Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A metabolic phenotype-associated prognostic signature reveals distinct immune landscapes and CST5-mediated tumor suppression in lung adenocarcinoma

Huiyan Liu
Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer, Tianjin's Clinical Research Center for Cancer

Peihan Zhao
Tianjin Medical University

Yantao Shao
Tianjin Medical University

Xubo Huang
Nanshan Hospital of Shenzhen University

Junjie Yu
Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer, Tianjin's Clinical Research Center for Cancer

Yantao Jiang
Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer, Tianjin's Clinical Research Center for Cancer

Jiayue Chen
Tianjin Medical University

Yi Lu
Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer, Tianjin's Clinical Research Center for Cancer

Luyao Tong
Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer, Tianjin's Clinical Research Center for Cancer

Tingting Qin
qintingting@tjmuch.com

Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer, Tianjin's Clinical Research Center for Cancer

Research Article

Keywords: Lung adenocarcinoma, Metabolic subtypes, Tumor immune microenvironment, Spatial transcriptomics, CST5

Posted Date: November 20th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11700292/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

A metabolic phenotype-associated prognostic signature reveals distinct immune landscapes and CST5-mediated tumor suppression in lung adenocarcinoma

Huiyan Liu[#1], Peihan Zhao[#2], Yantao Shao[#2], Xubo Huang[3], Junjie Yu[1], Yantao Jiang[1], Jiayue Chen[2], Yi Lu[1], Luyao Tong[1], Tingting Qin[1*]

1 Department of Thoracic Oncology, Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer, Tianjin Key Laboratory of Cancer Prevention and Therapy, Tianjin's Clinical Research Center for Cancer, Tianjin, China.

2 Department of Bioinformatics, Department of Cell Biology, School of Basic Medical Sciences, Tianjin Medical University, Tianjin, China

3 Nanshan Hospital of Shenzhen University, Shenzhen, China

# Contributed equally.

*Correspondence: Tingting Qin, Ph.D., Department of Thoracic Oncology, Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer, Tianjin Key Laboratory of Cancer Prevention and Therapy, Tianjin's Clinical Research Center for Cancer, Tianjin, China. Email: qintingting@tjmuch.com

Abstract

Lung adenocarcinoma (LUAD) is a leading cause of cancer-related mortality worldwide. Tumor metabolic reprogramming critically modulates the immune microenvironment, yet the clinical implications of metabolic-immune crosstalk remain poorly defined. We systematically quantified 75 metabolic pathways via single-sample gene set enrichment analysis and identified two robust metabolic subtypes with divergent immune landscapes through consensus clustering. By integrating machine learning algorithms, we constructed a metabolic phenotype-associated prognostic signature (MPAPS). MPAPS demonstrated robust prognostic performance across multiple cohorts, showing favorable discrimination compared with existing signatures. The high-risk group displayed a profoundly immunosuppressive microenvironment with elevated myeloid-derived suppressor cells and cancer-associated fibroblast infiltration, correlating with poor immunotherapy response across multiple cancer types. Conversely, the low-risk group exhibited abundant B cells, CD8+ T cells, and natural killer cells. Additionally, high-risk tumors showed increased sensitivity to chemotherapy and targeted agents. Mechanistically, the MPAPS gene CST5 suppressed tumor progression by inhibiting the PI3K/AKT/mTOR pathway and promoting M1 macrophage polarization, as validated by in vitro and in vivo experiments. Collectively, MPAPS reveals distinct immune landscapes in LUAD, with high-risk tumors exhibiting an immunosuppressive microenvironment and showing poor immunotherapy response across independent cohorts, and highlights CST5 as a tumor suppressor that inhibits the PI3K/AKT/mTOR pathway and promotes M1 macrophage polarization.

Keywords: Lung adenocarcinoma, Metabolic subtypes, Tumor immune microenvironment, Spatial transcriptomics, CST5

Introduction

Lung cancer remains the leading cause of cancer-related mortality worldwide and one of the most commonly diagnosed malignancies globally [1]. As the most prevalent histological subtype of lung cancer, LUAD contributes to approximately 40% of all lung cancer diagnoses [2, 3]. Although recent advances in surgical resection, molecular targeted therapy, and immunotherapy have improved outcomes for select patients, the overall 5-year survival rate remains below 20%, reflecting profound intratumoral heterogeneity and the frequent emergence of therapeutic resistance [4].

Metabolic reprogramming represents a well-established hallmark of cancer that enables malignant cells to meet the bioenergetic and biosynthetic demands of relentless proliferation [5]. In LUAD, malignant cells undergo profound metabolic rewiring to meet the elevated energy and biosynthetic demands imposed by rapid proliferation[6]. This adaptive metabolic remodeling further fuels uncontrolled tumor growth[7] and drives resistance to both conventional therapies and cellular stress[8]. Multiple core determinants shape this metabolic reprogramming, including the limited nutrient availability in the tumor microenvironment (TME) [9] and oncogenic signaling that initiates and sustains tumorigenesis[10]. Although extensive studies have delineated the global metabolic landscape of LUAD, it remains largely underexplored whether metabolism-based molecular stratification can reliably stratify patient prognosis and predict immunotherapy response to guide personalized treatment [11].

Mounting evidence indicates that the crosstalk between tumor metabolic pathways and the immune system is highly intricate and multifaceted [12]. The metabolic reprogramming of TME reshapes the function and distribution of major tumor-infiltrating immune populations, including tumor-associated macrophages (TAMs), regulatory T cells (Tregs), and cytotoxic CD8+ T cells, which in turn further modulates

anti-tumor immunity [13]. Tumor metabolic dysregulation can also hinder effector immune cell infiltration by promoting angiogenesis, ultimately facilitating tumor immune evasion [14]. While the landscape of metabolic-immune crosstalk in LUAD is well characterized, it remains unclear whether metabolic-based stratification can reliably predict prognosis and immunotherapy response in LUAD.

The advent of multi-omics technologies including transcriptomics, proteomics and metabolomics have brought unprecedented insights into cancer metabolic-immune crosstalk [15]. For LUAD specifically, single-cell and spatial transcriptomics have substantially advanced our understanding of intratumoral heterogeneity and TME complexity [16, 17], while cutting-edge tools like CRISPR screening and advanced imaging allow identification of specific metabolic signature genes for tumor classification and therapeutic target prioritization [18]. In this study, we systematically quantified 75 metabolic pathways across bulk transcriptomic datasets and identified two robust metabolic subtypes with markedly divergent immune infiltration patterns through consensus clustering. By integrating machine learning algorithms with transcriptomic and clinical data, we constructed the metabolic phenotype-associated prognostic signature MPAPS and validated its predictive performance across multiple independent cohorts. Furthermore, we comprehensively characterized the immune landscape across MPAPS-defined risk groups. Integrative single-cell and spatial transcriptomic analyses provided additional resolution into the tumor-immune architecture. Finally, we performed in-depth mechanistic validation of CST5, a key MPAPS gene, to elucidate its biological function in LUAD progression and immune modulation. Collectively, this study establishes MPAPS as a clinically actionable prognostic and immunological stratification tool, and functionally validates CST5 as a promising therapeutic target for LUAD intervention.

Materials and methods

Data resources and clinical samples

In this study, we compiled six bulk transcriptome datasets for LUAD, including the TCGA-LUAD RNA-seq cohort from TCGA and five independent microarray datasets from GEO (TCGA: https://portal.gdc.cancer.gov/; GEO: https://www.ncbi.nlm.nih.gov/geo/). The three GEO microarray datasets , including GSE13213, GSE31210, and GSE42127, were combined using the R packages “limma” and “sva”, with batch effects corrected by ComBat (from sva), to generate the merged dataset called "GEOmerge". We additionally included two immunotherapy-focused cohorts, GSE91061 and Melanoma_phs000452, alongside the IMvigor210 dataset, to validate the prognostic performance and immunotherapy response prediction capacity of our model.

Single-cell transcriptomics data were retrieved from the GEO database under accession numbers GSE189357 and GSE131907. The single-cell RNA-seq dataset GSE189357 was subjected to strict quality control filtering, with the following per-cell thresholds: 200–10,000 detected genes, total unique molecular identifier count > 500, mitochondrial gene fraction < 10%, and hemoglobin gene fraction < 10%. All qualified cells were retained for downstream analyses. The public GEO scRNA-seq dataset GSE131907 was pre-quality-controlled by the original study. The Visium spatial transcriptomics dataset was sourced from the GEO database under accession number GSE189487. We screened out six samples that met the requirements from this dataset. Using the Cottrazm (v0.1.1) and Scissor (v2.0.0) algorithms, these samples were successfully divided into two clusters (Cluster 1 and Cluster 2) and two risk groups (low-risk and high-risk), respectively.

Metabolic pathway collection and enrichment scoring

The metabolic pathway selection in this study was based on MSigDB. We retrieved a total of 75 metabolism-related pathways, and subsequently utilized the GSVA (V1.48.3) package to perform ssGSEA on these metabolic gene sets to calculate the enrichment scores of metabolic pathways for each sample.

Identification of LUAD metabolic subtypes

In this study, we identified metabolic subtypes of LUAD via consensus clustering based on the ssGSEA scores of prognosis-related metabolic pathways, and systematically evaluated the prognostic value of these pathways in two independent cohorts, including TCGA-LUAD and the merged GEO cohort (GEOmerge). Firstly, Kaplan–Meier survival analysis was performed for each pathway in both datasets. Pathways with a hazard ratio (HR) < 1 and p-value < 0.05 across both cohorts were defined as protective pathways, while those with HR > 1 and p-value < 0.05 were defined as risk pathways. Subsequently, consensus clustering of patients was performed based on the ssGSEA scores of the pathways integrated from the two datasets, with Euclidean distance as the distance metric and the K-means used for clustering. The optimal number of clusters was determined by the cumulative distribution function (CDF) curve, and all analyses were completed using the ConsensusClusterPlus R package (v1.64.0).

GSEA (Gene Set Enrichment Analysis)

GSEA was performed via the clusterProfiler R package (v4.8.3) to characterize functional differences between the two LUAD subtypes, with all input gene sets obtained from MSigDB. A pre-ranked gene list was constructed from the log2 fold change (log2FC) values derived from inter-subtype differential expression analysis conducted with the limma package. Cluster-specific enriched functions were identified based on adjusted p-value and normalized enrichment score (NES).

Assessment of immune cell infiltration and immunotherapy response

Immune infiltration analysis, a widely used approach for assessing the composition of diverse immune cell populations in the TME, quantifies the abundance of immune cell subtypes via support vector regression-based deconvolution of bulk gene expression matrices. Within this research, patient data underwent analysis utilizing the CIBERSORT, MCPcounter, EPIC, xCell, IPS, quanTlseq, ESTIMATE, ABIS, ConsensusTME and TIMER algorithms integrated in the tigeR package (v3.0.0). Furthermore, 23 commonly immunotherapy-related gene signatures, including TIDE-associated, TME-associated, and other well-documented signatures (IPS, CD8+Teffector, TLS, EMT, etc) were used for LUAD subtype analysis within the tigeR package.

Functional enrichment analysis

To decipher the biological functions and pathways associated with tumorigenesis and progression, we conducted Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis on the identified differentially expressed genes (DEGs). Min overlap ≥ 5 & $p \leq 0.05$ were considered statistically significant. Additionally, the Pathway RespOnsive GENes (PROGENy; v1.22.0) algorithm was applied for pathway activity inference.

Construction and evaluation of prognostic models

Initially, differential expression analysis was performed between the two previously identified LUAD metabolic subtypes to screen for DEGs meeting $log2FC > 1$ and $P < 0.05$. Univariate Cox regression analysis was then conducted to further select genes significantly associated with overall survival ($P < 0.05$). To construct a robust, generalizable prognostic model, we evaluated 10 distinct machine learning workflows, each consisting of a feature selection step followed by survival model construction. Among all tested workflows, the StepCox(forward)+ random survival forest (RSF)

approach achieved the highest average concordance index (C-index) across multiple independent validation cohorts, and was therefore selected as the final prognostic model. Specifically, StepCox(forward) was used for feature selection with AIC as the stopping criterion. The retained prognostic features were used to build the final prediction model via RSF, and RSF hyperparameters were optimized via 10-fold cross-validation. Risk scores were extracted from the model's prediction output, and model performance was validated in independent external cohorts. Patients were stratified into low- and high-risk groups based on the median risk score. The prognostic performance of the model was evaluated via Kaplan-Meier survival analysis and time-dependent ROC analysis.

A nomogram based on Cox regression coefficients was constructed to predict 1-, 3-, 5-year overall survival in LUAD patients, integrating prognostic risk score and clinical parameters (age, sex, metastatic status). All analyses were performed via R packages rms, regplot, Hmisc, and calibration curves plus ROC analysis assessed concordance between predicted and observed survival outcomes.

