Am J Cancer Res 2026;16(3):1171-1197
www.ajcr.us /ISSN:2156-6976/ajcr0172350

Original Article

Integrated bulk and single-cell transcriptomics with experimental validation to identify calmodulin-related prognostic genes in lung adenocarcinoma

Zhouyan Lin[1*], Minshu Ao[1*], Yue Hu[2*], Ziyan Yu[2*], Mengyu Lu[2], Yuchao Dong[1], Chong Bai[1], Meng Guo[2]

[1]Department of Respiratory and Critical Care Medicine, The First Affiliated Hospital of Naval Medical University, Shanghai, China; [2]National Key Laboratory of Immunity and Inflammation & Institute of Immunology, Naval Medical University, Shanghai, China. [*]Equal contributors.

Received March 1, 2026; Accepted March 24, 2026; Epub March 25, 2026; Published March 30, 2026

Abstract: Lung adenocarcinoma (LUAD) exhibits poor prognosis partly due to dysregulated calcium signaling. Here, we integrated bulk RNA-seq (TCGA/LUAD, n = 440) and single-cell sequencing (GSE131907) to identify calmodulin-related prognostic genes. Differential expression analysis and LASSO-Cox regression identified SPHK1 and ASPM as independent predictors of overall survival (HR = 1.52/1.47, $P < 0.001$). A risk model incorporating these genes stratified patients into high/low-risk groups with distinct clinical outcomes (3-year AUC = 0.83). High-risk patients showed elevated γδ T cells and TIDE scores, while SPHK1/ASPM expression correlated with an immunosuppressive myeloid phenotype and M2-like polarization. Single-cell analysis revealed $SPHK1^+$-$ASPM^+$ myeloid cells dominated in early differentiation stages, with pseudo-time trajectories indicating SPHK1 downregulation and ASPM upregulation during myeloid maturation. Multiplex immunofluorescence validated co-localization of SPHK1/ASPM with MPO^+ myeloid cells in tumor tissues. Importantly, the $MPO^+SPHK1^+/ASPM^+$ myeloid compartment was significantly enriched in anti-PD-1 non-responders ($P < 0.05$). Our study establishes a calcium signaling-based prognostic signature, uncovers myeloid SPHK1/ASPM as drivers of immunosuppression, and provides a potential biomarker for immunotherapy stratification.

Keywords: Lung adenocarcinoma, calmodulin, prognostic model, immune microenvironment, single-cell RNA-seq, pseudotime trajectory analysis, multiplex immunofluorescence

Introduction

Lung cancer is a prevalent malignancy worldwide. Lung adenocarcinoma (LUAD), a major subtype of non-small cell lung cancer (NSCLC), accounts for a substantial proportion of lung cancer cases and is a leading cause of cancer-related death [1, 2]. While therapies including chemotherapy, radiotherapy, targeted therapy, and immunotherapy have extended survival in a subset of patients, tumor heterogeneity results in marked individual variability in clinical outcomes, rendering current clinicopathological indicators inadequate for the accurate prediction of individualized treatment responses [3, 4]. Tumor progression is closely associated with aberrations in multiple molecular signaling pathways. The calcium ion signaling pathway serves as a pivotal intracellular signaling hub, governing key tumor cell biological behaviors, including proliferation, migration, and chemoresistance development [5, 6]. As a core effector of this pathway, calmodulin (CaM) senses fluctuations in intracellular Ca^{2+} concentrations and interacts with a wide range of target proteins, thereby regulating multiple stages of tumorigenesis and progression [7-9]. Thus, systematic analysis of the expression patterns and prognostic value of calmodulin-related genes (CRGs) in LUAD is critical for the development of risk stratification models and the guidance of personalized therapeutic strategies.

Emerging evidence has shown that CaM promotes lung adenocarcinoma progression through a multi-level regulatory network. In terms of therapeutic resistance, CaM directly modu-

https://doi.org/10.62347/SDLR5143

lates tumor cell resistance to tyrosine kinase inhibitors (TKIs) by stabilizing mutant epidermal growth factor receptor (EGFR) proteins and delaying their degradation [10]. The Ca^{2+}/CaM complex stabilizes drug-resistant mutants by preventing their degradation, representing a novel explanation for TKI resistance in the context of protein homeostasis [6]. In the context of tumor stemness and recurrence, Ca^{2+}/CaM-dependent protein kinase IIα (CaMKIIα), a key downstream effector of CaM, is aberrantly activated in LUAD. Aberrant CaMKIIα activation correlates with adverse clinical outcomes and enhances the self-renewal capacity of tumor-initiating cells, suggesting its potential role in promoting tumor recurrence by sustaining cancer cell stemness [11]. Furthermore, in transcriptional regulation and immune microenvironment remodeling, the Ca^{2+}-CaM-calcineurin-NFATc2 signaling axis promotes the acquisition of tumor-initiating cell phenotypes and enhances cell motility and therapeutic resistance [12]. NFATc2 further contributes to the establishment of an immunosuppressive tumor microenvironment by regulating glycolytic metabolism and driving M2 polarization in tumor-associated macrophages [13]. In metabolic adaptation and phenotypic transition, the CaMKK-AMPK signaling axis transduces calcium signal into adaptation to cellular metabolic and stress survival phenotype [14]. CaM-mediated regulation of MARCKS family proteins modulates cytoskeletal remodeling and the epithelial-mesenchymal transition (EMT) process [15]. CaM also regulates phosphodiesterase 1A (PDE1A), which functions as a critical mediator of tumor metastasis by serving as a signaling bridge linking calcium and STAT3 pathways [16]. These findings suggest that CRGs contribute to the development and progression of LUAD through several signaling cascades. To date, no prognostic risk assessment model based on CRGs has been established for clinical stratifications. The in-depth study of the role of CRGs in the tumor microenvironment and their relationship with immunotherapy responses remains unclear. Therefore, systematic screening and validation of the expression profiles, regulatory networks, and clinical prognostic value of CRGs in LUAD, through integrating multi-omics data, may facilitate the identification of novel biomarkers for precise risk stratification and the development of individualized treatment strategies.

Rapidly developed high-throughput transcriptomic technologies have allowed systematic investigations of CRGs in LUAD. Single-cell RNA sequencing (scRNA-seq) provides high-resolution data for understanding tumor mechanisms, allowing the identification of prognostic genes and the characterization of dynamics in cell subsets [17]. Based on large-scale transcriptomic and genomic data, prognosis-related signatures were constructed to predict LUAD, validating the correlation between candidate genes and patients' survival outcomes, and emphasizing the potential of molecular biomarkers in prognostic prediction and therapeutic regimen selection [18]. ScRNA-seq helps dissect cancer cell diversity and resolve heterogeneity [19, 20]. Integrating multi-omics to identify calcium signal-related tumor subtypes and regulatory networks is a key trend in LUAD research [21].

Against this background, we integrated bulk transcriptomic and scRNA-seq data to perform a systematic analysis of CRGs, aiming to develop a prognostic risk assessment model based on CRGs, systematically examine their expression patterns and regulatory networks, and assess their association with the tumor microenvironment, thereby informing precise risk stratification and personalized treatment strategies for patients with LUAD.

Data and methods

Data acquisition

Gene expression profiles and corresponding clinical information were retrieved from The Cancer Genome Atlas (TCGA) database (http://cancergenome.nih.gov/) via downloading the TCGA-LUAD dataset, serving as the training set. This dataset included 513 LUAD tumor samples and 58 control samples. For subsequent survival analysis, 440 tumor samples with available survival information were retained; for further clinical parameter analysis, 369 tumor samples with complete clinicopathological indicators (including age, sex, pathological stage, and tumor stage (T stage)) were included. Gene expression profiles and clinical data were obtained from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/) by downloading the GSE50081 dataset (platform: GPL570) [22, 23], serving as the validation set. Samples of other cancer types

were excluded, including 2 adenosquamous carcinoma samples, 7 large cell carcinoma samples, 1 NSCLC sample with a large cell carcinoma component, 1 NSCLC sample favoring adenocarcinoma, 42 squamous cell carcinoma samples, and 1 second case of squamous cell carcinoma. A total of 127 LUAD samples with survival information were selected for subsequent analyses. For single-cell transcriptomic analysis, the GSE131907 dataset (platform: GPL16791) was downloaded from the GEO database (https://www.ncbi.nlm.nih.gov/geo/). Non-LUAD tumor samples were excluded, including 10 normal lymph node samples, 10 metastatic brain tissue samples, 7 metastatic lymph node samples, and 5 pleural effusion samples from patients with LUAD complicated by malignant pleural effusion. Finally, 11 LUAD tumor tissue samples and 11 matched distant normal lung tissue samples were analyzed in this study. A total of 255 CRGs were obtained from previously published literature [24] (Table S1). All data were retrieved on July 8, 2025.

Differential expression analysis and candidate gene screening

To identify differentially expressed genes (DEGs) between LUAD and control samples, gene expression analysis was performed using the *DESeq2* package (v1.40.2) [25] (adjusted *P*-value (p.adj) < 0.05, |log2fold change (FC)| > 0.5). Intersection analysis between DEGs and CRGs was performed using the *ggvenn* package (v0.1.10) [26]. The overlapping genes were designated as candidate genes.

Functional enrichment analysis and protein-protein interaction (PPI) network construction

To elucidate the biological processes and signaling pathways associated with the candidate genes, functional enrichment analysis was performed. Gene Ontology (GO) enrichment analysis, encompassing biological processes (BPs), molecular functions (MFs), and cellular components (CCs), and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis were conducted using the *clusterProfiler* package (v4.10.1) [27]. The top 5 GO terms and top 10 significantly enriched KEGG pathways, ranked by ascending *p*-value, were selected for visualization. The protein-level interaction relationships between the candidate genes were explored by inputting the genes into the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database (http://string-db.org) to predict protein functions and PPI relationships (interaction score > 0.9). After screening to remove isolated proteins, the resulting PPI network was visualized using the *circlize* package (v0.4.10) [28].

Prognostic gene screening

To identify genes significantly associated with survival outcomes, univariate Cox regression analysis was performed on the candidate genes in the TCGA-LUAD cohort (n = 440 tumor samples with survival information) using the *survival* package (v3.8.3) [29] ($P < 0.05$, hazard ratio (HR) ≠ 1). Subsequently, the proportional hazards (PH) assumption testing was conducted. Genes that passed the PH assumption test were subjected to Least Absolute Shrinkage and Selection Operator (LASSO) regression using the *glmnet* package (v4.1.8) [30]. Genes selected by LASSO regression were defined as prognostic genes. Importantly, the LASSO-Cox procedure was used for feature selection to derive the prognostic gene set, rather than as the final prognostic model for risk scoring.

Prognostic model construction and validation

A random survival forest (RSF) model was constructed to develop the prognostic model based on the identified prognostic genes using the *randomForestSRC* package (v3.4.1) [31]. RSF is a non-parametric ensemble learning method that extends the random forest framework to right-censored survival data and does not rely on PHs or linearity assumptions. Bootstrap resampling was applied to generate B subsets, on which individual survival trees were generated. At each node, variables were randomly selected, and the optimal split was determined by maximizing survival differences via the log-rank test. Model parameters were optimized through out-of-bag (OOB) error estimation. After all survival trees were constructed, the cumulative hazard function (CHF) for each patient was calculated by averaging the CHF predictions from all trees in the ensemble. For a given patient x at time t, the ensemble hazard function was calculated as:

$$\hat{H}(t \mid x) = \frac{1}{B}\sum_{i=1}^{B}\hat{H}_i(t \mid x)$$

where $\hat{H}_i(t|x)$ represents the predicted hazard for individual x at time t from the i-th tree.

The RSF model for prognostic genes was constructed in the TCGA-LUAD cohort (n = 440 tumor samples with survival information) with optimized parameters (number of trees (ntree) = 1000, number of variables tried at each split (mtry) = 2). The RSF-derived ensemble hazard was used as the prognostic risk score (risk score) for each patient. A higher risk score indicated a greater predicted mortality risk. Based on the optimal cutoff value of the RSF-derived risk scores, samples in the TCGA-LUAD cohort (n = 440) were stratified into high- and low-risk groups. Furthermore, Kaplan-Meier (K-M) survival curves were generated to assess survival differences between the two groups using the *survminer* package (v0.5.0) [32]. Additionally, time-dependent receiver operating characteristic (ROC) curves for 3-, 5-, and 7-year survival rates were plotted using the *survivalROC* package (v1.0.3.1) [33], with an area under the curve (AUC) > 0.6 considered clinically informative. The trained RSF model was directly applied to the GSE50081 dataset to validate the generalizability of the prognostic model.

Comparison of risk scores across clinicopathological subgroups and identification of independent prognostic factors

To compare risk scores across different clinicopathological subgroups, 369 tumor samples with complete clinicopathological indicators (including age, sex, pathological stage, and T stage) from the TCGA-LUAD cohort were selected for analysis. Differences in the computed risk scores among subgroups stratified by these clinicopathological indicators were evaluated using the Kruskal-Wallis test or the Wilcoxon test. To identify independent prognostic factors and evaluate the clinical utility of a predictive nomogram, analyses were performed on the TCGA-LUAD cohort (n = 369 tumor samples with complete clinicopathological indicators). Univariate Cox regression analyses were performed first, using the *survminer* package (v0.5.0), for individual clinical variables predicting overall survival ($P < 0.05$, HR ≠ 1); the variables included in the analyses were age, sex, pathological stage, T stage, and our risk score. PH assumption testing was then performed on the statistically significant variables identified in the univariate Cox analysis. Variables satisfying the PH assumption ($P > 0.05$ in the PH test) were retained for subsequent analyses. Furthermore, the remaining variables were subjected to multivariate Cox regression ($P < 0.05$, HR ≠ 1) and a second round of PH assumption test ($P > 0.05$) to identify independent prognostic factors. Based on the confirmed independent prognostic factors, a nomogram was constructed to predict 3-, 5-, and 7-year overall survival probabilities of LUAD patients. The nomogram was developed using the *regplot* package (v1.1) [33] on the TCGA-LUAD cohort (n = 369). The performance and calibration accuracy of the nomogram were assessed using a calibration curve generated with the "rms" package (v6.8.1) [34]. Additionally, the diagnostic performance of the nomogram was evaluated through time-dependent ROC curve analysis with the *survivalROC* package (v1.0.3.1).

Analysis of immune cell infiltration and correlation with prognostic genes

Immune cell infiltration profiles in LUAD samples from the TCGA-LUAD cohort (n = 440 tumor samples with complete survival information) were assessed using the Single-Sample Gene Set Enrichment Analysis (ssGSEA) algorithm within the *GSVA* package (v1.46.0) [35] to quantify the relative abundances of 28 immune cell subsets [36]. Subsequently, Wilcoxon tests were applied to compare the infiltration levels of each immune cell subset between risk groups. Immune cell types exhibiting statistically significant differences were defined as differentially infiltrated immune cells. Furthermore, Spearman correlation analyses were conducted to investigate pairwise correlations among the differentially infiltrated immune cells across all samples in the TCGA-LUAD dataset (|correlation coefficient (cor)| > 0.3, $P < 0.05$). Additionally, correlations between the prognostic genes and differentially infiltrated immune cells were assessed (|cor| > 0.3, $P < 0.05$). All these correlation analyses were performed using the *psych* package (v 2.2.9) [37].

