Translational Oncology 69 (2026) 102800

Contents lists available at ScienceDirect

Translational Oncology

journal homepage: www.elsevier.com/locate/tranon

Single-cell and spatial transcriptomic profiling of POU5F1 in Lung Adenocarcinoma: Dynamics, spatial niche, and prognosis via multi-algorithm ML

Bingbing Yu [a,1], Meili Zhang [a,1], Haiyong Zhu [b,1], Zhu Wu [c], Hong Gao [d,*], Yang Bai [a,*]

[a] Department of Pathology, Qilu Hospital of Shandong University Dezhou Hospital, Dezhou, China
[b] Department of Thoracic Surgery, The Second Hospital of Hebei Medical University, Shijiazhuang, China
[c] The Hunan Provincial Key Laboratory of the TCM Agricultural Biogenomics, Changsha Medical University, Changsha, China
[d] Department of Oncology, The Second Hospital of Hebei Medical University, Shijiazhuang, China

ARTICLE INFO

Keywords:
POU5F1
Lung cancer
CAFs
Machine learning
Prognostic model

ABSTRACT

Background: POU5F1 (OCT4), a core regulator of pluripotency, plays an important role in tumor stemness and immune microenvironment remodeling, yet its systematic function and mechanisms in lung adenocarcinoma (LUAD) remain incompletely elucidated.

Methods: This study integrated genetic causality inference, single-cell and spatial transcriptomics, RNA velocity analysis, and multi-algorithm machine learning to systematically investigate the role of POU5F1 in LUAD. Summary-based Mendelian Randomization (SMR) was used to link GWAS with lung tissue eQTL data; single-cell and spatial analyses characterized POU5F1$^+$ cell states and their interactions with cancer-associated fibroblasts (CAFs); a machine-learning prognostic model was constructed based on POU5F1-related differentially expressed genes (LDEGs); and in vitro functional experiments were performed to validate its biological functions.

Results: POU5F1 was genetically associated with lung cancer risk and enriched in malignant, stem-like cell clusters. RNA velocity indicated progressive activation along pseudotime trajectories. Spatial transcriptomics confirmed the co-localization of POU5F1$^+$ tumor cells with CAFs. The LDEG-based prognostic model showed excellent predictive performance (AUC $>$ 0.9) and was correlated with immune suppression and targeted therapy response. Basic experiments further demonstrated that POU5F1 was specifically highly expressed in LUAD cell lines; knockdown of POU5F1 significantly inhibited cell proliferation, migration and invasion, and induced apoptosis; mechanistically, these effects involved regulation of the Bcl-2/caspase-3 apoptotic pathway, modulation of the epithelial-mesenchymal transition (EMT) process mediated by E-cadherin/Vimentin, and down-regulation of the stemness factor Nanog.

Conclusion: POU5F1 promotes LUAD progression by maintaining stemness, reprogramming CAFs, and shaping immune evasion, representing a promising biomarker and therapeutic target.

Introduction

Cancers are a major contributor to the global disease burden, which is projected to increase over the next two decades [1,2]. In 2022, there were an estimated 20 million new cancer cases and 9.7 million cancer-related deaths worldwide [3]. Early diagnosis and treatment are critical for effective cancer management, with surgical resection being the primary curative approach for most solid tumors [4]. However, recurrence and metastasis of solid cancer remain significant challenges, frequently leading to treatment failure and increased mortality. Lung cancer is the most common cancer globally and the leading cause of cancer-related deaths [3]. Lung adenocarcinoma (LUAD) is the predominant subtype of non-small cell lung cancer (NSCLC) in East Asia [5,6]. Although radical resections, such as anatomic lobectomy with systematic lymph node dissection, significantly improve patient prognosis [7], the complex nature of NSCLC metastasis challenges the conventional perspectives that primary tumors only metastasize after reaching a certain size. Notably, many NSCLC cases exhibit lymph node or distant

* Corresponding authors.
E-mail addresses: hongaza321@163.com (H. Gao), 13853493039@163.com (Y. Bai).
[1] These authors have contributed equally to this work and share first authorship.

https://doi.org/10.1016/j.tranon.2026.102800
Received 22 January 2026; Received in revised form 9 April 2026; Accepted 27 April 2026

metastasis even when the primary tumor is small [8,9]. Clinicians have also observed rapid recurrence in NSCLCs with small primary tumors following radical resection [10]. These cases underscore the complexity of recurrence and metastasis mechanisms in NSCLC.

The tumor microenvironment (TME) is composed of tumor cells and stromal cells. The malignant potential of tumors has long been thought to be entirely due to cancer cells [11]. However, the dynamic crosstalk between cancer cells and stromal cells has been shown to be involved in cancer progression [12]. The stroma consists of fibroblasts, pericytes, mesenchymal stem cells, and various types of immune cells, surrounded by fibrous structural proteins in the extracellular matrix [13]. Cancer-associated fibroblasts (CAFs) are important components of the TME which arise from bone marrow-derived mesenchymal stem cells, hematopoietic stem cells, adipocytes, and endothelial cells [8,14], as well as cancer cells [15]. CAFs have been observed in a majority of cancers, such as breast cancer, prostate cancers, and HCC [16,17], and their crosstalk with cancer cells has been revealed to be crucial for tumor progression [18]. CAFs secrete a variety of growth factors and cytokines, and degrade extracellular matrix proteins, thereby affecting tumor cell proliferation, metastasis and chemotherapy resistance [19–21]. CAFs can stably maintain tumor-promoting characteristics even without exposure to cancer cells [22].

Among the molecular drivers of tumor aggressiveness, POU5F1 (also known as OCT4), a core transcription factor governing embryonic stem cell pluripotency, has garnered increasing attention in cancer biology. Beyond its well-established role in maintaining cancer stemness and therapy resistance [29,30], emerging evidence suggests that POU5F1 may exert non-cell-autonomous functions by remodeling the tumor microenvironment. For instance, POU5F1 has been shown to regulate cytokine secretion and intercellular signaling that influence stromal fibroblasts and immune cells in gastric and breast cancers [31,32]. However, in LUAD, the systematic role of POU5F1—particularly its potential involvement in CAF reprogramming, spatial niche co-localization, and immune evasion—remains poorly understood. Furthermore, whether POU5F1 expression defines a distinct malignant cell state that actively communicates with CAFs within the tumor spatial architecture has not been systematically investigated.

Therefore, in this study, we integrated genetic causality inference (SMR), single-cell and spatial transcriptomics, RNA velocity analysis, and multi-algorithm machine learning to comprehensively characterize POU5F1 in LUAD. We aimed to: (i) determine the genetic association between POU5F1 and lung cancer risk; (ii) dissect the single-cell distribution and pseudotemporal dynamics of $POU5F1^+$ malignant cells; (iii) investigate the spatial co-localization and intercellular communication between $POU5F1^+$ cells and CAFs; (iv) construct a robust POU5F1-related prognostic signature using ensemble machine learning; and (v) experimentally validate the functional role of POU5F1 in LUAD cell proliferation, apoptosis, migration, invasion, and stemness regulation. By addressing these questions, we provide a multidimensional framework linking POU5F1 to stemness–stroma–immunity interactions, with potential implications for prognosis and targeted therapy in LUAD.

Materials and methods

Summary-data-based Mendelian Randomization (SMR) analysis

We performed a Summary-based Mendelian Randomization (SMR) analysis to integrate lung cancer GWAS data with lung tissue eQTL data. The analysis was carried out using the SMR software with the following parameters: minor allele frequency (MAF) > 0.01, HEIDI test p-value > 0.05 to exclude pleiotropic false positives, and a significance threshold of p_SMR $< 5 \times 10^{-8}$.

Single-cell RNA data processing and cell type annotation

Single-cell RNA sequencing data were processed using Seurat (V4). Quality control involved the exclusion of cells with $> 20\%$ mitochondrial gene content or UMI counts < 500. Data normalization was conducted using the logNormalize method, and the top 2000 highly variable genes were selected. Principal component analysis (PCA) was performed, followed by UMAP visualization. Clustering was conducted using the Louvain algorithm (resolution $= 0.8$). POU5F1 positivity was defined as normalized counts > 0.

Cell-Cell communication analysis

Cell-cell communication was analyzed using the CellChat (R) and NicheNet (R) packages. The input data consisted of the normalized single-cell expression matrix, with parameters set as: min.cells $= 10$ and p-value cutoff $= 0.05$. The output included sender-receiver interaction intensity matrices and key ligand-receptor pairs. Significant communication pathways were visualized and subjected to enrichment analysis.

Statistical significance was assessed using permutation tests ($n = 1000$ permutations) implemented in CellChat. Ligand-receptor pairs with a permutation-based p-value < 0.05 were considered statistically significant. Communication probability (ranging from 0 to 1) represents the likelihood of interaction between two cell types given the expression levels of the ligand and receptor. For each significant interaction, we reported the communication probability, p-value, and 95% confidence interval. All significant ligand-receptor pairs are provided in Supplementary Table S2.

RNA velocity analysis

RNA velocity analysis was conducted using velocyto.R. Spliced counts were computed from loom files, and cell development trajectories and transcriptional velocities were inferred using the scVelo dynamical model, with parameters: min shared counts $= 20$, n pcs $= 30$. The results were used to construct pseudotime trajectories to analyze the expression dynamics of POU5F1.

Spatial transcriptomics and deconvolution

Spatial transcriptomics data were analyzed using RCTD for cell-type decomposition, with reference matrices derived from single-cell annotations. Cell2location was run with default parameters, and the resulting cell composition for each spatial spot was outputted as a matrix. Spatial correlations between POU5F1-positive cells and CAF cells were calculated using Spearman correlation and permutation tests to validate statistical significance.

Differential expression analysis and machine learning modeling

Differential expression analysis between POU5F1-positive and negative samples was performed using the limma package with thresholds of FDR < 0.05 and $|\log_2 FC| > 0.3$, and the resulting LDEGs were used for prognostic model construction. We developed an ensemble framework incorporating 101 survival algorithms, including Cox-based methods, regularized regression, tree-based ensemble methods (Random Survival Forest, XGBoost, LightGBM), support vector machines, neural networks, k-nearest neighbors, stacking methods, and other survival models. The TCGA-LUAD cohort was randomly split into training (50%, $n = 267$) and internal validation (50%, $n = 267$) sets, with GSE50081 and GSE31210 as external validation cohorts. All algorithms were implemented using the caret and survival packages in R and scikit-learn in Python, with 10-fold cross-validation repeated 5 times and grid search for hyperparameter optimization. Models were evaluated using the concordance index (C-index), time-dependent AUC at 1, 3, and 5 years, integrated Brier score, and log-rank p-value for risk stratification. The optimal model was selected based on mean C-index rank across folds, stability (standard deviation of C-index), external validation consistency, and significant log-rank p-value (< 0.05) in all

three validation cohorts. The selected Random Survival Forest model was used to calculate risk scores by weighting the expression of prognostic genes with their variable importance scores, and patients were stratified into high- and low-risk groups using the median risk score as the cutoff.

