Cellular Oncology (2024) 47:2383–2405
https://doi.org/10.1007/s13402-024-01025-6

RESEARCH

Exploring tumor microenvironment interactions and apoptosis pathways in NSCLC through spatial transcriptomics and machine learning

Huimin Li[1] · Yuheng Jiao[2] · Yi Zhang[3] · Junzhi Liu[3] · Shuixian Huang[3]

Accepted: 2 December 2024 / Published online: 19 December 2024

Abstract

Background The most common type of lung cancer is non-small cell lung cancer (NSCLC), accounting for 85% of all cases. Programmed cell death (PCD), an important regulatory mechanism for cell survival and homeostasis, has become increasingly prominent in cancer research in recent years. As such, exploring the role of PCD in NSCLC may help uncover new mechanisms for therapeutic targets.

Methods We utilized the GEO database and TCGA NSCLC gene data to screen for co-expressed genes. To delve deeper, single-cell sequencing combined with spatial transcriptomics was employed to study the intrinsic mechanisms of programmed cell death in cells and their interaction with the tumor microenvironment. Furthermore, Mendelian randomization was applied to screen for causally related genes. Prognostic models were constructed using various machine learning algorithms, and multi-cohort multi-omics analyses were conducted to screen for genes. In vitro experiments were then carried out to reveal the biological functions of the genes and their relationship with apoptosis.

Results Cells with high programmed cell death activity primarily activate pathways related to apoptosis, cell migration, and hypoxia, while also exhibiting strong interactions with smooth muscle cells in the tumor microenvironment. Based on a set of programmed cell death genes, the prognostic model NSCLCPCD demonstrates strong predictive capabilities. Moreover, laboratory experiments confirm that SLC7A5 promotes the proliferation of NSCLC cells, and the knockout of SLC7A5 significantly increases tumor cell apoptosis.

Conclusions Our data indicate that programmed cell death is predominantly associated with pathways related to apoptosis, tumor metastasis, and hypoxia. Additionally, it suggests that SLC7A5 is a significant risk indicator for the prognosis of non-small cell lung cancer (NSCLC) and may serve as an effective target for enhancing apoptosis in NSCLC tumor cells.

Keywords Single-cell RNA sequencing · Spatial transcriptomics · Cell apoptosis · SLC7A5 · Biomarker

1 Introduction

The most common type of lung cancer is non-small cell lung cancer (NSCLC), accounting for 85% of all cases [1]. Despite advances in diagnosis and treatment, the prognosis for NSCLC patients remains poor, with high mortality rates primarily due to late-stage diagnosis and resistance to conventional therapies. Therefore, understanding the underlying molecular mechanisms that drive NSCLC progression and therapeutic response is crucial for developing more effective treatment strategies.

The onset, progression, and therapeutic outcomes of NSCLC are influenced by programmed cell death, such as apoptosis and autophagy [2]. The process of apoptosis

✉ Junzhi Liu
junzhil97@163.com

✉ Shuixian Huang
enthsx@163.com

[1] Department of Internal Medicine Residency Training Base, Gongli Hospital of Shanghai Pudong New Area, Shanghai 200135, China

[2] Department of Heart Failure, Shanghai East Hospital, School of Medicine, Tongji University, Shanghai 200120, China

[3] Department of Otolaryngology, Gongli Hospital of Shanghai Pudong New Area, Shanghai 200135, China

is one of the most highly regulated forms of cell death, and it is essential for maintaining tissue homeostasis and eliminating damaged or abnormal cells from the body. Dysregulation of the apoptotic pathway is a key contributor to NSCLC tumorigenesis, allowing cancer cells to evade death and proliferate uncontrollably [3]. Our study has identified an NSCLC-specific programmed cell death gene panel that is particularly enriched in tumor epithelial cells and closely associated with apoptotic pathways. Notably, tumor epithelial cells exhibiting high enrichment scores for this gene panel (PCDhighepi) also show upregulation of pathways related to cell cycle regulation, metastasis, and tumor hypoxia. This suggests a complex interplay between apoptosis and the cell cycle, where cell cycle abnormalities can trigger apoptosis, particularly through the involvement of the p53 protein, a critical mediator of the DNA damage response.