Drug sensitivity analysis

To predict antitumor drug response in LUAD patients, we performed drug sensitivity analysis using the `OncoPredict` R package developed by Maeser et al. A total of 198 antitumor drugs were included in the analysis, and differences in drug sensitivity between the groups were compared via unpaired t-tests. The half-maximal inhibitory concentration (IC50) of each antitumor drug was estimated for every individual LUAD patient based on the gene expression matrix. Additionally, Pearson correlation analysis was conducted to quantify the correlation between predicted drug sensitivity (IC50) and patients' risk scores.

scRNA-seq data processing

scRNA-seq data were imported into the `Seurat` R package (v5.1.0) for downstream

analysis and visualization. Uniform manifold approximation and projection (UMAP) was used to visualize all annotated cell populations, followed by dot plotting of cell type marker genes. We applied the Harmony method for batch correction across each single-cell samples. We used the `copyKAT` R package (v1.1.0) with default settings to infer single-cell copy number variations (CNVs) and identify tumor cells among epithelial cells: diploid cells were classified as normal and aneuploid cells as malignant tumor cells. Differential expression analysis was performed using the `FindMarkers` function in Seurat, and GSEA was performed for functional enrichment analysis between the two metabolic subtypes of tumor cells. The `Scissor` algorithm (v2.0.0) was used to link TCGA bulk labels to scRNA-seq tumor cells. In Scissor, the penalty balancing parameter α was set to 0.01 and the model family was set to "binomial". Scissor+ cells are associated with Cluster 2/low-risk, Scissor- cells are associated with Cluster 1/high-risk, and the remaining cells are intermediate-state cells.

Spatial transcriptomics data processing

Spatial transcriptomics data were preprocessed and analyzed with the `Seurat` R package, including data normalization, dimensionality reduction, and visualization. We then performed precise cell type mapping of the spatial transcriptomics data using the non-negative least squares-based `RCTD` method (v2.2.1), with our 13 annotated scRNA-seq cell clusters used as the reference dataset. This approach extracts cell type-specific expression signatures and deconvolutes spatial transcriptomics data to generate cell abundance maps for all cell types, and all analyses were run with default parameters. Next, we identified malignant, boundary, and non-malignant regions using the boundary define function from the `Cottrazm` package (v0.1.1). Simultaneously, we used the `Scissor` algorithm to annotate malignant regions into Cluster 1 (low-risk), Cluster 2 (high-risk) and intermediate state, with all parameters kept consistent with

those used in the previous single-cell transcriptomics analysis. Additionally, we applied `CellTrek` (v0.0.94) to map single-cell information onto spatial transcriptomics sections. The input scRNA-seq data for CellTrek was our 13 annotated scRNA-seq cell types, with malignant cells pre-labeled as Cluster 1 (low-risk), Cluster 2 (high-risk) and intermediate state by the `Scissor` algorithm. Based on cell type annotations and spatial coordinates, we used the kdist function to calculate the spatial distance between immune cells and each of the three cell states (Cluster 1/low-risk, Cluster 2/high-risk, and intermediate state). Differences in cell abundance and spatial distance among Cluster 1 (low-risk), Cluster 2 (high-risk), and the intermediate state were assessed via pairwise Wilcoxon rank-sum tests.

Cell culture and establishment of stable cell lines

Human LUAD cell lines (A549 and PC9) and human embryonic kidney HEK293T cells were purchased from the American Type Culture Collection (ATCC), and cultured in Dulbecco's modified Eagle's medium (DMEM; Gibco) supplemented with 10% fetal bovine serum (FBS; Gibco) and 1% penicillin–streptomycin (P/S; Biosharp). THP-1 cells purchased from ATCC were cultured in RPMI 1640 medium supplemented with 10% fetal bovine serum, 1% penicillin-streptomycin, and 0.05 mM 2-mercaptoethanol. To differentiate THP-1 cells into macrophages, cells were cultured with phorbol 12-myristate 13-acetate (PMA; 100 ng/mL; MCE) for 48 h. All cells were maintained in a humidified 5% CO2 atmosphere at 37 °C.

Stable CST5 overexpression cell lines (PC9-CST5-OE and A549-CST5-OE) were generated via lentiviral transduction. Briefly, the full coding sequence of CST5 for overexpression was cloned into the plenti transfer vector. Lentiviral particles were then packaged in HEK293T cells by co-transfecting the recombinant CST5-OE plasmid along with packaging plasmids PLP1, PLP2, and VSVG. Target cells were transduced

with the harvested lentiviral particles, followed by puromycin selection (1 μg/mL) for 3 days.

Western blotting (WB)

Total proteins were extracted from cells using prechilled RIPA lysis buffer supplemented with 1% protease and phosphatase inhibitor cocktail. Protein concentration was quantified with a BCA protein assay kit (Thermo Scientific). Equal amounts of protein samples were separated via SDS-PAGE, then transferred to polyvinylidene fluoride membranes. After blocking with 5% milk in TBST buffer, membranes were incubated with primary antibodies overnight at 4 °C, followed by 1 h incubation with horseradish peroxidase (HRP)-conjugated secondary antibodies at room temperature. Finally, blots were imaged on the GelView 6000Plus system. Primary antibodies used were as follows: GAPDH (1:1000, CST Cat# 3700), CST5 (1:2000, Proteintech Cat# 30988-1-AP), AKT (1:1000, CST Cat# 4691), p-AKT (Ser473) (1:2000, CST Cat# 4060). Western blot stripping buffer (P0025) was obtained from Beyotime Biotechnology (Shanghai, China). To detect multiple proteins on the same membrane, we performed sequential antibody hybridization with stripping buffer between incubations, with all operations conducted to preserve target protein integrity.

Gene expression analysis by real-time quantitative PCR (RT-qPCR)

Total RNA was extracted using TRIzol reagent (Invitrogen, catalog no. 15596026) following the manufacturer's protocol. Extracted total RNA was then reverse-transcribed into complementary DNA using the PrimeScript™ RT Reagent Kit (catalog no. RR420A, Takara, China). RT-qPCR was subsequently performed with the Universal Blue qPCR SYBR Green Master Mix (Yeasen, China). All primer sequences used for mRNA expression quantification are provided in Table S1.

Cell proliferation

PC9 and A549 cells stably transduced with control lentivirus (empty vector control) or CST5-OE lentivirus were seeded into 96-well plates at a density of 1500 cells per well. At 24, 48, 72, and 96 h after seeding, CCK-8 reagent (Zeta, France) was added to each well, followed by 2 h of incubation at 37 °C in the dark. Absorbance at 450 nm was measured using a microplate reader.

Additionally, the plate colony formation assay was performed to evaluate the colony-forming ability of tumor cells. Our established stable cell lines were seeded into six-well plates at a density of 1000 cells per well. After 14 days of continuous culture, formed colonies were fixed and stained with crystal violet for quantification and assessment.

Transwell assay

Transwell migration assays were performed using 24-well transwell culture inserts. Briefly, 5 × 10^4 cells were resuspended in 200 μL serum-free DMEM medium, then seeded into the upper chamber of the transwell insert. For migration assays, the upper chamber was left uncoated (no Matrigel extracellular matrix was added, BD Biosciences, NJ, USA). Next, 600 μL of DMEM supplemented with 10% FBS was added to the lower chamber as a chemoattractant. Transwell chambers were incubated for 24h at 37 °C to allow cells to migrate through the porous membrane. After incubation, non-migrated cells remaining on the upper surface of the membrane were carefully removed by gentle wiping with a cotton swab. Cells that migrated to the lower surface of the membrane were fixed with 4% paraformaldehyde (Solarbio, Beijing, China) and stained with crystal violet for subsequent quantification.

Wound healing assay

Wound-healing migration assays were performed as follows: cells were seeded into six-well plates and cultured until they reached 90% confluence. A linear wound was

generated by scraping the confluent cell monolayer with a 100 μL pipette tip. Wound images were captured immediately after scratching (0 h, baseline) and then at 24 h, 36 h, and 48 h of incubation to evaluate cell migration capacity.

In Vitro Cell Co-Culture

To establish an in vitro co-culture system of macrophages and LUAD cells, we used 24-mm Transwell chambers with 4 μm pore polycarbonate membranes (Corning, USA). Briefly, THP-1 monocytes (1 × 10^6 cells per well) were seeded into the lower chamber, and treated with 100 ng/mL PMA to induce differentiation into M0 macrophages. PC9 and A549 cells (1 × 10^6 cells per well) were then added to the upper chamber, and co-cultured with the M0 macrophages for 24 h. After co-culture, total RNA was extracted from macrophages in the lower chamber, and the expression of canonical M1 marker (e.g., CD86) and canonical M2 marker (e.g., CD206) was detected by RT-qPCR.

In vivo tumor xenograft assays

Control PC9 cells and CST5-OE PC9 cells were prepared for subcutaneous inoculation: 5 × 10^6 cells per mouse were resuspended in 0.1 mL of a 1:1 (v/v) mixture of Matrigel and sterile saline, then injected subcutaneously into the groins of 6-week-old female BALB/c nude mice. All mice were purchased from Jiangsu Gempharmatech, with 5 mice per experimental group.

Statistics analysis

All statistical analyses were performed using R statistical software (version 4.4.3). Kaplan–Meier survival curves were plotted, and differences between survival groups were assessed via the log-rank test. Continuous variables were compared with the unpaired Student's t-test for normally distributed data, and the Wilcoxon rank-sum test for non-normally distributed data; categorical data were analyzed with the chi-square test or Fisher's exact test as appropriate. For multiple comparison adjustment, p-values

were corrected using the Benjamini–Hochberg procedure. Correlations between variables were calculated with Pearson or Spearman correlation. All tests were two-sided, and a p-value < 0.05 was considered statistically significant.

Results

Widespread metabolic pathway dysregulation defines two distinct metabolic subtypes in LUAD

To systematically characterize the link between metabolic pathway dysregulation and LUAD, we used ssGSEA to quantify the enrichment scores of all Reactome metabolic pathways in the TCGA-LUAD discovery cohort, then compared pathway activity between tumor and paired adjacent normal tissues. We found widespread significant metabolic pathway remodeling in LUAD (Fig. 1A). Specifically, bile acid/bile salt metabolism and heparan sulfate/heparin glycosaminoglycan metabolism were markedly upregulated in tumors, while glucose, nucleotide and polyamine metabolism (linked to biomacromolecule synthesis) were notably downregulated. These findings were validated in three independent GEO cohorts (Fig. 1A).

To accurately stratify the metabolic heterogeneity of LUAD patients, we performed unsupervised clustering based on ssGSEA-derived metabolic pathway enrichment scores. We first screened prognostic metabolic features in two independent datasets (TCGA-LUAD and batch-corrected merged GEO cohort), and identified 12 protective and 3 risk pathways consistently replicated by Kaplan–Meier analysis across both cohorts (Fig. 1B). Using these prognostic features, we stably stratified LUAD patients into two molecular subtypes with distinct metabolic profiles (Figs. 1C, D). Cluster 1 showed consistently higher activity across most profiled pathways, including phenylalanine, arachidonic acid and glucose metabolism, while Cluster 2 had selective

upregulation of steroid and steroid hormone metabolism (Fig. 1E). Prognostic validation in two independent cohorts confirmed that Cluster 1 was significantly associated with worse clinical outcomes than Cluster 2 (Fig. 1F, Fig. S1A). Collectively, LUAD exhibits widespread metabolic pathway remodeling, and we identified two prognosis-associated metabolic subtypes with fundamentally distinct metabolic landscapes.

Divergent LUAD metabolic subtypes display heterogeneous functional features

To characterize biological differences between the two metabolic subtypes, we performed GSEA on the differential expression profile between Cluster 1 and Cluster 2 (Fig. 2A). Cluster 1 was preferentially enriched for tumor proliferation and cell cycle progression pathways (Fig. 2B). In contrast, Cluster 2 showed significant enrichment of immune-related biological processes, consistent with its activated immune microenvironment and better prognosis (Fig. 1F).

Immune infiltration analysis confirmed that Cluster 2 contained higher proportions of multiple immune subsets, including B cells, T cells, mast cells, and dendritic cells (Fig. 2C). Cluster 2 also had elevated expression of multiple chemokines, including XCL1, which recruits cross-priming cDC1 dendritic cells to initiate adaptive immunity[19], along with CCL22, CXCL4, and CX3CL1 that promotes effector CD8+ T cell and NK cell retention via CX3CR1 binding[20, 21]. Multiple immune-related receptors and immunomodulatory factors, including CCR4, CCR2, and CX3CR1, were also consistently upregulated in this cluster (Fig. 2D).