Survival analysis and nomogram construction

Using the median risk score, TCGA training and internal validation cohorts were divided into high-risk and low-risk groups. Kaplan-Meier survival curves were generated using the "survminer" R package to assess significant differences in overall survival (OS) between the two groups (log-rank test, $p < 0.05$). Receiver operating characteristic (ROC) curve analysis was performed using the "timeROC" package to predict the sensitivity and specificity of OS in LUAD patients. The area under the curve (AUC) was compared with other clinical features. Clinical factors such as age, gender, stage, T, M, N, and grade were also explored for correlation. Univariate and multivariate Cox regression analyses were performed on TCGA-LUAD data to determine whether LDEGs are independent prognostic factors for LUAD patient survival. To improve the prognostic accuracy and predictive power, a nomogram was developed combining LDEGs and clinical features to quantify the expected survival of LUAD patients. The nomogram's precision, discrimination, and accuracy were evaluated using ROC curves, C-index, calibration curves, and decision curve analysis (DCA).

Immune features and response to immune checkpoint inhibitor therapy

To examine the relationship between LDEGs and immune cell infiltration in the LUAD tumor microenvironment, the CIBERSORT algorithm was used to quantify the infiltration of 22 immune cell types. To validate the results from CIBERSORT, we additionally employed the ESTIMATE, ssGSEA, and XCell algorithms. The cancer-immunity cycle, a critical component of cancer immunotherapy, includes seven key steps: (1) release of tumor antigens, (2) antigen presentation, (3) immune cell activation, (4) immune cell trafficking to the tumor, (5) immune cell infiltration into the tumor, (6) T cell recognition of cancer cells, and (7) cancer cell killing. These steps together form the cancer immunity cycle. Activity scores for each of these steps were obtained from the Tumor Immune Phenotype (TIP) platform.

To assess immunogenicity based on immune modulators, immunesuppressive cells, MHC molecules, and effector cells, we used the Immune Phenotype Score (IPS) algorithm. This algorithm uses machine learning methods to compute IPS scores based on unbiased gene expression in representative cell types. Higher IPS scores indicate better responses to immunotherapy. IPS scores for TCGA-LUAD samples were obtained from the Cancer Immunome Atlas (TCIA) database.

LDEGs and drug sensitivity

To facilitate personalized treatment, we used the "prorophetic" R package to predict chemotherapy sensitivity in LUAD patients with different LDEGs risk scores. Prorophetic matches patient gene expression profiles to cancer cell line expression profiles and calculates the half-maximal inhibitory concentration (IC50). Wilcoxon tests were used to detect differences in IC50 between high-risk and low-risk groups, with $p < 0.05$ considered statistically significant.

Fundamental experimental methods

Cell culture

The following cell lines from the American Type Culture Collection (ATCC) were used: BEAS-2B (normal bronchial epithelium), A549 (lung adenocarcinoma), HCC827 (lung adenocarcinoma, EGFR mutant), NCI-H1975 (lung adenocarcinoma, EGFR T790M mutant), NCI-H2228 (lung adenocarcinoma, EML4-ALK fusion), and Calu-1 (lung squamous cell carcinoma). Cell culture conditions were strictly standardized: A549 and Calu-1 were cultured in DMEM medium (Hyclone), BEAS-2B in Bronchial Epithelial Cell Growth Medium (BEGM, Lonza), and the remaining cells in RPMI-1640 medium (Gibco). All media were supplemented with 10% fetal bovine serum (FBS, Hyclone) and 1% penicillin-streptomycin solution (Keygen). Cells were maintained in a standard incubator (37 °C, 5% CO_2, saturated humidity), ensuring logarithmic growth phase throughout the experiments.

Transfection

siRNA transfection was performed on NCI-H1975 and HCC827 cells. siRNA reagents from Sangon Biotech were used for transient knockdown of POU5F1 gene expression, with a negative control siRNA (si-NC) established. The procedure was as follows: When cells in six-well plates reached 80% confluence, lipo3000 transfection reagent and siRNA were diluted in Opti-MEM medium (Thermo Fisher). After 5 min, the mixtures were combined, allowed to stand for 20 min, and then added to the wells. Fresh complete medium was replaced 5 h post-transfection.

Total RNA extraction and RT-qPCR

Following digestion and centrifugation, cells were lysed with 950 μL of Trizol reagent (Takara) to inhibit RNase activity. After 5 min of lysis, 150 μL of chloroform (Sinopharm Chemical Reagent Co., Ltd.) was added, followed by vigorous vortexing until a milky white emulsion formed. After centrifugation, the supernatant was collected and mixed with an equal volume of isopropanol (Sinopharm) for RNA precipitation. The pellet was retained after another centrifugation, washed thoroughly with 1 mL of 75% ethanol, and air-dried. Genomic DNA was removed using the PrimeScript RT kit (Takara), and cDNA was synthesized. Quantitative real-time PCR was performed using SYBR GreenER Supermix (Takara) on a Roche 480 PCR system. Three technical replicates were used per group, with β-actin as the internal reference.

CCK-8 cell proliferation assay

Twenty-four hours post-transfection, cell suspensions were seeded into 96-well plates at a density of 4000 cells per well, with three replicate wells per condition. After cell adhesion, CCK-8 reagent (KeyGEN) was mixed with basal medium according to the manufacturer's instructions, and 100 μL was added to each well. Plates were incubated in the dark for 1.5 h, and absorbance was measured at 450 nm using a microplate reader. Measurements were taken at multiple designated time points.

Apoptosis detection

Cell apoptosis was analyzed by flow cytometry using Annexin V-FITC/PI double staining. NCI-H1975 cells transfected with si-NC or si-POU5F1 were harvested, washed with PBS, and stained with Annexin V-FITC and PI in the dark for 15 min at room temperature. Stained cells were immediately analyzed using a BD FACSCalibur flow cytometer. The percentage of apoptotic cells (early: Annexin V^+/PI^-; late: Annexin V^+/PI^+) was calculated using FlowJo software. All experiments were performed in triplicate.

Transwell migration and invasion assay

Cell migration and invasion abilities were assessed using Transwell chambers. After transfection, 1×10^5 NCI-H1975 cells in serum-free medium were seeded into the upper chamber. The lower chamber contained RPMI-1640 medium with 10% FBS as a chemoattractant. For invasion assays, the membrane was pre-coated with Matrigel. After 24 h of incubation, non-migrated/non-invaded cells were removed. Migrated/invaded cells were fixed, stained with crystal violet, and counted under a microscope in five random fields per well.

Western blot analysis

Total cellular proteins were extracted using RIPA lysis buffer containing protease and phosphatase inhibitors on ice for 30 min, followed

by centrifugation at 12,000 × g for 15 min at 4 °C. Protein concentration was determined by the BCA method. Equal amounts of protein (30–50 μg per lane) were separated by 10% SDS-PAGE and transferred to PVDF membranes using a semi-dry transfer system. Membranes were blocked with 5% non-fat milk in TBST for 1 h at room temperature, then incubated overnight at 4 °C with the following primary antibodies: anti-cleaved caspase-3 (1:1000, #9661, CST), anti-Bcl-2 (1:2000, ab182858, Abcam), anti-E-cadherin (1:1000, ab1416, Abcam), anti-Vimentin (1:1000, ab92547, Abcam), anti-Nanog (1:1000, ab21624, Abcam), and anti-β-actin (1:1000, ab179511, Abcam) as a loading control. After washing with TBST, membranes were incubated with HRP-conjugated secondary antibody for 1 h at room temperature. Protein bands were visualized using enhanced chemiluminescence reagent and imaged with a ChemiDoc XRS+ system. Band intensity was quantified using ImageJ software.

Statistical analysis

All statistical analyses were conducted using R software (version: R 4.3.1). Chi-square tests were used to compare clinical features between the training and internal validation cohorts. Wilcoxon tests, as a non-parametric method, were employed to estimate differences between two non-normally distributed variables. In DEGs analysis, FDR-corrected p-values were used to evaluate significant DEGs. Kaplan-Meier survival analysis and log-rank tests were used to compare overall survival (OS) between different patient subgroups. Univariate and multivariate Cox regression analyses were performed to identify independent prognostic factors. The performance of the model was evaluated by ROC curve analysis and AUC calculation using the "timeROC" R package. Spearman correlation analysis was used to explore the relationship between risk scores and immune cell infiltration.

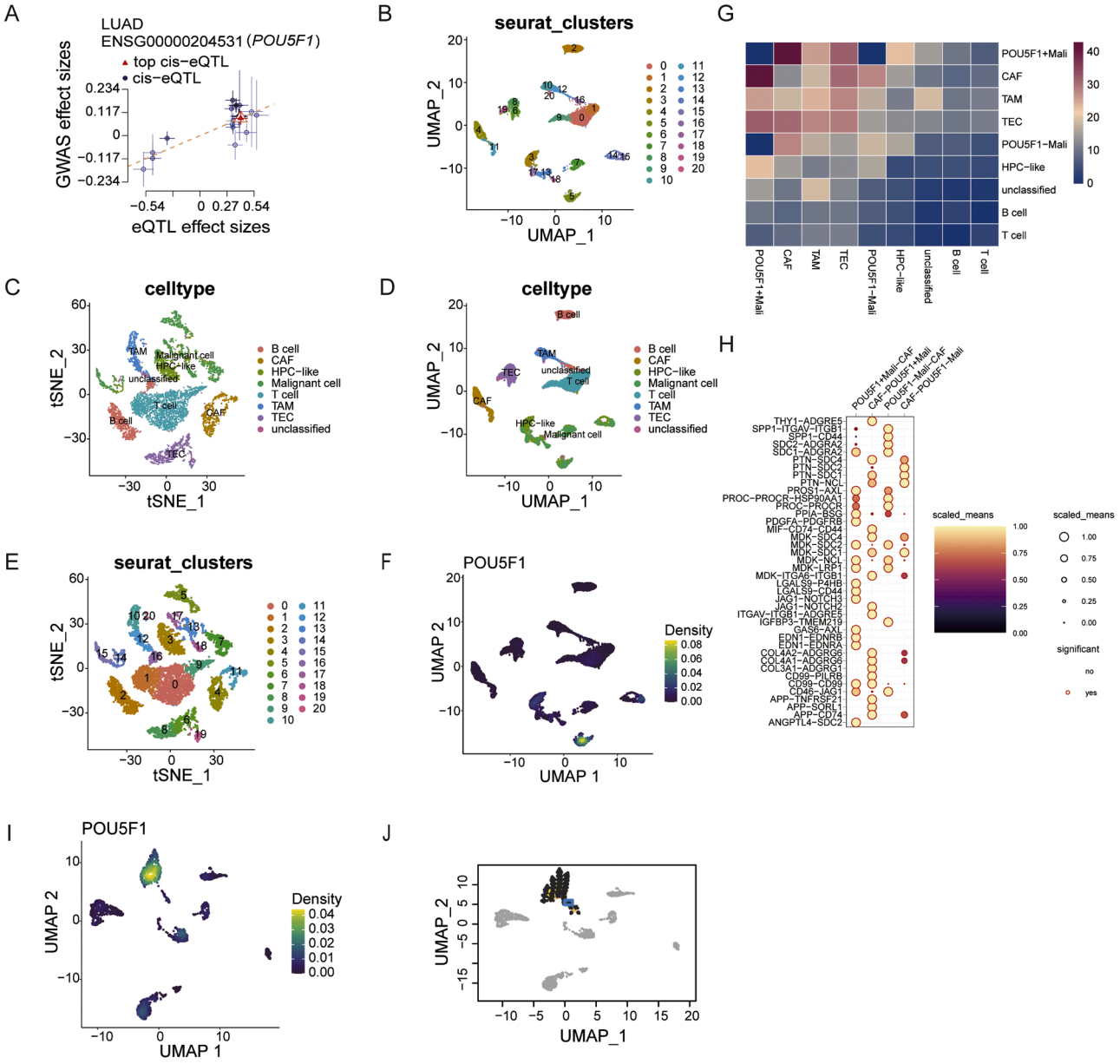


Fig. 1. Identification and characterization of POU5F1 in lung cancer.
(A) SMR analysis integrating GWAS and lung eQTL data identified POU5F1 as a senescence-related candidate gene. (B–F) Single-cell RNA-seq revealed POU5F1 enrichment in malignant and stem-like tumor cell clusters. (G, H) Cell–cell communication analysis showed stronger interactions between POU5F1⁺ cells and CAFs via PROC–PROCR–HSP90AA1 and related ligand–receptor pairs. (I, J) RNA velocity analysis indicated progressive POU5F1 upregulation along pseudotime, suggesting activation during dedifferentiation.