In addition to apoptosis, autophagy plays a multifaceted role in NSCLC. Autophagy can act as a tumor suppressor by degrading damaged organelles and proteins, thus maintaining cellular function in the early stages of cancer [4]. However, during tumor progression, particularly under adverse conditions like hypoxia and nutrient deprivation, autophagy may facilitate tumor cell survival and metastasis. Our research indicates that PCDhighepi tumor epithelial cells may evolve into cells with low enrichment scores for the gene panel (PCDlowepi), with autophagy contributing to the maintenance of metabolic balance and survival in tumor stem cells under stressful conditions. This finding highlights the dual role of autophagy in cancer progression and its potential as a therapeutic target.

Moreover, autophagy significantly influences the interaction between tumor cells and smooth muscle cells within the tumor microenvironment [5]. This interaction is crucial for tumor angiogenesis, invasion, and metastasis, with autophagy playing a key role in regulating smooth muscle cell response to hypoxia and other stressors. Our study suggests that targeting autophagy could modulate these interactions, potentially reducing tumor-associated angiogenesis and the invasive capabilities of tumor cells.

Furthermore, our research has identified the NSCLC-specific programmed cell death gene SLC7A5 as a key player in promoting cell proliferation and inhibiting apoptosis. SLC7A5 is highly expressed in various cancers, including NSCLC, and is associated with poor prognosis [6, 7]. In particular, its overexpression is particularly prevalent in squamous cell carcinoma, where it correlates with worse clinical outcomes. Given its critical role in cancer progression, SLC7A5 presents a promising therapeutic target. Inhibitors such as JPH203, which specifically target SLC7A5, could offer new avenues for NSCLC treatment, particularly in cases resistant to current therapies.

As a result, NSCLC progression and therapeutic resistance are primarily related to dysregulation of programmed cell death processes, such as apoptosis and autophagy. It provides valuable insights into the molecular mechanisms driving this disease and highlights potential therapeutic targets through the identification of NSCLC-specific gene panels and key regulators such as SLC7A5. However, our study has certain limitations. The exclusion of the large cell carcinoma subtype from our sample introduces some constraints to the generalizability of our findings. Additionally, our prognostic model, NSCLCPCD, shows insufficient accuracy in predicting patient outcomes and lacks robustness in applying personalized treatment strategies. We further elaborate on the theoretical basis for developing a programed cell death gene panel specific to NSCLC (NSCLCPCD), highlighting how it addresses some of the shortcomings of existing models. Nevertheless, we also acknowledge that this panel has limitations, such as the need for larger-scale, multicenter validation studies to confirm its robustness and broad applicability in clinical settings.

2 Methods and materials

2.1 Data acquisition

The single-cell RNA (scRNA) sequencing data (EMTAB6149, GSE127465, GSE131907, GSE149655, GSE153935) was procured from the GEO database. The spatial transcriptomics and information were downloaded from the ArrayExpress dataset. We carried out bulk transcriptomic analyses using TCGA lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) cohorts (https://xena.ucsc.edu). Functional pathways were found according to the Gene Set Enrichment Analysis (GSEA) and CancerSEA databases. Single-cell metabolism was analyzed using a website (https://github.com/changwn/scFEA). The programmed cell death gene panel was constructed by intersecting differentially expressed genes from five bulk RNA-seq datasets with 18 curated programmed cell death-related gene sets, followed by meta-analysis filtering with meta_FDR < 0.05 and $|\text{meta\_Hedges}| > 0.5$.

2.2 scRNA-seq analysis

We processed the available scRNA-seq data using the Seurat R package (v4.0.5) and (v4.0.2) [8]. Quality control filters included: (a) excluding genes detected in fewer than three cells, (b) excluding cells with fewer than 50 genes, and (c) excluding cells with 5% or more mitochondrial gene expression. No additional filtering was done as previous studies had already performed quality control. Data

were normalized with SCTransform [9], and batch effects were corrected using the harmony method (v0.1.0). Subsequently, Seurat objects were integrated into a dataset [10]. Dimensionality was reduced using Principal Component Analysis (PCA). Cells with similar attributes were grouped using FindNeighbors and FindClusters. Cell-cycle scores were determined with Seurat's CellCycleScoring function, considering cell cycle phase effects. Finally, data visualization was done using UMAP.

2.3 Recognition of cell type

Using Seurat's FindAllMarkers function, we identified cluster-specific markers based on differential expression in cell clusters. The thresholds were set at P-values adjusted to 0.05, log2(fold change) absolute > 0.25, and expression percentage above 0.25. Additionally, diverse cell subclusters were identified and labeled using the singleR package based on composition patterns. All findings were manually verified and corrected using the CellMarker database [11].