We further quantified immunotherapy-related signature activity, and found Cluster 2 had significantly higher angiogenesis, stronger myeloid inflammatory responses, more prominent tertiary lymphoid structure (TLS) enrichment, and lower DNA damage

response (DDR) activity (Fig. 2E). Finally, integrated single-cell and spatial transcriptomic analyses confirmed that Cluster 2 is associated with both increased immune infiltration and significantly closer spatial proximity between tumor and immune cells (Figs. S2 and S3).

We next applied PROGENy to compare signaling pathway activity between the two metabolic subtypes, and found Cluster 2 exhibited higher p53 activity, while the pro-tumorigenic PI3K pathway was significantly downregulated (Fig. S1B). Mutation analysis showed Cluster 1 harbored higher TP53 mutation frequency, intratumoral heterogeneity and tumor mutation burden, although none of these differences reached statistical significance, while KRAS mutations were more frequent in Cluster 2 (Fig. S1C). The two subtypes also differed significantly in T-stage (tumor size, Fig. S1F). Collectively, our multi-omic analyses identified two metabolic subtypes with distinct traits, where Cluster 1 is characterized by enhanced proliferation and therapy resistance, while Cluster 2 has an activated anti-tumor immune microenvironment and better prognosis.

Development and validation of a prognostic model for LUAD

To explore the clinical value of metabolic subtype-related genes in LUAD, we first performed univariate Cox regression on DEGs from the TCGA-LUAD and GEOmerge cohorts, identifying 12 high-risk and 17 favorable prognostic genes after intersection (Fig. 3A, Fig. S4A). Within the TCGA-LUAD training cohort, we developed a robust predictive model by integrating 111 machine learning algorithms. The average C-index across the training and validation cohorts served as the criterion for model selection. Ultimately, the StepCox(forward)+ random survival forest algorithm emerged as the optimal model, and the final prognostic signature was constructed via random forest-

based importance quantification (Figs. 3B–C). The risk score effectively stratified patient prognosis across all validation cohorts (Fig. 3D). Patients in the high-risk score group exhibited significantly worse outcomes compared with those in the low-risk score group. Pooled analysis across five datasets confirmed that our model achieved a higher C-index than all individual clinical features, including age, stage, and key mutations (Fig. 4A). ROC curves further validated that risk scores reliably predicted LUAD patient prognosis across all datasets, with 1-, 3-, and 5-year AUC values generally exceeding 0.6 (Fig. 4B). Subsequently, we compared our risk score against 32 previously published LUAD signatures and found that it consistently demonstrated superior performance across all evaluated datasets (Fig. 4C). Collectively, our metabolic phenotype-associated prognostic signature (MPAPS) accurately predicts survival outcomes for LUAD patients.

MPAPS-defined risk groups of LUAD harbor distinct molecular and clinical characteristics

To identify DEGs between the two MPAPS-defined risk groups, we performed differential expression analysis (Fig. S4B), followed by GO functional annotation and KEGG pathway enrichment. GO analysis showed that DEGs were mainly enriched in extracellular matrix remodeling and cell proliferation-related processes (Fig. S4C), whereas KEGG analysis confirmed significant enrichment of the PI3K-Akt signaling pathway and cytokine-cytokine receptor interaction (Fig. S4D). PROGENy pathway analysis further revealed higher MAPK and PI3K activity and lower p53 and TRAIL activity in the high-MPAPS-risk group (Fig. S4E), which was validated by independent Hallmark and BioCarta gene set analyses (Fig. S4F). Somatic mutation analysis demonstrated that the high-risk group exhibited a higher TP53 mutation frequency,

increased intratumoral heterogeneity, and elevated tumor mutation burden (Figs. 5A–B; Fig. S4G). Prognostic analysis confirmed that high signature expression was significantly associated with increased mortality, advanced TNM stage, and a higher distant metastasis rate (Fig. S4H). In summary, the high-MPAPS-risk LUAD group is characterized by activated proliferation and oncogenic signaling, an elevated TP53 mutation burden, and is significantly associated with advanced TNM stage and inferior survival.

MPAPS-derived risk score in LUAD is associated with immune response

To characterize the immune landscape stratified by the MPAPS-derived risk score (hereinafter referred to as the risk score), we first analyzed the GSE161537 cohort, a LUAD immunotherapy dataset, and found that patients in the high-risk group had significantly worse overall survival (Fig. 5C). This prognostic association was further validated across two independent non-LUAD immunotherapy cohorts (melanoma: GSE91061; bladder cancer: IMvigor210), wherein higher risk scores were consistently correlated with poorer clinical outcomes (Figs. 5D–E). Notably, within the GSE91061 cohort, risk scores progressively increased with deteriorating immunotherapy response, with CR patients showing significantly lower scores than PD patients ($P = 5e\text{-}09$; Fig. 5F). TIDE analysis revealed that the high-risk group exhibited significantly higher levels of cancer-associated fibroblast (CAF) infiltration ($P = 0.013$), myeloid-derived suppressor cell (MDSC) accumulation ($P < 0.001$), and T cell exclusion ($P < 0.001$) compared with the low-risk group. Notably, T cell dysfunction scores were paradoxically lower in the high-risk group ($P < 0.001$), suggesting a complex immunosuppressive landscape. Collectively, these findings indicate that an elevated risk score is associated with a profoundly immunosuppressive tumor microenvironment

characterized by enhanced CAF infiltration, MDSC accumulation, and T cell exclusion (Fig. 5G). Furthermore, IPSs calculated using the TCIA website also suggested that the low MPAPS group is more appropriate for treatment with PD-1, CTLA-4, or their combination therapy (Fig. 5H). Analysis with ten immune infiltration algorithms consistently demonstrated that the high-risk group had higher tumor purity, lower overall immune scores, and reduced infiltration of most immune subsets, including B cells, CD4+ T cells, CD8+ T cells, and macrophages (Fig. S5A). Correlation analysis further confirmed the consistent negative associations between the risk score and the majority of immune cell populations (Figs. S5B and S7).

We next examined correlations between the risk score and chemokine expression and identified significant associations with multiple chemokines (Fig. S6A). The risk score was significantly negatively correlated with key T cell and myeloid cell chemotactic factors, including CX3CL1 and CX3CR1, and positively correlated with immunosuppressive factors, including TGFBR1 and IL-6 [22, 23] (Fig. S6B). Taken together, these results establish that an elevated risk score is intimately linked to a profoundly suppressed immune response, characterized by diminished productive immune infiltration, enhanced immunosuppressive signaling, and impaired immunotherapy efficacy.

Single-cell and Spatial Transcriptomic analyses reveal distinct immune landscapes across MPAPS-Defined risk groups

To dissect the molecular characteristics and immunomodulatory features of MPAPS-defined risk groups at single-cell resolution, we analyzed six LUAD scRNA-seq samples and integrated bulk-derived MPAPS signatures with single-cell transcriptomes using Scissor. We identified 3844 Scissor+ (low-risk) and 2277 Scissor- (high-risk)

tumor cells, with remaining unclassified cells designated as intermediate cells (Figs. 6A–D). GSEA confirmed clear functional differences between the two risk groups. The high-risk group was preferentially enriched for cell division pathways, whereas the low-risk group showed significant enrichment of B cell- and T cell-mediated immune responses (Fig. 6E). Overall, the functional profiles of these single-cell risk groups were highly consistent with those from bulk transcriptomes.

To further validate the reduced immune activation and infiltration in the high-risk group, we analyzed six LUAD spatial transcriptomics samples. After segmenting tissues into malignant, tumor boundary, and non-malignant compartments, we applied Scissor to integrate bulk features with spatial data, annotating malignant spots as low-risk (Scissor+), high-risk (Scissor-), or intermediate (remaining unannotated spots). Using RCTD, we estimated cell-type proportions per spatial spot and confirmed that high-risk malignant spots harbored significantly lower abundances of B cells, CD4+ T cells, and NK cells (Figs. 6F–I; Fig. S8A). CellTrek-based co-localization analysis further showed that all tested immune cells were consistently located at significantly greater distances from high-risk malignant spots (Figs. 7A–B, Fig. S8B; Figs. 7C–D, Fig. S9A). Additionally, AUCell scoring of risk-associated spatial spots confirmed that the distinct functional profiles of high- and low-risk groups were robustly recapitulated in most samples (Fig. 7E). Taken together, our integrated single-cell and spatial transcriptomic analyses demonstrate that the high-risk group has substantially reduced immune cell infiltration.

MPAPS-derived risk score predicts drug sensitivity in LUAD

To explore the clinical relevance of risk stratification, we analyzed the association between the risk score and the IC50 of multiple anti-tumor drugs and identified

consistent positive correlations for all tested agents, including uprosertib, epirubicin, afuresertib, and ribociclib (Fig. 7F). We further confirmed that the high-risk group had significantly higher IC50 values for uprosertib, ribociclib, and oxaliplatin, indicating reduced overall drug sensitivity (Fig. S9B). Collectively, these results indicate that the high-risk group exhibits an intrinsic multi-drug resistance phenotype.

CST5 suppresses LUAD progression via the PI3K/AKT/mTOR pathway

Among the signature genes included in our previously established LUAD metabolic risk model, several genes such as EXO1, CHEK1, and SCNN1B have been well-documented for their functional roles in LUAD development and progression[24-26]. However, the critical model gene CST5 remains poorly characterized. Although reduced CST5 expression has been linked to poor tumor differentiation in colorectal cancer [27], its prognostic value, immunomodulatory role, and underlying mechanisms in LUAD remain to be systematically explored. In this study, we first found that CST5 expression was significantly negatively correlated with metabolic risk score in LUAD (Fig. 8A). Consistently, CST5 expression was significantly lower in LUAD tumor tissues than in matched adjacent normal tissues (Fig. 8B). Survival analysis further confirmed that low CST5 expression was associated with worse overall survival in LUAD patients (Fig. 8C, Fig. S10A). To functionally validate the tumor-suppressive role of CST5 in LUAD, we generated stably overexpressing CST5 PC9 and A549 cell lines via lentiviral transduction, and verified overexpression efficiency at both mRNA and protein levels (Figs. 8D, E). Functional assays including CCK8, colony formation, Transwell, and wound healing revealed that CST5 overexpression significantly impaired the proliferative, colony-forming, and migratory capacities of LUAD cells (Figs. 8F-I). Collectively, Collectively, these results confirm the tumor-suppressive

function of CST5 in LUAD progression.

To further explore the downstream mechanism of CST5, we performed enrichment analysis on differentially expressed genes between CST5 high- and low-expression groups using integrated TCGA and GEO datasets, which revealed significant enrichment of the PI3K/AKT/mTOR pathway (Fig. 9A, Figs. S10B and S10C). Consistent with bioinformatic predictions, Western blotting in control and CST5-overexpressing PC9/A549 cells confirmed that CST5 overexpression substantially reduced p-AKT and p-mTOR levels and suppressed PI3K/AKT/mTOR pathway activity (Fig. 9B). Furthermore, CST5 overexpression activated the apoptotic pathway, as evidenced by downregulation of the anti-apoptotic protein BCL-2, and upregulation of the pro-apoptotic protein Bax and cleaved caspase-3, a key activated effector of apoptosis (Figs. 9C-E) . To confirm these findings in vivo, we validated the tumor-suppressive role of CST5 using a subcutaneous xenograft model in BALB/c nude mice. After injecting control or CST5-overexpressing PC9 cells into nude mice, we found that CST5 overexpression significantly reduced final tumor weight (Figs. 9F, G). Collectively, based on in vitro and in vivo evidence, we demonstrated that CST5 acts as a prognostic tumor suppressor in LUAD, and its low expression is consistently associated with unfavorable clinical outcomes.

CST5 remodels the tumor immune microenvironment and promotes M1 macrophage polarization in LUAD

We next explored the relationship between CST5 expression and the tumor immune microenvironment in LUAD. Correlation analysis showed that CST5 expression was positively correlated with stromal score, immune score, and ESTIMATE score (Fig. 10A, Figs. S10D and S10E), indicating higher immune infiltration in the CST5 high-

expression group. We further compared immune-related gene expression between groups, we found CST5 was positively correlated with most chemokines, receptors, MHC molecules, and immune co-stimulatory/co-inhibitory molecules (Fig. 10B). Consistent with the above findings, quanTIseq deconvolution revealed that the proportion of M1 macrophages was significantly higher in the CST5 high-expression group (Fig. 10C). Correlation analysis further confirmed that CST5 expression was positively correlated with the M1 macrophage signature ($r = 0.209$, $P = 1.22\times10^{-6}$) and negatively correlated with the pro-tumor M2 macrophage signature ($r = -0.277$, $P = 9.32\times10^{-11}$; Fig. 10D). Finally, we verified this finding through a co-culture assay. We induced THP-1 monocytes to differentiate into M0 macrophages, then co-cultured them with empty vector control and CST5-overexpressing PC9/A549 cells. qRT-PCR confirmed that M1 macrophage markers were significantly upregulated in macrophages co-cultured with CST5-overexpressing cells, while M2 macrophage markers were downregulated (Figs. 10E, F). Collectively, these results indicated that CST5 could promote M1 macrophage polarization and remodel the immune microenvironment of LUAD.