Results

SMR screening and intersection with senescence-related genes

Using SMR with GWAS and lung tissue-specific eQTL datasets ($p\_SMR < 0.05$), we identified candidate genes and intersected them with a curated senescence-associated gene set. This analysis pinpointed POU5F1 as a significant candidate (Fig. 1A, Table S1).

Single-cell annotation and POU5F1-based stratification

Single-cell RNA-seq data from GSE131907 were processed through quality control, normalization, dimensionality reduction, and clustering. Based on canonical marker gene expression, we identified the following major cell types: epithelial/malignant cells (EPCAM$^+$/KRT7$^+$), cancer-associated fibroblasts (CAFs; ACTA2$^+$/FAP$^+$/COL1A1$^+$/PDGFRA$^+$, with exclusion of EPCAM, PECAM1, and PTPRC), T cells (CD3D$^+$/CD8A$^+$/CD4$^+$), B cells (CD79A$^+$/MS4A1$^+$), myeloid cells (CD68$^+$/CD14$^+$), endothelial cells (PECAM1$^+$/VWF$^+$), NK cells (NCAM1$^+$/KLRF1$^+$), mast cells (TPSAB1$^+$/KIT$^+$), and plasma cells (MZB1$^+$/SDC1$^+$). POU5F1 was enriched in malignant clusters and a subset of stem-like tumor cells (Fig. 1B–F). Cells were then stratified into POU5F1$^+$ and POU5F1$^-$ groups based on detectable expression.

Cell–cell communication: strengthened interactions between POU5F1$^+$ cells and CAFs

Cell–cell communication networks were inferred using CellChat (or NicheNet). Compared with POU5F1$^-$ cells, POU5F1$^+$ cells exhibited markedly enhanced signaling with CAFs. Statistical evaluation using permutation tests ($n = 1000$ permutations) revealed that the interaction between POU5F1$^+$ cells and CAFs was significantly stronger than baseline (communication probability $= 0.042$, $p = 0.018$; Supplementary Table S2). The key ligand–receptor pairs mediating this interaction included PROC–PROCR–HSP90AA1 and PROC–PROCR (Fig. 1G–H; $p < 0.05$ indicated by asterisks). These findings suggest that POU5F1$^+$ cells may regulate CAF function via the protein C pathway, a key anticoagulant and cytoprotective cascade within the coagulation system. Complete lists of significant ligand-receptor pairs, including communication probabilities and p-values, are provided in Supplementary Table S2.

RNA velocity: progressive upregulation of POU5F1 along pseudotime

RNA velocity analysis with velocyto.R revealed a gradual increase in POU5F1 expression along inferred trajectories (Fig. 1I–J), supporting its activation during cellular dedifferentiation or transdifferentiation.

Spatial transcriptomics deconvolution revealed the co-localization of POU5F1$^+$ malignant cells and CAFs

After rigorous quality control and normalization, spatial

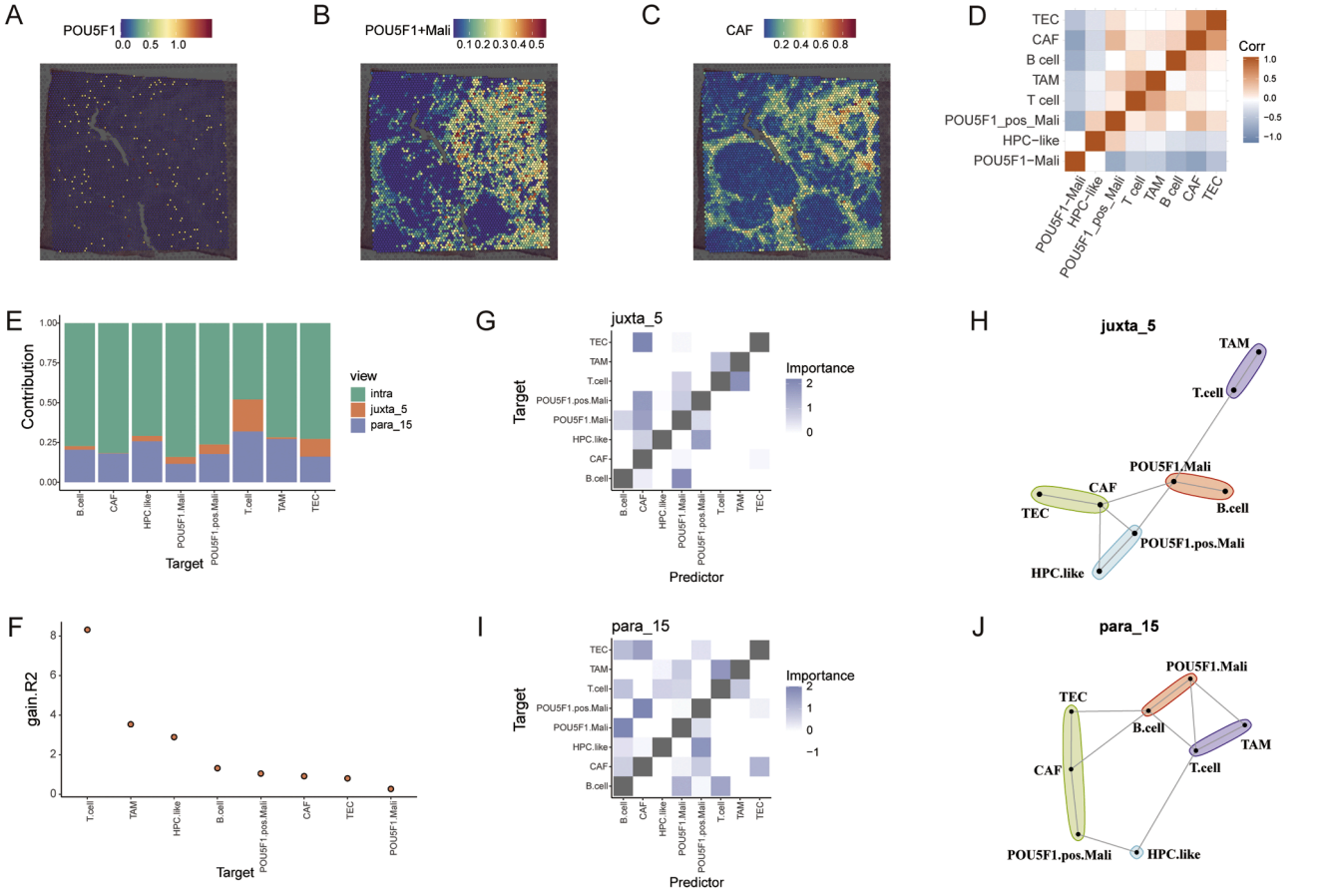


Fig. 2. Spatial co-localization and interaction between POU5F1$^+$ malignant cells and CAFs.
(A–C) Spatial transcriptomics maps showing the expression and localization of POU5F1, POU5F1$^+$ malignant cells, and CAFs in tumor tissue. (D–F) Spatial deconvolution using RCTD revealed marked co-localization of POU5F1$^+$ malignant cells with CAFs across spatial spots. (G–J) Multilevel spatial analyses with mistyR and SPOTlight demonstrated strong correlations and interaction networks between POU5F1$^+$ malignant cells and CAFs, identifying CAFs as key hubs within the tumor microenvironment.

transcriptomics data were deconvoluted using RCTD, mapping single-cell reference annotations onto spatial spots. Spatial transcriptomics mapping revealed the expression patterns and spatial distribution of POU5F1, POU5F1^{+} malignant cells, and CAFs within tumor tissues (Fig. 2A–C). This analysis demonstrated a significant spatial co-localization of POU5F1^{+} malignant cells with CAFs (Fig. 2D–F). Furthermore, multilevel spatial analyses performed with mistyR revealed that POU5F1^{+} malignant cells exhibited a strong correlation with CAFs in the main view and significant interactions in both the juxta_5H and para_15 views, identifying CAFs as central hubs within the

