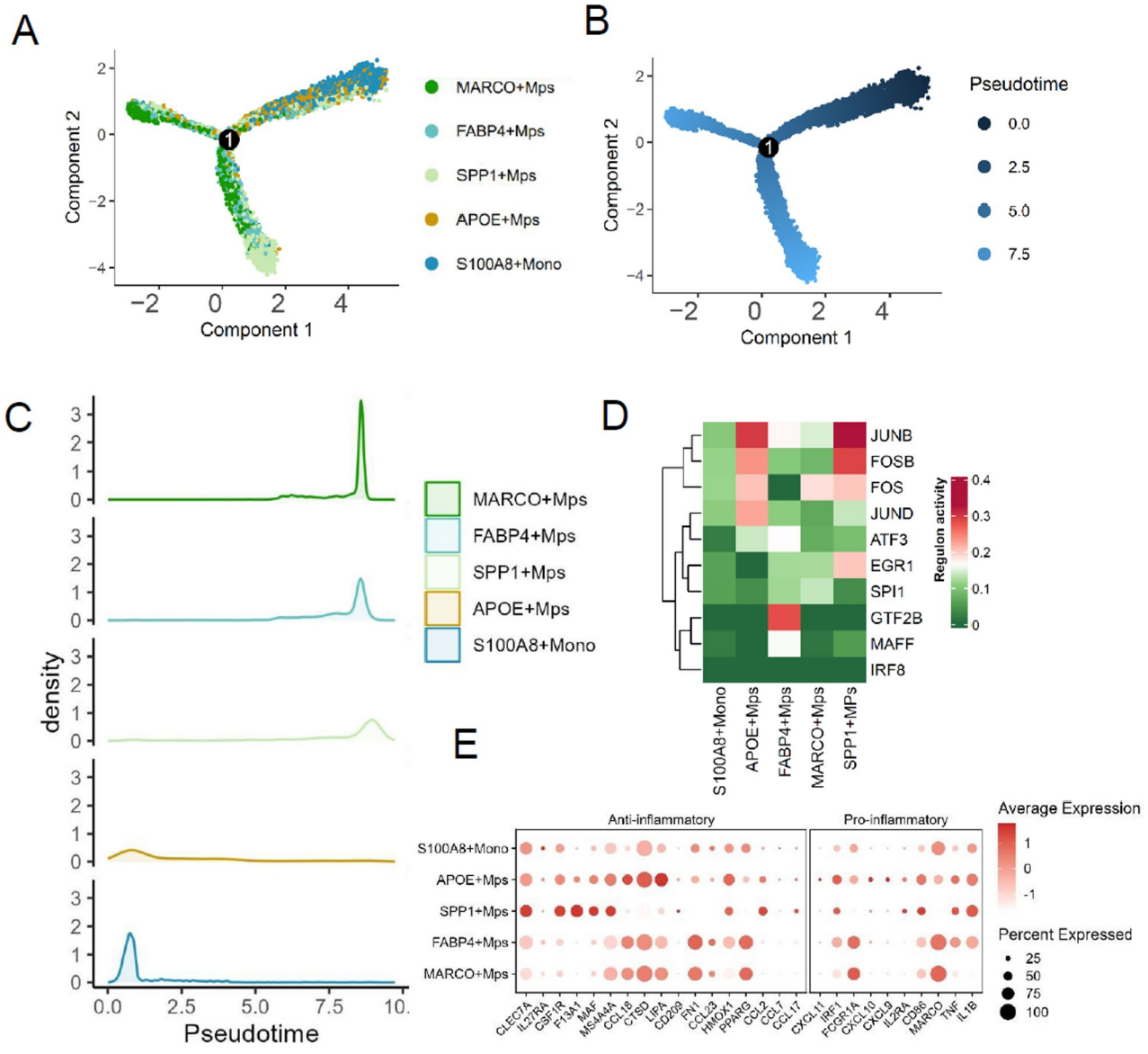


Fig. 3 Functional properties of macrophage subsets. **A** Pseudotime trajectory analysis showing differentiation of monocyte–macrophage lineages. **B**, **C** Distribution of subsets along pseudotime, with S100A8 + monocytes at the origin, and MARCO + and SPP1 + Mps enriched at terminal states. **D** Transcription factor activity inferred by SCENIC, highlighting JUNB, FOSB, FOS, and EGR1 as key regulators in SPP1 + Mps. **E** Expression of pro-inflammatory (M1-like) and anti-inflammatory (M2-like) genes across macrophage subsets, showing the enrichment of M2-associated signatures in SPP1 + Mps

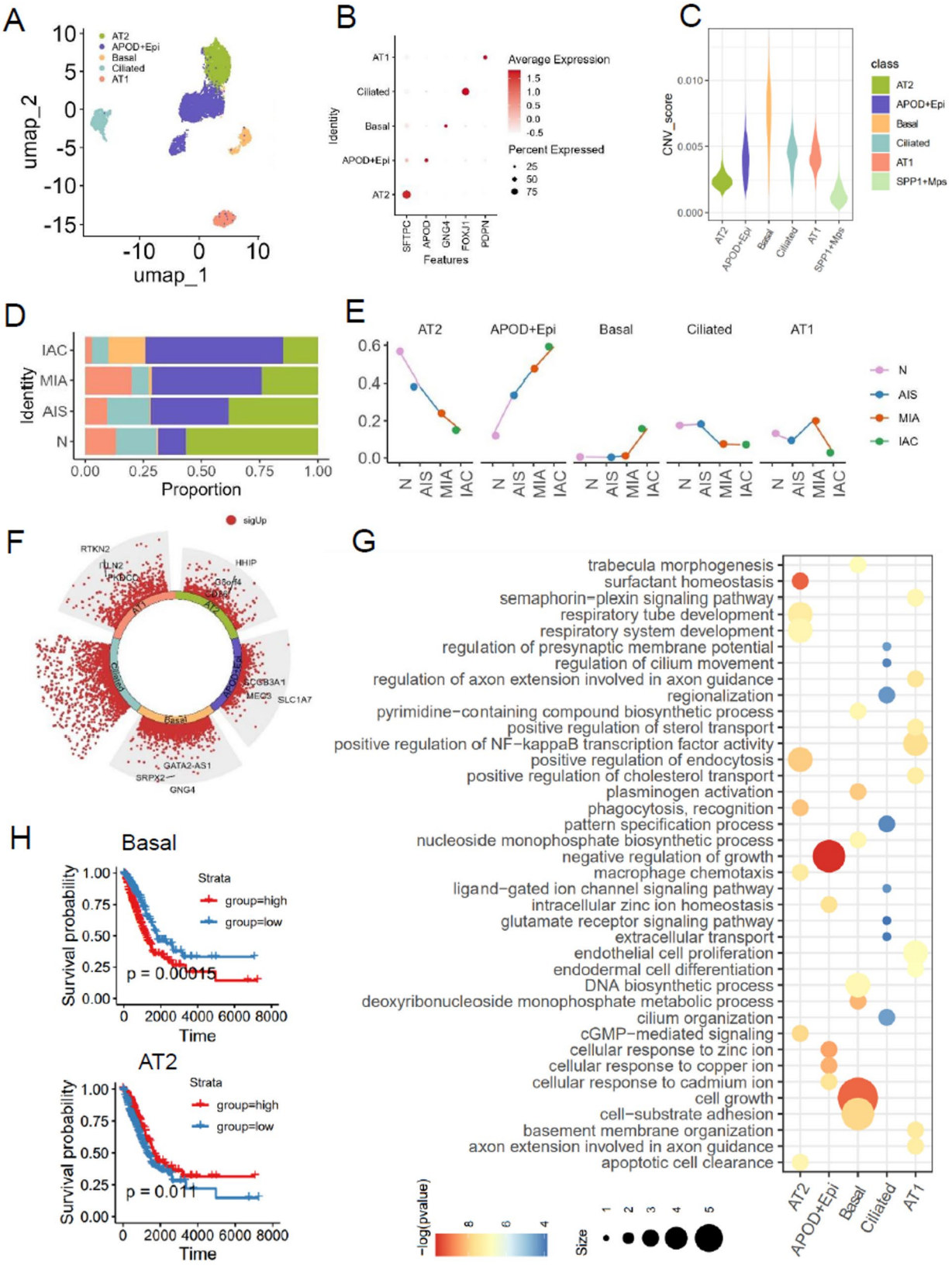


Fig. 4 (See legend on next page.)