Discussion

Metabolic pathways are fundamental to cellular homeostasis, and centrally govern cell proliferation, development, and adaptation to microenvironmental perturbations[5]. Beyond supporting cell growth, metabolic pathways provide the energy and biosynthetic backbone for immune cell function, and directly shape anti-tumor immune responses via bioactive metabolites[28]. In this study, we systematically profiled all MSigDB-annotated metabolic pathways in LUAD and identified two robust metabolic subtypes with divergent immune landscapes. Cluster 1 was characterized by enhanced

oncogenic proliferation, elevated intratumoral heterogeneity, and hyperactivated PI3K signaling, converging to drive worse clinical outcomes. Conversely, Cluster 2 exhibited dominant immune-related biological processes, increased immune cell infiltration, and upregulated p53 tumor suppressor activity, corresponding to a favorable prognosis. These findings underscore that metabolic heterogeneity in LUAD is not merely a reflection of bioenergetic adaptation, but a determinant of immune contexture and clinical behavior.

To translate these metabolic-immune insights into clinically actionable tools, we constructed a metabolic phenotype-associated prognostic signature MPAPS by integrating machine learning algorithms. MPAPS demonstrated robust predictive performance across multiple independent cohorts, achieving a higher concordance index than individual clinical variables and outperforming 32 previously published LUAD prognostic signatures. Notably, the risk score remained independently associated with unfavorable prognosis after adjusting for clinical tumor stage, indicating that metabolic heterogeneity captures prognostic information beyond conventional clinicopathological parameters.

Immune profiling revealed that patients in the high-MPAPS-risk group harbored a profoundly immunosuppressive TME, characterized by elevated myeloid-derived suppressor cell accumulation, increased cancer-associated fibroblast infiltration, and T cell exclusion, alongside reduced tertiary lymphoid structure activity. These features collectively suggest compromised adaptive immune surveillance in the high-risk group. Regarding immunotherapy response, validation within the LUAD immunotherapy cohort confirmed that patients with elevated risk scores experienced significantly worse overall survival. While the mechanistic relevance of metabolic-immune crosstalk to checkpoint blockade efficacy is supported by our LUAD data, we acknowledge that the

generalizability of MPAPS to other malignancies requires further prospective validation before any cross-cancer application can be asserted.

Recent advances in single-cell and spatial transcriptomics have greatly improved our understanding of spatial co-localization and reciprocal crosstalk between tumor and immune cell populations [29]. To dissect the molecular characteristics of MPAPS-defined risk groups at single-cell resolution, we employed the Scissor algorithm and identified that high-risk tumor cells were functionally enriched for cell division programs, whereas low-risk cells exhibited activated B cell- and T cell-mediated immune responses. More importantly, spatial transcriptomic analysis provided critical topological evidence that tumor regions classified as high-risk not only harbored significantly lower abundances of B cells, CD4+ T cells, and NK cells, but also featured immune cells localized at substantially greater spatial distances from malignant spots. This spatial segregation suggests that the immunosuppressive microenvironment in high-risk LUAD is not solely attributable to reduced immune cell recruitment, but also to physical exclusion and disrupted tumor-immune spatial organization. Such spatial decoupling may impair antigen presentation and effector cell engagement, representing a structural basis for immune evasion that cannot be fully captured by bulk deconvolution methods alone. These integrative findings reinforce that metabolic risk stratification reflects genuine TME architectural differences with functional consequences for anti-tumor immunity.

Drug sensitivity analysis additionally revealed that higher risk scores correlated with increased IC50 values for multiple antitumor agents, including uprosertib and ribociclib. We interpret this association primarily as an extension of the aggressive biological phenotype captured by MPAPS rather than a direct basis for treatment selection, and thus consider these findings as complementary to the prognostic stratification

framework.

Among the signature genes, CST5 emerged as a key protective factor whose functional role in LUAD had remained poorly characterized. Cystatin D, the core prognostic protein encoded by CST5 in our risk model, is an endogenous inhibitor of cysteine proteases, including the well-characterized cathepsins S, H, and L [30]. The active vitamin D metabolite 1α,25-dihydroxyvitamin D_3 (1α,25$(OH)_2D_3$) and its analogs have emerged as promising preventive and therapeutic anticancer agents, attracting growing research interest, with multiple ongoing clinical trials evaluating their clinical efficacy [31]. Prior work in colorectal cancer demonstrated that calcitriol-induced upregulation of CST5 suppresses proliferation, migration, anchorage-independent growth, and xenograft tumor formation, and that CST5 potentiates the anti-tumor efficacy of 1α,25$(OH)_2D_3$[27]. Our bioinformatic and experimental analyses now establish CST5 as a tumor suppressor in LUAD that concurrently inhibits malignant progression and remodels anti-tumor immunity. Mechanistically, CST5 overexpression significantly attenuated LUAD cell proliferation, colony formation, and migratory capacity in vitro, and suppressed subcutaneous xenograft growth in vivo. These anti-tumorigenic effects were mediated, at least in part, through inhibition of the PI3K/AKT/mTOR oncogenic axis and activation of apoptotic signaling, as evidenced by reduced p-AKT and p-mTOR levels alongside downregulation of BCL-2 and upregulation of Bax and cleaved caspase-3. Simultaneously, CST5 exhibited significant immunomodulatory capacity: CST5-high tumors displayed elevated immune scores and increased M1 macrophage infiltration, while in vitro co-culture assays confirmed that CST5-overexpressing LUAD cells promoted a phenotypic shift from M0 toward M1 macrophages and suppressed M2 polarization. Thus, CST5 appears to function as a dual-action metabolic-immune node that constrains tumor cell-intrinsic malignancy while fostering

a pro-inflammatory microenvironment. The integration of these tumor cell-autonomous and non-cell-autonomous effects positions CST5 as a particularly attractive candidate for therapeutic intervention, though the direct molecular mechanism by which CST5 regulates PI3K/AKT/mTOR signaling—such as potential interactions with specific receptors, phosphatases, or upstream regulators—remains to be fully elucidated.

This study has several limitations that warrant acknowledgment. First, all transcriptomic data analyzed herein were obtained from retrospective public datasets; consequently, prospective multi-center clinical cohorts are needed to further validate the clinical utility and generalizability of MPAPS. Second, while we demonstrated that CST5 suppresses PI3K/AKT/mTOR signaling and modulates macrophage polarization, the precise direct molecular interactions underlying these effects remain undefined and require mechanistic dissection in future studies. Third, although our spatial and single-cell analyses provide compelling associative evidence, orthogonal experimental approaches such as multiplex immunohistochemistry and flow cytometry are needed to directly validate the observed immune spatial distributions. Fourth, as an emerging technology, spatial metabolomics enables profiling of cellular metabolic states across distinct intratumoral regions [32], and will deepen our understanding of heterogeneous metabolism within the TME; however, its integration with our current framework awaits future investigation. Finally, whether targeting CST5 or its downstream signaling can enhance anti-tumor immunity and sensitize LUAD to immunotherapy remains to be tested in preclinical immunocompetent models.

In conclusion, we developed a robust machine learning-based metabolic prognostic model for lung adenocarcinoma that effectively stratifies patients into subgroups with divergent survival outcomes, distinct tumor microenvironment profiles, and differential immunotherapy responses. Our integrative single-cell and spatial transcriptomic

analyses further demonstrate that metabolic risk stratification reflects fundamental differences in tumor-immune spatial organization. We functionally validated CST5 as a key protective metabolic driver that suppresses LUAD malignancy via PI3K/AKT/mTOR pathway inhibition and promotes anti-tumor M1 macrophage polarization. While MPAPS holds promise as a clinically useful prognostic tool, its prospective validation is essential prior to clinical implementation. Similarly, CST5 represents a promising actionable target for personalized LUAD management, yet further investigation into its precise molecular mechanisms and therapeutic applicability in immunocompetent settings is warranted.

Ethical statement and consent to participate

All animal experiments were conducted in accordance with institutional, national, and international guidelines for the care and use of laboratory animals, including the Basel Declaration. Animal studies and procedures were carried out in accordance with an approved protocol from the Institutional Animal Care and Use Committee (IACUC) of Tianjin Medical University Cancer Institute and Hospital (approval number: IACUC-2025029). No anesthetic agent was used in this study. Euthanasia was performed by carbon dioxide (CO_2) inhalation. This study does not involve direct participation of human subjects or individuals, and the ethics committee has approved the study protocol. Consent to participate: Not applicable. This study did not involve direct patient contact, intervention, or collection of identifiable human data. Clinical trial number: Not applicable.

Consent for publication

Not applicable. All data used in this work are publicly available, and no individual

person's identifiable information is presented. All authors confirm that this manuscript has not been published elsewhere, is not under consideration for publication in any other journal, and that its publication has been approved by all co-authors as well as the responsible authorities at the institutions where the work was carried out.

Author contributions

H.L., P.Z. and Y.S.: investigation, methodology, validation, original draft preparation, bioinformatics analysis, editing, and revision of the manuscript. X.H., J.Y., Y.J., J.C., Y.L. and L.T.: bioinformatics analysis. T.Q.: Supervision, design, editing, and revision of the manuscript. All authors have approved the final version of this manuscript. H.L., P.Z. and Y.S. contributed equally to this manuscript and should be considered as co-first authors.

Conflict of Interest

Conflict of interest relevant to this article was not reported.

Funding

This work was supported in part by grants from National Natural Science Foundation of China [Grant No.82573471 and Grant No.82203628].

Data availability

All data supporting the findings of this study are provided in the article and its Supplementary Materials. This study did not generate new original public datasets. The gene expression and clinical data of lung adenocarcinoma analyzed in this study are

publicly available from the TCGA-LUAD cohort (https://portal.gdc.cancer.gov/) and GEO (https://www.ncbi.nlm.nih.gov/geo/), with the following GEO accession numbers: GSE13213, GSE427, GSE30219, GSE31210, GSE189487, GSE131907, GSE189357. The 75 metabolic pathway gene sets were obtained from MSigDB (https://www.gsea-msigdb.org/gsea/msigdb/). This study did not develop any novel algorithms. All custom analysis code generated in this work can be obtained from the corresponding.

References

1. Siegel RL, Kratzer TB, Giaquinto AN, Sung H, Jemal A. Cancer statistics, 2025. CA Cancer J Clin. 2025;75(1):10–45. doi: 10.3322/caac.21871.
2. Noguchi M, Morikawa A, Kawasaki M, Matsuno Y, Yamada T, Hirohashi S, et al. Small adenocarcinoma of the lung. Histologic characteristics and prognosis. Cancer. 1995;75(12):2844–52. doi: 10.1002/1097-0142(19950615)75:12<2844::aid-cncr2820751209>3.0.co;2-#.
3. Denisenko TV, Budkevich IN, Zhivotovsky B. Cell death-based treatment of lung adenocarcinoma. Cell Death Dis. 2018;9(2):117. doi: 10.1038/s41419-017-0063-y.
4. Allemani C, Matsuda T, Di Carlo V, Harewood R, Matz M, Nikšić M, et al. Global surveillance of trends in cancer survival 2000-14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. Lancet. 2018;391(10125):1023–75. doi: 10.1016/s0140-6736(17)33326-3.
5. Hanahan D. Hallmarks of Cancer: New Dimensions. Cancer Discov. 2022;12(1):31–46. doi: 10.1158/2159-8290.Cd-21-1059.
6. Keibler MA, Wasylenko TM, Kelleher JK, Iliopoulos O, Vander Heiden MG, Stephanopoulos G. Metabolic requirements for cancer cell proliferation. Cancer Metab. 2016;4:16. doi: 10.1186/s40170-016-0156-6.
7. Pavlova NN, Thompson CB. The Emerging Hallmarks of Cancer Metabolism. Cell Metab. 2016;23(1):27–47. doi: 10.1016/j.cmet.2015.12.006.
8. Kimmelman AC, White E. Autophagy and Tumor Metabolism. Cell Metab. 2017;25(5):1037–43. doi: 10.1016/j.cmet.2017.04.004.
9. Sullivan MR, Danai LV, Lewis CA, Chan SH, Gui DY, Kunchok T, et al. Quantification of microenvironmental metabolites in murine cancers reveals determinants of tumor nutrient availability. Elife. 2019;8. doi: 10.7554/eLife.44235.
10. Liu J, Zhang C, Hu W, Feng Z. Tumor suppressor p53 and its mutants in cancer metabolism. Cancer Lett. 2015;356(2 Pt A):197–203. doi: 10.1016/j.canlet.2013.12.025.
11. Huang Y, Wu G, Bi G, Cheng L, Liang J, Li M, et al. Unveiling chemotherapy-induced immune landscape remodeling and metabolic reprogramming in lung adenocarcinoma by scRNA-sequencing. Elife. 2024;13. doi: 10.7554/eLife.95988.
12. Xia L, Oyang L, Lin J, Tan S, Han Y, Wu N, et al. The cancer metabolic reprogramming and immune response. Mol Cancer. 2021;20(1):28. doi: 10.1186/s12943-021-01316-8.