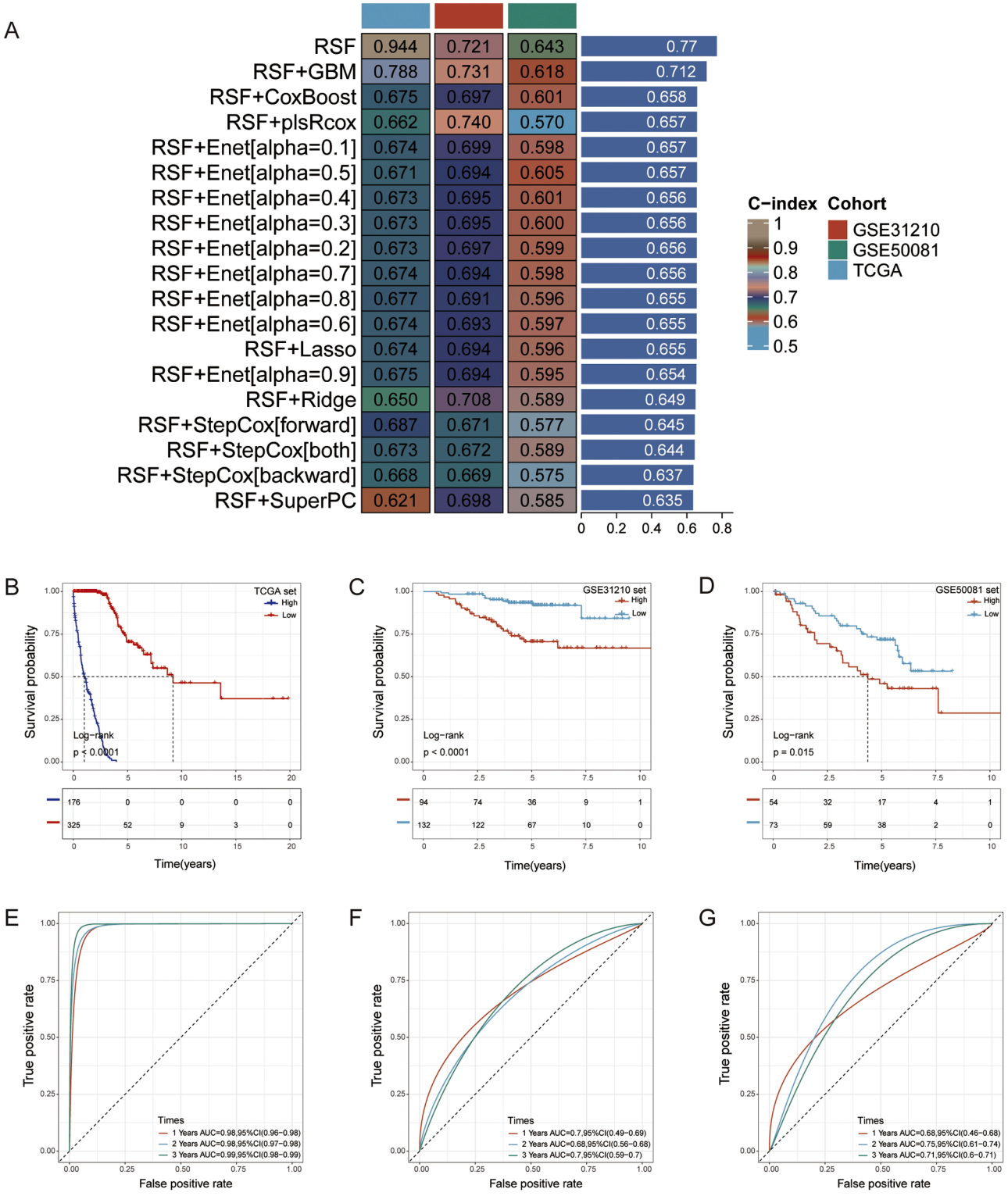


Fig. 3. Construction and validation of the LDEGs signature.
(A) Integrated machine learning with 101 algorithms identified the random survival forest (RSF) as the optimal model for LDEGs construction. (B–D) Kaplan–Meier analysis showed that high-risk patients had significantly worse overall survival across all cohorts. (E–G) Time-dependent ROC curves confirmed the strong predictive performance of the LDEGs signature in TCGA-LUAD, GSE50081, and GSE31210 cohorts.

tumor microenvironment. Interaction networks reconstructed with SPOTlight further validated the close interplay between POU5F1^{+} malignant cells and CAFs (Fig. 2G–J). Collectively, these findings suggest that POU5F1^{+} malignant cells may reshape the tumor stroma and potentiate tumor–stroma interactions through local co-localization and intercellular communication with CAFs.

Construction of a predictive signature based on integrated machine learning

To develop a robust tumor heterogeneity–related signature, we applied an ensemble framework incorporating 101 machine learning algorithms to evaluate 21 prognostic genes identified by univariate Cox regression analysis. The TCGA-LUAD cohort was randomly split into training and internal validation sets at a 1:1 ratio. Within the training set, 101 predictive models were trained using tenfold cross-validation, and C-indices were computed for both training and validation sets (Fig. 3A). Among these, the random survival forest (RSF) model consistently ranked among the top five with the highest mean C-index and was therefore selected for final signature construction. The RSF-based signature exhibited robust predictive performance across the training, internal validation, and external validation cohorts. Risk scores for individual patients were calculated by weighting the expression of the 21 genes with their Cox regression coefficients, allowing stratification into high- and low-risk groups based on the median risk score. Kaplan–Meier survival analysis demonstrated that high-risk patients had significantly worse OS compared to low-risk patients ($p < 0.001$, Fig. 3B–D).

Validation of the prognostic and clinical relevance of LDEGs

ROC curve analysis demonstrated that the LDEGs achieved remarkable predictive accuracy in the TCGA-LUAD cohort, with AUCs of 0.98, 0.98, and 0.99 at 1-, 3-, and 5-year intervals, respectively. In the external validation cohorts, the AUCs were 0.68, 0.75, and 0.71 in GSE50081, and 0.70, 0.68, and 0.70 in GSE31210 (Fig. 3E–G), underscoring the strong discriminative ability of LDEGs.

To further assess its clinical relevance, we examined the association between LDEGs and established clinicopathological features in LUAD. In the TCGA-LUAD dataset, significant differences were observed in grade, stage, T, N, and M distributions between the high- and low-risk groups ($p < 0.05$, Fig. 4A, F). Moreover, patients with advanced

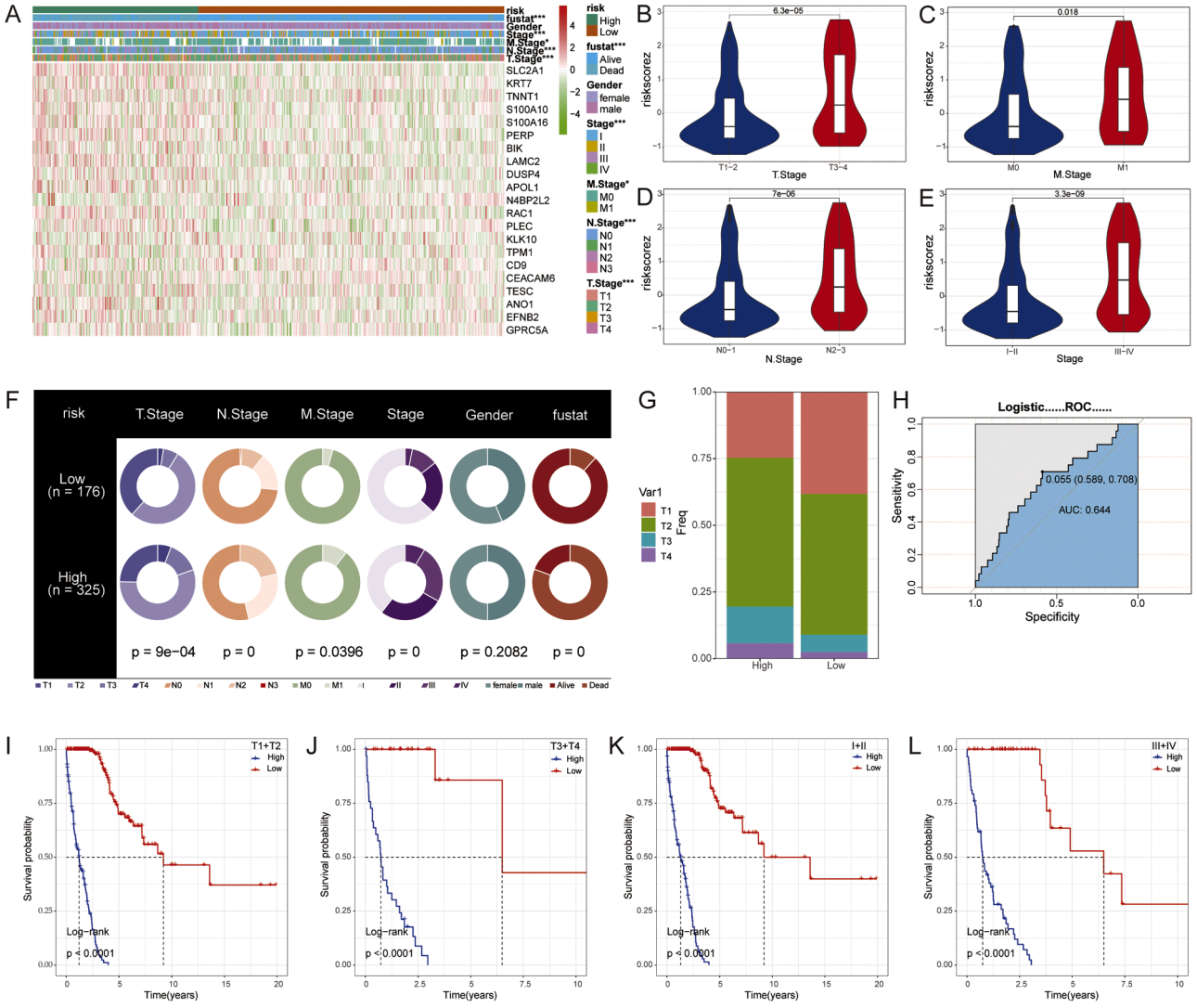


Fig. 4. Clinical correlation and prognostic significance of the LDEGs signature in LUAD.
(A, F) Distribution of clinicopathological features (grade, stage, T, N, M) between high- and low-risk groups. (B–E) Patients with advanced disease (M1, stage III–IV, N2–3, T3–4) exhibited higher risk scores. (G, H) ROC analysis demonstrating the diagnostic accuracy of LDEGs for metastasis prediction. (I–L) Kaplan–Meier survival analyses showing that high-risk patients had significantly poorer overall survival across clinical subgroups.

disease—including M1, stage III–IV, N2–3, and T3–4—had significantly higher risk scores than those with M0, stage I–II, N0–1, and T1–2 disease ($p < 0.05$, Fig. 4B–E), suggesting that LDEGs are associated with poor prognosis in LUAD. Notably, LDEGs also demonstrated the ability to predict metastatic status. Diagnostic ROC curve analysis yielded an AUC of 0.644 for predicting M stage (Fig. 4G–H), highlighting its potential clinical utility in forecasting metastatic progression. Furthermore, Kaplan–Meier survival analyses across stratified subgroups (stage and T classification) confirmed the robust prognostic performance of LDEGs in LUAD (Fig. 4I–L).

Construction and validation of a clinical nomogram

To assess whether LDEGs represent an independent prognostic factor in LUAD, we performed univariate and multivariate Cox regression analyses of OS in the TCGA-LUAD cohort. In univariate analysis, LDEGs were identified as a significant risk factor for OS (HR = 9.124, 95% CI: 7.273–11.447, $p < 0.001$). Notably, multivariate analysis confirmed that LDEGs remained an independent prognostic indicator (HR = 9.55, 95% CI: 7.214–12.643, $p < 0.001$), underscoring its strong prognostic value in LUAD patients (Fig. 5A, B).

To improve clinical applicability, we constructed a nomogram incorporating LDEGs with conventional clinical features (Fig. 5F). Calibration plots demonstrated excellent concordance between predicted and observed OS (Fig. 5D). Furthermore, C-index analysis indicated that the nomogram provided superior accuracy and stability in predicting 1–10 year OS compared with individual clinical features (Fig. 5C). DCA further confirmed that the nomogram offered greater clinical net benefit than other clinical variables (Fig. 5E). Collectively, these results establish the LDEGs -based nomogram as a reliable and precise tool for personalized prognostic assessment in LUAD.

Potential molecular mechanisms of LDEGs across large-scale transcriptomes

To further elucidate the molecular mechanisms linking LDEGs with LUAD prognosis, we performed functional enrichment analyses. In GSEA based on GO gene sets, the low-risk group was enriched for immune activation, inflammatory responses, oxidative stress, and positive regulation of inflammation-driven tumor progression (Fig. 6A), whereas the high-risk group was enriched for cell proliferation, metabolic reprogramming, cell cycle progression, and aberrant tumor cell activation (Warburg effect) (Fig. 6B). GSVA analysis revealed that high-risk patients exhibited increased activity in pathways related to GLYCOLYSIS, MYC_TARGETS and E2F_TARGETS, while the low-risk group showed higher activity in BILE_ACID_METABOLISM, MYOGENESIS and HEME_METABOLISM (Fig. 6C). Correlation analyses between LDEGs and hallmark pathway scores further supported these findings (Fig. 6D), highlighting the close association of LDEGs with cancer-related biological processes and metabolic pathways.