(See figure on previous page.)

Fig. 4 Epithelial cell heterogeneity in LUAD. **A** UMAP visualization of epithelial cell subsets. **B** Expression of marker genes distinguishing AT2 (SFTPC), AT1 (PDPN), basal (GNG4), ciliated (FOXJ1), and APOD + epithelial cells. **C** CNV profiles of epithelial subsets compared with SPP1 + macrophages as the reference. Basal cells showed the highest CNV scores, suggesting malignant potential. **D**, **E** Distribution and dynamic changes of epithelial subsets across N, AIS, MIA, and IAC groups. The presence of basal and APOD + epithelial cells increased with progression, whereas AT2 cells decreased. **F** Subset-specific feature genes. **G** GO enrichment analysis of epithelial subsets. **H** Kaplan–Meier survival analysis of LUAD patients stratified by GSVA scores of the basal and AT2 subsets. High basal cell enrichment correlated with poor survival, whereas enrichment with AT2 cells correlated with a favorable outcome ($p < 0.05$, log-rank test)

MARCO + Mps and SPP1 + Mps predominate at the terminal differentiation site (Fig. 3A, B, and C). This strategy allowed subsequent macrophage subsets to be ordered along a biologically plausible differentiation continuum from monocytes toward mature macrophage states.

The differentiation of Mps is orchestrated by a sophisticated network of transcription factors (TFs) that regulate each other and their effectors by interacting with cofactors and downstream genes. Therefore, we evaluated the top 10 specifically expressed TFs by pySCENIC, which infers TFs, gene regulatory networks, and cell types from scRNA-seq data. JUNB, FOSB, FOS, and EGR1 were found to have specific activity levels in the regulatory network of SPP1 + Mps, and may represent the master TF driving oncogenic effects (Fig. 3D). Among them, high expression of EGR1 (early growth response factor 1) has previously been observed in various tumor types. As a downstream protein of the MAPK signaling pathway, EGR1 can induce tumor cell proliferation by upregulating cell-cycle-related proteins.

In terms of macrophage activation, these cells can be polarized into M1 and M2 phenotypes. M2 Mps exhibit anti-inflammatory characteristics, whereas M1 Mps feature pro-inflammatory characteristics. To investigate these pro- and anti-inflammatory characteristics, we examined the expression pattern of a panel of marker genes. SPP1 + Mps showed upregulated expression of anti-inflammatory marker genes, thus resembling an M2 phenotype. The expression of pro-inflammatory marker genes in SPP1 + Mps was weaker than that of anti-inflammatory markers, suggesting incomplete differentiation and resembling a subtype of the M2 phenotype (Fig. 3).

In summary, pseudotime trajectory analysis and TF activity profiling indicate that SPP1 + Mps represent a terminally differentiated, transcriptionally distinct population with immunosuppressive features. These findings suggest a potential role for SPP1 + Mps in shaping tumor–epithelial interactions, warranting analysis of their crosstalk with malignant epithelial cell subsets. Therefore, we subsequently focused on SPP1 + Mps and their relationship with epithelial cell populations.

The diversity of components and pathways within epithelial cells in LUAD

We next analyzed epithelial cell populations in LUAD, leading to the identification of five distinct subpopulations: alveolar type 2 cells (AT2), APOD + epithelial cells (APOD + Epi), basal cells, ciliated epithelial cells, and alveolar type 1 cells (AT1) (Fig. 4A). The alveolar epithelial cell types, AT1 and AT2, were distinguished by their expression of Surfactant Protein C (SFTPC) and Podoplanin (PDPN), respectively. Ciliated epithelial cells were characterized by expression of the marker gene FOXJ1, and basal cells by expression of the canonical basal cell marker GNG4. APOD + epithelial cells subpopulations were identified based on their expression of specific cluster marker genes (Fig. 4B).

To distinguish malignant cells from epithelial populations, we conducted Copy Number Variations (CNVs) analysis using the inferCNV package. This involved averaging the expression of neighboring genes in the genome to eliminate gene-specific expression variability, thereby generating profiles that accurately represent chromosomal CNVs. SPP1 + macrophage immune cells were designated as reference cells, and all epithelial cells were presumed to be malignant cells. The reference sample was then subtracted from the epithelial sample to derive the final CNV profile for each gene and cell. The CNV score for each cell was calculated as the quadratic sum of CNV regions within each epithelial subset (Fig. 3C). This analysis revealed that all epithelial subsets exhibited higher CNV scores than the reference immune cluster. Notably, basal cells demonstrated the highest CNV score, indicating they are likely to be malignant.

The abundance of subsets across various groups is illustrated in Fig. 4D, while alterations in epithelial subsets across the groups are shown in Fig. 4E. A marked decrease in AT2 cells was noted with the sequential progression of LUAD. Conversely, the presence of APOD + epi and basal cells was found to increase with this progression. Basal cells were predominantly enriched in the IAC group, highlighting their tumor-specific origin.

Next, cluster-specific feature genes were identified (Fig. 4F), and Gene Ontology (GO) analysis was conducted to reveal the fundamental biological processes within each subpopulation (Fig. 4G). Of note, APOD + epi cells were enriched in pathways related to the regulation of growth and the cellular response to metal ions, while basal cells showed enrichment in pathways associated with cell growth and cell-substrate adhesion.

Similar to the analysis of macrophage subpopulations, we also evaluated the impact of epithelial cell populations on the prognosis of patients with LUAD. Greater

enrichment with basal cells was associated with a poorer prognosis, whereas enrichment with AT2 cells was linked to a better prognosis (Fig. 4H).

Taken together, our scRNA-seq analysis revealed significant heterogeneity among the identified epithelial cell populations. Particularly noteworthy was the presumptive malignant nature of basal cells, which appear to play a critical role in LUAD. Consequently, we explored the relationship between basal cells and SPP1 + Mps in further detail.

Cellular interaction module composed of SPP1 + macrophages and basal cells

We next investigated complex cell crosstalk to uncover the regulatory mechanisms within the LUAD TME. Cell-cell interactions (CCIs) between Mps and epithelial cell subsets were identified. Notably, Mps exhibited strong CCI activities with multiple epithelial cell populations (Fig. 5A). Subsequently, our focus shifted to the interaction between SPP1 + Mps and epithelial populations. SPP1 + Mps were found to have the highest number of CCIs with basal cells, with a significant correlation between these cell subsets (Fig. 5B).

Following on from the previous investigations of clinical relevance (Figs. 2G and 4H), a significant relationship was observed between basal cells and poor prognosis among lung cancer patients in the TCGA-LUAD database (Fig. 4H). These results suggest the malignant cell subset exhibiting the most interactions with SPP1 + Mps could potentially indicate a poorer prognosis for lung cancer patients.

To pinpoint essential mediators of the interaction between SPP1 + Mps and basal cells in LUAD patients, we delved further into the intricacies of cell-cell communication mechanisms. This involved assessment of putative crosstalk with the R package "NicheNet," which examines the expression profiles and downstream targets of ligand-receptor pairs.

SPP1 + Mps were found to engage with basal cells through the secretion of cytokines such as IL1A, IL1B, IL6, and TNF (Fig. 5C). This signaling cascade may trigger NF-kB pathway activation, consequently driving cell cycle progression, tumor growth, and basal cell resistance to therapy. In addition to cytokines, SPP1 + Mps can specifically support basal cell proliferation via insulin-like growth factor signaling (IGF1/IGFBP3) (Fig. 5C).