Comprehensive analysis of tumor microenvironment, TIDE score, and immune checkpoint expression

TCGA-LUAD cohort (n = 440 tumor samples with complete survival information) was further analyzed to comprehensively explore the

differences in tumor microenvironment (TME) composition, immune escape potential, and responsiveness to immunotherapy between high- and low-risk LUAD patients. The ESTIMATE algorithm was utilized to quantify the abundance of stromal and immune cells in tumor tissues, yielding immune scores, stromal scores, and ESTIMATE scores (sum of immune and stromal scores); inter-group differences in these three scores were assessed using the Wilcoxon test. Tumor Immune Dysfunction and Exclusion (TIDE) analysis was conducted to evaluate the propensity for tumor immune escape, and TIDE scores for the 440 samples were retrieved from the online TIDE platform (http://tide.dfci.harvard.edu/) (accessed on July 8, 2025). Spearman correlation analysis between TIDE scores and risk scores was performed using the "psych" package (v2.2.9), while inter-group differences in TIDE scores were compared via the Wilcoxon rank-sum test. Additionally, differences in the expression of 79 immune checkpoint molecules derived from published literature [38] (Table S2) were compared between high- and low-risk groups using the Wilcoxon test.

Gene set enrichment analysis (GSEA)

To delineate the biological functional disparities between high- and low-risk LUAD patients, GSEA was performed using the TCGA-LUAD cohort (n = 440). The gene set collection "c2.cp.kegg_legacy.v2025.1.Hs.symbols" was retrieved from the Molecular Signatures Database (MSigDB) (https://www.gsea-msigdb.org/) for this analysis. First, differential expression analysis between the high- and low-risk groups in the TCGA-LUAD cohort was performed using the "DESeq2" package (v1.40.2). This analysis generated $\log_2$FC values for all genes, which were then ranked in descending order of $\log_2$FC values. GSEA was subsequently conducted using the *clusterProfiler* package (v4.10.1) (|normalized enrichment score (NES)| > 1, $P < 0.05$).

Somatic mutation and drug sensitivity analyses

Somatic mutation burden data from the TCGA-LUAD cohort (n = 440) were analyzed using the *maftools* package (v2.18.0) [39] to characterize mutation landscapes across high- and low-risk groups. To further explore the predictive value of the risk score for chemotherapy responsiveness, the half-maximal inhibitory concentration (IC50) values of a panel of therapeutic agents were estimated using the *pRRophetic* package (v0.5) [40] based on the 440 samples. The drug sensitivity reference data were retrieved from the Genomics of Drug Sensitivity in Cancer (GDSC) database (https://www.cancerrxgene.org/). Inter-group differences in predicted IC50 values were assessed using the Wilcoxon test to compare chemotherapy responsiveness between high- and low-risk LUAD patients.

Processing of single-cell RNA sequencing (scRNA-seq) data and cell annotation

Single-cell expression data from the GSE131907 dataset were retrieved and processed using the "Seurat" package (v5.30) [41]. Quality control was performed using the *PercentageFeatureSet* function with the following criteria: (1) number of RNA molecules detected per cell (nCount_RNA): 500 < nCount_RNA < 30,000; (2) number of unique genes detected per cell (nFeature_RNA): 200 < nFeature_RNA < 5,000; (3) mitochondrial gene content (percent.mt): cells with percent.mt < 10% were retained. Highly variable genes (HVGs) were identified using the *FindVariableFeatures* function with the variance stabilization transformation (VST) method, and the top 2,000 HVGs were selected for downstream analysis. Gene expression values were normalized across all cells using the *ScaleData* function. Subsequently, principal component analysis (PCA) was performed on the 2,000 HVGs using the *RunPCA* function. The number of PCs for downstream analysis was selected using a combination of approaches: statistical validation using the JackStraw permutation test ($P < 0.05$) and the elbow criterion (choosing PCs corresponding to the inflection point where the scree plot 'flattens'). Then, an unsupervised clustering analysis was conducted using the *FindNeighbors* and *FindClusters* functions on the selected PCs to define the respective cell clusters (resolution = 0.2). The clustering results were visualized with UMAP dimensionality reduction plots. Cell clusters were annotated by referencing the marker gene information from the literature [42].

Identification of key cell populations

A chi-squared test was performed comparing the proportionate distributions of annotated

cell types between the LUAD group and the corresponding control group, to identify significant difference in cell-type abundance. The expression patterns of prognostic genes across annotated cell populations were subsequently analyzed. Based on the abundance profiles of differentially represented cell populations and the expression levels of the prognostic genes, cell populations that were differentially abundant and exhibited consistent expression of all prognostic genes were designated as key cell populations.

Cell-cell communication and metabolic pathway analysis

To characterize intercellular communication among annotated cell populations, cell-cell communication analysis was performed on LUAD and control samples from the GSE131907 dataset using the *CellChat* package (v1.6.1) [43]. Potential ligand-receptor interaction pairs were inferred to assess signaling crosstalk between key cell populations and other annotated cells. Additionally, metabolic activity at the single-cell level was quantified using the VISION algorithm implemented in the *scMetabolism* package (v0.2.1) [44].

Functional enrichment and pseudotime trajectory analyses of key cell subclusters

Key cell populations were re-clustered using the top PCs identified previously, with the resolution parameter set to 0.1. To investigate the biological functions and pathways associated with key cell subclusters, functional enrichment analysis was performed using the *ReactomeGSA* package (v3.2.0) [45]. Furthermore, developmental trajectories of cellular states were reconstructed using the *monocle* package (v2.26) [46]. Pseudotime analysis was performed to infer the dynamic progression of cellular states across the single-cell atlas, and the expression dynamics of prognostic genes along pseudotime trajectory were computationally profiled.

Patient specimens

Formalin-fixed, paraffin-embedded (FFPE) LUAD tissue sections used in this study were obtained from residual clinical specimens following clinical diagnosis at the Department of Pathology, the First Affiliated Hospital of the Naval Medical University. The study protocol was reviewed and approved by the Shanghai Changhai Hospital Ethics Committee (Approval No.: CHEC2024-109). All procedures were conducted in strict compliance with the Declaration of Helsinki and relevant Chinese regulations governing biomedical research ethics.

Multicolor immunofluorescence

Formalin-fixed, FFPE LUAD tissue sections were used for multiplex immunofluorescence (mIF) analysis. Tumor regions were annotated according to pathological evaluation and/or matched hematoxylin and eosin (H&E) staining. Sections were deparaffinized, rehydrated, and subjected to antigen retrieval, followed by blocking of nonspecific binding. Tyramide signal amplification (TSA)-based sequential staining was performed to generate two three-color panels: (i) MPO + SPHK1 + DAPI and (ii) MPO + ASPM + DAPI. Briefly, the target antigen (SPHK1 or ASPM) was stained first using primary antibody (SPHK1: Proteintech, Cat# 10670-1-AP, 1:600; ASPM: Proteintech, Cat# 26223-1-AP, 1:600), followed by a species-specific HRP-conjugated secondary antibody (ABclonal, Cat# AS014) and subsequent TSA fluorophore deposition. Antibody complexes were then stripped by heat-induced epitope retrieval to allow subsequent staining. MPO was stained in the second round using the same TSA workflow with an anti-MPO primary antibody (ZSGB-BIO, Cat# ZA-0197) and the corresponding HRP-conjugated secondary antibody. Nuclei were counterstained with DAPI, and sections were mounted for imaging. Representative images were acquired under identical microscope settings for each marker. Statistical analysis was performed as described below.

Statistical analysis

Statistical analyses were performed using R software (version 4.3.3) and GraphPad Prism (version 9.5.1). The Wilcoxon rank-sum test was used to assess differences in gene expression between groups. Differential expression analysis between LUAD and control samples was conducted using DESeq2 (v1.40.2), with thresholds set at $|\log2\text{fold change}| > 0.5$ and P-value < 0.05. A threshold of $P < 0.05$ was considered statistically significant for Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses. Univariate Cox regression (survival package, v3.8.3) was performed to identify survival-associated genes, followed by LASSO-Cox regression (*glmnet* package, v4.1.8) for feature

selection. A random survival forest model (*randomForestSRC* package, v3.4.1) was subsequently constructed, with 1000 bootstrap iterations and mtry = 2. Multivariate Cox regression (survival package) was performed to determine independent prognostic factors.

Immune cell infiltration was quantified using ssGSEA (GSVA package, v1.46.0), based on 28 predefined immune cell signatures. GSEA was performed using the clusterProfiler package (v4.10.1) with KEGG pathways (c2.cp.kegg_legacy. v2025.1.Hs. symbols).

Somatic mutation analysis was conducted using *maftools* package (v2.18.0), and drug sensitivity was predicted via the *pRRophetic* package (v0.5) based on GDSC data. All statistical tests were two-sided, with $P < 0.05$ considered statistically significant.

Results

Identification of CRG-related candidate genes and their underlying biological mechanisms in LUAD

Comparative transcriptome analysis between LUAD and control samples identified 9,404 DEGs in LUAD (5,934 upregulated and 3,470 downregulated genes) (**Figure 1A**, **1B**). Intersection of these DEGs with 255 CRGs further yielded 136 candidate genes (**Figure 1C**). GO enrichment analysis identified 925 significantly enriched terms ($P < 0.05$), including 684 BPs, 132 CCs, and 109 MFs (**Figure 1D** and Table S3). Enriched BP terms included muscle contraction, calcium ion transport, and striated muscle contraction (**Figure 1D**); CC terms included myosin complex, contractile fiber, myofibril, myosin filament, and sarcomere; and MF terms included calmodulin binding, microfilament motor activity, and actin binding (**Figure 1D** and Table S3). KEGG pathway analysis identified 77 significantly enriched pathways ($P < 0.05$), including the calcium signaling pathway, cGMP-PKG signaling pathway, vascular smooth muscle contraction, motor proteins, and cytoskeleton-related pathways (**Figure 1E** and Table S4). Further, analysis of PPI network revealed complex functional interactions among candidate genes, with representative nodes including *CALML3* and *NOS1* (**Figure 1F**). Overall, dysregulation of CRGs in LUAD provides critical information on the molecular mechanisms underlying LUAD and potential targets for diagnostic and therapeutic (theranostic) applications.

Identification of independent prognostic biomarkers for risk stratification in LUAD

Univariate Cox regression analysis and PH assumption testing identified 32 genes significantly associated with overall survival in the TCGA-LUAD cohort ($P < 0.05$, HR ≠ 1; PH $P > 0.05$) (**Figure 2A1**, **2A2**). Notably, 15 genes (e.g., *ARPP21*) functioned as protective factors (HR < 1), while another 15 genes (e.g., *PCP4*) acted as risk factors (HR > 1). To refine the candidate genes, LASSO regression was applied to these survival-associated genes. Based on the minimum cross-validation error (lambda. min = 0.11293785), SPHK1 and ASPM retained non-zero regression coefficients, and were defined as prognostic genes (**Figure 2B**). A prognostic model was constructed using the identified prognostic genes (SPHK1 and ASPM) to stratify patients in the TCGA-LUAD cohort (n = 440) into high- and low-risk groups. Using the optimal risk score cutoff (67.92034), 154 patients were classified as high-risk and 286 as low-risk (**Figure 2C**). Kaplan-Meier survival analysis demonstrated significantly poorer overall survival in the high-risk group compared with the low-risk group ($P < 0.0001$) (**Figure 2D**). Time-dependent ROC analysis showed good predictive performance, with 3-, 5-, and 7-year AUCs were 0.81, 0.82, and 0.83, respectively (**Figure 2E**). To further validate the robustness and generalizability of the model, it was applied to the GSE50081 cohort (N = 127). Using the optimal cut-off specific for that cohort (68.42191), patients were assigned to high-risk (n = 48) and low-risk (n = 79) groups (**Figure 2F**). The high-risk group exhibited significantly worse overall survival ($P < 0.0001$) and a higher incidence of mortality (**Figure 2G**). The 3-, 5-, and 7-year AUCs were 0.78, 0.76, and 0.72, respectively (**Figure 2H**), indicating stable predictive performance across datasets. These findings demonstrate that SPHK1 and ASPM serve as robust prognostic biomarkers with potential clinical utility for risk stratification in LUAD.

Risk score integrating SPHK1 and ASPM stratifies prognosis in LUAD

Given the reliable prognostic performance of the SPHK1- and ASPM-based risk model, its

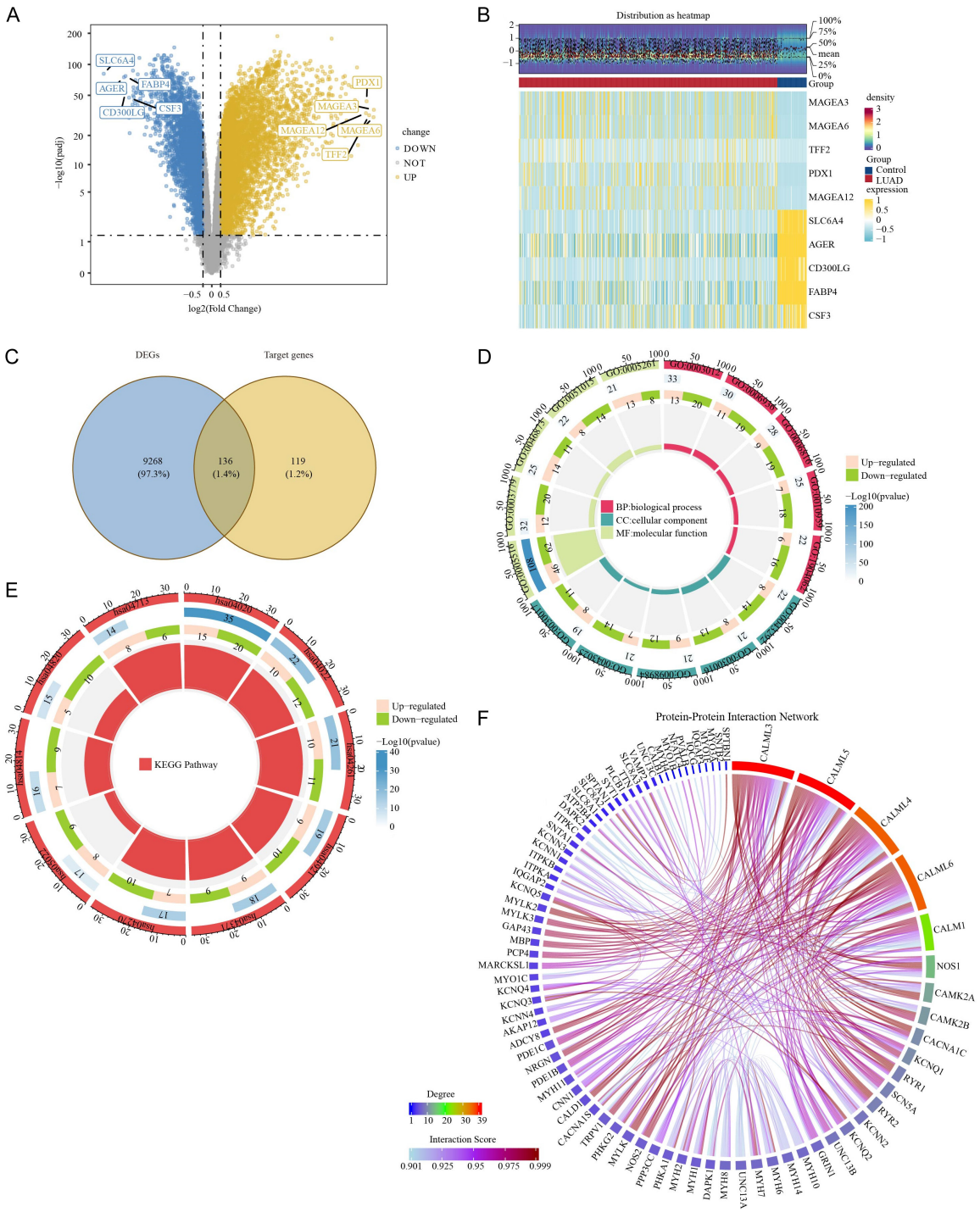

Figure 1. Identification and functional characterization of candidate genes in LUAD. A. Volcano plot of DEGs between LUAD and control samples, with the top 5 up- and downregulated genes labeled. B. Heatmap showing expression profiles of the top 10 DEGs. C. Venn diagram intersecting DEGs with CRGs to identify candidate genes. D, E. GO and KEGG enrichment analyses of candidate genes. F. PPI network of candidate genes.