Correlation between LDEGs and the immune microenvironment

To assess the immune landscape of LUAD samples, we applied the ESTIMATE algorithm to calculate immune, stromal, and ESTIMATE scores, along with tumor purity. StromalScore, ImmuneScore and ESTIMATEScore were significantly lower in the high-risk group compared with controls (Fig. 7A). Using the ssGSEA algorithm, we further assessed immune-related pathway activity, showing that the high-risk group exhibited markedly higher activity in T_CELL_RECEPTOR, CHEMOKINE, and LEUKOCYTE_TRANSENDOTHELIAL_MIGRATION (Fig. 7B).

To further dissect differences in immune cell infiltration between groups, we employed the CIBERSORT algorithm to quantify immune cell abundance in each sample (Fig. 7C). High-risk patients had higher infiltration of T_cells_CD4_memory_activated, NK_cells_resting,

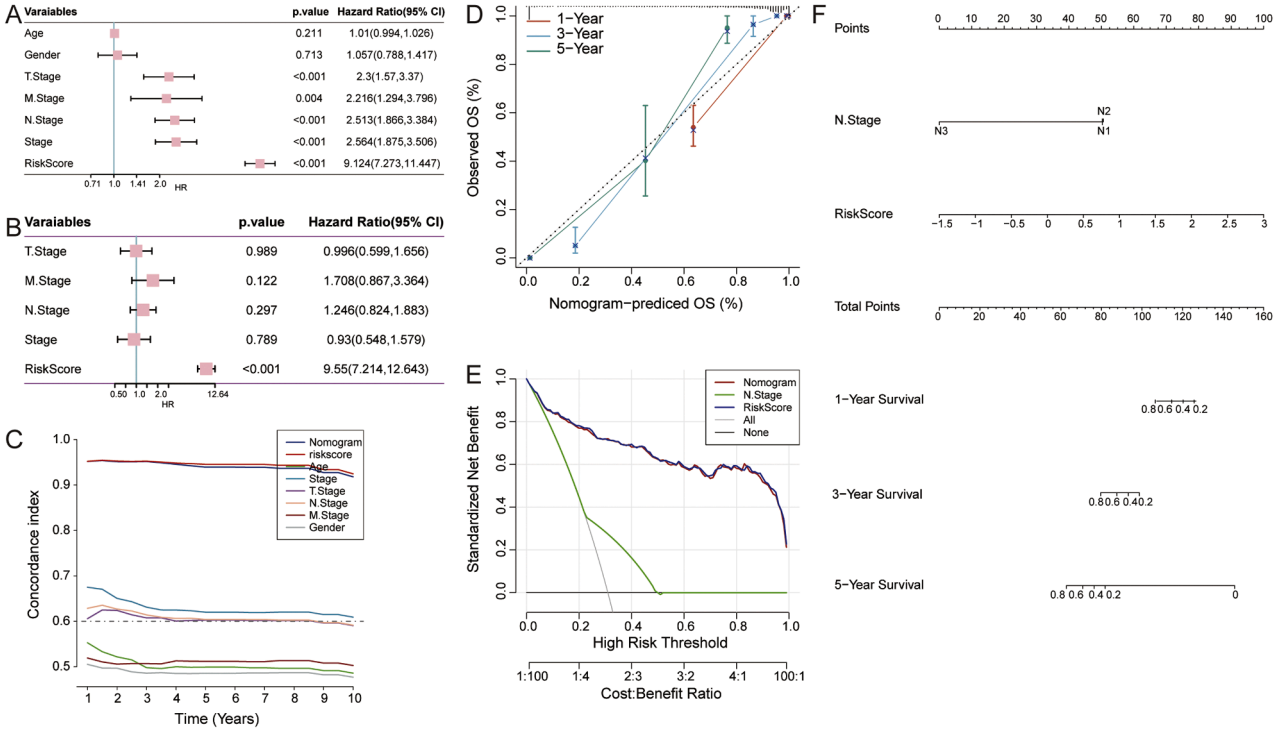


Fig. 5. Construction and validation of an LDEGs-based clinical nomogram.
(A, B) Univariate and multivariate Cox regression analyses confirming LDEGs as an independent prognostic factor. (C) Comparison of concordance indices (C-index) among LDEGs, clinical variables, and the nomogram. (D) Calibration plot showing good agreement between predicted and observed OS. (E) Decision curve analysis (DCA) demonstrating higher clinical benefit of the nomogram. (F) Nomogram integrating LDEGs and clinical features for individualized survival prediction.

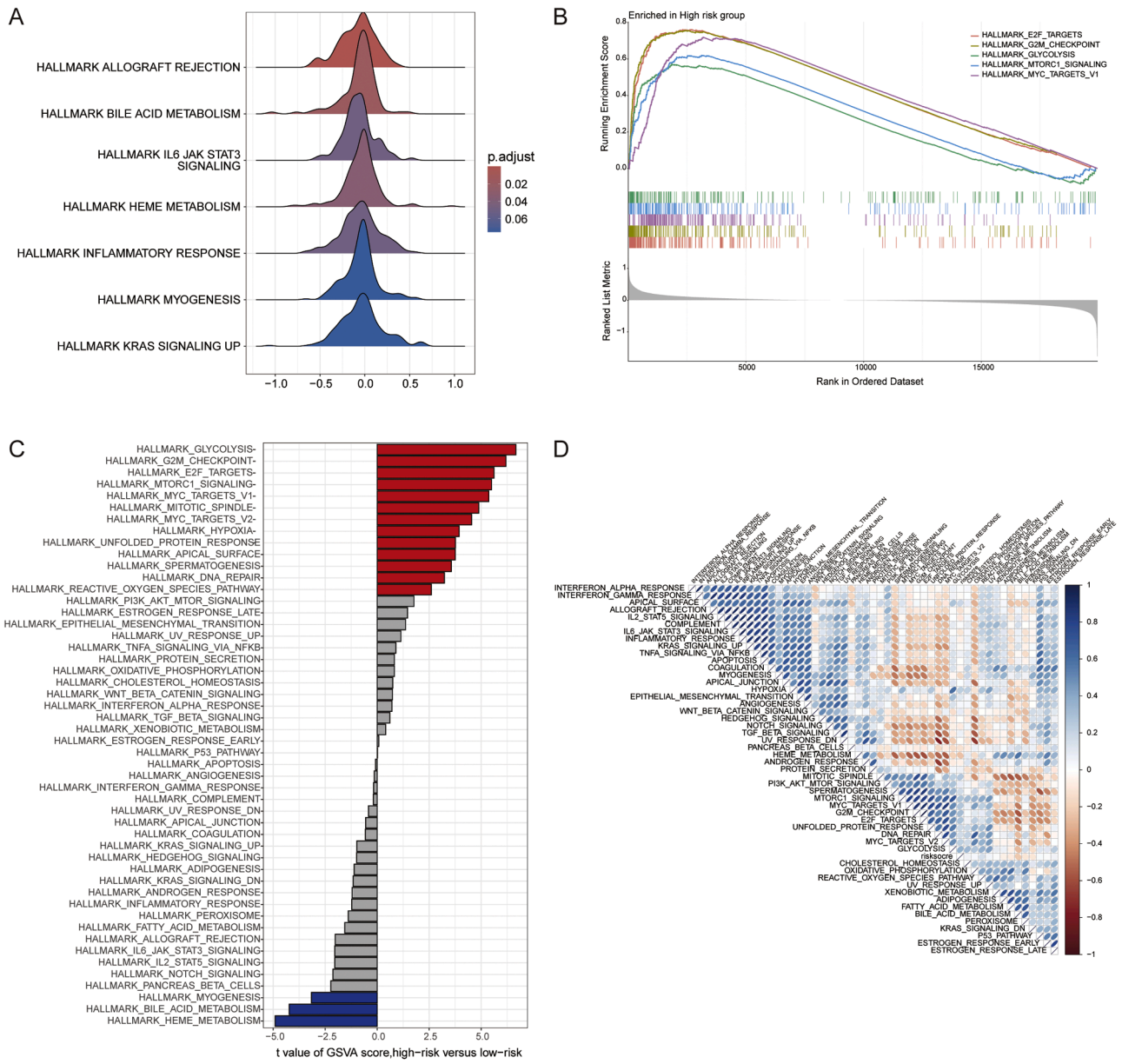


Fig. 6. Functional and pathway enrichment analyses of LDEGs.
(A, B) GSEA showing that the low-risk group was enriched in immune and inflammatory pathways, while the high-risk group showed enrichment in proliferation and metabolic reprogramming pathways. (C) GSVA demonstrating differential pathway activities between groups. (D) Correlation between LDEGs expression and hallmark pathway scores.

Macrophages_M0, and Macrophages_M1, while cell types with weaker antitumor activity, such as T_cells_CD4_memory_resting, B_cells_memory, Monocytes, and Dendritic_cells_resting, were enriched in the low-risk group. These findings were validated by ssGSEA and Xcell algorithms (Fig. 7D). Additionally, several LDEGs demonstrated strong correlations with tumor-infiltrating immune cells, including SLC2A1 positively associated with T_cells_CD4_memory_activated, Macrophages_M0, and Mast_cells_activated, and genes such as APOL1, TNNT1, LAMC2, DUSP4, N4BP2L2, RAC1, and PLEC positively associated with Macrophages_M0 (Fig. 7E).

LDEGs and their association with the cancer-immunity cycle and immunotherapy response

Given the complexity of tumor immune processes and the microenvironment, immune cell abundance alone is insufficient to fully capture the state of immune activation or exhaustion. By evaluating the activity of steps in the cancer-immunity cycle, we gained a more comprehensive understanding of the antitumor immune response and its implications for immunotherapy guidance [23]. As shown in Fig. 8A, significant differences were observed between LDEGs risk groups in steps 2, 3, 4, and 6 of the cycle. The high-risk group exhibited enhanced activity in antigen presentation and activation (Step 3), immune cell trafficking to tumors (Step 4), and T-cell recognition of cancer cells (Step 6). Further refinement of Step 4 revealed stronger recruitment capacity for immune cells, including T cells, CD8+ T cells, dendritic cells, and B cells, in the high-risk group (Fig. 8B).