Based on the integrative analysis of scRNA-seq and bulk RNA-seq datasets, we identified a cellular interaction module (CIM). This was defined as a coordinated multicellular unit composed of interacting SPP1 + macrophages and basal epithelial cells, along with their associated ligand–receptor signaling programs. The elevated

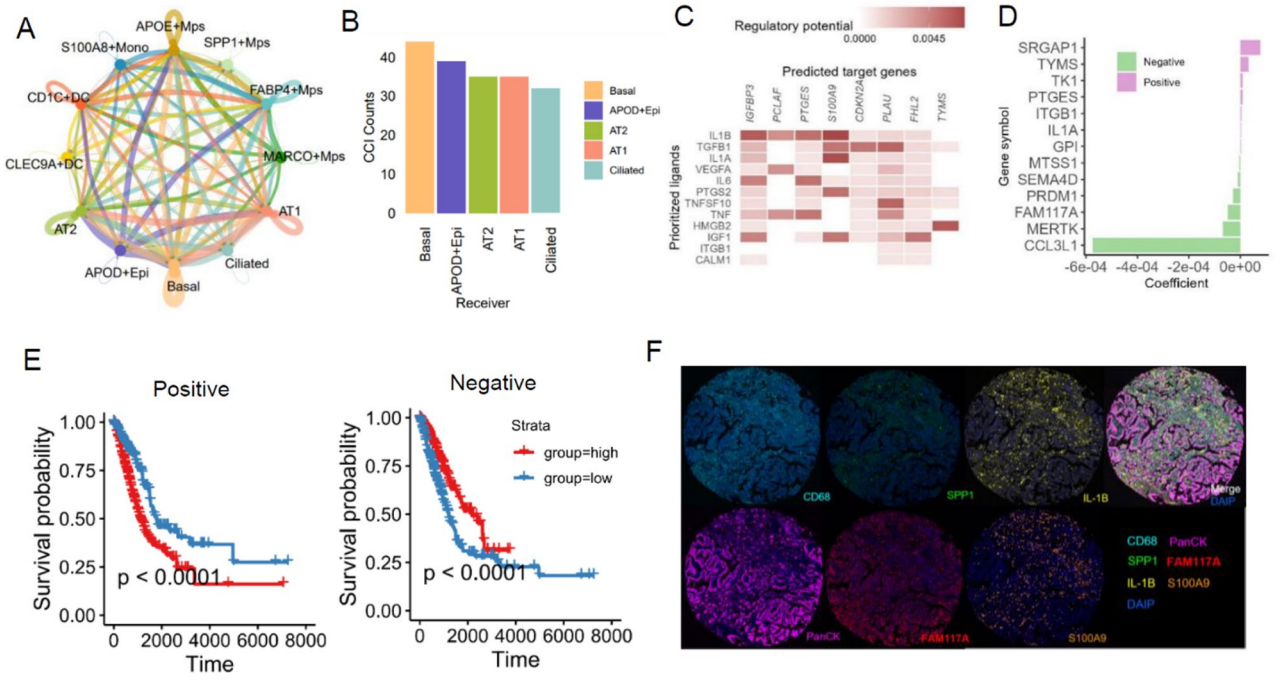


Fig. 5 Establishment of the SPP1 + macrophage–basal cell cellular interaction module (CIM). **A** Network visualization of predicted cell–cell interactions among macrophage and epithelial subsets. **B** Quantification of interactions showing the strongest connectivity between SPP1 + macrophages and basal cells. **C** Ligand–receptor analysis using NicheNet identified pro-inflammatory cytokines (IL1A, IL1B, IL6, TNF) and IGF signaling (IGF1/IGFBP3) as mediators of SPP1 + macrophage–basal cell crosstalk. **D** LASSO regression coefficients for the CIM gene set. **E** Kaplan–Meier survival analysis stratified by positive and negative CIM scores in the TCGA cohort. High CIM scores predicted worse survival ($p < 0.05$). **F** Multiplex immunohistochemistry (mIHC) staining showing spatial co-localization of CD68 (cyan), IL1B (yellow), PanCK (magenta), SPP1 (green), S100A9 (orange), and DAPI (blue) in LUAD tissues

concentration of CIM entities indicates their strong influence on cancer cell proliferation.

A CIM scoring mechanism was devised to categorize LUAD patients into distinct prognostic groups. The top 20 genes characteristic of SPP1 + Mps and basal cells, along with their intercellular ligand-receptor pairs, were utilized as inputs for the LASSO Cox regression model. Following regression analysis, the final 13-gene CIM signature was selected based on non-zero coefficients derived from the LASSO Cox model at the optimal cross-validated λ value. This reflected genes having the strongest and most stable association with overall survival, as depicted in Fig. 5D. A positive CIM score suggests that an increase in gene expression is linked to more severe disease. Conversely, a decrease in gene expression is linked to less severe disease. By leveraging the coefficients and specific expression profiles of the 13 genes, positive and negative CIM scores were computed for each patient. Patients with higher positive CIM scores were found to exhibit worse survival compared to those with lower CIM scores. Conversely, patients with negative CIM scores showed the opposite trend, as illustrated in Fig. 5E. The topological crosstalk pattern of SPP1 + Mps and basal cells was validated by mIHC technology using specific markers (Fig. 5F).

In summary, we uncovered a specific cellular composition involving SPP1 + Mps and basal cells through integration analysis of scRNA-seq and bulk RNA-seq. Termed the CIM module, this cellular arrangement has a critical role in maintaining cancer cell proliferation, thereby impacting the progression and unfavorable prognosis of lung cancer.

The CIM component is implicated in cell proliferation

Prior research has established FAM117A as a prognostic marker for LUAD [7]. In the present study, FAM117A was identified within the CIM module. To further investigate the components of CIM, we conducted validation studies of FAM117A functionality using patient tissue. In our scRNA-seq dataset, the expression of FAM117A progressively increased from normal tissue to the IAC stage (Fig. 6A). A total of 57 LUAD patients with complete clinical data and available tissue specimens were screened. Postoperative pathology staging was performed according to the 8th edition of the TNM staging system for lung cancer, with detailed information shown in Table S1.

The protein and mRNA expression levels of FAM117A, p21 (cell cycle-related inhibitory protein), and Cyclin D (cell cycle-promoting protein) are shown in Fig. 6B and C, respectively. The protein expression of p21 was consistent with that of FAM117A, whereas an inverse trend was observed between Cyclin D and FAM117A protein expression. The detection of mRNA through real-time fluorescence quantitative PCR technology showed that expression levels of FAM117A, p21, and Cyclin D mRNA were essentially consistent with the protein levels. The IHC results also showed that the FAM117A protein was downregulated in tumor tissues (Fig. 6D).

The above results demonstrate that expression of FAM117A is significantly reduced in LUAD tissue compared with para-cancerous tissue, implying a potential role in cell cycle regulation. The association between FAM117A and cell cycle control implicates the CIM module as a key regulator of tumor cell heterogeneity in the development of LUAD.

FAM117A interacts with DYRK1A to regulate the G1/S transition

To investigate the molecular mechanism by which FAM117A regulates cell cycle progression, we conducted an immunopurification-mass spectrometry screen for proteins that interact with FAM117A. This screen identified multiple candidate binding partners, with a dual-specificity kinase, DYRK1A, amongst the binding candidates with the highest score (Fig. 7A). Because DYRK1A is a well-characterized G_1/S regulator, we proceeded to validate the interaction between FAM117A and DYRK1A [15]. Co-immunoprecipitation experiments confirmed a specific interaction of FAM117A with DYRK1A in LUAD cells. In A549 and NCI-H1975 cells transfected with the Flag-tagged FAM117A plasmid, DYRK1A was readily detected in the Flag immunoprecipitates, but almost undetectable in control IgG precipitations (Fig. 7B). This indicates that FAM117A and DYRK1A interact physically to form a complex that can potentially regulate cell cycle progression.