correlations with clinical features in LUAD patients were further explored. The risk score was significantly associated with sex and T stage in LUAD patients. Specifically, male LUAD patients exhibited significantly higher risk scores than female patients ($P < 0.01$) (**Figure**

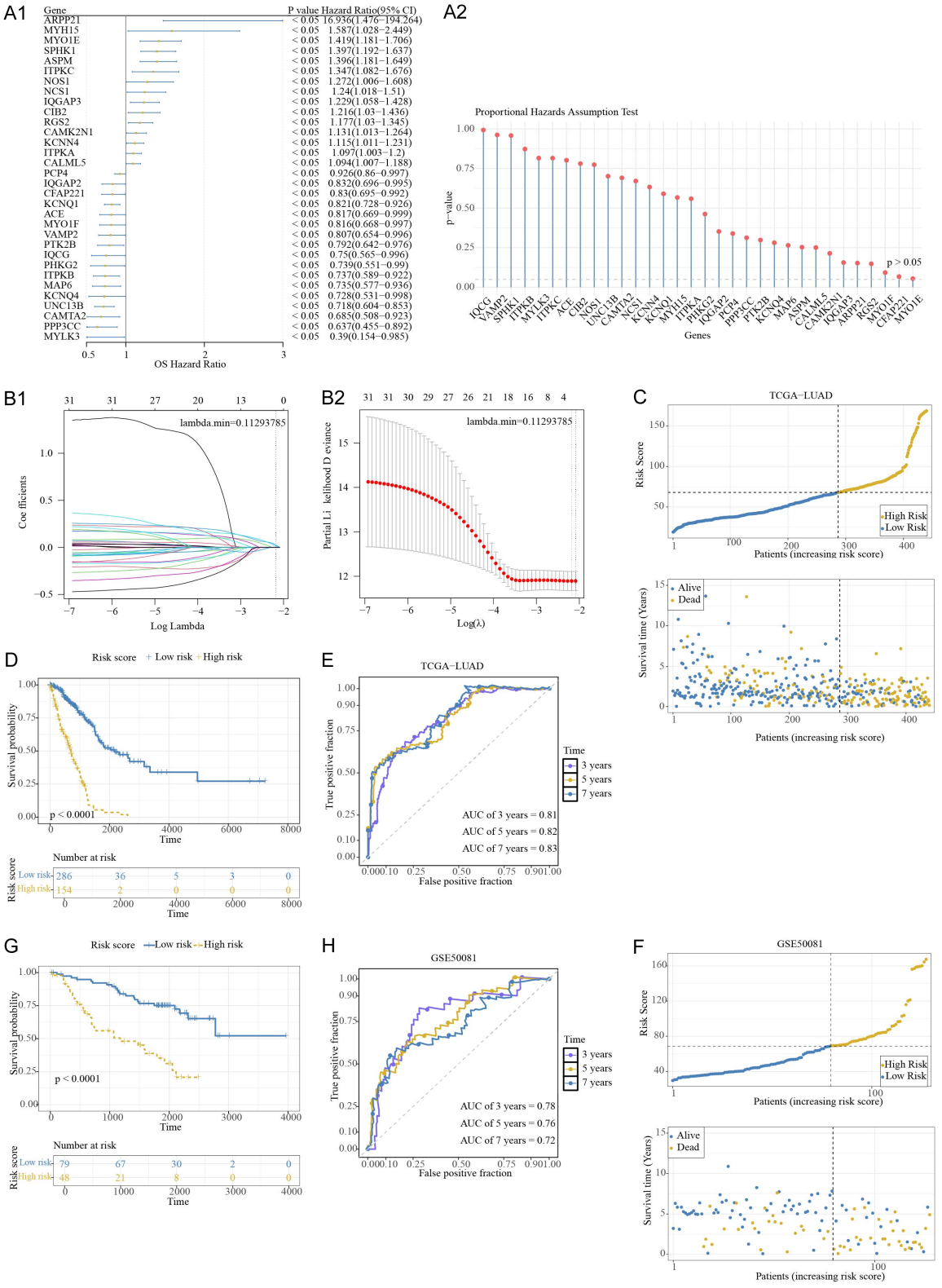

Figure 2. Screening of prognostic genes and construction and validation of prognostic model. A. Forest plot and PH assumption test of univariate Cox regression. B. LASSO regression coefficient profiles and tuning parameter (lambda) selection. C. Risk score and survival status regularization path plot in the TCGA-LUAD cohort. D. K-M survival curve in the TCGA-LUAD cohort. E. ROC curve of the risk model in the TCGA-LUAD cohort. F. Risk score distribution plot and survival time scatter plot in the GSE50081 cohort. G. K-M survival curve in the GSE50081 cohort. H. ROC curve of the risk model in the GSE50081 cohort.

3A1). Patients with T1 tumors showed significantly lower risk scores than those with T2 ($P < 0.01$) and T3/4 ($P < 0.01$) tumors (**Figure 3A2**). These findings suggest that the SPHK1-ASPM-derived risk score is associated with disease progression, as it increases systematically with male sex and advancing T stage.

To further assess its prognostic value and independence from clinical parameters, univariate and multivariate Cox regression analyses were performed in conjunction with PH assumption testing. In univariate Cox regression analysis, risk score, stage (III vs I/II; IV vs I/II), T stage (T2 vs T1; T3/4 vs T1) were significantly associated with overall survival (HR ≠ 1, $P < 0.05$) (**Figure 3B1**). PH assumption testing confirmed that the Cox models satisfied the proportional hazards assumption (**Figure 3B2**). Subsequent multivariate Cox analysis identified risk score (HR = 1.02, 95% CI: 1.02-1.03), stage (III vs I/II) (HR = 1.99, 95% CI: 1.26-3.14), stage (IV vs I/II) (HR = 3.28, 95% CI: 1.67-6.44), and T stage (T3/T4 vs T1) (HR = 1.85, 95% CI: 1.05-3.30) as independent prognostic factors ($P < 0.05$; PH $P > 0.05$) (**Figure 3C1**, **3C2**). A nomogram incorporating these independent prognostic variables was constructed to estimate individual survival probabilities (**Figure 3D**). For a total score of 114 points, the estimated 3-, 5-, and 7-year overall survival rates were 50.8%, 20.5%, and 17.0%, respectively. Calibration curves demonstrated good agreement between predicted and observed overall survival probabilities (**Figure 3E**). Furthermore, time-dependent ROC analysis validated the discriminative capacity of the nomogram, with AUC values of 0.81, 0.81, and 0.85 at 3, 5, and 7 years, respectively (**Figure 3F**). These findings indicate that the risk score, stage, and T stage are independent prognostic factors for LUAD, and their integration into a nomogram provides a reliable tool for individualized survival prediction.

High-risk LUAD defined by SPHK1/ASPM displays immune infiltration remodeling and an immunosuppressive tumor microenvironment

Given the critical role of the tumor immune microenvironment in LUAD progression, immune cell infiltration patterns between the high- and low-risk groups were further analyzed. Evaluation of 28 immune cell subsets revealed relatively high enrichment scores for central memory $CD4^+$ T cells and myeloid-derived suppressor cells (MDSCs) (**Figure 4A**). Comparative analysis revealed significant variations in the infiltration of 16 immune cell types between the two groups: activated B cells, activated dendritic cells (DCs), eosinophils, immature B cells, immature DCs, mast cells, monocytes, plasmacytoid DCs, T follicular helper (Tfh) cells, and type 17 T helper (Th17) cells were decreased, while activated $CD4^+$ T cells, effector memory $CD4^+$ T cells, γδ T cells, memory B cells, natural killer T (NKT) cells, and type 2 T helper (Th2) cells were increased in the high-risk group ($P < 0.05$) (**Figure 4B**). These findings indicate substantial differences between the high-risk and low-risk groups of LUAD in immune microenvironment composition, with the high-risk group exhibiting features consistent with an immunosuppressive phenotype. Correlation analysis of differentially infiltrated immune cells demonstrated that immature B cells showed the strongest positive correlation with activated B cells (cor = 0.88, $P < 0.001$) (**Figure 4C**). Additionally, the expression levels of SPHK1 and ASPM were significantly positively correlated with the infiltration of activated $CD4^+$ T cells (ASPM: r = 0.50, $P < 0.001$; SPHK1: r = 0.47, $P < 0.001$) and memory B cells (ASPM: r = 0.49, $P < 0.001$; SPHK1: r = 0.37, $P < 0.001$) (**Figure 4D** and Table S5). Specifically, ASPM was also positively correlated with Th2 cells (r = 0.36, $P < 0.001$) and negatively correlated with eosinophils (r = -0.37, $P < 0.001$) and monocytes (r = -0.32, $P < 0.001$). In contrast, SPHK1 showed the strongest positive correlations with NKT cells (r = 0.50, $P < 0.001$) and γδ T cells (r = 0.40, $P < 0.001$). In summary, the SPHK1- and ASPM-based prognostic signature is closely associated with immune microenvironment remodeling and may contribute to LUAD progression and poor prognosis by regulating the infiltration of specific immune cell subsets.

Immunosuppressive tumor microenvironment and immune checkpoint activation characterize the high-risk LUAD subtype defined by SPHK1/ASPM dysregulation

Immune scores differed significantly between the high- and low-risk groups ($P < 0.05$), with

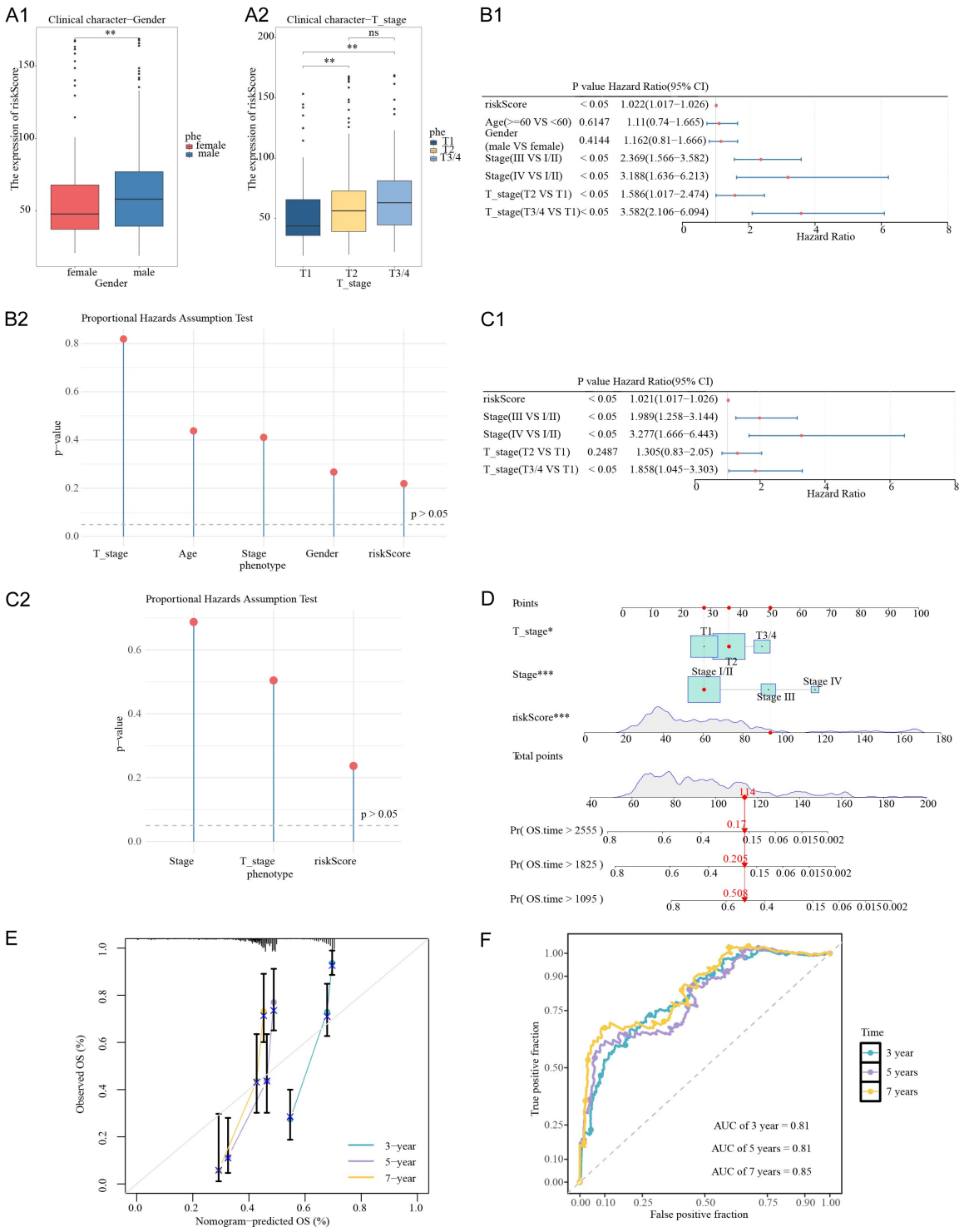

Figure 3. Correlation between risk score and clinical features and construction of prognostic nomogram. A. Box plot comparing risk scores between patients of different genders and T stages. B. Forest plot and PH assumption test of univariate Cox regression. C. Forest plot and PH assumption test of multivariate Cox regression. D. Nomogram predicting LUAD survival outcome. E. Calibration curve of the nomogram. F. ROC curve analysis of the nomogram.

the low-risk group exhibiting markedly higher immune scores (**Figure 4E**). This finding suggests a more active immune response and abundant immune cell infiltration within the