Notably, previous studies have reported that higher expression of immune checkpoint molecules is associated with improved responses to immune checkpoint inhibitor (ICI) therapy [24–26]. We therefore examined checkpoint expression between LDEG-defined risk groups. CD27 and CTLA4 were significantly upregulated in the low-risk group

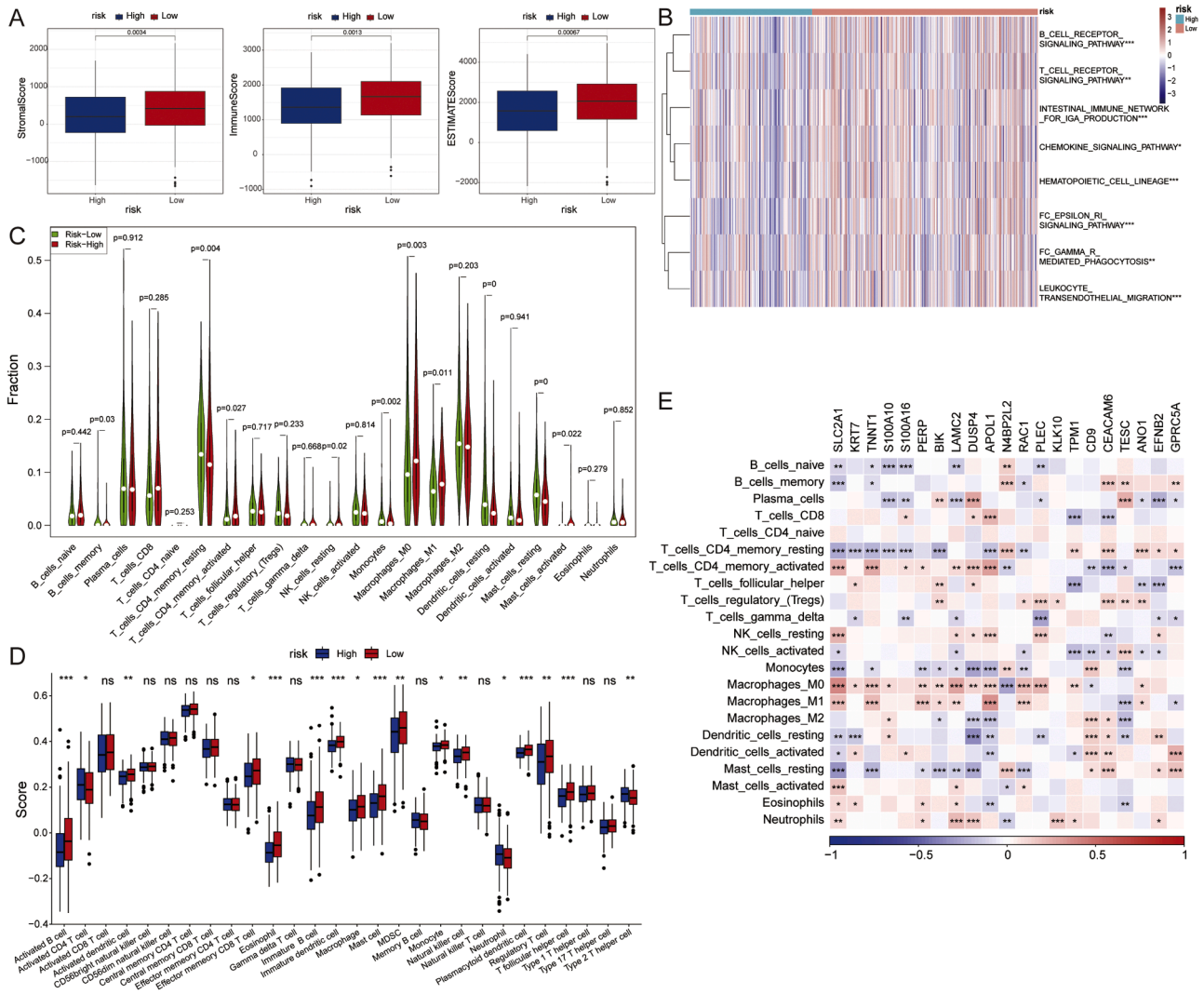


Fig. 7. Association of LDEGs with the immune microenvironment.
(A) ESTIMATE algorithm showing lower StromalScore, ImmuneScore, and ESTIMATEScore in the high-risk group. (B) ssGSEA indicating higher immune-related pathway activity in the high-risk group. (C) CIBERSORT analysis of immune cell infiltration patterns. (D) Validation of immune infiltration differences by ssGSEA and Xcell. (E) Correlation heatmap between LDEGs expression and immune cell abundance.

(Fig. 8C). To further validate the predictive power of LDEGs for immunotherapy response, we analyzed the IMvigor210 cohort of patients treated with atezolizumab. Risk scores were calculated using the LDEG model, stratifying patients into high- and low-risk subgroups. Importantly, the low-risk group exhibited a higher tumor mutational burden (TMB), a well-established predictor of immunotherapy response (Fig. 8D). Chi-square analysis revealed that the high-risk group had a significantly lower proportion of complete/partial response (CR/PR) and a higher proportion of stable/progressive disease (SD/PD) compared to the low-risk group (Fig. 8E). Consistently, patients achieving CR/PR exhibited significantly lower risk scores than those with SD/PD (Fig. 8F). Collectively, these findings demonstrate the predictive potential of LDEGs for immunotherapy response and suggest that patients in the high-risk group may derive greater clinical benefit from immune checkpoint blockade.

Correlation of LDEGs with drug sensitivity and validation of gene expression

In LUAD, frontline treatment strategies commonly involve multi-target tyrosine kinase inhibitors (TKIs) and mTOR inhibitors. Nevertheless, therapeutic resistance remains a major challenge due to the highly dynamic and heterogeneous tumor microenvironment [27,28]. To further explore potential therapeutic implications, we assessed the sensitivity of LDEG-defined risk subgroups to representative TKIs, including pazopanib and axitinib. Notably, the half-maximal inhibitory concentration (IC50) of pazopanib was significantly reduced in the high-risk group, whereas the IC50 of axitinib was markedly increased in the low-risk group (Fig. 8G, I). Consistently, the risk score showed a positive correlation with axitinib IC50 but a negative correlation with pazopanib IC50 (Fig. 8H, J). Collectively, these results suggest that patients classified into the low-risk group may benefit more from pazopanib, while those in the high-risk group could be more responsive to axitinib.

Expression and functional validation of POU5F1 in lung adenocarcinoma cells

We first detected the expression level of POU5F1 in multiple cell lines by qRT-PCR. The results showed that POU5F1 mRNA expression was significantly higher in NCI-H1975 and HCC827 cells compared to BEAS-2B and other cancer cell lines (Fig. 9A). To investigate the function of POU5F1, siRNA-mediated knockdown was performed in these two cell lines. Compared with the control group, POU5F1 mRNA levels were