We next examined whether the FAM117A–DYRK1A complex functionally modulates the G_0/G_1-to-S phase transition. A549 and NCI-H1975 cells underwent RNA interference-mediated knockdown of FAM117A or DYRK1A alone, or knockdown of FAM117A and DYRK1A simultaneously. Flow cytometric analysis of the cell cycle revealed that silencing of FAM117A or DYRK1A caused significant accumulation of cells in the G_0/G_1 phase with a concomitant drop in the S-phase fraction (relative to controls), indicating a slowdown in the G_1/S transition. In line with this observation, Western blotting of cell cycle regulators revealed that silencing of FAM117A or DYRK1A promoted higher levels of Cyclin D1 protein and increased the phosphorylation of CDK4^Thr172 (Fig. 7C). This was accompanied by a significant reduction in the CDK inhibitors p21^Cip1 and p27^Kip1. The impact of silencing FAM117A alone was comparable to the silencing of DYRK1A alone, with no significant differences observed between the two single knockdown conditions. Moreover, the co-silencing of FAM117A and DYRK1A did not result in any additive effects on the expression levels of Cyclin D1,

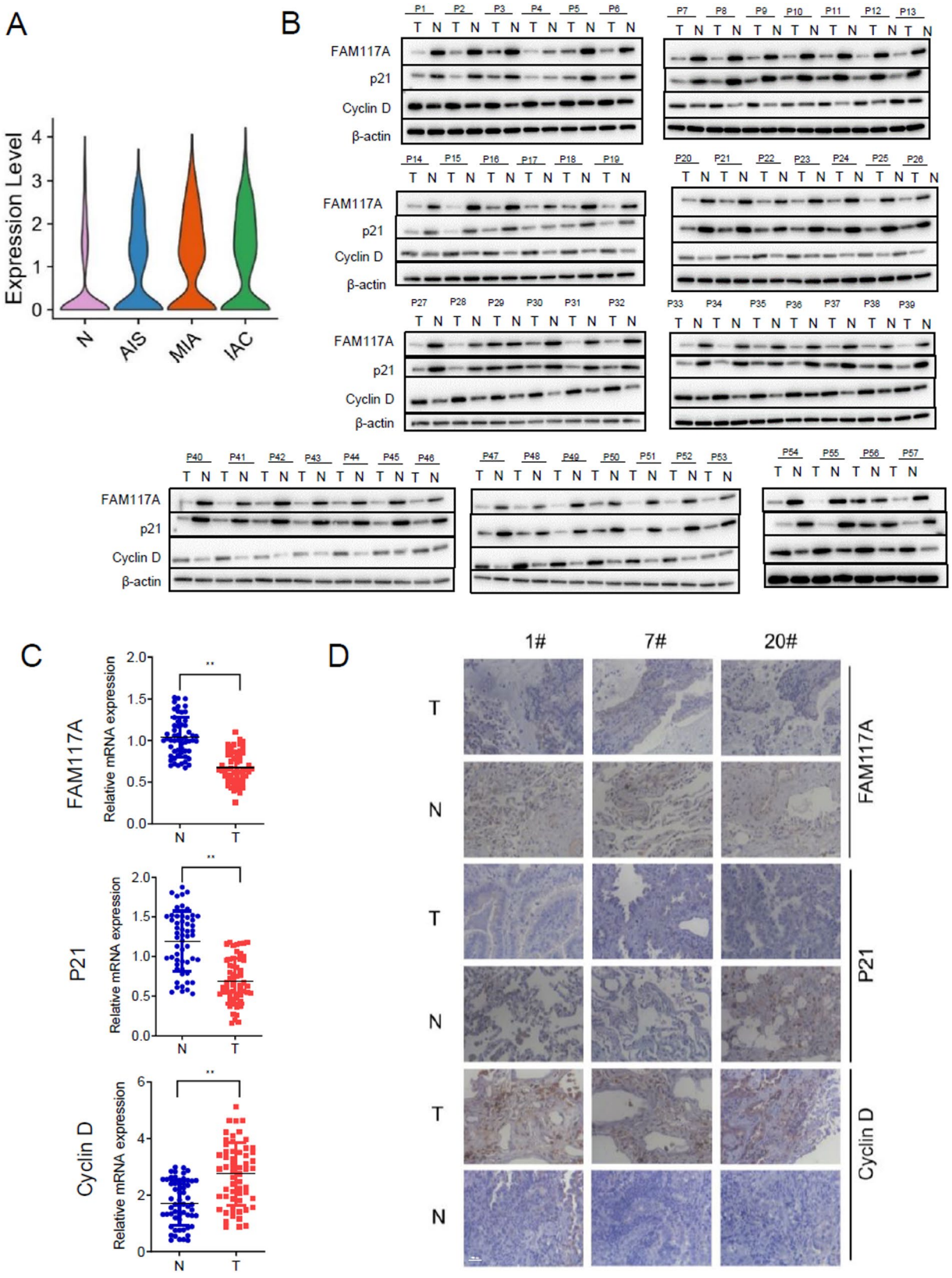
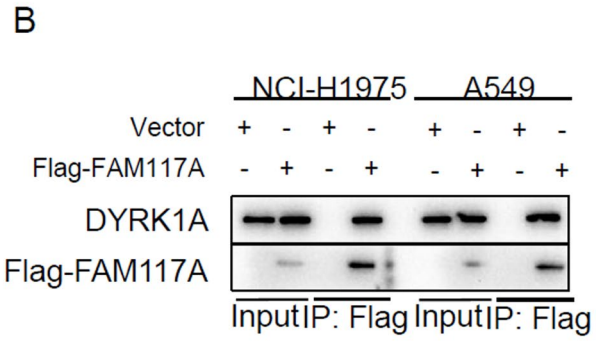
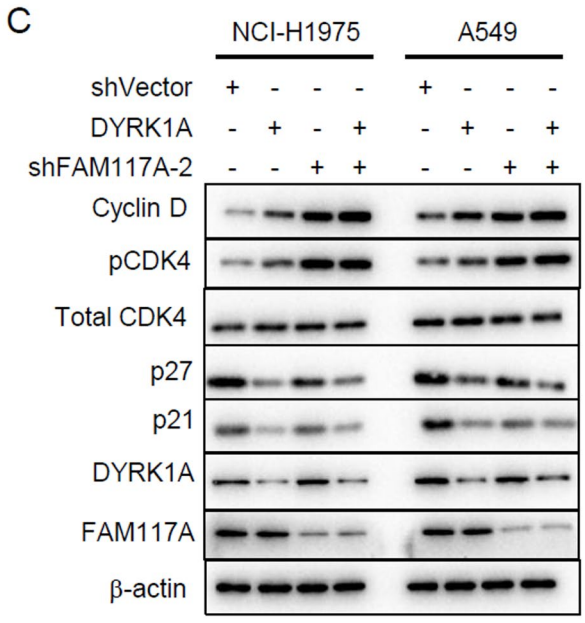


Fig. 6 FAM117A expression and association with cell cycle regulation in LUAD. **A** Expression of FAM117A across N, AIS, MIA, and IAC groups based on scRNA-seq data. **B** Western blots of FAM117A, p21, and Cyclin D1 protein expression in LUAD patient tissues. **C** Quantitative PCR validation of FAM117A, p21, and Cyclin D1 mRNA expression ($p < 0.05$, $p < 0.01$, Student's *t*-test). **D** Immunohistochemistry (IHC) staining showing reduced FAM117A expression in LUAD tissues compared with adjacent normal tissues

A

Name	Matches	Seqences	mPAI
FAM117A	6715	49030	197(154)
EFTU	1579	49852	46(37)
RS3	1423	26842	48(34)
DYRK1A	1269	86043	26(21)
1433E	1069	29326	22(20)
AL1B1	874	57626	35(26)
TOPA	627	60819	32(21)
EPIA3	583	50495	29(18)
MCM7	479	81884	20(14)