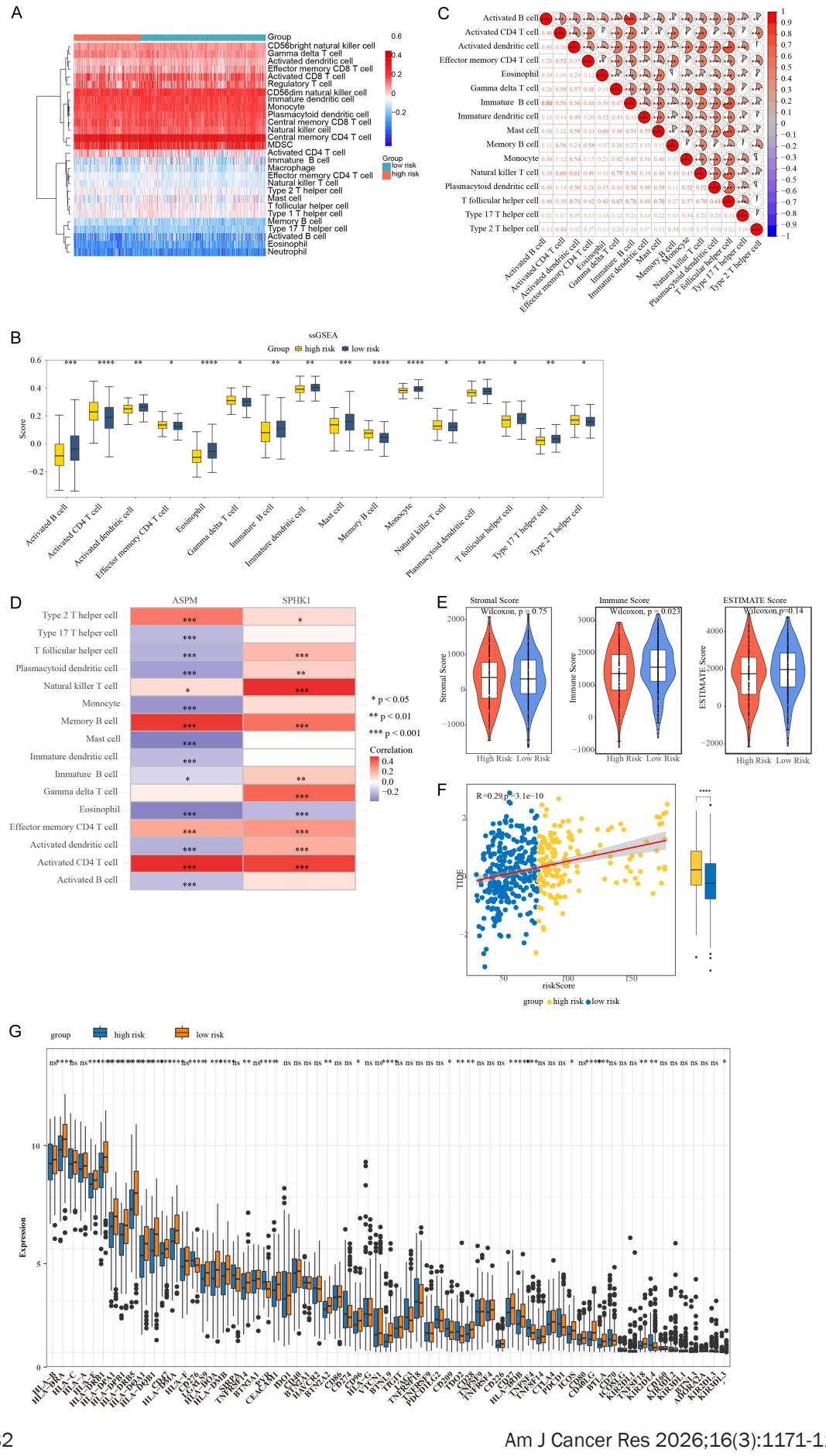
A
Group
CD56bright natural killer cell
Gamma delta T cell
Activated dendritic cell
Effector memory CD8 T cell
Activated CD8 T cell
Regulatory T cell
CD56dim natural killer cell
Immature dendritic cell
Monocyte
Plasmacytoid dendritic cell
Central memory CD8 T cell
Natural killer cell
Central memory CD4 T cell
MDSC
Activated CD4 T cell
Immature B cell
Macrophage
Effector memory CD4 T cell
Natural killer T cell
Type 2 T helper cell
Mast cell
T follicular helper cell
Type 1 T helper cell
Memory B cell
Type 17 T helper cell
Activated B cell
Eosinophil
Neutrophil
Group
low risk
high risk
C
Activated B cell
Activated CD4 T cell
Activated dendritic cell
Effector memory CD4 T cell
Eosinophil
Gamma delta T cell
Immature B cell
Immature dendritic cell
Mast cell
Memory B cell
Monocyte
Natural killer T cell
Plasmacytoid dendritic cell
T follicular helper cell
Type 17 T helper cell
Type 2 T helper cell
B
ssGSEA
Group high risk low risk
Score
D
ASPM
SPHK1
* p < 0.05
** p < 0.01
*** p < 0.001
Correlation
E
Stromal Score
Wilcoxon, p = 0.75
Immune Score
Wilcoxon, p = 0.023
ESTIMATE Score
Wilcoxon,p=0.14
High Risk Low Risk
F
R=0.29,p=3.1e−10
TIDE
riskScore
group high risk low risk
G
group high risk low risk
Expression

Figure 4. Differences in immune cell infiltration, tumor microenvironment, TIDE score, and immune checkpoint expression between high- and low-risk groups. A. Relative abundance of 28 immune cell subsets between high- and low-risk groups. B. Differentially infiltrated immune cells between high- and low-risk groups. C. Correlation network among differentially infiltrated immune cells. D. Correlation plot showing associations between prognostic genes and differentially infiltrated immune cells. E. Comparison of immune scores between high- and low-risk groups. F. Comparison of TIDE scores between high- and low-risk groups and their correlation with risk scores. G. Immune checkpoint expression analysis between risk groups.

tumor microenvironment, a characteristic that correlated with a favorable prognosis and potential responsiveness to immunotherapy. In contrast, TIDE scores were significantly higher in the high-risk group ($P < 0.0001$) and positively correlated with the risk score ($R = 0.29$, $P < 0.0001$) (**Figure 4F**). Elevated TIDE scores indicate increased immune evasion, suggesting that the high-risk group may exhibit reduced T-cell infiltration and impaired antitumor immune function. Immune checkpoint analysis identified 32 differentially expressed checkpoint-related genes between the two groups ($P < 0.05$). Notably, HLA-DRA, CD276, PVR, and TDO2 were significantly upregulated in the low-risk tumors, whereas HLA-E, HLA-DMA, LGALS9, and CEACAM1 were upregulated in the high-risk tumors (**Figure 4G**). These results indicate distinct immune regulatory patterns between the two risk groups.

SPHK1/ASPM-based risk stratification is associated with cell cycle, DNA replication, and tumor progression-related pathways in LUAD

To investigate the potential biological mechanisms underlying the prognostic model, GSEA was performed between high- and low-risk groups. A total of 42 pathways were markedly enriched ($|NES| > 1$, $P < 0.05$), including cell cycle (NES = 2.36), asthma (NES = -2.20), intestinal immune network for IgA production (NES = -2.10), DNA replication (NES = 2.08), p53 signaling pathway (NES = 1.57), proteasome (NES = 1.57), regulation of actin cytoskeleton (NES = 1.31), and ECM receptor interaction (NES = 1.47) (**Figure 5A** and Table S6).

SPHK1/ASPM-defined high-risk LUAD exhibits increased mutational burden and distinct drug sensitivity

To further characterize molecular differences between high- and low-risk patients, somatic mutation profiles were analyzed. In the low-risk group, the most frequently mutated genes included TP53 (45%), TTN (43%) and MUC16 (37%). In the high-risk group, the top mutated genes included TP53 (65%), TTN (59%), MUC16 (50%), and *CSMD3* (50%) (**Figure 5B1**, **5B2**). The unique mutation profiles between the two risk groups suggest greater genomic instability in the high-risk group, which may contribute to its more aggressive phenotype and poorer prognosis. Drug sensitivity analysis identified 67 compounds with significantly different predicted response between the two groups ($P < 0.05$) (Table S7). Among these, AZD7762 (*PLK1*), BI-2536 (*PLK1*), docetaxel, doramapimod (*MAPK*), and MK-1775 (*CDC2*) showed the most significant differences ($P < 0.0001$). Notably, the high-risk group exhibited increased predicted sensitivity to doramapimod (**Figure 5C**).

Single-cell transcriptomic profiling reveals cellular heterogeneity and major cell populations in LUAD

To characterize the extensive cellular heterogeneity in LUAD, single-cell transcriptomic profiling was performed on 88,144 cells from LUAD and control samples derived from the GSE131907 dataset. After quality control filtering, 82,816 high-quality cells expressing 27,578 unique genes were retained for downstream analyses (Figure S1A, S1B). A total of 2,000 highly variable genes (HVGs), including representative genes such as IGLC2 and SCGB1A1, were identified to capture biological heterogeneity across cell populations (Figure S1C). PCA based on these HVGs yielded 30 PCs, which captured the major sources of transcriptomic variation (Figure S1D). UMAP visualization identified 22 transcriptionally distinct cell clusters (**Figure 6A**). Cell type annotation was based on canonical marker genes with cluster-specific expression patterns (**Figure 6B**). For example, CD68, MARCO, and LYZ were highly expressed in myeloids, supporting their annotation (**Table 1**). Ultimately, eight major cell populations were identified:

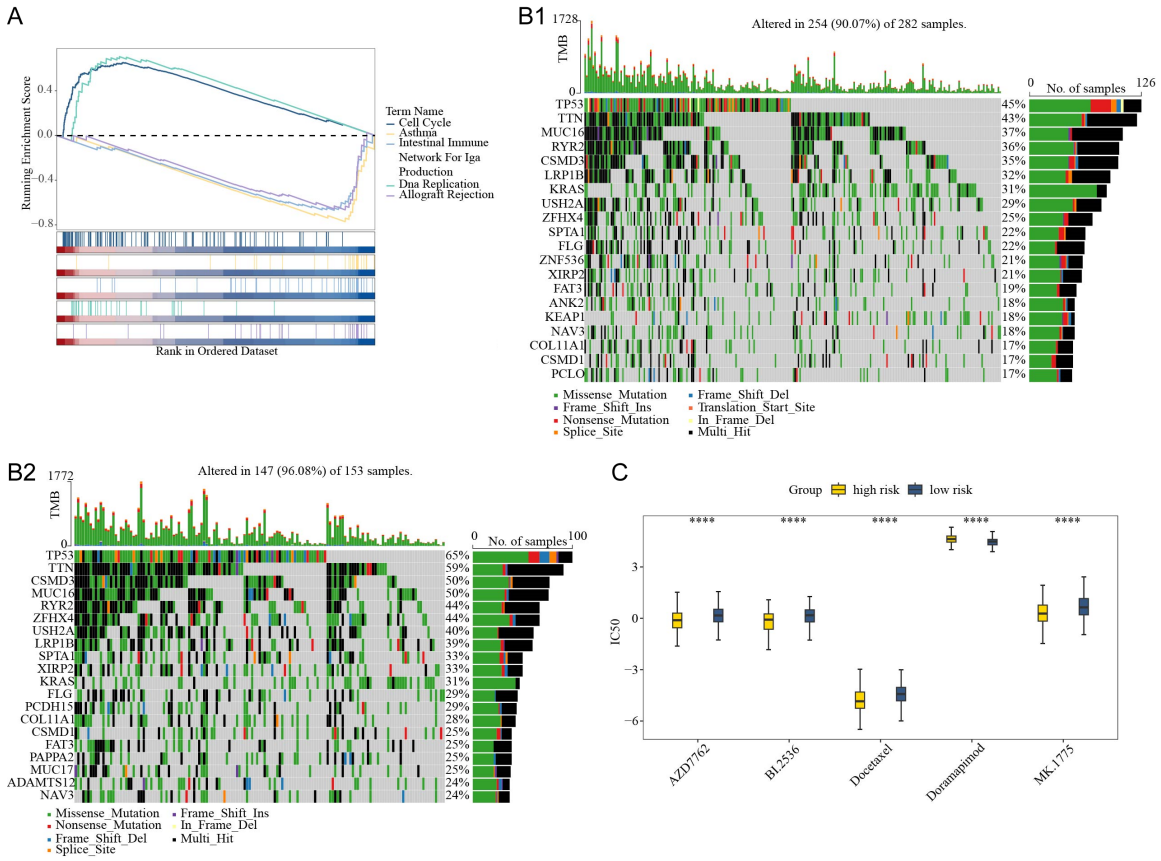

Figure 5. Functional enrichment analysis, somatic mutation profiles and drug sensitivity between high- and low-risk groups. A. GSEA results showing significantly enriched pathways between high- and low-risk groups. B. Somatic mutation landscape of the low- and high-risk groups. C. Comparison of IC50 values of the top 5 differentially sensitive drugs between high- and low-risk groups.

myeloids, natural killer (NK) cells, endothelial cells, fibroblasts, mast cells, epithelial cells, T lymphocytes, and B lymphocytes (**Figure 6C**).

Differentially abundant myeloids expressing SPHK1 and ASPM are identified as key cells in LUAD

Among the 8 annotated cell populations, B lymphocytes, NK cells, T lymphocytes, and myeloids showed marked differences in abundance between LUAD and control groups (**Figure 6D**). Further comparative analysis confirmed that these four cell populations were differentially abundant between groups ($P < 0.05$) (**Figure 6E**). Specifically, B lymphocytes and T lymphocytes were significantly increased in the LUAD group, while myeloid cells and NK cells were decreased. Prognostic gene expression profiling demonstrated that SPHK1 and ASPM were both expressed in myeloids (**Figure 6F**). Based on this observation, myeloids were defined as the key cell populations.

Myeloids in LUAD display altered ligand-receptor interactions and metabolic remodeling

To investigate alterations in intercellular communication networks in LUAD, the dynamics of ligand-receptor interactions across the annotated cells were systematically analyzed. In both control and LUAD groups, myeloids exhibited frequent and strong interactions with fibroblasts and epithelial cells (**Figure 7A**, **7B**), with the interaction strength between myeloids and fibroblasts being significantly increased in the LUAD samples.

Analysis of ligand-receptor pairs indicated that, in the control group, epithelial-myeloid communication was primarily mediated by SCGB3A2-MARCO interactions; in the LUAD

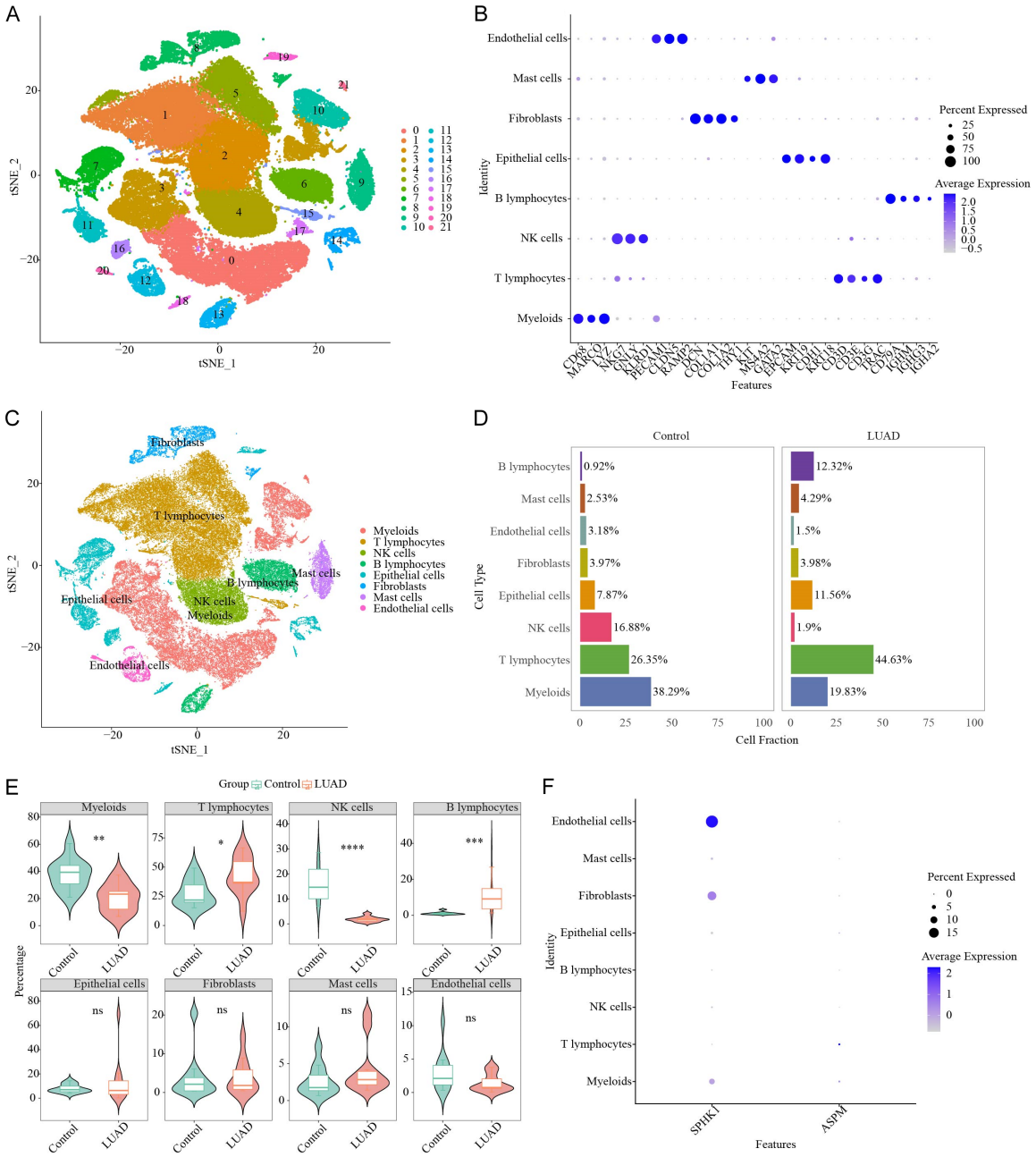

Figure 6. Cell annotation of single-cell RNA sequencing data and screening of key cells. A. UMAP dimensionality reduction clustering plot of single-cell sequencing data from the GSE131907 dataset. B. Dotplot of marker gene expression in each cell cluster. C. UMAP distribution map of 8 major cell types. D. Proportion grouped bar plot of 8 cell types between the LUAD group and the control group. E. Violin plot comparison of proportions of 8 cell types. F. Dotplot of the expression of prognostic genes *SPHK1* and *ASPM* in each cell type.

group, epithelial-myeloid communication was predominantly mediated by the MIF-(CD74+ CD44) axis ($P < 0.01$) (**Figure 7C1**, **7C2**), indicating a shift in key signaling pathways at the single-cell level. Metabolic activity analysis showed that several metabolic pathways exhibited high activity in myeloids, including the citrate cycle (TCA cycle), fructose and mannose metabolism, glutathione metabolism, glycolysis/gluconeogenesis, and oxidative phosphorylation (**Figure 7D**). These findings suggest that metabolic reprogramming may support the functional remodeling of myeloid cells in LUAD.