13. Park J, Hsueh PC, Li Z, Ho PC. Microenvironment-driven metabolic adaptations guiding CD8(+) T cell anti-tumor immunity. Immunity. 2023;56(1):32–42. doi: 10.1016/j.immuni.2022.12.008.
14. Fang JH, Chen JY, Zheng JL, Zeng HX, Chen JG, Wu CH, et al. Fructose Metabolism in Tumor Endothelial Cells Promotes Angiogenesis by Activating AMPK Signaling and Mitochondrial Respiration. Cancer Res. 2023;83(8):1249–63. doi: 10.1158/0008-5472.Can-22-1844.
15. Zhang Z, Zhai M, Bao S, Sun X, Chen R, Wang B, et al. Integrative multi-omics profiling deciphers tumor microenvironment heterogeneity and immunotherapy vulnerabilities in lung neuroendocrine carcinomas. J Adv Res. 2026;81:797–808. doi: 10.1016/j.jare.2025.06.017.
16. Williams CG, Lee HJ, Asatsuma T, Vento-Tormo R, Haque A. An introduction to spatial transcriptomics for biomedical research. Genome Med. 2022;14(1):68. doi: 10.1186/s13073-022-01075-1.
17. Guo W, Zhou B, Bie F, Huai Q, Xue X, Guo L, et al. Single-cell RNA sequencing analysis reveals transcriptional heterogeneity of multiple primary lung cancer. Clin Transl Med. 2023;13(10):e1453. doi: 10.1002/ctm2.1453.
18. Wang X, Tokheim C, Gu SS, Wang B, Tang Q, Li Y, et al. In vivo CRISPR screens identify the E3 ligase Cop1 as a modulator of macrophage infiltration and cancer immunotherapy target. Cell. 2021;184(21):5357–74.e22. doi: 10.1016/j.cell.2021.09.006.
19. Zhang K, Wuri Q, Cai Z, Qu X, Zhang S, Wu H, et al. The XCL1-Mediated DNA Vaccine Targeting Type 1 Conventional Dendritic Cells Combined with Gemcitabine and Anti-PD1 Antibody Induces Potent Antitumor Immunity in a Mouse Lung Cancer Model. Int J Mol Sci. 2024;25(3). doi: 10.3390/ijms25031880.
20. Yang Z, Wang X, Fu Y, Wu W, Hu Z, Lin Q, et al. YTHDF2 in peritumoral hepatocytes mediates chemotherapy-induced antitumor immune responses through CX3CL1-mediated CD8(+) T cell recruitment. Mol Cancer. 2024;23(1):186. doi: 10.1186/s12943-024-02097-6.
21. Robinson LA, Nataraj C, Thomas DW, Cosby JM, Griffiths R, Bautch VL, et al. The chemokine CX3CL1 regulates NK cell activity in vivo. Cell Immunol. 2003;225(2):122–30. doi: 10.1016/j.cellimm.2003.09.010.
22. Tong Y, Wang Z, Qiao N, Huang R, Wu C, Wang H, et al. Decoding adipocyte heterogeneity through single-nucleus transcriptomics unveils subtype-specific adipocytes orchestrate immunosuppressive niches in breast cancer. J Immunother Cancer. 2026;14(1). doi: 10.1136/jitc-2025-012711.
23. Jeong H, Koh J, Kim S, Yim J, Song SG, Kim H, et al. Cell-intrinsic PD-L1 signaling drives immunosuppression by myeloid-derived suppressor cells through IL-6/Jak/Stat3 in PD-L1-high lung cancer. J Immunother Cancer. 2025;13(3). doi: 10.1136/jitc-2024-010612.
24. Li Z, Lin X, Yang Y, Tian M, Zhang L, Huang F, et al. EXO1 is a key gene for lung-resident memory T cells and has diagnostic and predictive values for lung adenocarcinoma. Sci Rep. 2025;15(1):4002. doi: 10.1038/s41598-025-88126-w.
25. Tan Z, Chen M, Wang Y, Peng F, Zhu X, Li X, et al. CHEK1: a hub gene related to poor prognosis for lung adenocarcinoma. Biomark Med. 2022;16(2):83–100. doi: 10.2217/bmm-2021-0919.
26. Lu A, Shi Y, Liu Y, Lin J, Zhang H, Guo Y, et al. Integrative analyses identified ion channel genes GJB2 and SCNN1B as prognostic biomarkers and therapeutic targets for lung adenocarcinoma. Lung Cancer. 2021;158:29–39. doi: 10.1016/j.lungcan.2021.06.001.
27. Alvarez-Díaz S, Valle N, García JM, Peña C, Freije JM, Quesada V, et al. Cystatin D is a candidate tumor suppressor gene induced by vitamin D in human colon cancer cells. J Clin Invest. 2009;119(8):2343–58. doi: 10.1172/jci37205.

28. Martínez-Reyes I, Chandel NS. Cancer metabolism: looking forward. Nat Rev Cancer. 2021;21(10):669–80. doi: 10.1038/s41568-021-00378-6.
29. Yan Y, Sun D, Hu J, Chen Y, Sun L, Yu H, et al. Multi-omic profiling highlights factors associated with resistance to immuno-chemotherapy in non-small-cell lung cancer. Nat Genet. 2025;57(1):126–39. doi: 10.1038/s41588-024-01998-y.
30. Freije JP, Abrahamson M, Olafsson I, Velasco G, Grubb A, López-Otín C. Structure and expression of the gene encoding cystatin D, a novel human cysteine proteinase inhibitor. J Biol Chem. 1991;266(30):20538–43.
31. Krishnan AV, Feldman D. Mechanisms of the anti-cancer and anti-inflammatory actions of vitamin D. Annu Rev Pharmacol Toxicol. 2011;51:311–36. doi: 10.1146/annurev-pharmtox-010510-100611.
32. Alexandrov T. Spatial metabolomics: from a niche field towards a driver of innovation. Nat Metab. 2023;5(9):1443–5. doi: 10.1038/s42255-023-00881-0.

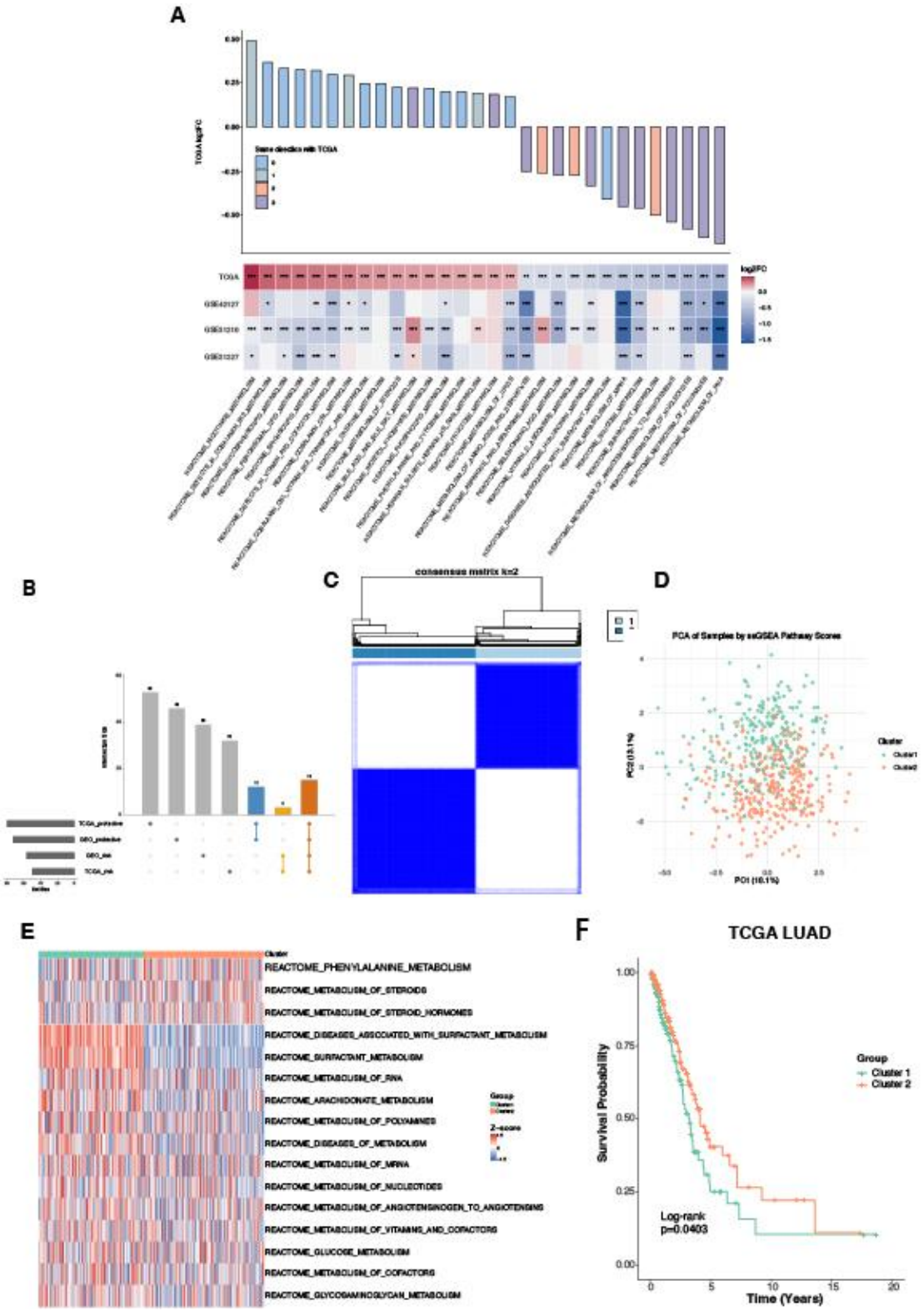


Fig. 1. Metabolic remodeling and subtype identification in LUAD. (A) The upper lollipop plot shows the top 30 altered metabolic pathways in TCGA-LUAD, with bar color indicating the number of validation datasets sharing a concordant regulatory

direction. The lower heatmap displays pathway remodeling across four independent LUAD datasets, colored by log2FC. (B) Overlap of prognostic protective and risk factors from Kaplan-Meier analysis in TCGA-LUAD and a merged GEO dataset. (C) Consensus clustering heatmap. (D) 2D PCA plot distinguishing the two metabolic subtypes. (E) Heatmap of significantly upregulated metabolic pathways within each subtype. (F) Kaplan-Meier survival curves for Cluster 1 and Cluster 2. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