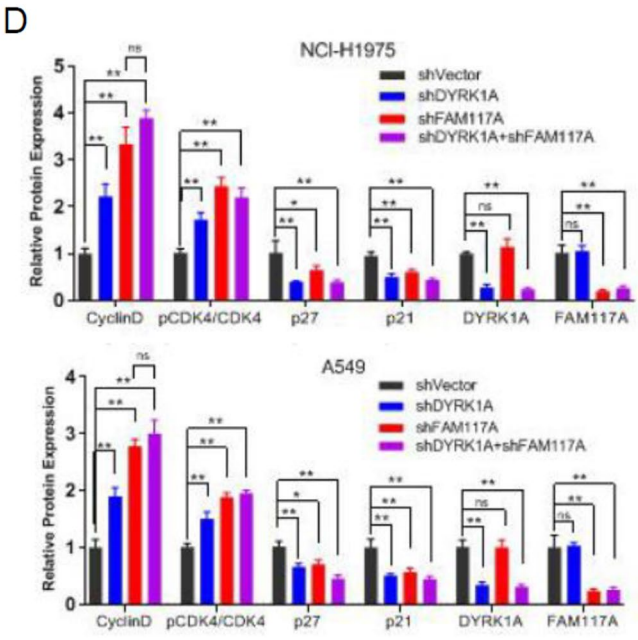


Fig. 7 FAM117A interacts with DYRK1A to regulate G1/S cell cycle transition in lung adenocarcinoma cells. **A** Mass spectrometry analysis of proteins co-purified with Flag-FAM117A in A549 cells identified DYRK1A as one of the top-scoring candidate interactors. **B** Co-immunoprecipitation (co-IP) assays in A549 and NCI-H1975 cells confirmed the interaction between Flag-FAM117A and DYRK1A, as DYRK1A was detected in Flag-FAM117A immunoprecipitates but not in controls. **C** Western blot analysis of cell cycle regulators following siRNA-mediated knockdown of FAM117A or DYRK1A in A549 and NCI-H1975 cells. Silencing of either gene led to increased Cyclin D1 expression and CDK4 phosphorylation, along with marked reductions in the levels of p21 and p27. **D** Comparison of single versus combined knockdown of FAM117A and DYRK1A. Co-silencing of both genes did not result in additive effects compared to single knockdowns, suggesting that FAM117A and DYRK1A function interdependently within the same pathway to regulate G1/S progression. Data are representative of three independent experiments. Quantification of protein expression was normalized to GAPDH and analyzed using Student's *t*-test. Error bars represent the mean ± SD. $p < 0.05$, $p < 0.01$

phospho-CDK4, p21, or p27 compared to single knockdowns (Fig. 7D). Taken together, these data suggest that FAM117A and DYRK1A function within the same pathway or protein complex to curb the G_1/S transition. The absence of an enhanced phenotype upon dual knockdown suggests that FAM117A and DYRK1A function interdependently and form a partnership in the G_1-phase progression of LUAD cells.

FAM117A knockdown enhances tumor growth in vivo

After demonstrating a role for FAM117A–DYRK1A in regulating cell cycle progression in vitro, we next assessed the effect of FAM117A loss on tumor growth in vivo. A subcutaneous xenograft model was established in BALB/c nude mice using A549 control cells and A549 cells exhibiting FAM117A knockdown. Similar to in vitro results, tumors from FAM117A-silenced cells showed accelerated growth compared to tumors from control cells. By the end of the observation period, tumors from the FAM117A knockdown group had significantly greater volume and weight compared to the control group (Fig. 8A–C). This in vivo data supports the conclusion that FAM117A acts as a negative regulator of cell proliferation. The loss of FAM117A results in a growth

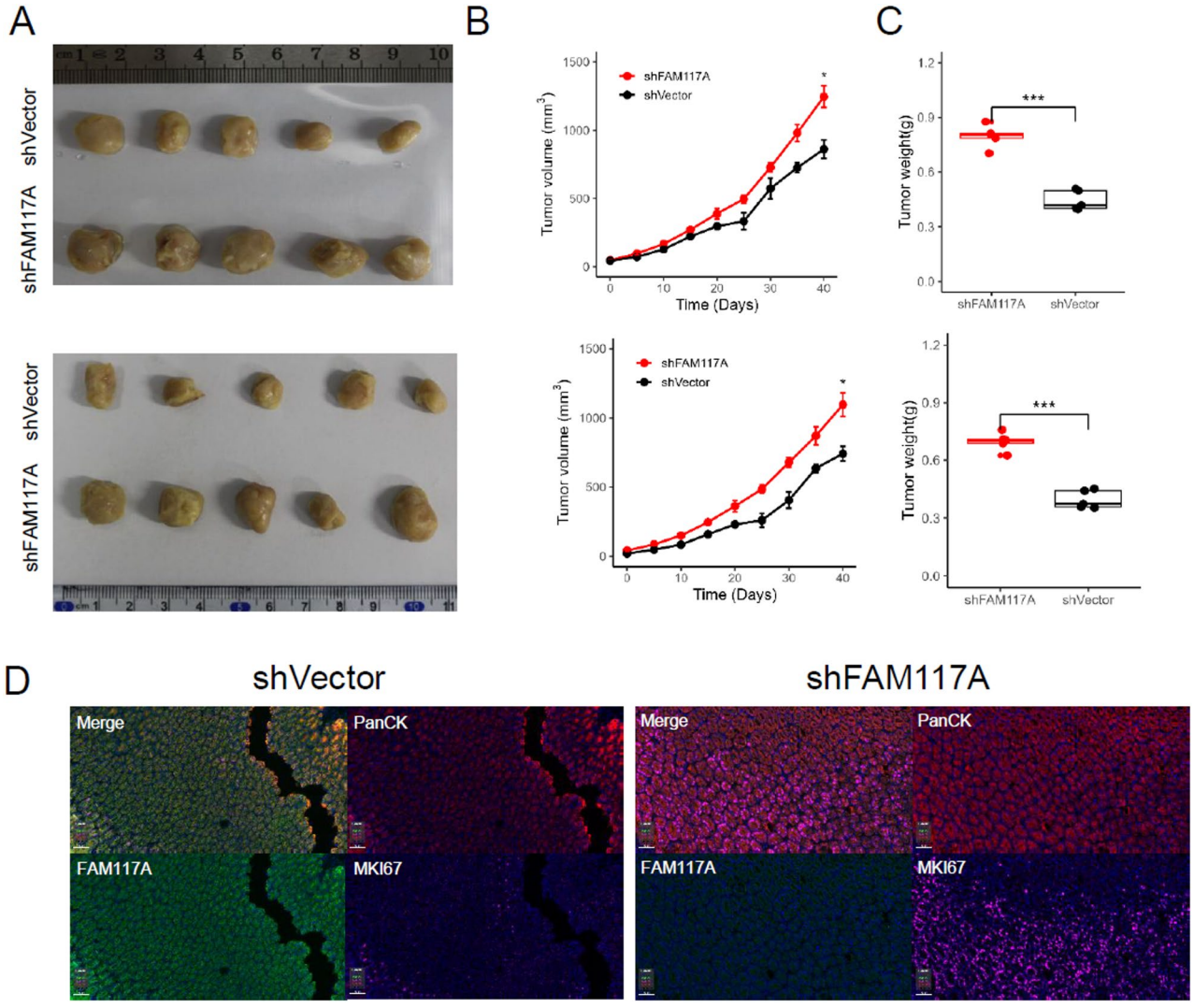


Fig. 8 FAM117A knockdown promotes tumor growth in vivo. **A** Representative images of xenograft tumors derived from A549 cells transduced with control shRNA or FAM117A shRNA#2. **B** Growth curves of xenograft tumors in BALB/c nude mice ($n = 6$ per group). Tumor volumes were measured twice weekly with calipers and calculated using the formula: volume = (width^2 × length) / 2. Tumors from the FAM117A knockdown group showed significantly faster growth compared to controls. **C** Final tumor weights at the study endpoint. Tumors from the FAM117A knockdown group were significantly heavier than those from the control group. Data are presented as the mean ± SD. Statistical comparisons were performed using Student's *t*-test (two-tailed). $p < 0.05$, $p < 0.01$. **D** mIHC showing spatial association between reduced FAM117A expression and tumor regions in LUAD tissues

advantage for LUAD cells and faster tumor growth in the nude mouse model. Additionally, mIHC analyses (Fig. 8D) demonstrate the spatial association between reduced FAM117A expression and regions enriched for proliferating tumor cells. In summary, FAM117A depletion increases the proliferation of LUAD cells in culture conditions and in an animal model, providing evidence of the important role of FAM117A in hindering tumor growth.