Table 1. Marker genes of major cell types and their enriched cell clusters

Cluster	marker gene	Cell
0, 3, 10, 17, 21	CD68, MARCO, LYZ	Myeloids
4	NKG7, GNLY, KLRD1	NK cells
12, 20	PECAM1, CLDN5, RAMP2	Endothelial cells
8	DCN, COL1A1, COL1A2	Fibroblasts
9	KIT, MS4A2, GATA2	Mast cells
7, 11, 14, 16, 18, 19	EPCAM, KRT19, CDH1, KRT18	Epithelial cells
1, 2, 5, 15	CD3D, CD3E, CD3G, TRAC	T lymphocytes
6, 13	CD79A, IGHM, IGHG3, IGHA2	B lymphocytes

Myeloid subclusters in LUAD exhibit functional specialization, distinct differentiation trajectories, and dynamic expression of SPHK1/ASPM

To further characterize the functional heterogeneity of myeloids, subclustering analysis identified 10 transcriptionally distinct subpopulations: MDSC, neutrophil, monocyte-macrophage, cDCs, mast cells, monocyte, macrophage, pDCs and monocyte-derived DCs (**Figure 8A**). Enrichment analysis showed that these subclusters were collectively enriched in 1744 pathways (Table S8) ($P < 0.05$), with the top 20 differentially enriched pathways displayed in **Figure 8B**. The analysis revealed marked differences in pathway enrichment profiles across myeloid subclusters. Specifically, pathways such as TWIK-related acid-sensitive K^+-channel (TASK), proton-coupled neutral amino acid transporters, and Sterols are 12-hydroxylated by *CYP8B1* exhibited distinct enrichment patterns across subclusters. Specifically, pathways including antigen processing and presentation (related to adaptive immunity), adherens junction (maintaining tissue mechanical strength and cell polarity), and aldosterone-regulated sodium reabsorption (regulating systemic sodium balance, blood volume, and blood pressure stability) exhibited distinct enrichment patterns across subclusters, reflecting the functional specialization of different myeloid subclusters in LUAD. In addition, pseudotime trajectory analysis based on transcriptional profiling identified nine cellular states corresponding to progressive differentiation stages (**Figure 8C**, **8D**). Mapping of the 9 subclusters onto these states indicated that MDSCs, neutrophils and monocytes were predominantly distributed in early- to-intermediate stages, whereas macrophages and monocyte-macrophages occupied intermediate-to-late stages, and cDCs were enriched in late-stage differentiation (**Figure 8E**). Notably, the LUAD group exhibited a reduced proportion of myeloid cells in early differentiation states, suggesting that disease status influences myeloid developmental trajectories (**Figure 8F**). Expression analysis of SPHK1 and ASPM along the pseudotime axis revealed that SPHK1 expression gradually decreased, whereas ASPM showed an increasing trend during the myeloid differentiation (**Figure 8G**). Collectively, these findings demonstrate the functional specialization, differentiation trajectories, and dynamic regulation of prognostic genes within myeloid subclusters in LUAD, providing insights into their heterogeneous roles in tumor progression.

SPHK1/ASPM-marked MPO^+ myeloid compartment is enriched in anti-PD-1 non-responders in LUAD

To validate the spatial tissue-level evidence supporting the myeloid-associated expression patterns suggested by the single-cell analysis, TSA-based sequential multiplex immunofluorescence (mIF) was performed on tumor FFPE sections using MPO as a myeloid marker (primarily reflecting granulocytic/neutrophil-related myeloid populations). Two three-color panels were applied: (i) MPO + SPHK1 + DAPI and (ii) MPO + ASPM + DAPI. Representative images showed that SPHK1 and ASPM signals were detectable in a subset of MPO^+ cells and exhibited partial co-localization or spatial proximity to MPO, supporting their localization within MPO^+ myeloid compartments in the tumor microenvironment (**Figure 9A**, **9B**).

To further evaluate the clinical relevance of this SPHK1/ASPM-marked MPO^+ myeloid niche, an independent cohort of patients treated with anti-PD-1 therapy was analyzed and stratified into responders and non-responders according to clinical response (**Figure 9C**). Pretreatment FFPE tumor sections were subjected to TSA-mIF using the same staining panels, and digital pathology quantification was performed using the HALO platform to assess MPO^+ASPM^+ and

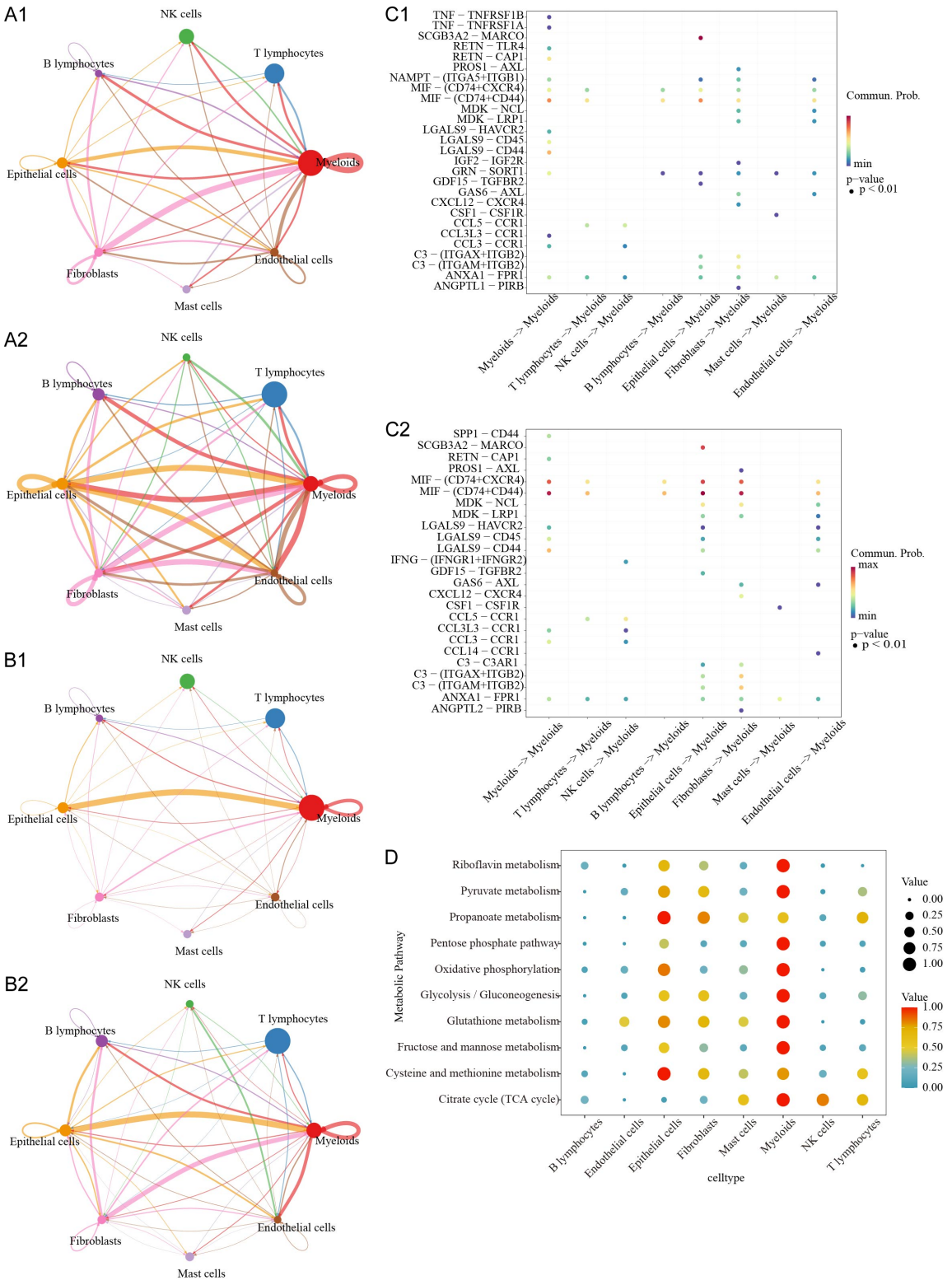

Figure 7. Intercellular communication and metabolic activity analysis. A, B. Analysis of intercellular communication frequency and intensity across 8 cell types in the LUAD group and the control group. C. Analysis of ligand-receptor pairs of myeloids in the control and LUAD groups. D. Bubble plot of high-activity metabolic pathways in 8 cell types.

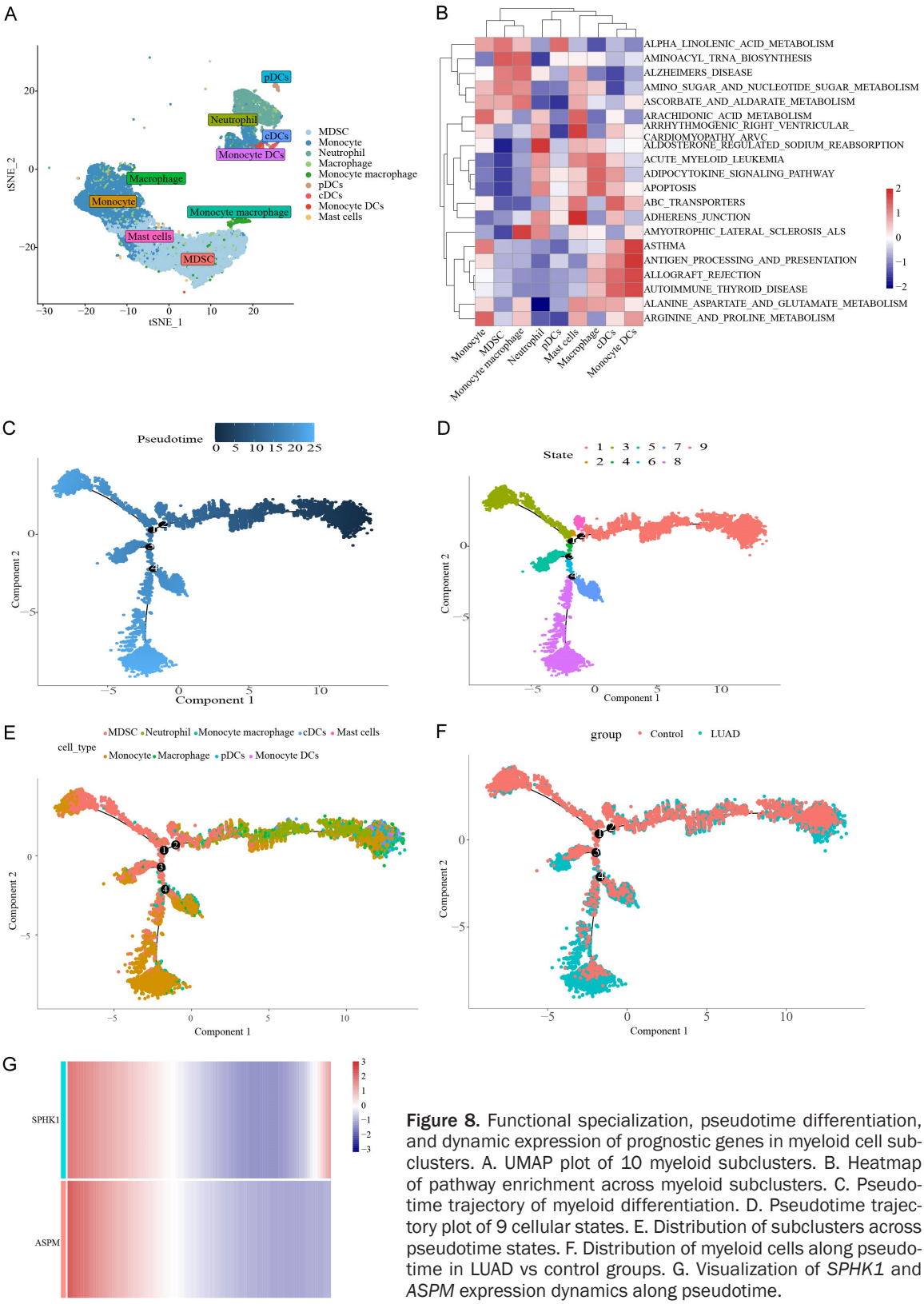

Figure 8. Functional specialization, pseudotime differentiation, and dynamic expression of prognostic genes in myeloid cell subclusters. A. UMAP plot of 10 myeloid subclusters. B. Heatmap of pathway enrichment across myeloid subclusters. C. Pseudotime trajectory of myeloid differentiation. D. Pseudotime trajectory plot of 9 cellular states. E. Distribution of subclusters across pseudotime states. F. Distribution of myeloid cells along pseudotime in LUAD vs control groups. G. Visualization of *SPHK1* and *ASPM* expression dynamics along pseudotime.