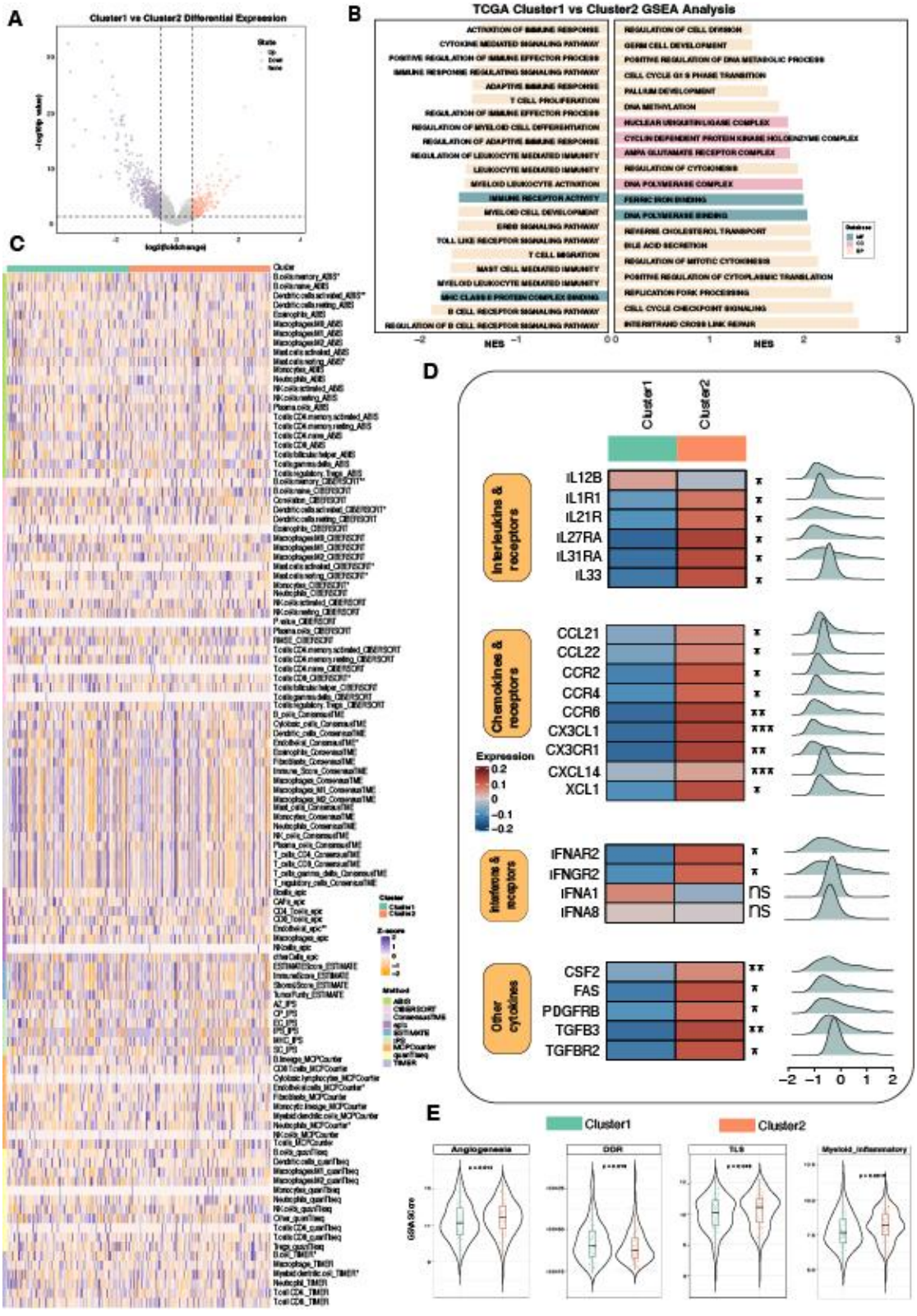
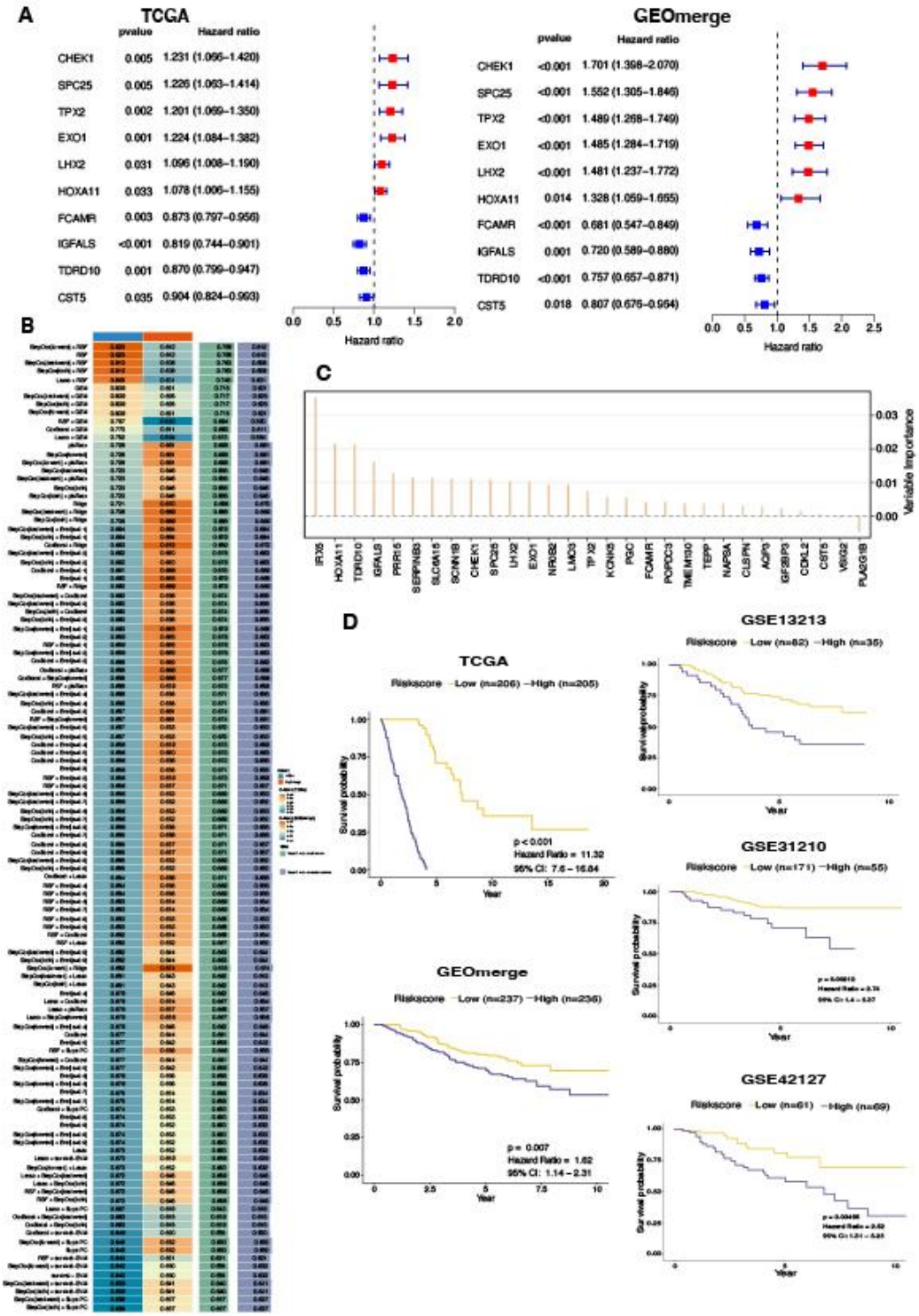


Fig. 2. Multi-dimensional characterization of two metabolic subtypes in LUAD. (A) Volcano plot of differentially expressed genes between the two metabolic subtypes. (B) Comparison of functional features between Cluster 1 and Cluster 2. (C) Correlation

heatmap between metabolic subtypes and tumor-infiltrating immune cells. (D) Expression analysis of immune-related genes between the two groups. (E) Expression differences of antigen presentation molecules, chemokines, and their receptors between Cluster 1 and Cluster 2. All data are expressed as mean±standard deviation, *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.



836
837
838 Fig. 3. Development and validation of a metabolic risk model in LUAD. (A) Univariate
839 Cox regression on TCGA and a merged GEO dataset identified high-risk ($p<0.05$,
840 $HR>1$) and favorable ($p<0.05$, $HR<1$) prognostic genes. (B) C-indexes of prognostic
841 models constructed by 10 individual machine learning algorithms and their
842 combinations in the TCGA training set. (C) Prognostic risk scoring model built based

on gene variable importance. (D) MPAPS effectively stratifies prognosis in LUAD patients across five independent datasets, with the low MPAPS group showing significantly better clinical outcomes.

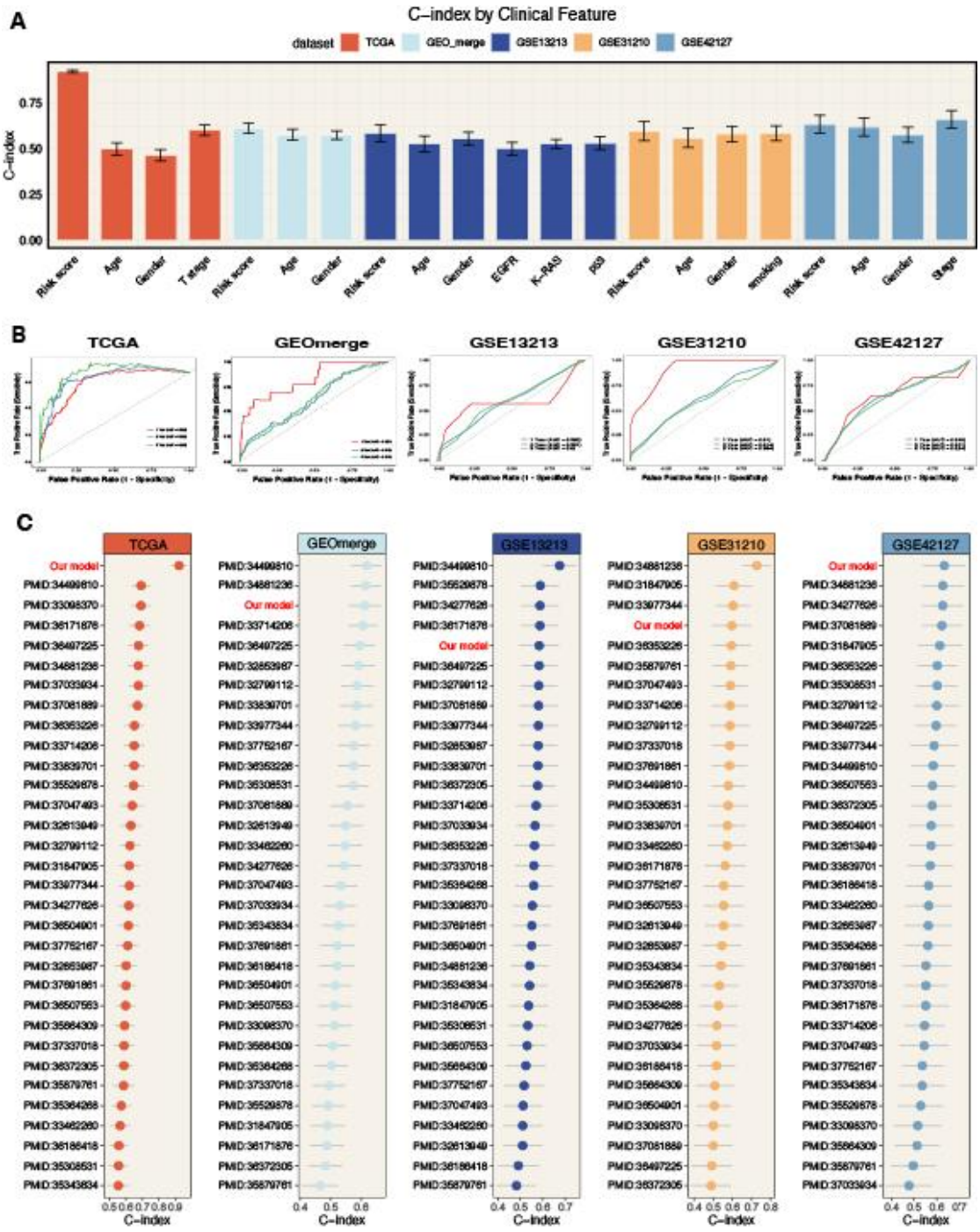


Fig. 4. Prognostic comparison of our model with clinical factors and existing models.

(A) Comparison of the metabolic risk model with other clinical indicators for prognostic prediction. (B) ROC curves evaluating the prognostic prediction performance of MPAPS across five independent datasets. (C) Comparison of our model with published LUAD-related prognostic signatures demonstrating its superior prognostic capability.