Discussion

Although recent advances in scRNA-seq technology have revolutionized our understanding of the TME in LUAD [16–21], in-depth knowledge of the invasive progression of this cancer subtype is still lacking. By conducting integrated analyses of publicly available scRNA-seq and TCGA bulk RNA-seq datasets, the present study has obtained comprehensive data at single-cell resolution on the TME in invasive LUAD. Additionally, our research sheds light on the potential role of the interaction between SPP1 + Mps and basal cells in shaping cancer cell proliferation during the invasive process. This interaction could potentially contribute to tumor cell resistance against immunotherapy.

Tumor-intrinsic regulatory networks, including coordinated miRNA programs, can simultaneously enhance PD-L1 expression and suppress antigen presentation, thereby driving immune evasion and resistance to PD-1/

PD-L1 blockade in NSCLC [22]. However, the heterogeneity within Mps plays a crucial role in modulating the tumor immune microenvironment. Recent evidence points towards a connection between the presence of circulating SPP1 and adverse outcomes in NSCLC patients [23–29]. Furthermore, the secretion of SPP1 by Mps is a significant contributor to the immunosuppression of various cancer types [30–35]. Insights gained from scRNA-seq have revealed a strong connection between the substantial presence of SPP1 + Mps and worse clinical outcomes. Additionally, a potential link between the SPP1/CD44 axis and cell proliferation has been raised, alongside the association of intratumoral heterogeneity with resistance to anticancer therapies in gastric cancer [31], glioma [32], and hepatocellular carcinoma [33]. Importantly, SPP1 + Mps in LUAD patients have been reported to correlate with resistance to immunotherapy and poorer prognostic outcomes [35]. These research findings highlight the crucial involvement of SPP1 + Mps in the diagnosis and treatment of diverse cancers. However, the specific role of SPP1 + Mps in the invasive progression of LUAD remains largely undefined. The present study revealed a positive association between SPP1 + Mps and the invasive development of LUAD, prompting us to focus on the role of these cells in the TME of LUAD. Our results show that SPP1 + Mps express anti-inflammatory genes, and that a greater abundance of these cells correlates with worse overall survival in LUAD patients. Consequently, both the existing body of literature and the current investigation underscore the importance of SPP1 + Mps in the TME, suggesting they may be potential therapeutic targets in LUAD. These outcomes highlight the importance of further research into the functional roles of SPP1 + Mps within the TME.

The TME is comprised of cancer cells, immune cells, mesenchymal stem cells (MSCs), fibroblasts, endothelial cells, and various ECM components [5, 36]. The histological documentation of almost all tumor types originating from epithelial cells consistently indicates a strong correlation between malignancy and the disruption of epithelial organization [37–39].

Our comprehensive analysis of epithelial cells identified a specific subset, basal cells, that display distinct tumor-specific variations in abundance. Significantly, the presence of basal cells was associated with malignancy and predicted a poorer clinical prognosis. Based on our findings, we further examined the interactions between basal cells and SPP1 + Mps using data from both single-cell and bulk transcriptomics. The goal of this work was to provide a novel perspective on the nature of SPP1 + macrophage interactions within the TME.

Emerging evidence highlights the potential role of tumor-associated macrophages (TAMs) in supporting the proliferation of cancer cells and shaping the TME so that it becomes conducive to the survival of cancer cells [40]. TAMs are able to release cytokines such as IL-6, IL-8, IL-1b, and TNF, as well as soluble factors including CCL5, pleiotrophin (PTN), globule-epidermal growth factor-VIII (MFG-E8), and CCL8. Additionally, through direct interactions such as the binding of sinusoidal endothelial cell C-type lectin (abundantly expressed on TAMs) with the BTN3A3 receptor on cancer cells, TAMs contribute to increased tumor cell expression of uPAR and Ki67, thereby maintaining cancer cell proliferation [41].

Our research revealed that SPP1 + Mps trigger an increase in the proliferation of basal cells, recognized as tumor cells, through the release of cytokines. This mechanism may contribute to the invasive behavior of LUAD. SPP1 + Mps have the potential to enhance the EMT of cancer cells and/or secrete an array of cytokines associated with cancer cell proliferation (e.g., IL1A, IL1B, IL6, and TNF) via the NF-kB pathway. Additionally, the involvement of IGF signaling (IGF1/IGFBP3) was noted in the crosstalk between SPP1 + Mps and basal cells. This pathway is recognized for its crucial role in sustaining cancer cell proliferation [42]. Consistent with the results of prior studies, our findings further elucidate how SPP1 + Mps impact the proliferation of LUAD cells and contribute to the emergence of drug resistance. Earlier studies primarily described the presence or prognostic association of SPP1 + macrophages. In contrast, our work advances the field by integrating epithelial cell states and defining a coordinated cellular interaction module (CIM) that links macrophage-derived signals to malignant basal cell proliferation.

To conduct a detailed assessment of clinical outcomes, we defined the CIM module as a fusion of feature genes derived from SPP1 + macrophages and basal cells, in addition to their ligand-target pairs. Through the integration of single-cell and bulk transcription data, a 13-gene set was identified within the CIM using the LASSO model. This gene set demonstrated a notable association with unfavorable prognosis. Our findings underscore the critical importance of SPP1 + macrophages and basal cells in the progression of invasive LUAD. Moreover, our deeper analysis of the CIM led to results that were consistent with earlier reports, highlighting FAM117A as a significant tumor antigen in the development of LUAD [7, 10, 11, 43, 44].

Importantly, our mechanistic studies further established a connection between FAM117A and cell cycle control. We found DYRK1A to be a high-confidence binding partner of FAM117A and showed that the two proteins formed a complex to regulate the G1/S transition. Silencing of either FAM117A or DYRK1A perturbed this checkpoint, resulting in enhanced Cyclin D1 and CDK4 phosphorylation, and diminished p21 and p27

levels. Furthermore, co-silencing of both genes did not give rise to additive effects, suggesting that FAM117A and DYRK1A operate in tandem in a common pathway for G1 suppression. Importantly, these molecular findings were confirmed in vivo, where FAM117A knockdown markedly accelerated tumor growth in xenograft models in nude mice. Collectively, these data support a dual role for FAM117A in LUAD: first by acting as a tumor antigen involved in invasive progression (delineated by the CIM module), and second by functioning as a cell cycle regulator in partnership with DYRK1A to delay proliferation.

From a translational perspective, therapeutic strategies that target TAMs and cell-cycle regulators are becoming increasingly feasible in clinical settings. Several macrophage-modulating approaches, including CSF1R inhibitors (such as pexidartinib and emactuzumab), CCR2/CCR5 axis inhibitors, and CD47–SIRPα blockade, have entered early-phase clinical trials and demonstrated the capacity to reshape the composition or function of Mps within the TME [45]. In parallel, CDK4/6 inhibitors such as palbociclib, ribociclib, and abemaciclib have been approved for clinical use in breast cancer and show activity in subsets of lung cancer, particularly in tumors with intact RB signaling [46]. Notably, previous studies have suggested that FAM117A expression correlates with sensitivity to CDK4/6 inhibition, supporting the biological rationale for combining macrophage-targeted therapies with CDK4/6 inhibitors. Together, these data indicate that the combined targeting strategy proposed in this study is clinically plausible and may be readily adaptable to future biomarker-guided trials in LUAD.

Several limitations of the present study should be acknowledged. First, the integration of retrospective scRNA-seq datasets obtained from different cohorts and platforms may introduce unavoidable technical and clinical heterogeneity, despite the application of batch-correction algorithms. Second, cell–cell interaction analyses, including CellChat and NicheNet, rely on computational inference of ligand–receptor signaling and do not constitute direct functional evidence of intercellular communication. Third, although the in vivo xenograft experiments provide strong evidence of a tumor-suppressive role for FAM117A, immune-related effects and macrophage–epithelial interactions were not functionally tested using immune-competent models. These limitations highlight the need for prospective and mechanistically-focused studies to further validate and extend our findings.