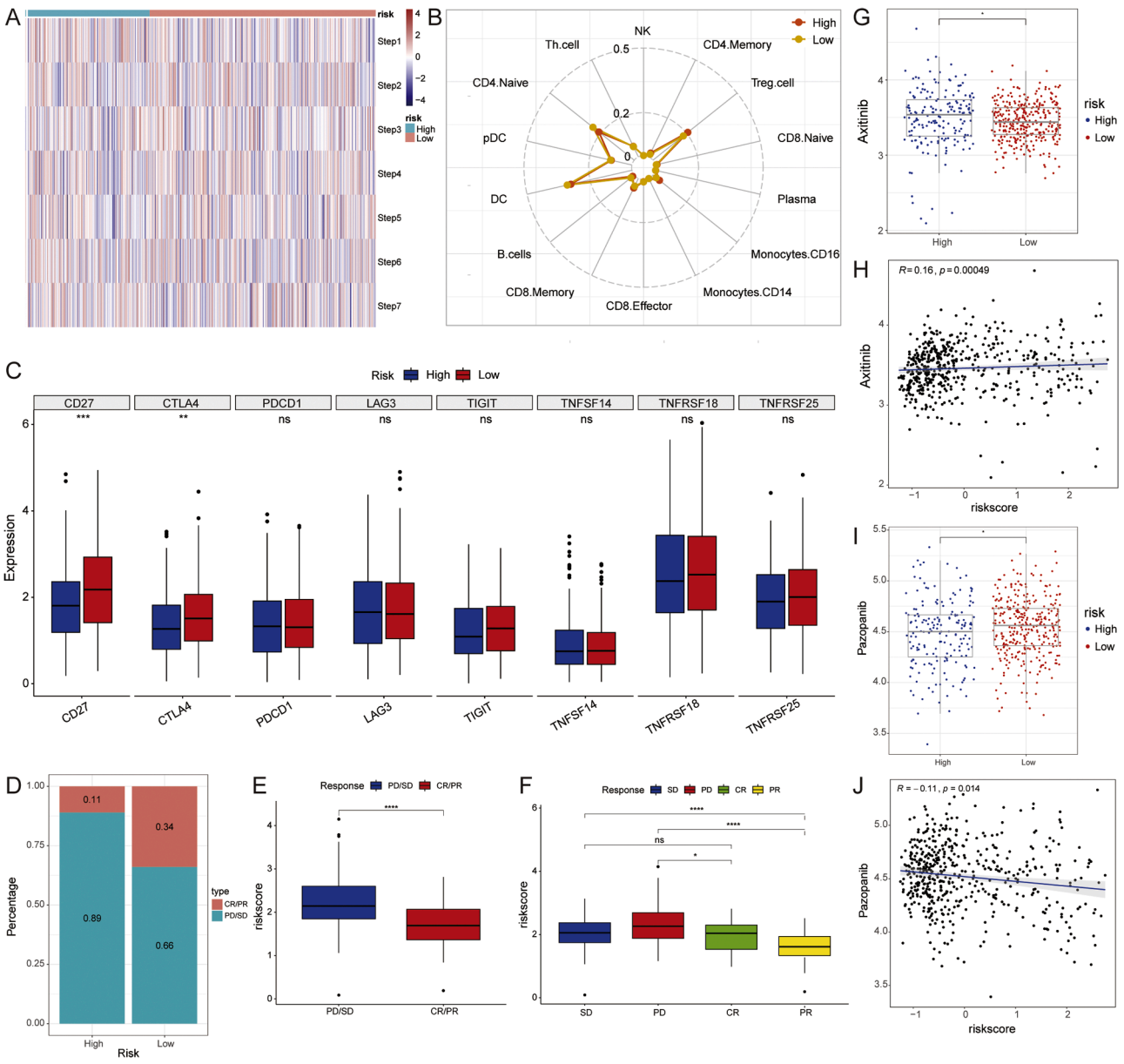


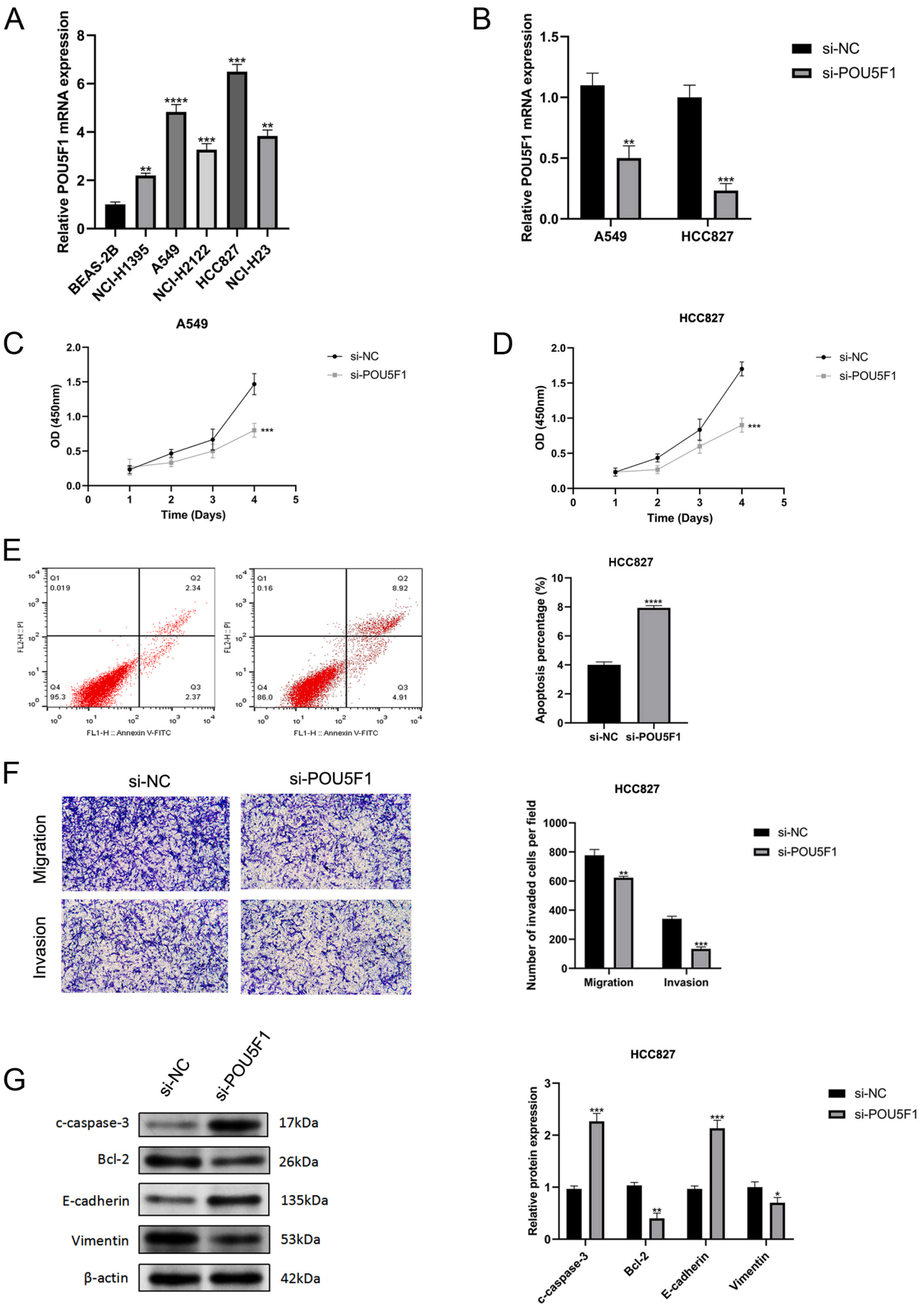
Fig. 8. LDEGs and their associations with the cancer-immunity cycle, immunotherapy response, and drug sensitivity.
(A, B) Activity differences across steps of the cancer-immunity cycle between risk groups. (C) Differential expression of immune checkpoint molecules (CD27, CTLA4).
(D–F) Validation in the IMvigor210 cohort showing TMB and response differences between risk groups. (G–J) Correlation between LDEGs-defined risk scores and drug sensitivity, indicating pazopanib benefit in low-risk and axitinib responsiveness in high-risk patients.

effectively suppressed in the knockdown groups (Fig. 9B). Next, we evaluated the effect of POU5F1 on cell proliferation. The CCK-8 assay demonstrated that knockdown of POU5F1 significantly attenuated the proliferative capacity of both cell lines, with growth curves markedly declining by day four (Fig. 9C, D). Flow cytometry analysis revealed that POU5F1 knockdown significantly increased the apoptosis rate (Fig. 9E). Furthermore, Transwell migration and invasion assays indicated a pronounced inhibitory effect of POU5F1 knockdown on tumor metastatic potential, with both migrated and invaded cell numbers significantly reduced (Fig. 9F). Western blot results showed that in NCI-H1975 cells, knockdown of POU5F1 led to upregulation of cleaved caspase-3 and E-cadherin, while expression of Bcl-2, Vimentin, and the stemness marker Nanog was downregulated (Fig. 9G). This suggests that POU5F1 not only influences apoptosis via modulation of the Bcl-2/cleaved caspase-3 axis but also participates in epithelial-mesenchymal transition (EMT) through regulation of E-cadherin/Vimentin expression, and affects cancer stem cell properties via Nanog.

Discussion

This study systematically integrates genetic causality, single-cell transcriptomics, spatial mapping, cell–cell communication, RNA velocity, and multi-algorithm machine learning to comprehensively characterize POU5F1 in LUAD. Our findings suggest that POU5F1 is not only genetically associated with lung cancer risk but also drives stemness maintenance, dedifferentiation, tumor–stroma crosstalk, and immune suppression, making it a promising biomarker and therapeutic target.

At the molecular level, POU5F1 (OCT4) is a master regulator of pluripotency. Aberrant reactivation of POU5F1 has been reported in multiple cancers [29,30]. Yang et al[29] demonstrated in gastric cancer that POU5F1 reduces ubiquitination of TRAF6, leading to sustained NF-κB signaling and enhanced proliferation and invasion [31]. This



(caption on next page)

Fig. 9. POU5F1 promotes proliferation, migration, and invasion of lung adenocarcinoma cells, while its knockdown markedly suppresses tumor phenotypes and induces apoptosis.
(A) qRT-PCR shows that POU5F1 mRNA levels are elevated in various lung adenocarcinoma cell lines, with the most pronounced increase in NCI-H1975 and HCC827. (B) After transfecting si-POU5F1 into NCI-H1975 and HCC827 cells, qRT-PCR confirms effective downregulation of POU5F1 expression. (C, D) CCK-8 assays demonstrate that POU5F1 knockdown significantly inhibits the proliferative activity of NCI-H1975 (C) and HCC827 (D) cells. (E) Flow cytometry reveals that si-POU5F1 treatment markedly elevates the apoptosis rate of NCI-H1975 cells (left panel shows flow plots, right panel shows bar graphs). (F) Transwell assays indicate that POU5F1 knockdown substantially reduces the migration and invasion capabilities of NCI-H1975 cells (upper images are microscopy pictures, lower graphs are quantitative bar charts). (G) Western blot analysis shows that POU5F1 knockdown leads to upregulation of cleaved caspase-3, downregulation of Bcl-2, upregulation of E-cadherin, downregulation of vimentin and Nanog, suggesting induction of apoptosis, reversal of epithelial-mesenchymal transition (EMT), and inhibition of stemness. β-actin serves as the loading control.

mechanism raises the hypothesis that in LUAD, POU5F1 may also activate NF-κB to promote inflammation-driven tumor progression. Similarly, Su et al[30] reported that OCT4 upregulates VCC-1 in lung cancer, which enhances TGF-β secretion and invasion [32]. Combined with our observation of spatial colocalization between POU5F1-positive tumor cells and CAFs, these findings suggest that POU5F1 may indirectly activate CAFs via TGF-β-related pathways, reinforcing stromal remodeling and immunosuppression.

CAFs are major regulators of the tumor microenvironment. Recent reviews highlight their role in ECM remodeling, angiogenesis, immune suppression, and therapeutic resistance [33–35]. Melchionna et al. (2023) emphasized transcription factors as critical drivers of CAF plasticity [36], suggesting that transcription factors like POU5F1 could modulate CAF states via paracrine or exosomal signaling. Jia et al. (2025) further pointed out that CAF heterogeneity and plasticity result in divergent tumor-promoting versus tumor-restraining functions [37]. Therefore, POU5F1–CAF interactions may not only foster immunosuppression but also dictate patient prognosis depending on CAF subtype specificity.

RNA velocity analysis showed progressive upregulation of POU5F1 along pseudotime, consistent with its role in dedifferentiation and lineage plasticity. Prior studies linked high POU5F1 expression with therapy resistance [38,39]. In colorectal cancer, Fujino et al. (2023) identified POU5F1-positive circulating tumor cells (CTCs) as highly stem-like and metastatic, driving relapse [40]. Analogously, in LUAD, POU5F1-positive subpopulations may act as “seed cells” for recurrence and therapy resistance.

Spatial transcriptomics revealed strong colocalization of POU5F1-positive tumor regions with CAF-rich areas. Liu et al. (2025) showed that CAFs form conserved spatial subtypes across cancers [41], while Cords et al. (2024) classified 11 CAF phenotypes with prognostic spatial arrangements [42]. These studies underscore the importance of investigating CAF subtype-specific spatial niches enriched in POU5F1-positive tumor cells.

Immunologically, CAFs are described as “architects” of the tumor niche [43,44]. In lung cancer, CAF infiltration correlates with M2 macrophage polarization and immune suppression via CXCL12–CXCR4 signaling [45]. Our findings on immune checkpoint expression and cancer-immunity cycle support the predictive potential of POU5F1-associated features for immunotherapy response. Milosevic et al. (2024) summarized CAF-mediated T cell suppression mechanisms, including metabolic deprivation, checkpoint factor expression, and spatial exclusion [46], which may be modulated by POU5F1-driven CAF reprogramming.

Clinically, our POU5F1-related prognostic model demonstrated robust predictive performance across cohorts and correlated with immunotherapy response and drug sensitivity. Recent studies also highlighted CAF-related features in predicting immunotherapy outcomes [47,48]. Shen et al developed CAF-targeting nanodrugs that improved drug penetration and suppressed tumor growth in colorectal cancer [49], providing a rationale for combinatorial strategies in LUAD. Designing therapies that simultaneously inhibit POU5F1 and target CAFs may synergistically enhance immune checkpoint blockade and chemotherapy.

Functional studies demonstrate that the transcription factor POU5F1 plays a central oncogenic role in LUAD. Basic experiments confirm that POU5F1 is specifically highly expressed in lung adenocarcinoma cell lines; its knockdown significantly inhibits cell proliferation, migration, and invasion, while inducing apoptosis. Mechanistically, POU5F1 promotes tumor survival, metastasis, and stemness maintenance by regulating the Bcl-2/caspase-3 apoptotic pathway, influencing the EMT process mediated by E-cadherin/Vimentin, and downregulating the core stemness factor Nanog. In summary, our work integrates single-cell and spatial transcriptomics to decode the dynamic interactive spatial niche between $POU5F1^{+}$ tumor cells and the CAF-enriched stromal microenvironment; combined with the above experimental validation, we systematically elucidate the multidimensional functions of POU5F1 in driving malignant progression; and finally, we construct a robust machine-learning-based prognostic model. These findings provide a new molecular framework and experimental evidence for prognosis assessment and precision therapy targeting the “stemness-stroma-immunity” interaction network in LUAD.

Several limitations of this study should be acknowledged. First, while our spatial transcriptomics and cell-cell communication analyses revealed significant co-localization and enriched ligand-receptor signaling between $POU5F1^{+}$ malignant cells and cancer-associated fibroblasts, these findings are correlational in nature. Spatial co-localization alone does not demonstrate functional interaction or causal remodeling of the tumor stroma, and the proposed model in which $POU5F1^{+}$ cells actively reprogram the microenvironment remains hypothetical until validated by direct functional experiments, such as conditioned medium transfer assays or in vivo co-injection models. Second, our in vitro functional experiments primarily addressed the cell-autonomous functions of POU5F1 (proliferation, apoptosis, migration, invasion, EMT, and stemness), but did not directly test its proposed non-cell-autonomous roles in CAF activation or immune modulation. Future studies using co-culture systems and immunocompetent animal models are needed to establish causality. Third, our machine learning prognostic model showed extremely high predictive performance in the TCGA cohort (AUC $>$ 0.98) but substantially lower performance in external validation cohorts (AUC approximately 0.70), suggesting potential overfitting. While this drop is partly attributable to differences in gene expression platforms, sample sizes, and patient populations, the external validation performance likely represents a more realistic estimate of model generalizability, and further validation in additional independent, multi-center LUAD cohorts is required before any clinical application. Despite these limitations, our multi-omics integration provides a comprehensive framework for understanding POU5F1 in LUAD and generates testable hypotheses for future mechanistic and translational studies.