$MPO^{+}SPHK1^{+}$ co-localized cells (reported using a consistent output metric, e.g., counts/density/proportion, as applicable). Notably, compared with responders, non-responders ex-

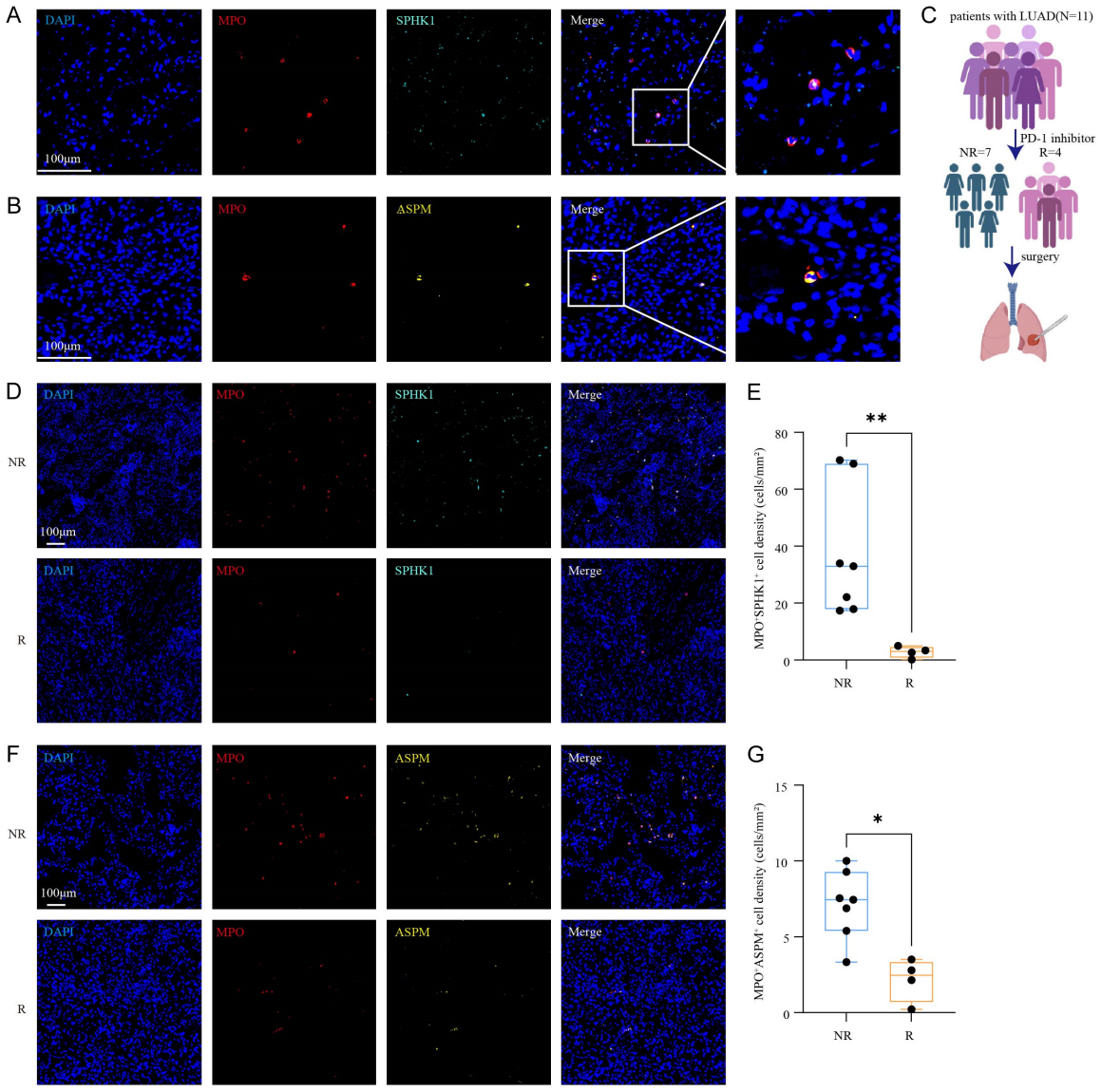

Figure 9. TSA-mIF and HALO-based quantification in a clinical anti-PD-1 cohort. A. Representative co-localization image (MPO + SPHK1 + DAPI; include a zoomed-in view; original magnification, ×20). B. Representative co-localization image (MPO + ASPM + DAPI; include a zoomed-in inset; original magnification, ×20). C. Schematic illustration of the clinical anti-PD-1-treated LUAD cohort, including patient enrollment, response stratification into pathological responders (R, n = 4) and non-responders (NR, n = 7), and surgical resection for tissue analysis. D. Representative images from the anti-PD-1-treated clinical cohort showing SPHK1 and MPO co-staining in responders (n = 4) and non-responders (n = 7) (original magnification, ×6.4). E. HALO-based quantitative analysis of MPO^+SPHK1^+ co-localized cells in responders (n = 4) versus non-responders (n = 7). F. Representative images from the anti-PD-1-treated clinical cohort showing ASPM and MPO co-staining in responders (n = 4) and non-responders (n = 7) (original magnification, ×6.4). G. HALO-based quantitative analysis of MPO^+ASPM^+ co-localized cells in responders (n = 4) versus non-responders (n = 7).

hibited a significant increase in abundance of MPO^+ASPM^+ and MPO^+SPHK1^+ co-localized cells (both $P < 0.05$) (**Figure 9D-G**). Collectively, these findings extend the integrative evidence from transcriptomic and single-cell analyses to a clinically relevant context, suggesting that the SPHK1/ASPM-marked MPO^+ myeloid compartment is associated with poor response to PD-1 blockade in LUAD.

Discussion

LUAD, the most prevalent subtype of NSCLC, poses a critical need for novel prognostic bio-

markers due to its high incidence, mortality, and significant molecular heterogeneity [1, 47]. Calcium, the highly regulated intracellular messenger, and its primary effector protein CaM play central roles in cancer cell proliferation, migration, and apoptosis [48]. However, the overall role of CRGs in calcium signaling within LUAD remains incompletely understood. In this study, we identified *SPHK1* and *ASPM* as calcium signaling-related prognostic biomarkers in LUAD and constructed a risk signature based on these two genes. This model stratified LUAD patients into different survival subgroups, and risk score, pathological stage, and T stage were validated as independent adverse prognostic factors. Furthermore, integration of these variables with TME status enabled improved prognostic assessment and potential therapeutic stratification. A nomogram incorporating these factors demonstrated robust performance in individualized survival prediction. Within the immunosuppressed TME, characterized by elevated γδ T cells and upregulation of immune-checkpoint receptors such as PD-1, high-risk patients exhibited substantial immune modulatory alterations. Importantly, the dynamic expression pattern of *SPHK1* downregulation and *ASPM* upregulation during cellular differentiation was predominantly observed in myeloid cells within the TME. To validate these findings, multiplex immunofluorescence further demonstrated increased enrichment of MPO^+SPHK1^+ and MPO^+ASPM^+ myeloid compartments in biopsies from patients who did not respond to anti-PD-1 therapy, suggesting a potential association between this distinct myeloid niche and immunotherapy resistance. In conclusion, the SPHK1-ASPM-based signature represents a promising tool for LUAD risk stratification and probably serves as a basis for therapeutic decision-making in LUAD.

We identified 136 candidate genes closely associated with both LUAD and CaM. Functional enrichment analysis showed that these genes were enriched in biological processes such as muscle contraction, calcium ion transmembrane transport, and actin cytoskeleton organization, as well as molecular functions including CaM binding and the calcium signaling pathway [5, 49]. Previous reports state that calcium signaling and its downstream effectors play essential roles in tumor proliferation, migration, and therapeutic resistance, particularly through the regulation of cytoskeletal remodeling and actin dynamics involved in cancer cell invasion and metastasis [50]. The gene enrichment patterns observed in our study are consistent with these findings, suggesting that LUAD cells may modulate calcium homeostasis and cytoskeletal organization through the identified candidate genes.

Two prognostic genes, *SPHK1* and *ASPM*, were identified in this study. *SPHK1* has a recognized CaM-binding motif, and its activity is regulated by CaM. Furthermore, there is a positive feedback loop between *SPHK1* and Ca^{2+} signaling [51], supporting its classification as a CRG. Functionally, *SPHK1* catalyzes the production of the bioactive lipid, sphingosine-1-phosphate (S1P), thus activating multiple signaling pathways that promote proliferation, inhibit apoptosis, and enhance migration; it is widely expressed in malignant tumors and is correlated with poor prognosis [52]. Mechanistically, in LUAD, *HOXC11* directly binds to the *SPHK1* promoter and upregulates *SPHK1* transcription, thereby promoting LUAD cell proliferation, migration, and invasion [53]. In this study, SPHK1 was significantly upregulated in LUAD tissues and identified as a prognostic risk factor. Patients in high-risk group exhibited significantly poorer OS, consistent with previous reports. SPHK1 may also contribute to tumor progression through modulation of the TME. Immune infiltration analysis demonstrated positive correlations between SPHK1 expression and abundance of NKT cells and γδ T cells, suggesting that *SPHK1* may be involved in reshaping the immune microenvironment in this patient population [54].

ASPM, a mitotic spindle assembly factor, promotes tumor proliferation and stemness by modulating microtubule dynamics, cell cycle progression, and DNA damage repair [50]. Several microtubule-associated proteins (MAPs) are known to act as downstream effectors of Ca^{2+}/CaM signaling pathway, linking calcium signaling to cytoskeletal and cell division [49]. *ASPM* has been reported to bind CaM, functioning as an effector linking Ca^{2+}/CaM signaling to spindle microtubule dynamics [55, 56]. This justifies the inclusion of *ASPM* in the CRGs from the perspective of structure and function. Previous studies have demonstrated that *ASPM* is highly expressed in LUAD tissues

and is associated with increased tumor size, invasiveness, and poorer prognosis. Silencing *ASPM* has been shown to inhibit LUAD proliferation and migration [57]. Consistent with these findings, in the current study, *ASPM* was identified as a prognostic risk gene associated with adverse outcome. Immune correlation analysis showed that ASPM expression was positively correlated with Th2 cells, and negatively correlated with eosinophils and monocytes, suggesting a potential link between *ASPM* expression and immune microenvironment remodeling. In addition, aberrant *ASPM* expression has been reported to influence tumor cell radiosensitivity via microtubule destabilization and chromosome missegregation [54].

Importantly, although both *SPHK1* and *ASPM* are associated with CaM, they act via different pathways. Specifically, *SPHK1* primarily mediates lipid-calcium signaling crosstalk and is involved in the regulation of the inflammatory and immune processes within the tumor microenvironment, with associations observed for NKT cells and γδ T cells. In contrast, *ASPM* modulates cell cycle and genomic stability via modulation of calcium-microtubule dynamics, potentially associated with immune features such as Th2 polarization and innate immune suppression. This functional complementarity likely underpins the strong predictive performance of the combined SPHK1-ASPM risk model. We postulate that SPHK1 links S1P-mediated lipid signaling to the Ca^{2+} pathway; whereas ASPM mediates calcium-associated regulation of cell proliferation and genomic stability in LUAD through the mitotic apparatus and cytoskeletal dynamics. Together, these molecules may represent key components of CaM-related signaling and potential targets for combinatorial intervention strategies targeting the “lipid metabolism-calcium signal-cell cycle” axis. The SPHK1-ASPM-based model enables robust risk stratification and provides a framework for further mechanistic investigation into CaM-related pathways in LUAD.

To further elucidate the biological basis for risk stratification, we compared the two risk groups in terms of immune microenvironment characteristics, signaling pathway activity, and predicted treatment response. Immune infiltration analysis demonstrated higher immune scores and greater overall immune cell infiltration in the low-risk group, consistent with a lymphocyte-rich immune-inflammatory phenotype [58, 59]. Specifically, activated B cells, Tfh cells, and plasma cells were significantly enriched in the low-risk group, whereas these populations were reduced in the high-risk group, indicative of disruption of the Tfh-germinal center B cell-TLS axis, which is important for effective antitumor immunity [60, 61]. In addition, eosinophils were enriched in the low-risk group but depleted in the high-risk group, suggesting impairment of the innate immune response in high-risk tumors [62, 63]. Reduced infiltration of monocytes in the high-risk group further contribute to an immunosuppressive microenvironment, although previous reports indicate that MDSCs can promote immunosuppression in NSCLC patients [64, 65]. Interestingly, certain immune cell types, such as Th2 cells, NKT cells, and γδ cells, were enriched in high-risk group. This counter-intuitive finding highlights the complexity of immune remodeling within the TME. The enrichment of Th2 cells suggests a shift from antitumor Th1-type immunity toward a protumor Th2-dominant response [66]. Meanwhile, although NKT and γδ T cells possess intrinsic antitumor potential, their increased abundance in the high-risk group may reflect functional impairment or exhaustion within a chronically immunosuppressive TME [67]. TIDE analysis confirms T cell dysfunction in high-risk groups. Upregulation of immune checkpoint-related molecules, including HLA-E, LGALS9, CEACAM1, may facilitate immune escape, while alterations in *HLA* class II-related genes may disrupt antigen presentation. In contrast, higher expression of *HLA-DRA*, *CD276*, *PVR* in the low-risk group indicates robust antigen presentation and controlled immune activation [68-71].

Based on the GSE131907 dataset, we constructed a single-cell atlas of LUAD. Analysis revealed an increased proportion of adaptive immune cells (B/T lymphocytes) and a decreased proportion of innate immune cells (myeloid cells, NK cells) within tumor tissues, reflecting reprogramming of the innate immune compartment [72, 73]. Myeloid cells (especially TAMs) play a pivotal role in maintaining immune homeostasis and therapeutic response. In this context, the co-expression of SPHK1 and ASPM in myeloid populations represents a key biolo-

gical feature [74-76]. Spatial multiplex immunofluorescence confirmed the co-localization of SPHK1/ASPM proteins with MPO^+ myeloid cells, validating the single-cell RNA sequencing findings at the tissue level. Importantly, analysis of an independent cohort of patients treated with anti-PD-1 therapy demonstrated a significant enrichment of MPO^+ASPM^+ and MPO^+SPHK1^+ myeloid cells in non-responders, suggesting an association between this myeloid compartment and resistance to immune checkpoint blockade. Consistently, high-risk patients exhibited elevated tumor immune dysfunction scores and aberrant immune checkpoint profiles (decreased *HLA-DRA* and increased *HLA-E*) - a profile consistent with T-cell exclusion rather than effective T-cell activation [68-71]. Mechanistically, these observations may be linked to signaling pathways identified in our integrative analyses. Single-cell communication analysis identified epithelial-myeloid crosstalk mediated by the *MIF* - (CD74+CD44) axis, which has been implicated in the induction of immunosuppressive myeloid populations [77-79]. In parallel, metabolic profiling even suggested that myeloid cells simultaneously display glycolytic and oxidative phosphorylation features, favoring their survival and function within hypoxic and immunosuppressive tumor niches [80-82]. As MPO labels most granulocytic lineage cells, these findings are in line with emerging evidence implicating neutrophil-like myeloid populations in immunosuppression and reduced responsiveness to immune checkpoint inhibitors [83]. Nevertheless, it should be emphasized that our findings are primarily associative rather than causal. The density of $MPO^+ASPM^+/SPHK1^+$ myeloid cells may act as a potential tissue-based biomarker. Integration of this marker into clinical pathology workflows may facilitate the identification of patients who may benefit from myeloid-targeted combination therapies rather than PD-1 inhibitor monotherapy.

This study is the first to systematically define the roles of CRG in LUAD by integrating large-scale transcriptomic data with single-cell atlases. We identified two core prognostic genes, SPHK1 and ASPM, and developed a robust risk stratification model. The risk score was validated as an independent prognostic factor. Multidimensional analyses revealed distinct phenotypes between risk groups. The low-risk group displayed immune-activated “hot tumor” features, whereas the high-risk group exhibited an immunosuppressive microenvironment, enhanced proliferative-related pathway activation (e.g., cell cycle and DNA replication; NES > 2.0), and a higher frequency of TP53 mutations (65%). Single-cell analysis further demonstrated co-expression of SPHK1 and ASPM in myeloid cells, with opposite differentiation trajectories, suggesting their potential involvement in shaping the tumor microenvironment via metabolic reprogramming. Importantly, multiplex immunofluorescence in an anti-PD-1-treated cohort demonstrated selective enrichment of MPO^+ASPM^+ and MPO^+SPHK1^+ myeloid cells in non-responders, implicating an association between this compartment and treatment resistance. Collectively, our findings provide a molecular basis for prognostic assessment and risk-stratified therapeutic strategies in LUAD, with potential implications for identifying patients who are more likely to benefit from immunotherapy.