Taken together, our results highlight two converging mechanisms that may drive the progression of LUAD: (i) crosstalk between SPP1 + Mps and basal cells promotes proliferation and immunotherapy resistance, and (ii) the FAM117A–DYRK1A axis enforces G1/S control and restrains tumor expansion. Both mechanisms underscore the intricate cellular interactions within the LUAD microenvironment and suggest therapeutic opportunities. Targeting of SPP1 + macrophage signaling and restoration of FAM117A–DYRK1A function may synergistically suppress invasive tumor growth and overcome drug resistance in LUAD.

Conclusion

This research provides a comprehensive view of the immune and non-immune cell composition within the TME of LUAD. Furthermore, it highlights the potential value of developing therapeutic strategies that target SPP1 + Mps and basal cells, or the molecules involved in their crosstalk, thereby hindering cancer cell proliferation and enhancing the response of tumors to immunotherapies.

Abbreviations

NSCLC	Non-small cell lung cancer
LUAD	Lung adenocarcinoma
AIS	Adenocarcinoma in situ
MIA	Microinvasive adenocarcinoma
IAC	Invasive adenocarcinoma
TME	Tumor microenvironment
TAMs	Tumor-associated macrophages
scRNA-seq	Single-cell RNA sequencing
RNA-seq	RNA sequencing
CIM	Cellular interaction module
SPP1+	Osteopontin-positive
TCGA	The cancer genome atlas
LASSO	Least absolute shrinkage and selection operator
AI	Artificial intelligence
FAM117A	Family with sequence similarity 117 Member A
DYRK1A	Dual-specificity tyrosine-regulated kinase 1 A
PLA	People's liberation army
RPMI	Roswell Park memorial institute (medium)
PBS	Phosphate-buffered saline
BSA	Bovine serum albumin
DNase I	Deoxyribonuclease I
mIHC	Multiplex immunohistochemistry
TSA-DIG	Tyramide signal amplification with digoxigenin
H&E	Hematoxylin and eosin
RIPA	Radioimmunoprecipitation assay
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
PBST	Phosphate-buffered saline with tween
IHC	Immunohistochemistry
qPCR	Quantitative polymerase chain reaction
RT-PCR	Reverse transcription polymerase chain reaction
SYBR	SYBR green (fluorescent dye for qPCR)
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
Seurat	R package for single-cell RNA-seq analysis
CCA	Canonical correlation analysis
UMAP	Uniform Manifold Approximation and Projection
GSVA	Gene set variation analysis
BCA	Bicinchoninic acid (protein assay)
GO	Gene Ontology
OS	Overall survival
TF	Transcription factor
EGR1	Early growth response 1
CNV	Copy Number Variation
AT1	Alveolar type 1 cell
AT2	Alveolar type 2 cell
APOD+	Apolipoprotein D positive
IGF1	Insulin-like growth factor 1
IGFBP3	Insulin-like growth factor binding protein 3
NF-κB	Nuclear factor kappa-light-chain-enhancer of activated B cells
CDK	Cyclin-Dependent Kinase

p21 (CDKN1A)	Cyclin-Dependent Kinase Inhibitor 1 A
p27 (CDKN1B)	Cyclin-Dependent Kinase Inhibitor 1B
Cyclin D1	G1/S-specific Cyclin D1

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40164-026-00745-9.

Supplementary Material 1.

Supplementary Material 2.

Acknowledgements
Not applicable.

Author contributions
CW, ZR and GC conceptualized the study, designed the experiments, and drafted the manuscript. CW, ZR and GC contributed equally to this work. HL and HT performed experimental work. DL organized the clinical trial and obtained the clinical samples. CW, ZR and GC contributed to data analysis and figure preparation. CW secured funding. JW, XH and YH supervised the project. All authors reviewed and approved the final manuscript.

Funding
This work was supported by the General Program of the Beijing Natural Science Foundation (Grant No. 7232164) to CW.

Data availability
The datasets generated and/or analyzed during the current study are available in the Gene Expression Omnibus (GEO) repository, with the following accession numbers: GSE131907 and GSE189357. The data and materials can be made available upon reasonable request to the corresponding author.

Declarations

Ethics approval and consent to participate
Ethical approval for the study was granted by the Ethics Committee of the First Medical Center, PLA General Hospital. Written informed consent was obtained from all individual participants included in the study.

Consent for publication
All authors have read and approved the final manuscript, and consent for publication was obtained from all participants involved in the study.

Competing interests
The authors declare no competing interests.

Received: 9 December 2025 / Accepted: 9 January 2026
Published online: 15 February 2026

References

1. Myers DJ, Wallen JM. Lung adenocarcinoma. In StatPearls, (StatPearls Publishing). 2023.
2. Shi H, Shao W, Song J, Che G, Zhao Y, Shi Y, Zhang J, Wan P, Zhu Q, Zhang D, Sheng J. Graph attention network enables multipurpose prediction of imaging mass cytometry in a hepatocellular carcinoma clinical trial. Intelli Oncol. 2025.
3. Wang Z, Li Z, Zhou K, Wang C, Jiang L, Zhang L, Yang Y, Luo W, Qiao W, Wang G. Deciphering cell lineage specification of human lung adenocarcinoma with single-cell RNA sequencing. Nat Commun. 2021;12:6500.
4. Lavin Y, Kobayashi S, Leader A, Amir E-aD, Elefant N, Bigenwald C, Remark R, Sweeney R, Becker CD, Levine JH. Innate immune landscape in early lung adenocarcinoma by paired single-cell analyses. Cell. 2017;169:750–65. e717.
5. Anderson NM, Simon MC. The tumor microenvironment. Curr Biol. 2020;30:R921–5.
6. Zhang X, Wu L, Zhang X, Xu Y. Identifying the tumor-associated macrophage of lung adenocarcinoma reveals immune landscape through omics data integration. Heliyon 2024;10.
7. Wu C, Zhang J, Wang K, Fan M, Hu Y. FAM117A is a new prognostic marker of lung adenocarcinoma and predicts sensitivity to PD0332991. Evidence-Based Complement Altern Med. 2022;2022:3945446.
8. Luo H, Li Q, Wang RT, Zhang L, Zhang W, Deng MS, Luo YY, Ji X, Wen Y, Zhou XR, Xu B. Downregulation of pro-surfactant protein B contributes to the recurrence of early-stage non-small cell lung cancer by activating PGK1-mediated Akt signaling. Exp Hematol Oncol. 2023;12(1):94.
9. Luo H, Shan J, Zhang H, Song G, Li Q, Xu CX. Targeting the epigenetic processes to enhance antitumor immunity in small cell lung cancer. In: Seminars in Cancer Biology, Vol. 86. 2022. pp. 960–70.
10. Bai M, Liu X, Wang L. Identification of potential tumor antigens and immune subtypes for lung adenocarcinoma. Med Oncol. 2023;40:100.
11. Wu CH, Hwang MJ. Risk stratification for lung adenocarcinoma on EGFR and TP53 mutation status, chemotherapy, and PD-L1 immunotherapy. Cancer Med. 2019;8:5850–61.
12. Aibar S, González-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselmans G, Rambow F, Marine J-C, Geurts P, Aerts J. SCENIC: single-cell regulatory network inference and clustering. Nat Methods. 2017;14:1083–6.
13. Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Ramos R, Kuan C-H, Myung P, Plikus MV, Nie Q. Inference and analysis of cell-cell communication using cellchat. Nat Commun. 2021;12:1088.
14. Bao X, Shi R, Zhao T, Wang Y. Mast cell-based molecular subtypes and signature associated with clinical outcome in early-stage lung adenocarcinoma. Mol Oncol. 2020;14:917–32.
15. Chen J-Y, Lin J-R, Tsai F-C, Meyer T. Dosage of Dyrk1a shifts cells within a p21-cyclin D1 signaling map to control the decision to enter the cell cycle. Mol Cell. 2013;52:87–100.
16. Wang G, Qiu M, Xing X, Zhou J, Yao H, Li M, Yin R, Hou Y, Li Y, Pan S. Lung cancer scRNA-seq and lipidomics reveal aberrant lipid metabolism for early-stage diagnosis. Sci Transl Med. 2022;14:eabk2756.
17. Maynard A, McCoach CE, Rotow JK, Harris L, Haderk F, Kerr DL, Elizabeth AY, Schenk EL, Tan W, Zee A. Therapy-induced evolution of human lung cancer revealed by single-cell RNA sequencing. Cell 2020;182:1232–1251. e1222.
18. Zhang L, Zhang Y, Wang C, Yang Y, Ni Y, Wang Z, Song T, Yao M, Liu Z, Chao N. Integrated single-cell RNA sequencing analysis reveals distinct cellular and transcriptional modules associated with survival in lung cancer. Signal Transduct Target Therapy. 2022;7:9.
19. Kim J, Xu Z, Marignani PA. Single-cell RNA sequencing for the identification of early-stage lung cancer biomarkers from Circulating blood. NPJ Genomic Med. 2021;6:87.
20. Aran D, Looney AP, Liu L, Wu E, Fong V, Hsu A, Chak S, Naikawadi RP, Wolters PJ, Abate AR. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. Nat Immunol. 2019;20:163–72.
21. Schupp JC, Adams TS, Cosme C Jr, Raredon MSB, Yuan Y, Omote N, Poli S, Chioccioli M, Rose K-A, Manning EP. Integrated single-cell atlas of endothelial cells of the human lung. Circulation. 2021;144:286–302.
22. Luo H, Hu B, Gu XR, Chen J, Fan XQ, Zhang W, Wang RT, He XD, Guo W, Dai N, Jian D. The miR-23a/27a/24– 2 cluster drives immune evasion and resistance to PD-1/PD-L1 blockade in non-small cell lung cancer. Mol Cancer. 2024;23(1):285.
23. Rud AK, Boye K, Øijordsbakken M, Lund-Iversen M, Halvorsen AR, Solberg SK, Berge G, Helland Å, Brustugun OT, Mælandsmo GM. Osteopontin is a prognostic biomarker in non-small cell lung cancer. BMC Cancer. 2013;13:1–10.
24. Mack PC, Redman MW, Chansky K, Williamson SK, Farneth NC, Lara Jr PN, Franklin WA, Le Q-T, Crowley JJ, Gandara DR. Lower osteopontin plasma levels are associated with superior outcomes in advanced non–small-cell lung cancer patients receiving platinum-based chemotherapy: SWOG study S0003. J Clin Oncol. 2008;26:4771–6.
25. Isa S-I, Kawaguchi T, Teramukai S, Minato K, Ohsaki Y, Shibata K, Yonei T, Hayashibara K, Fukushima M, Kawahara M. Serum osteopontin levels are highly prognostic for survival in advanced non-small cell lung cancer: results from JMTO LC 0004. J Thorac Oncol. 2009;4:1104–10.
26. Han S-S, Lee S-J, Kim WJ, Ryu DR, Won JY, Park S, Cheon MJ. Plasma osteopontin is a useful diagnostic biomarker for advanced non-small cell lung cancer. Tuberc Respir Dis. 2013;75:104–10.
27. Ostheimer C, Bache M, Güttler A, Kotzsch M, Vordermark D. A pilot study on potential plasma hypoxia markers in the radiotherapy of non-small cell lung cancer. Strahlenther Onkol. 2014;190:276.