Conclusion

POU5F1 promotes LUAD progression through cell-autonomous mechanisms including maintaining stemness and driving aggressive tumor phenotypes. Computational evidence further suggests that POU5F1 may contribute to reprogramming CAFs and shaping immune evasion, but these non-cell-autonomous functions require direct experimental validation in future studies.

Funding

This research received support from the Medical Science Research Project of Hebei (20260350).

Ethics statement

All methods are carried out in accordance with relevant guidelines and regulations. The datasets generated during and/or analyzed during the current study are publicly available. The patients involved in the database have obtained ethical approval. Gene verification only involves cell lines and animal, does not involve human studies and experiments.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

CRediT authorship contribution statement

Bingbing Yu: Writing – original draft, Formal analysis, Data curation. **Meili Zhang:** Writing – original draft, Visualization, Software. **Haiyong Zhu:** Writing – original draft, Visualization, Software. **Zhu Wu:** Writing – original draft, Methodology, Investigation. **Hong Gao:** Writing – review & editing, Validation, Project administration, Funding acquisition. **Yang Bai:** Writing – review & editing, Supervision, Resources, Project administration, Conceptualization.

Declaration of competing interest

The authors declare no potential competing interests.

Acknowledgements

Not applicable.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.tranon.2026.102800.

References

[1] H. Sung, et al., Global Cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA Cancer J. Clin. 71 (3) (2021) 209–249.
[2] F. Bray, et al., Global cancer transitions according to the human development index (2008-2030): a population-based study, Lancet Oncol. 13 (8) (2012) 790–801.
[3] F. Bray, et al., Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA Cancer J. Clin. 74 (3) (2024) 229–263.
[4] R. Sullivan, et al., Global cancer surgery: delivering safe, affordable, and timely cancer surgery, Lancet Oncol. 16 (11) (2015) 1193–1224.
[5] P. Wang, et al., Systemic inflammation influences the prognosis of patients with radically resected non-small cell lung cancer and correlates with the immunosuppressive microenvironment, Int. J. Cancer 153 (4) (2023) 826–842.
[6] P. Chen, et al., Non-small cell lung cancer in China, Cancer Commun. 42 (10) (2022) 937–970.
[7] K. Chen, et al., Propensity-matched comparison of video-assisted thoracoscopic with thoracotomy lobectomy for locally advanced non-small cell lung cancer, J. Thorac. Cardiovasc. Surg. 153 (4) (2017) 967–976.e2.
[8] R. Rami-Porta, et al., Lung cancer - major changes in the American joint committee on cancer eighth edition cancer staging manual, CA Cancer J. Clin. 67 (2) (2017) 138–155.
[9] G. Martinez-Zayas, et al., Predicting lymph node metastasis in non-small cell lung cancer: prospective external and temporal validation of the HAL and HOMER models, Chest 160 (3) (2021) 1108–1120.
[10] M.V. Brock, et al., DNA methylation markers and early recurrence in stage I lung cancer, N. Engl. J. Med. 358 (11) (2008) 1118–1128.
[11] K. Shiga, et al., Cancer-associated fibroblasts: their characteristics and their roles in tumor growth, Cancers 7 (4) (2015) 2443–2458.
[12] W. Zhang, P. Huang, Cancer-stromal interactions: role in cell survival, metabolism and drug sensitivity, Cancer Biol. Ther. 11 (2) (2011) 150–156.
[13] Z. Werb, P. Lu, The role of stroma in tumor development, Cancer J. 21 (4) (2015) 250–253.
[14] P. Gascard, T.D. Tlsty, Carcinoma-associated fibroblasts: orchestrating the composition of malignancy, Genes Dev. 30 (9) (2016) 1002–1019.
[15] P. Cirri, P. Chiarugi, Cancer associated fibroblasts: the dark side of the coin, Am. J. Cancer Res. 1 (4) (2011) 482–497.
[16] M. Musa, Single-cell analysis on stromal fibroblasts in the microenvironment of solid tumours, Adv. Med. Sci. 65 (1) (2020) 163–169.
[17] P. Gandellini, et al., Complexity in the tumour microenvironment: cancer associated fibroblast gene expression patterns identify both common and unique features of tumour-stroma crosstalk across cancer types, Semin. Cancer Biol. 35 (2015) 96–106.
[18] C.H. Sukowati, et al., The role of multipotent cancer associated fibroblasts in hepatocarcinogenesis, BMC. Cancer 15 (2015) 188.
[19] T. Terada, et al., Alpha-smooth muscle actin-positive stromal cells in cholangiocarcinomas, hepatocellular carcinomas and metastatic liver carcinomas, J. Hepatol. 24 (6) (1996) 706–712.
[20] Y. Yamamura, et al., Akt-girdin signaling in cancer-associated fibroblasts contributes to tumor progression, Cancer Res. 75 (5) (2015) 813–823.
[21] J. Song, et al., Hepatic stellate cells activated by acidic tumor microenvironment promote the metastasis of hepatocellular carcinoma via osteopontin, Cancer Lett. 356 (2 Pt B) (2015) 713–720.
[22] L.E. Littlepage, M. Egeblad, Z. Werb, Coevolution of cancer and stromal cellular responses, Cancer Cell. 7 (6) (2005) 499–500.
[23] D.S. Chen, I. Mellman, Oncology meets immunology: the cancer-immunity cycle, Immunity 39 (1) (2013) 1–10.
[24] Z.Y. Xu-Monette, J. Zhou, K.H. Young, PD-1 expression and clinical PD-1 blockade in B-cell lymphomas, Blood 131 (1) (2018) 68–83.
[25] S. Qin, et al., Novel immune checkpoint targets: moving beyond PD-1 and CTLA-4, Mol. Cancer 18 (1) (2019) 155.
[26] J.M. Chauvin, H.M. Zarour, TIGIT in cancer immunotherapy, J. ImmunOther Cancer. 8 (2) (2020) e000957.
[27] P. Makhov, et al., Resistance to systemic therapies in clear cell renal cell carcinoma: mechanisms and management strategies, Mol. Cancer Ther. 17 (7) (2018) 1355–1364.
[28] M.B. Atkins, N.M. Tannir, Current and emerging therapies for first-line treatment of metastatic clear cell renal cell carcinoma, Cancer Treat. Rev. 70 (2018) 127–137.
[29] W. Yang, et al., POU5F1 promotes the proliferation, migration, and invasion of gastric cancer cells by reducing the ubiquitination level of TRAF6, Cell Death. Dis. 14 (12) (2023) 802.
[30] B.H. Su, et al., OCT4 promotes lung cancer progression through upregulation of VEGF-correlated chemokine-1, Int. J. Med. Sci. 22 (3) (2025) 680–695.
[31] X. Cui, et al., Single-cell and spatial transcriptomic analyses revealing tumor microenvironment remodeling after neoadjuvant chemoimmunotherapy in non-small cell lung cancer, Mol. Cancer 24 (1) (2025) 111.
[32] A.N. Kazakova, et al., Exploring the diversity of cancer-associated fibroblasts: insights into mechanisms of drug resistance, Front. Cell Dev. Biol. 12 (2024) 1403122.
[33] K. Xu, et al., Distinct fibroblast subpopulations associated with bone, brain or intrapulmonary metastasis in advanced non-small-cell lung cancer, Clin. Transl. Med. 14 (3) (2024) e1605.
[34] S. Li, et al., Pan-cancer single cell and spatial transcriptomics analysis deciphers the molecular landscapes of senescence related cancer-associated fibroblasts and reveals its predictive value in neuroblastoma via integrated multi-omics analysis and machine learning, Front. Immunol. 15 (2024) 1506256.
[35] H. Deng, et al., Spatial transcriptomics uncovers immune-cell plasticity and dedifferentiation signatures in aggressive lung adenocarcinoma subtypes, Front. Immunol. 16 (2025) 1620886.
[36] R. Melchionna, et al., Transcription factors in fibroblast plasticity and CAF heterogeneity, J. Exp. Clin. Cancer Res. 42 (1) (2023) 347.
[37] H. Jia, et al., Cancer associated fibroblasts in cancer development and therapy, J. Hematol. Oncol. 18 (1) (2025) 36.
[38] S. Cai, et al., POU5F1/Oct-4 expression in breast cancer tissue is significantly associated with non-sentinel lymph node metastasis, BMC. Cancer 16 (2016) 175.
[39] H. Wang, et al., Ferritinophagy mediates adaptive resistance to EGFR tyrosine kinase inhibitors in non-small cell lung cancer, Nat. Commun. 15 (1) (2024) 4195.
[40] S. Fujino, et al., Metastases and treatment-resistant lineages in patient-derived cancer cells of colorectal cancer, Commun. Biol. 6 (1) (2023) 1191.
[41] Y. Liu, et al., Conserved spatial subtypes and cellular neighborhoods of cancer-associated fibroblasts revealed by single-cell spatial multi-omics, Cancer Cell. 43 (5) (2025) 905–924.e6.
[42] L. Cords, et al., Cancer-associated fibroblast phenotypes are associated with patient outcome in non-small cell lung cancer, Cancer Cell. 42 (3) (2024) 396–412.e5.
[43] M. Sarkar, et al., Cancer-associated fibroblasts: the chief architect in the tumor microenvironment, Front. Cell Dev. Biol. 11 (2023) 1089068.
[44] R.A. Glabman, P.L. Choyke, N. Sato, Cancer-associated fibroblasts: tumorigenicity and targeting for Cancer therapy, Cancers 14 (16) (2022) 3906.
[45] R. Zhu, J. Huang, F. Qian, The role of tumor-associated macrophages in lung cancer, Front. Immunol. 16 (2025) 1556209.
[46] X. Mao, et al., Crosstalk between cancer-associated fibroblasts and immune cells in the tumor microenvironment: new findings and future perspectives, Mol. Cancer 20 (1) (2021) 131.

[47] J. Zhao, et al., Cancer-associated fibroblasts and metabolic reprogramming predict pathologic response to neoadjuvant PD-1 blockade in resected non-small cell lung cancer, Cell Oncol. 48 (4) (2025) 1105–1119.
[48] M. Fu, et al., Overcoming tyrosine kinase inhibitor resistance in lung cancer brain metastasis with CTLA4 blockade, Cancer Cell. 42 (11) (2024) 1882–1897.e7.
[49] W. Shen, et al., Cancer-associated fibroblast-targeted nanodrugs reshape colorectal tumor microenvironments to suppress tumor proliferation, metastasis and improve drug penetration, J. Mater. Chem. B 11 (9) (2023) 1871–1880.