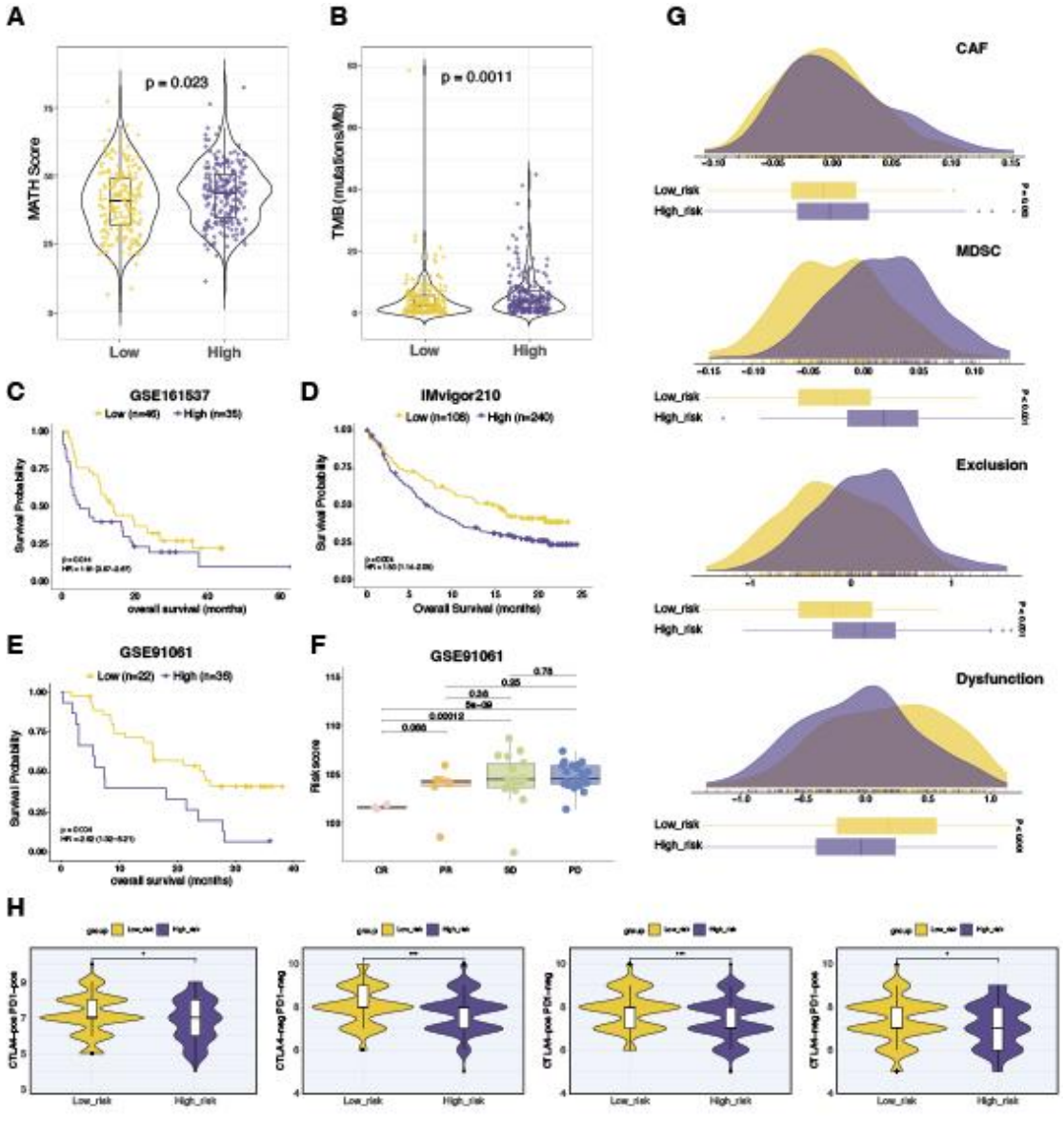


Fig. 5. Analysis of the metabolic risk model in LUAD and its implications for immunotherapy response. (A) Bar chart summarizing mutation frequencies between two subgroups. (B) Heterogeneity index scores of high- and low-metabolic risk

subgroups. (C-E) Kaplan–Meier survival analysis of low- and high-metabolic-risk patients across three cohorts demonstrating distinct survival outcomes. (F) Box plot comparing metabolic risk scores across different immunotherapy response categories in the GSE91061 cohort. Metabolic risk score differences across clinical subgroups in LUAD. (G) TIDE analysis of immunotherapy response differences between two risk groups. (H) IPS prediction of TCGA-LUAD patients from The Cancer Immunome Atlas, consistently revealing higher IPS and increased sensitivity to immunotherapy in the low MPAPS group. CR, complete response; PR, partial response; SD, stable disease; PD, progressive disease; CAF, cancer-associated fibroblast; MDSC, myeloid-derived suppressor cell; IPS, immunophenoscore. All data are expressed as mean±standard deviation, *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

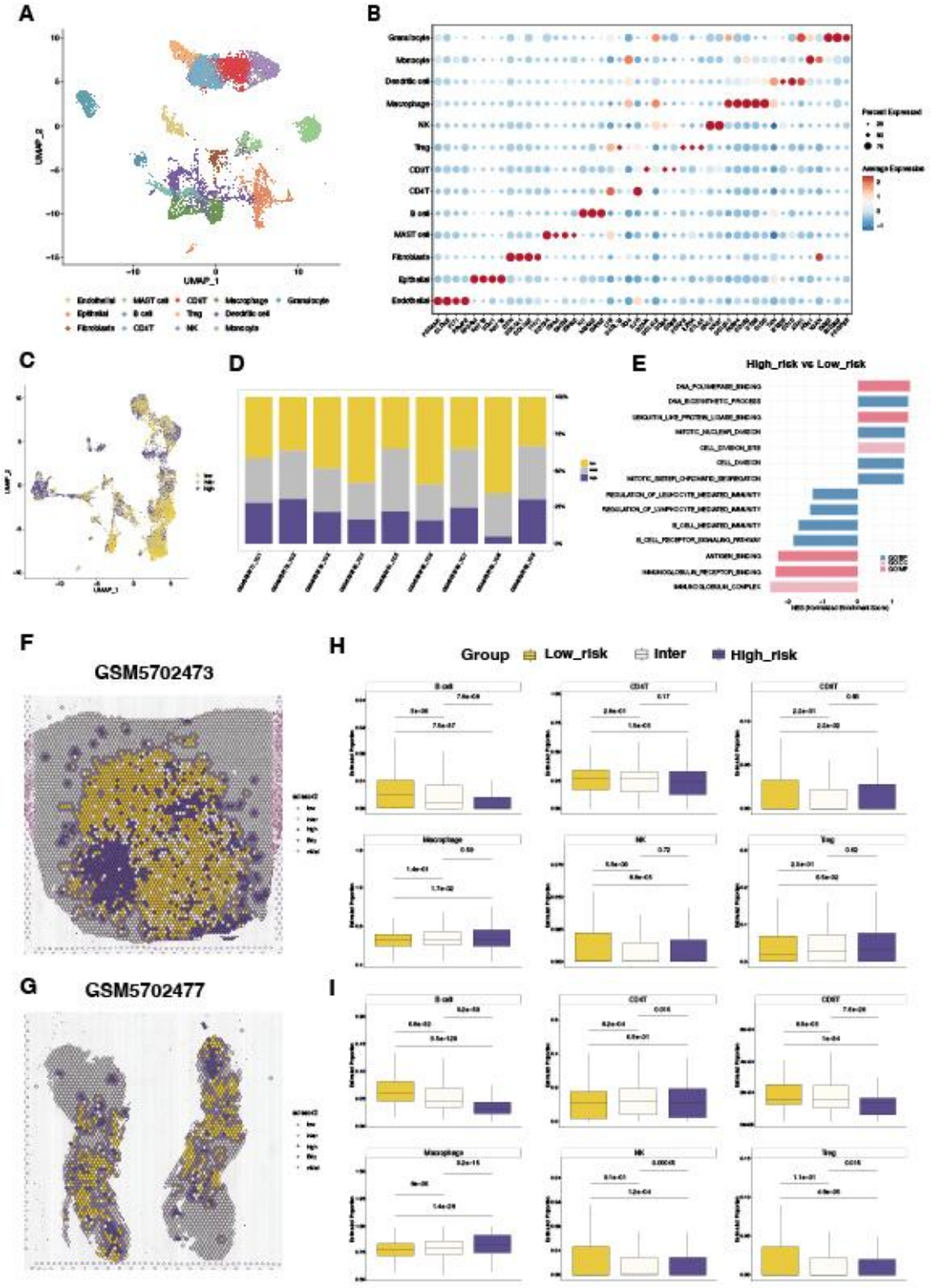


Fig. 6. Single-cell and spatial characterization of LUAD metabolic risk subgroups. (A) UMAP plot of annotated cell types in LUAD single-cell data. (B) Bubble plot of cell type-specific marker expression in annotated LUAD single-cell clusters. (C) UMAP of Scissor-selected tumor cells by metabolic risk group. (D) Stacked bar plot of risk

subgroup proportion across different samples. (E) Specifically upregulated functions in low- and high-metabolic risk subgroups. (F, G) Spatial distribution of risk-stratified malignant, tumor boundary and non-malignant regions across two LUAD spatial transcriptomic samples. (H, I) Box plots showing infiltration differences of six immune cell subsets across three metabolic risk groups.

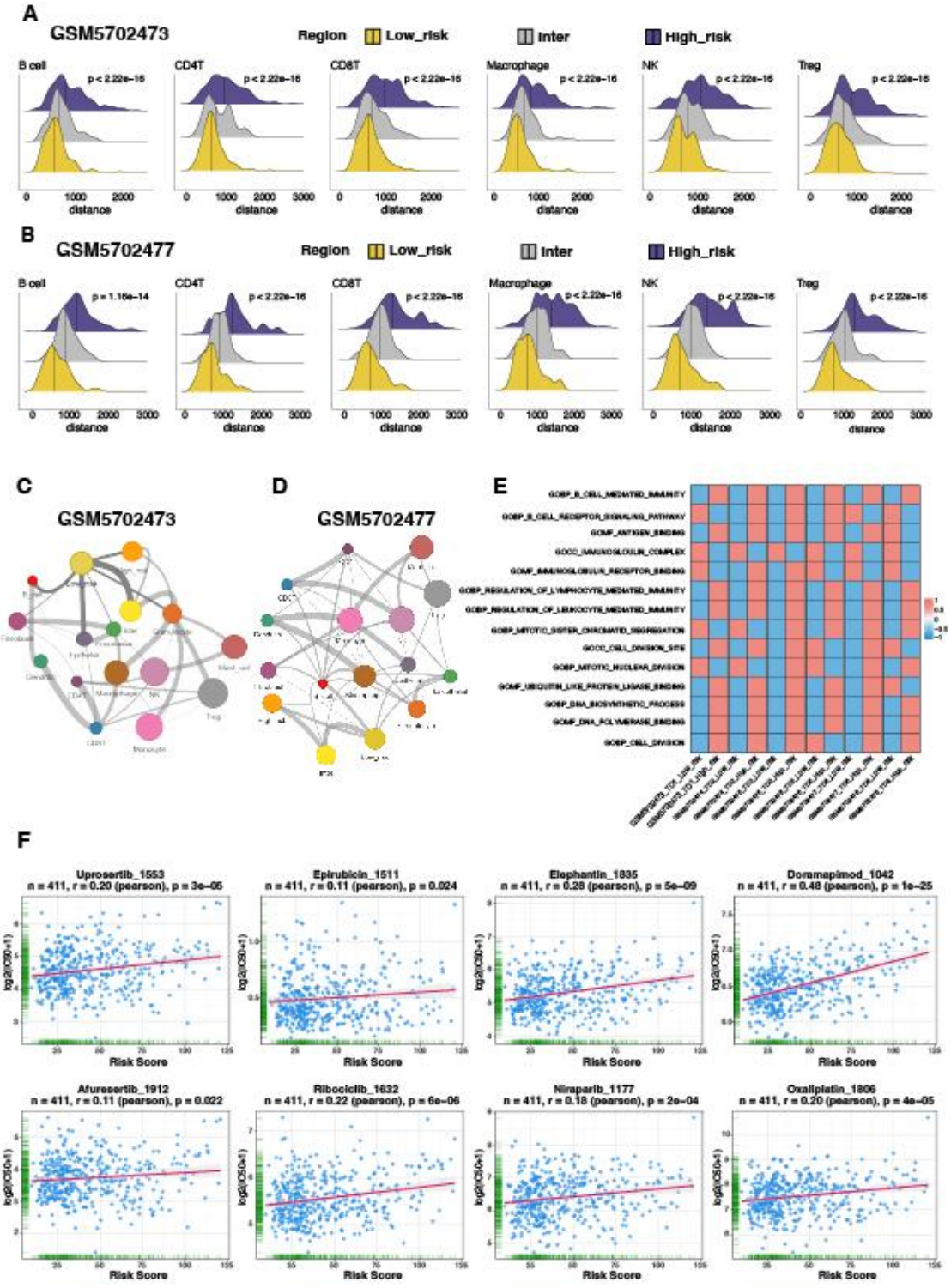


Fig. 7. Spatial immune infiltration analysis and drug sensitivity prediction for LUAD metabolic risk subgroups. (A, B) CellTrek-derived spatial K-distance between immune cells and risk-stratified tumor cells (low-risk, intermediate-risk, high-risk). Statistical significance shown in the ridge plots corresponded to the low-risk vs. high-risk

comparison. (C, D) Spatial colocalization plots of cell types across two samples. (E) Functional characteristics of spatial spots in low-risk and high-risk metabolic subgroups. (F) Correlation between metabolic risk score and predicted anticancer drug IC50 values.