Several limitations should be acknowledged. First, the transcriptomic analyses were primarily based on a single public dataset, lacking multi-center, prospective clinical validation. Second, the mechanistic studies largely relied on bioinformatic analyses and were not supported by in vitro or in vivo validation of SPHK1/ASPM-mediated regulation of myeloid cell function. Third, the single cell analysis was also based on a single small-sample dataset. Future studies should address these gaps by incorporating larger, multi-center cohorts, functional validation experiments, and integrative multi-omics approaches to enhance the robustness and clinical applicability of these findings.

Conclusion

SPHK1 and *ASPM* can serve as key prognostic genes in LUAD. A risk model based on these genes demonstrated stable predictive performance across independent cohorts. The risk score was validated as an independent prognostic factor (HR = 1.02). The low-risk group was characterized by an immune-activated phenotype with enrichment of Tfhc cells and B cells, whereas the high-risk group exhibited an immunosuppressive state, elevated TIDE scores, and activation of proliferation-related pathways (NES > 2.0), along with increased sensitivity to cell cycle-targeting agents.

Single-cell analyses revealed co-expression of *SPHK1* and *ASPM* in myeloid cells, with divergent expression patterns along differentiation trajectories, suggesting their involvement in tumor progression through modulation of myeloid cell states. Notably, $MPO^{+}ASPM^{+}$ and $MPO^{+}SPHK1^{+}$ cells were significantly enriched in anti-PD-1 non-responders, indicating a potential role of this myeloid compartment in both treatment response and associated poor prognosis.

Acknowledgements

This work was supported by grants from National Natural Science Foundation of China (82471811, 82071799, 81972683, 82388201) and Shanghai Municipal Health Commission Collaborative Innovation Cluster Project (Grant No. 2020CXJQ03).

Disclosure of conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Abbreviations

LUAD, Lung Adenocarcinoma; CRGs, Calmodulin-Related Genes; TCGA, The Cancer Genome Atlas; GEO, Gene Expression Omnibus; OS, Overall Survival; AUC, Area Under the Curve; TIDE, Tumor Immune Dysfunction and Exclusion; MDSCs, Myeloid-Derived Suppressor Cells; DEGs, Differentially Expressed Genes; KEGG, Kyoto Encyclopedia of Genes and Genomes; GO, Gene Ontology; PPI, Protein-Protein Interaction; scRNA-seq, Single-Cell RNA Sequencing; GSEA, Gene Set Enrichment Analysis; ssGSEA, Single-Sample Gene Set Enrichment Analysis; IC50, Half Maximal Inhibitory Concentration; GDSC, Genomics of Drug Sensitivity in Cancer; NSCLC, Non-Small Cell Lung Cancer; HR, Hazard Ratio; FFPE, Formalin-Fixed Paraffin-Embedded; RSF, random survival forest; TME, tumor microenvironment; mIF, multiplex immunofluorescence.

Address correspondence to: Yuchao Dong and Chong Bai, Department of Respiratory Medicine Changhai Hospital Affiliated to Naval Medical University, No. 168 Changhai Road, Shanghai, China. E-mail: Dongyc1020@aliyun.com (YCD); ChongBai@smmu.edu.cn (CB); Meng Guo, National Key Laboratory of Immunity and Inflammation & Institute of Immunology, Naval Medical University, No. 800 Xiangyin Road, Shanghai, China. E-mail: Guo918meng@163.com

References

[1] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A and Bray F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021; 71: 209-249.

[2] Herbst RS, Morgensztern D and Boshoff C. The biology and management of non-small cell lung cancer. Nature 2018; 553: 446-454.

[3] Miller KD, Nogueira L, Mariotto AB, Rowland JH, Yabroff KR, Alfano CM, Jemal A, Kramer JL and Siegel RL. Cancer treatment and survivorship statistics, 2019. CA Cancer J Clin 2019; 69: 363-385.

[4] Lahiri A, Maji A, Potdar PD, Singh N, Parikh P, Bisht B, Mukherjee A and Paul MK. Lung cancer immunotherapy: progress, pitfalls, and promises. Mol Cancer 2023; 22: 40.

[5] Song KJ, Choi S, Kim K, Hwang HS, Chang E, Park JS, Shim SB, Choi S, Heo YJ, An WJ, Yang DY, Cho KC, Ji W, Choi CM, Lee JC, Kim H, Yoo J, Ahn HS, Lee GH, Hwa C, Kim S, Kim K, Kim MS, Paek E, Na S, Jang SJ, An JY and Kim KP. Proteogenomic analysis reveals non-small cell lung cancer subtypes predicting chromosome instability, and tumor microenvironment. Nat Commun 2024; 15: 10164.

[6] Li L, Lian X, Ding L, Guo R, Xu J, Bai R, Yi Y, Li X, Chen X, Zheng H, Gao J, Jim Xiao ZX and Niu M. Calcium triggers calmodulin degradation to induce EGF receptor instability and overcome non-small cell lung cancer resistance to tyrosine kinase inhibitors. J Biol Chem 2025; 301: 110305.

[7] Hussey JW, Limpitikul WB and Dick IE. Calmodulin mutations in human disease. Channels 2023; 17: 2165278.

[8] Thiel G and Rössler OG. Calmodulin regulates transient receptor potential TRPM3 and TRPM8-induced gene transcription. Int J Mol Sci 2023; 24: 7902.

[9] Feng J, Chen Z, Wang G, Yao Y, Min X, Luo J and Xie K. Prognostic significance of calcium-related genes in lung adenocarcinoma and the role of TNNC1 in macrophage polarization and erlotinib resistance. Front Immunol 2025; 16: 1509222.

[10] Brzozowski JS and Skelding KA. The multifunctional calcium/calmodulin stimulated protein kinase (CaMK) family: emerging targets for anti-cancer therapeutic intervention. Pharmaceuticals (Basel) 2019; 12: 8.

[11] Wang SQ, Liu J, Qin J, Zhu Y, Tin VP, Yam JWP, Wong MP and Xiao ZJ. CAMK2A supported tumor initiating cells of lung adenocarcinoma by upregulating SOX2 through EZH2 phosphorylation. Cell Death Dis 2020; 11: 410.
[12] Xiao ZJ, Liu J, Wang SQ, Zhu Y, Gao XY, Tin VP, Qin J, Wang JW and Wong MP. NFATc2 enhances tumor-initiating phenotypes through the NFATc2/SOX2/ALDH axis in lung adenocarcinoma. Elife 2017; 6: e26733.
[13] Wang L, Ma Y, Zhang S, Yang Y and Huang B. NFATc2 promotes lactate and M2 macrophage polarization through USP17 in lung adenocarcinoma. Anticancer Drugs 2024; 35: 385-396.
[14] Tokumitsu H and Sakagami H. Molecular mechanisms underlying Ca^{2+}/calmodulin-dependent protein kinase kinase signal transduction. Int J Mol Sci 2022; 23: 11025.
[15] Liang W, Gao R, Yang M, Wang X, Cheng K, Shi X, He C, Li Y, Wu Y, Shi L, Chen J and Yu X. MARCKSL1 promotes the proliferation, migration and invasion of lung adenocarcinoma cells. Oncol Lett 2020; 19: 2272-2280.
[16] Zhang C, Zhang Z, Wu Y, Wu Y, Cheng J, Luo K, Li Z, Zhang M, Wang J, Zhang X and Li Y. Phosphodiesterase 1A physically interacts with YTHDF2 and reinforces the progression of non-small cell lung cancer. Elife 2025; 13: RP98903.
[17] Liu A, Liu G, Wang X, Yan D, Zhang J and Wei L. Comprehensive analysis of single-cell RNA and bulk RNA sequencing based on M2 tumor-associated macrophage and angiogenesis-related genes to assess prognosis and therapeutic response in lung adenocarcinoma. Heliyon 2024; 10: e34784.
[18] Wei H, Teng F, Wang X, Hou X, Wang H, Wang H, Sun H and Zhou X. Identification of a prognosis-related gene signature and ceRNA regulatory networks in lung adenocarcinoma. Heliyon 2024; 10: e28084.
[19] Bischoff P, Trinks A, Obermayer B, Pett JP, Wiederspahn J, Uhlitz F, Liang X, Lehmann A, Jurmeister P, Elsner A, Dziodzio T, Rückert JC, Neudecker J, Falk C, Beule D, Sers C, Morkel M, Horst D, Blüthgen N and Klauschen F. Single-cell RNA sequencing reveals distinct tumor microenvironmental patterns in lung adenocarcinoma. Oncogene 2021; 40: 6748-6758.
[20] Deng Y, Xia L, Zhang J, Deng S, Wang M, Wei S, Li K, Lai H, Yang Y, Bai Y, Liu Y, Luo L, Yang Z, Chen Y, Kang R, Gan F, Pu Q, Mei J, Ma L, Lin F, Guo C, Liao H, Zhu Y, Liu Z, Liu C, Hu Y, Yuan Y, Zha Z, Yuan G, Zhang G, Chen L, Cheng Q, Shen S and Liu L. Multicellular ecotypes shape progression of lung adenocarcinoma from ground-glass opacity toward advanced stages. Cell Rep Med 2024; 5: 101489.
[21] Zhan X and Li N. Editorial: new molecular targets involved in lung adenocarcinoma. Front Endocrinol 2023; 14: 1138849.
[22] Cheng Y, Hou K, Wang Y, Chen Y, Zheng X, Qi J, Yang B, Tang S, Han X, Shi D, Wang X, Liu Y, Hu X and Che X. Identification of prognostic signature and gliclazide as candidate drugs in lung adenocarcinoma. Front Oncol 2021; 11: 665276.
[23] Zheng S, Yao L, Li F, Huang L, Yu Y, Lin Z, Li H, Xia J, Lanuti M and Zhou H. Homologous recombination repair rathway and RAD54L in early-stage lung adenocarcinoma. PeerJ 2021; 9: e10680.
[24] Cao S, Lun S, Duan L, Gao Z, Wang X, Li Y and Zhang Y. Harnessing calmodulin-related genes to build a prognostic model in esophageal squamous cell carcinoma for a comprehensive analysis of single-cell immune characteristics and drug efficacy. J Immunother 2025; 48: 244-257.
[25] Love MI, Huber W and Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014; 15: 550.
[26] Zhou W, Li H, Zhang J, Liu C, Liu D, Chen X, Ouyang J, Zeng T, Peng S, Ouyang F, Long Y and Li Y. Identification and mechanism analysis of biomarkers related to butyrate metabolism in COVID-19 patients. Ann Med 2025; 57: 2477301.
[27] Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, Feng T, Zhou L, Tang W, Zhan L, Fu X, Liu S, Bo X and Yu G. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. Innovation (Camb) 2021; 2: 100141.
[28] Gu Z, Gu L, Eils R, Schlesner M and Brors B. *circlize* implements and enhances circular visualization in R. Bioinformatics 2014; 30: 2811-2812.
[29] Zhang Y, Liang S, Zhang Y, Liu M and Zhang K. Identification of a novel endocytosis-associated gene signature for prognostic prediction in lung adenocarcinoma. Oncol Lett 2023; 26: 511.
[30] Friedman J, Hastie T and Tibshirani R. Regularization paths for generalized linear models via coordinate descent. J Stat Softw 2010; 33: 1-22.
[31] Zhao P, Zhen H, Zhao H, Huang Y and Cao B. Identification of hub genes and potential molecular mechanisms related to radiotherapy sensitivity in rectal cancer based on multiple datasets. J Transl Med 2023; 21: 176.
[32] Liu TT, Li R, Huo C, Li JP, Yao J, Ji X and Qu YQ. Identification of CDK2-related immune forecast model and ceRNA in lung adenocarcinoma, a pan-cancer analysis. Front Cell Dev Biol 2021; 9: 682002.

[33] Sui Z, Wu X, Du L, Wang H, Yuan L, Zhang JV and Yu Z. Characterization of the immune cell infiltration landscape in esophageal squamous cell carcinoma. Front Oncol 2022; 12: 879326.
[34] Sachs MC. plotROC: a tool for plotting ROC curves. J Stat Softw 2017; 79: 2.
[35] Hänzelmann S, Castelo R and Guinney J. GSVA: gene set variation analysis for microarray and RNA-Seq data. BMC Bioinformatics 2013; 14: 7.
[36] Chen Z, Cao W, Luo J, Abdelrahman Z, Lu Q, Wang H and Wang X. Gene set enrichment analysis identifies immune subtypes of kidney renal clear cell carcinoma with significantly different molecular and clinical properties. Front Immunol 2023; 14: 1191365.
[37] Revelle W. psych: Procedures for Psychological, Psychometric, and Personality Research (Version 2.2.9) [Computer software]. Evanston, Illinois: Northwestern University; 2022.
[38] Hu FF, Liu CJ, Liu LL, Zhang Q and Guo AY. Expression profile of immune checkpoint genes and their roles in predicting immunotherapy response. Brief Bioinform 2021; 22: bbaa176.
[39] Mayakonda A, Lin DC, Assenov Y, Plass C and Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. Genome Res 2018; 28: 1747-1756.
[40] Geeleher P, Cox N and Huang RS. pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels. PLoS One 2014; 9: e107468.
[41] Hao Y, Stuart T, Kowalski MH, Choudhary S, Hoffman P, Hartman A, Srivastava A, Molla G, Madad S, Fernandez-Granda C and Satija R. Dictionary learning for integrative, multimodal and scalable single-cell analysis. Nat Biotechnol 2024; 42: 293-304.
[42] Kim N, Kim HK, Lee K, Hong Y, Cho JH, Choi JW, Lee JI, Suh YL, Ku BM, Eum HH, Choi S, Choi YL, Joung JG, Park WY, Jung HA, Sun JM, Lee SH, Ahn JS, Park K, Ahn MJ and Lee HO. Single-cell RNA sequencing demonstrates the molecular and cellular reprogramming of metastatic lung adenocarcinoma. Nat Commun 2020; 11: 2285.
[43] Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Ramos R, Kuan CH, Myung P, Plikus MV and Nie Q. Inference and analysis of cell-cell communication using CellChat. Nat Commun 2021; 12: 1088.
[44] Griss J, Viteri G, Sidiropoulos K, Nguyen V, Fabregat A and Hermjakob H. ReactomeGSA - efficient multi-omics comparative pathway analysis. Mol Cell Proteomics 2020; 19: 2115-2125.
[45] Li J, Shyr Y and Liu Q. aKNNO: single-cell and spatial transcriptomics clustering with an optimized adaptive k-nearest neighbor graph. Genome Biol 2024; 25: 203.
[46] Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA and Trapnell C. Reversed graph embedding resolves complex single-cell trajectories. Nat Methods 2017; 14: 979-982.
[47] Satpathy S, Clark NM, Chen YJ, Hosseini N, Chang YH, Hsiao Y, Lei JT, Petralia F, Chen JS, Geffen Y, Heiman DI, Paul I, Cho H, Hollenberg M, Marino GB, Lin KT, Mannan R, White CJ, Allen J, Avanessian SC, Kane MH, Wolfe A, Kinarivala M, Liu W, Anand S, Lin MW, Haines M, Bergstrom EJ, Hussey G, Li GX, Mani DC, Fang H, Jaehnig EJ, Keshishian H, Miller B, Su KY, Hsiao YJ, Hsu HH, Hsieh MS, Hsu KH, Monovoukas A, Gohsman S, Thorup JR, Deng Y, Akiyama Y, Deng E, Sheng-Wen Chen E, Krek A, Espinoza R, Ma W, Charytonowicz D, Sebra R, Lin JH, Chen YS, Hsu YC, Lin ZS, Chen KC, Yeh CW, Wang YT, Lazar AJ, Mesri M, An E, Zhang X, Clauser KR, Fenyö D, Chinnaiyan AM, Zhang B, Ding L, Ruggles K, Newton C, Zhang H, Wang P, Hostetter G, Omenn GS, Kumar-Sinha C, Thiagarajan M, Govindan R, Paik P, Parolia A, Li QK, Ma'ayan A, Getz GA, Dhanasekaran SM, Robles AI, Chang GC, Yang PC, Yu SL, Chen HY, Nesvizhskii AI, Carr SA, Mani DR, Cieslik MP, Chen YJ, Gillette MA, Lu CW, Chu CH, Shen CY, Han CL, Lin CC, Lin CY, Chen CW, Lin CH, Hsu HE, Tsai HJ, Wu JJ, Lin JW, Waniwan J, Liao KH, Chen PH, Huang PR, Huang SM, Lin TC, Lai WC, Chiu WT, Chiang XH, Chang YL, Chen YM, Chen YL, Wang YW, Lin YW, Chang YC, Huang YT, Lien Y, Zheng ZR, Hashimi AS, Mohan A, Pandey A, Pilozzi A, Webster A, Paulovich AG, Dagar A, Godwin AK, Pruetz B, Williams BO, Druker BJ, Rohrer DC, Chan DW, Petrov D, Chesla D, Davaar D, Duffy E, Wilson GD, Zhao G, Kołodziejczak I, Lubinski J, Huang J, Hafron J, Tyner JW, Koomen J, Zaalishvili K, Ketchum KA, Wiznerowicz M, Domagalski MJ, Anurag M, Borucki M, Edwards NJ, Vatanian N, Grady P, Piehowski PD, Bogdan P, Li Q, Fonseca R, Madan R, Thangudu RR, Crispen R, Matteotti R, Bremner R, Cerda S, Cottingham SL, Tsang S, Cai S, Liu T, Bauer T, Maggio WW, Jing X, Zhang Y, Shutack Y and Andric Z. Integrative analysis of lung adenocarcinoma across diverse ethnicities and exposures. Cancer Cell 2025; 43: 1731-1757, e17.
[48] Qin Z, Di Y, Ma T, Zeng W, Liu X and He W. The calcium homeostasis in tumor and the mechanism involving progression and metastasis. Cancer Lett 2025; 630: 217908.
[49] Kölling M, Kumari P and Bürstenbinder K. Calcium- and calmodulin-regulated microtubule-associated proteins as signal-integration hubs at the plasma membrane-cytoskeleton nexus. J Exp Bot 2019; 70: 387-396.