28. Takenaka M, Hanagiri T, Shinohara S, Yasuda M, Chikaishi Y, Oka S, Shimokawa H, Nagata Y, Nakagawa M, Uramoto H. Serum level of osteopontin as a prognostic factor in patients who underwent surgical resection for non–small-cell lung cancer. Clin Lung Cancer. 2013;14:288–94.
29. Carbone F, Grossi F, Bonaventura A, Vecchié A, Minetti S, Bardi N, Elia E, Ansaldo AM, Ferrara D, Rijavec E. Baseline serum levels of osteopontin predict clinical response to treatment with nivolumab in patients with non-small cell lung cancer. Clin Exp Metastasis. 2019;36:449–56.
30. Liu Y, Xun Z, Ma K, Liang S, Li X, Zhou S, Sun L, Liu Y, Du Y, Guo X. Identification of a tumour immune barrier in the HCC microenvironment that determines the efficacy of immunotherapy. J Hepatol. 2023;78:770–82.
31. Xie W, Cheng J, Hong Z, Cai W, Zhuo H, Hou J, Lin L, Wei X, Wang K, Chen X. Multi-transcriptomic analysis reveals the heterogeneity and tumor-promoting role of SPP1/CD44-mediated intratumoral crosstalk in gastric cancer. Cancers. 2022;15:164.
32. He C, Sheng L, Pan D, Jiang S, Ding L, Ma X, Liu Y, Jia D. Single-cell transcriptomic analysis revealed a critical role of SPP1/CD44-mediated crosstalk between macrophages and cancer cells in glioma. Front Cell Dev Biology. 2021;9:779319.
33. Liu L, Zhang R, Deng J, Dai X, Zhu X, Fu Q, Zhang H, Tong Z, Zhao P, Fang W. Construction of TME and identification of crosstalk between malignant cells and macrophages by SPP1 in hepatocellular carcinoma. Cancer Immunol Immunother. 2022;71:121–36.
34. Nakajima T, Uehara T, Iwaya M, Matsuda K, Wada M, Nagaya T, Ehara T, Ota H. Osteopontin expression in the invasive front stroma of colorectal adenocarcinoma is associated with tumor budding and prognosis. Pathology-Research Pract. 2022;240:154190.
35. Matsubara E, Komohara Y, Esumi S, Shinchi Y, Ishizuka S, Mito R, Pan C, Yano H, Kobayashi D, Fujiwara Y. SPP1 derived from macrophages is associated with a worse clinical course and chemo-resistance in lung adenocarcinoma. Cancers. 2022;14:4374.
36. Sun Y. Tumor microenvironment and cancer therapy resistance. Cancer Lett. 2016;380:205–15.
37. Lee M, Vasioukhin V. Cell Polarity and cancer–cell and tissue Polarity as a non-canonical tumor suppressor. J Cell Sci. 2008;121:1141–50.
38. Hugo H, Ackland ML, Blick T, Lawrence MG, Clements JA, Williams ED, Thompson EW. Epithelial—mesenchymal and mesenchymal—epithelial transitions in carcinoma progression. J Cell Physiol. 2007;213:374–83.
39. Moreno-Bueno G, Portillo F, Cano A. Transcriptional regulation of cell Polarity in EMT and cancer. Oncogene. 2008;27:6958–69.
40. Tu D, Dou J, Wang M, Zhuang H, Zhang X. M2 macrophages contribute to cell proliferation and migration of breast cancer. Cell Biol Int. 2021;45:831–8.
41. Alexander TL, Hedbrant A, Ramberg J, Wijkander A, Solterbeck. Effect of macrophages on breast cancer cell proliferation, and on expression of hormone receptors, uPAR and HER-2. International Journal of Oncology. 2017.
42. Liu F, Ye S, Zhao L, Niu Q. The role of IGF/IGF-1R signaling in the regulation of cancer stem cells. Clin Transl Oncol, 2024;1–11.
43. Wang Y, Zhang Q, Gao Z, Xin S, Zhao Y, Zhang K, Shi R, Bao X. A novel 4-gene signature for overall survival prediction in lung adenocarcinoma patients with lymph node metastasis. Cancer Cell Int. 2019;19:1–9.
44. Zhang L, Zhang X, Guan M, Zeng J, Yu F, Lai F. Identification of a novel ADCC-related gene signature for predicting the prognosis and therapy response in lung adenocarcinoma. Inflamm Res. 2024;73:841–66.
45. Mantovani A, Allavena P, Marchesi F, Garlanda C. Macrophages as tools and targets in cancer therapy. Nat Rev Drug Discovery. 2022;21:799–820.
46. Morrison L, Loibl S, Turner NC. The CDK4/6 inhibitor revolution—a game-changing era for breast cancer treatment. Nat Reviews Clin Oncol. 2024;21:89–105.

Publisher's Note