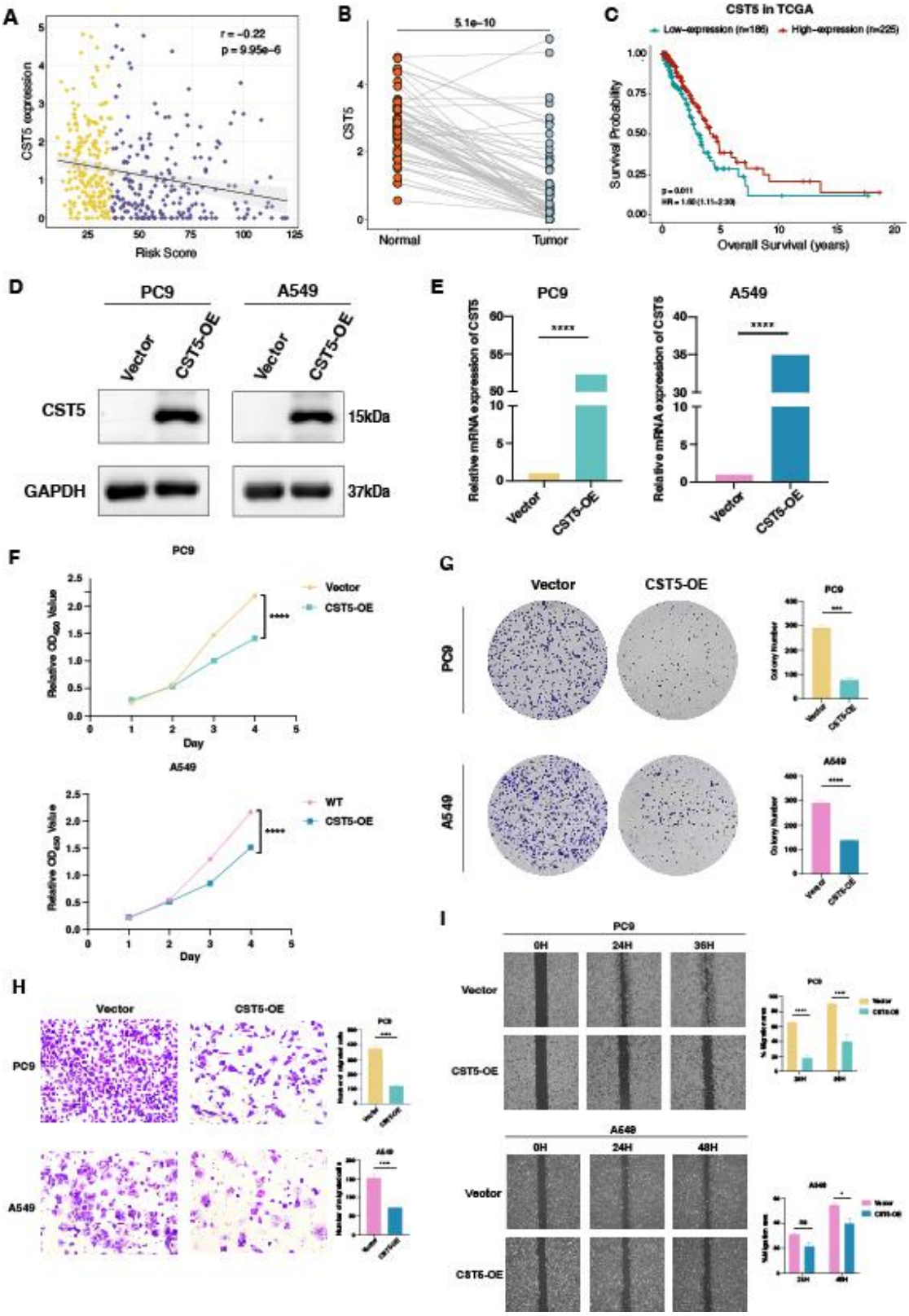


Fig. 8. Clinical prognostic and functional analysis of CST5 in LUAD. (A) Correlation analysis of CST5 expression and metabolic risk score. (B) Comparison of CST5 expression between normal and LUAD tumor tissues. (C) Kaplan-Meier survival curves

of overall survival grouped by CST5 expression in TCGA-LUAD cohort. (D, E) WB and qRT-PCR validated the overexpression efficiency of CST5 at the protein and mRNA levels, respectively. (F, G) CCK-8 and colony formation assays comparing proliferation between wild-type and CST5-overexpressing PC9 and A549 cell lines. (H, I) Transwell and wound healing assays detected invasive and migratory abilities of wild-type and CST5-overexpressing PC9 and A549 cell lines. Data represents mean± standard error of mean, *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$.

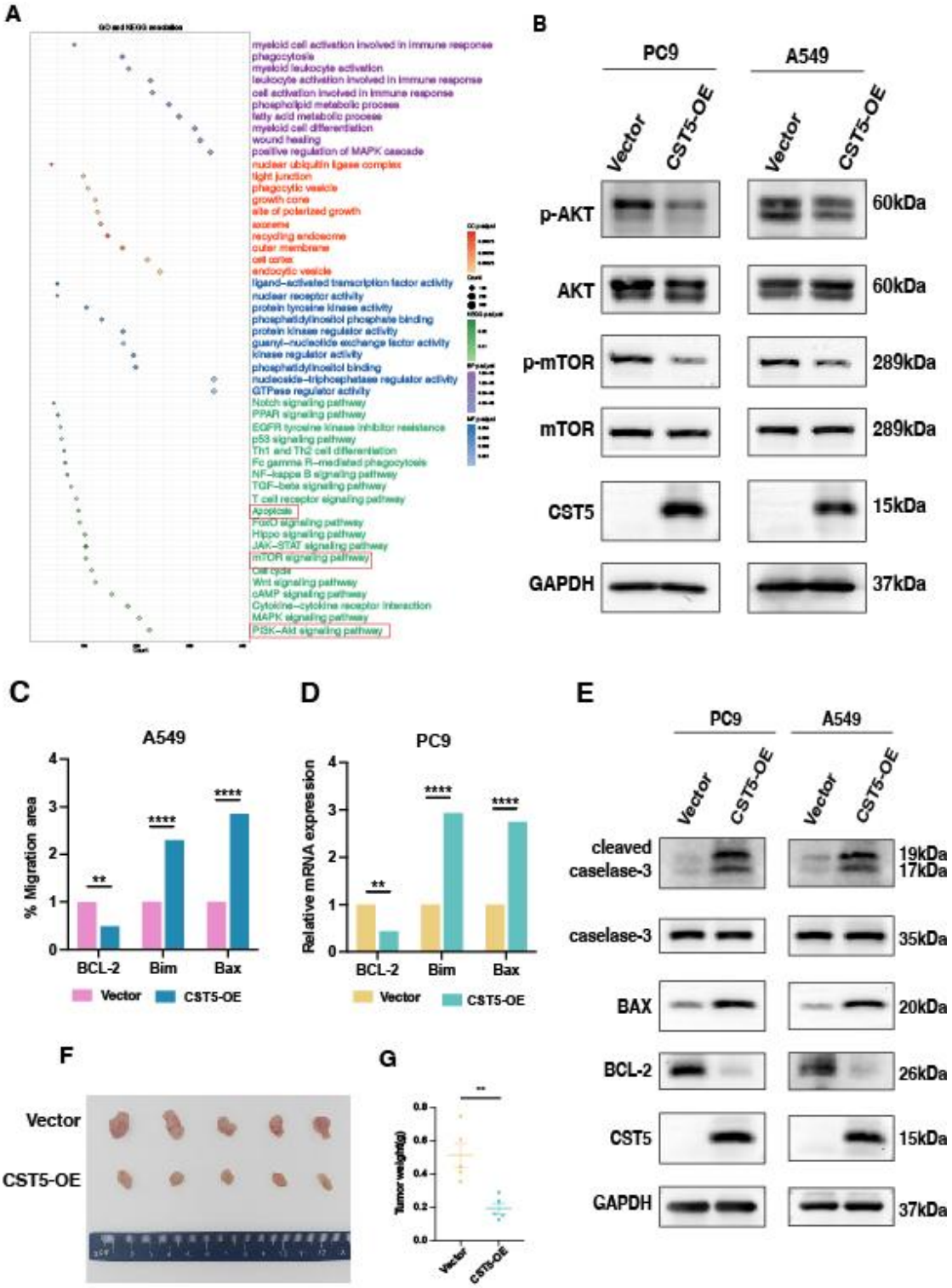


Fig. 9. Functional and mechanistic validation of CST5 in driving LUAD progression. (A) KEGG and GO Enrichment of DEGs between high- and low-CST5 groups in LUAD. (B) WB Analysis of key molecules in the PI3K/AKT/mTOR pathway between wild-type and CST5-overexpressing cell lines. (C-E) qRT-PCR and WB analysis of key

914 apoptotic pathway-related molecules in wild-type and CST5-overexpressing cell lines.
915 (F) 5×10^6 control or CST5-overexpressing PC9 cells were subcutaneously implanted
916 into BALB/c nude mice. (G) Tumor weight was compared between the two groups.
917 Data represents mean± standard error of mean, *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$;
918 ****$p < 0.0001$.
919

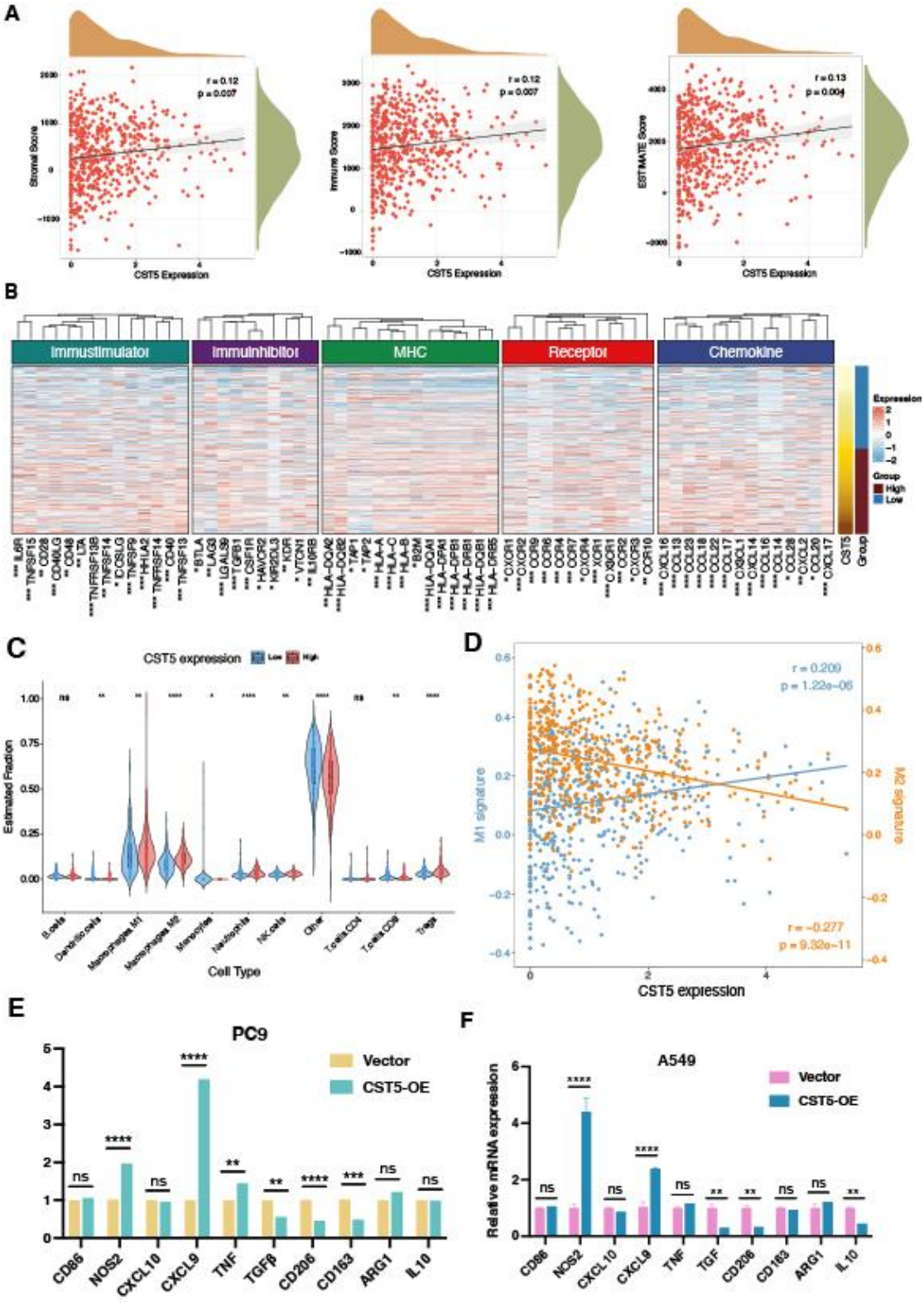


Fig. 10. Association of CST5 expression with the tumor immune microenvironment in LUAD. (A) Correlations between CST5 expression and stromal score, immune score, estimate score. (B) Heatmap showing the association between CST5 expression and

924 immune-related gene expression. (C) Differences in tumor-infiltrating immune cells
925 between CST5 high- and low-expression groups, quantified by the quanTIseq algorithm.
926 (D) CST5 expression correlates with M1/M2 macrophage polarization. (E, F) Effects
927 of CST5 overexpression on macrophage polarization-related gene expression in LUAD
928 cell lines. Data represents mean± standard error of mean, *p < 0.05; **p < 0.01; ***p
929 < 0.001; ****p < 0.0001.
930

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMaterial1.docx
- SupplementaryMaterial2.docx