[50] Tsai KK, Bae BI, Hsu CC, Cheng LH and Shaked Y. Oncogenic ASPM is a regulatory hub of developmental and stemness signaling in cancers. Cancer Res 2023; 83: 2993-3000.
[51] Bazzazi H and Popel AS. Computational investigation of sphingosine kinase 1 (SphK1) and calcium dependent ERK1/2 activation downstream of VEGFR2 in endothelial cells. PLoS Comput Biol 2017; 13: e1005332.
[52] Zheng X, Li W, Ren L, Liu J, Pang X, Chen X, Kang D, Wang J and Du G. The sphingosine kinase-1/sphingosine-1-phosphate axis in cancer: potential target for anticancer therapy. Pharmacol Ther 2019; 195: 85-99.
[53] Peng X, Liu X, Hu W, Zhou Y, Ouyang L, Peng X, Long Y, Sun J, Tao T, Chen L, Shi Y, Tao Y, Xiao D and Liu S. HOXC11 drives lung adenocarcinoma progression through transcriptional regulation of SPHK1. Cell Death Dis 2023; 14: 153.
[54] Zhong T, Liu N, Wang J, Xie S, Liu L, Wang M, Wu F, Chen X, Xiao C, Gongye X, Wu M, Wen L, Yu J and Chen D. ASPM induces radiotherapy resistance by disrupting microtubule stability leading to chromosome malsegregation in non-small cell lung cancer. Exploration (Beijing) 2025; 5: e20230024.
[55] Jiang K, Rezabkova L, Hua S, Liu Q, Capitani G, Altelaar AFM, Heck AJR, Kammerer RA, Steinmetz MO and Akhmanova A. Microtubule minus-end regulation at spindle poles by an ASPM-katanin complex. Nat Cell Biol 2017; 19: 480-492.
[56] Xu XL, Ma W, Zhu YB, Wang C, Wang BY, An N, An L, Liu Y, Wu ZH and Tian JH. The microtubule-associated protein ASPM regulates spindle assembly and meiotic progression in mouse oocytes. PLoS One 2012; 7: e49303.
[57] Feng Z, Zhang J, Zheng Y, Liu J, Duan T and Tian T. Overexpression of abnormal spindle-like microcephaly-associated (ASPM) increases tumor aggressiveness and predicts poor outcome in patients with lung adenocarcinoma. Transl Cancer Res 2021; 10: 983-997.
[58] Ren C, Li J, Zhou Y, Zhang S and Wang Q. Typical tumor immune microenvironment status determine prognosis in lung adenocarcinoma. Transl Oncol 2022; 18: 101367.
[59] Qiu L, Yang Z, Jia G, Liang Y, Du S, Zhang J, Liu M, Zhao X and Jiao S. Clinical significance and immune landscape of a novel immune cell infiltration-based prognostic model in lung adenocarcinoma. Heliyon 2024; 10: e33109.
[60] Wu X, Huang Q, Chen X, Zhang B, Liang J and Zhang B. B cells and tertiary lymphoid structures in tumors: immunity cycle, clinical impact, and therapeutic applications. Theranostics 2025; 15: 605-631.
[61] Liu W, You W, Lan Z, Ren Y, Gao S, Li S, Chen WW, Huang C, Zeng Y, Xiao N, Wang Z, Xie H, Ma H, Chen Y, Wang G, Chen C and Li H. An immune cell map of human lung adenocarcinoma development reveals an anti-tumoral role of the Tfh-dependent tertiary lymphoid structure. Cell Rep Med 2024; 5: 101448.
[62] Davoine F and Lacy P. Eosinophil cytokines, chemokines, and growth factors: emerging roles in immunity. Front Immunol 2014; 5: 570.
[63] Grisaru-Tal S, Jacobsen EA and Munitz A. Evolving role for eosinophils in cancer: from bench to bedside. Trends Cancer 2025; 11: 862-876.
[64] Gao Z, Liu S, Xiao H, Li M, Ren W, Xu L and Peng Z. IRF8 deficiency-induced myeloid-derived suppressor cell promote immune evasion in lung adenocarcinoma. J Transl Med 2024; 22: 678.
[65] Lin S, Zhang X, Huang G, Cheng L, Lv J, Zheng D, Lin S, Wang S, Wu Q, Long Y, Li B, Wei W, Liu P, Pei D, Li Y, Wen Z, Cui S, Li P, Sun X, Wu Y and Yao Y. Myeloid-derived suppressor cells promote lung cancer metastasis by CCL11 to activate ERK and AKT signaling and induce epithelial-mesenchymal transition in tumor cells. Oncogene 2021; 40: 1476-1489.
[66] Wang LH, Wang LL, Zhang J, Zhang P and Li SZ. Th1/Th2 and Treg/Th17 cell balance in peripheral blood of patients with ovarian cancer. Nan Fang Yi Ke Da Xue Xue Bao 2017; 37: 1066-1070.
[67] Dieli F. Tumor-infiltrating γδ T lymphocytes: pathogenic role, clinical significance, and differential programing in the tumor microenvironment. Front Immunol 2014; 5: 607.
[68] Liu X, Song J, Zhang H, Liu X, Zuo F, Zhao Y, Zhao Y, Yin X, Guo X, Wu X, Zhang H, Xu J, Hu J, Jing J, Ma X and Shi H. Immune checkpoint HLA-E:CD94-NKG2A mediates evasion of circulating tumor cells from NK cell surveillance. Cancer Cell 2023; 41: 272-287, e9.
[69] Liu X, Zuo F, Song J, Tang L, Wang X, Liu X, Zhang H, Yang Z, Jing J, Ma X and Shi H. Immune checkpoints HLA-E:CD94-NKG2A and HLA-C:KIR2DL1 complementarily shield circulating tumor cells from NK-mediated immune surveillance. Cell Discov 2024; 10: 16.
[70] Lv Y, Ma X, Ma Y, Du Y and Feng J. A new emerging target in cancer immunotherapy: Galectin-9 (LGALS9). Genes Dis 2023; 10: 2366-2382.
[71] Jiang N, Yu Y, Wu D, Wang S, Fang Y, Miao H, Ma P, Huang H, Zhang M, Zhang Y, Tang Y and Li N. HLA and tumour immunology: immune escape, immunotherapy and immune-related adverse events. J Cancer Res Clin Oncol 2023; 149: 737-747.

[72] Lu H, Qian J, Cheng L, Shen Y, Chu T and Zhao C. Single-cell RNA-sequencing uncovers the dynamic changes of tumour immune microenvironment in advanced lung adenocarcinoma. BMJ Open Respir Res 2023; 10: e001878.
[73] Guo X, Deng Y, Jiang W, Li H, Luo Y, Zhang H and Wu H. Single cell transcriptomic analysis reveals tumor immune infiltration by macrophage cells gene signature in lung adenocarcinoma. Discov Oncol 2025; 16: 261.
[74] Jing ZQ, Luo ZQ, Chen SR and Sun ZJ. Heterogeneity of myeloid cells in common cancers: single cell insights and targeting strategies. Int Immunopharmacol 2024; 134: 112253.
[75] Park WY, Nguyen T, Park HJ, Lee DS, Shah N and Mandzhieva B. Pan-cancer single-cell RNA sequencing analysis refines multi-origin monocyte and macrophage lineages. 2023, https://www.researchsquare.com/article/rs-3229780/v1.
[76] Ma RY, Black A and Qian BZ. Macrophage diversity in cancer revisited in the era of single-cell omics. Trends Immunol 2022; 43: 546-563.
[77] Shi X, Leng L, Wang T, Wang W, Du X, Li J, McDonald C, Chen Z, Murphy JW, Lolis E, Noble P, Knudson W and Bucala R. CD44 is the signaling component of the macrophage migration inhibitory factor-CD74 receptor complex. Immunity 2006; 25: 595-606.
[78] Youness RA, Elemam NM, Abdelhamid AM, Mohamed AH, Elsherbiny LM, Ramzy A and Assal RA. Macrophage migration inhibitory factor (MIF) and the tumor ecosystem: a tale of inflammation, immune escape, and tumor growth. Front Immunol 2025; 16: 1636839.
[79] Cao L, Wang X, Liu X, Meng W, Guo W, Duan C, Liang X, Kang L, Lv P, Lin Q, Zhang R, Zhang X and Shen H. Tumor necrosis factor α-dependent lung inflammation promotes the progression of lung adenocarcinoma originating from alveolar type II cells by upregulating MIF-CD74. Lab Invest 2023; 103: 100034.
[80] Ramel E, Lillo S, Daher B, Fioleau M, Daubon T and Saleh M. The metabolic control of myeloid cells in the tumor microenvironment. Cells 2021; 10: 2960.
[81] Garlatti V, Consonni FM, Ballerini G, Incerti M, Balboni A, Cassatella M and Sica A. Myeloid metabolism and its role in immunotherapy of cancer. J Immunother Cancer 2025; 13: e012127.
[82] Hu C, Pang B, Lin G, Zhen Y and Yi H. Energy metabolism manipulates the fate and function of tumour myeloid-derived suppressor cells. Br J Cancer 2020; 122: 23-29.
[83] Kargl J, Zhu X, Zhang H, Yang GHY, Friesen TJ, Shipley M, Maeda DY, Zebala JA, McKay-Fleisch J, Meredith G, Mashadi-Hossein A, Baik C, Pierce RH, Redman MW, Thompson JC, Albelda SM, Bolouri H and Houghton AM. Neutrophil content predicts lymphocyte depletion and anti-PD1 treatment failure in NSCLC. JCI Insight 2019; 4: e130850.

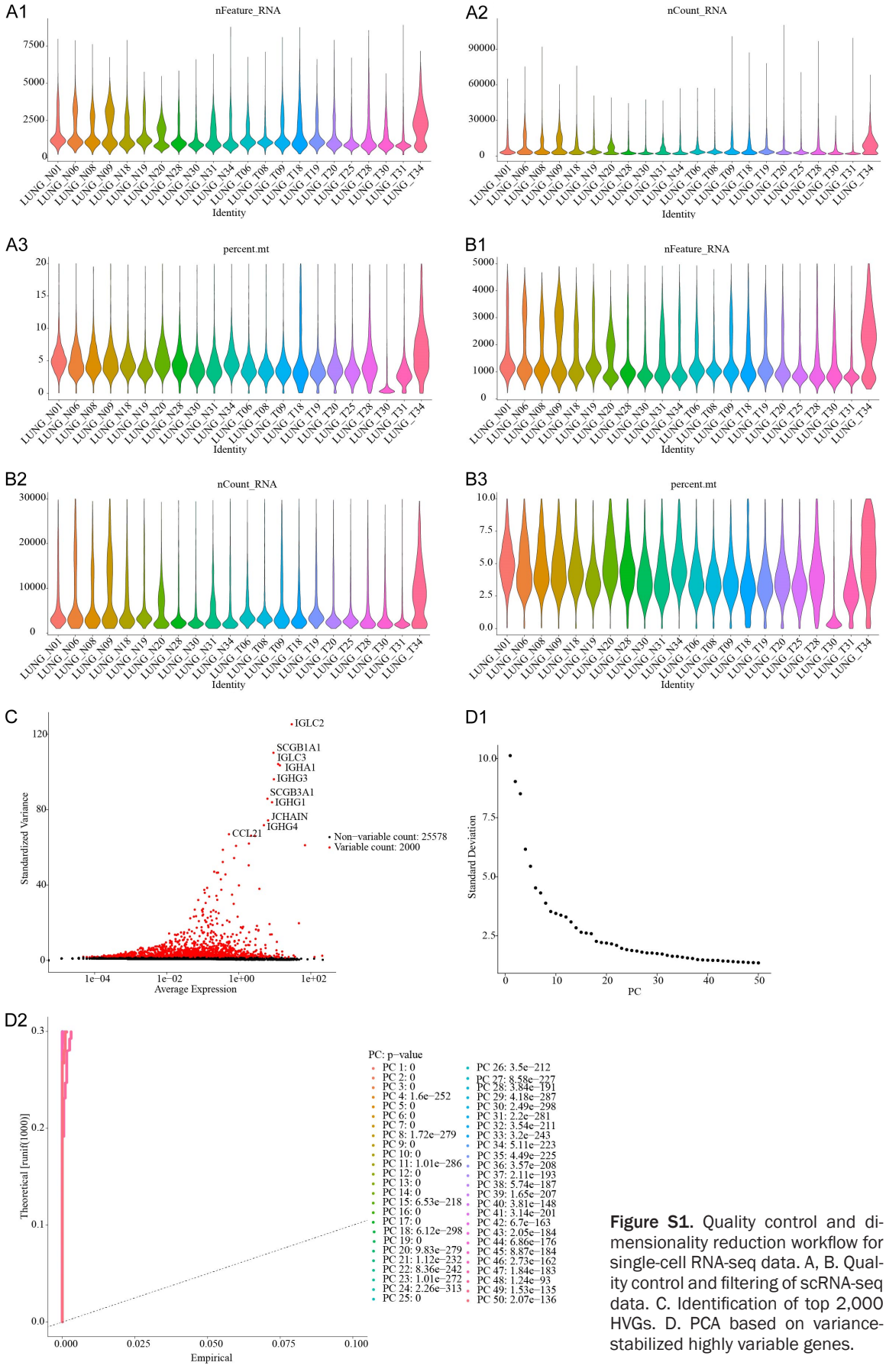

Figure S1. Quality control and dimensionality reduction workflow for single-cell RNA-seq data. A, B. Quality control and filtering of scRNA-seq data. C. Identification of top 2,000 HVGs. D. PCA based on variance-stabilized highly variable genes.