2.4 The genetset score

We used irGSEA (https://github.com/chuiqin/irGSEA) to integrate 13 different scoring methods on single-cell datasets, with the irGSEA.score function parameters set to min.feature = 0, min.cells = 3, kcdf = 'Gaussian', species = "Homo sapiens", msigdb = true, and custom = false.

2.5 Intercellular communication analysis

To identify undiscovered links amongst diverse cell subtypes within the tumor microenvironment (TME), a pivotal intercellular communication analysis was conducted employing Cellcall. It offers a comprehensive and available reflection of uniquely curated ligands, receptors, and their subsequent interactions [12]. Inferring receptor-ligand interactions among various cellular subtypes was based on receptor expression by one cellular subtype and corresponding ligand expression by another cell type. It facilitated the identification of significant cell type-specific ligand-receptor interactions, considering only those expressed in more than 10% of cells within the respective subclusters.

2.6 Constructing trajectories at single-cell level

Trajectory analyses at the single-cell level were conducted using the Monocle2 R package (v2.16.0) to uncover cell-state transitions through a unidimensional 'time' parameter [13]. Epithelial cell subclusters were included using the newCellDataSet function to create an object with expressionFamily set to negbinomial.size. In pseudo-time via trajectory analysis, cells were classified by mean_expression 0.1 and dispersion_empirical dispersion_fit (Monocle2). Reducing dimensions with reduceDimension(), DDRTree as the reduction method, and max_components as two enabled visualization via plot_cell_trajectory. Using the "differentialGeneTest" function, genes with pseudotime variations were identified and displayed on the plot_pseudotime_heatmap. They were further classified into subgroups based on their expression patterns.

2.7 Regulon activity of transcription factors using SCENIC

In SCENIC, the log-normalized expression matrices of Seurat are used as input to identify regulatory networks involving TFs and their target genes [14]. Using the hg19-tss-centered-10 kb-7species.mc9nr.feather motif dataset, regulons for each TF were established within the SCENIC framework. A gene regulatory network (regulon) was created by running the "runSCENIC" procedure on each gene co-expressed with each of the transcription factors. AUCell software then assessed regulon activity, binarizing them at a default threshold where "0" indicated inactive TFs and "1" indicated active TFs.

2.8 Cellular (Cyto) trajectory reconstruction (CytoTRACE) analysis

The CytoTRACE algorithm simplifies and quantifies gene expression levels in scRNA-Seq data, correlating them with gene counts. It assigns each cell a score indicating its stemness. CytoTRACE excels in predicting differentiation states using scRNA-seq data, outperforming existing stemness evaluation methods and validated on large-scale datasets [15]. The CytoTRACE R package v0.3.3 assigns scores from 0 to 1 to tumor cells, with higher scores indicating greater stemness and lower scores indicating less stemness.

2.9 ST data analysis

We processed and visualized the spatial transcriptomics (ST) data using the Seurat R package. The SCT method standardized the data. For integration, we used PrepSCTIntegration, FindIntegrationAnchors, SelectIntegrationFeatures, and IntegrateData functions. Subsequently, similar ST regions were grouped via unsupervised clustering. Cell population annotations utilized HE staining sections and highly variable genes per cluster. A spatial dim plot and a spatial feature plot were used to determine cell expression levels in the ST data.

2.10 Robust cell type decomposition (RCTD) analysis of ST data

Cell types in the reference scRNA-seq dataset were aligned with the ST data using RCTD [16]. Unique marker genes for each cell type were identified using Seurat's FindAllMarkers function, focusing on markers with positive log2 fold changes. By using both reference and Visium ST data in full doublet mode, we strictly followed the standard RCTD analysis pathway.

2.11 Estimating functional information from spatial data

Progeny's [17, 18] model matrix (based on the top 1,000 genes in each transcriptional dataset, normalized) was used to estimate signaling pathway engagements for individual sites.

2.12 Spatial map of cell dependencies

In mistyR [19], MISTy determines the prevalence of other cell subtypes based on the abundance of each primary cell subtype. Three spatial contexts were used to combine the RCTD cell-type estimates from all slides: (i) intrinsic, assessing correlations within a specific area, (ii) juxta, considering nearby neighbors within a distance of 5, and (iii) para, incorporating more distant neighbors within a radius of 15 spots. In different spatial contexts, cell types are associated with colocalization or mutual exclusion across slides, but these relationships do not suggest causation. Predictors with an $R^2 < 10\%$ were excluded before aggregation.

In order to connect tissue structures with functions, a MISTy model was created to show the distribution of PROGENY pathway activity scores. This multi-view model included these predictors: (i) intrinsic perspective for local pathway interactions, (ii) juxta perspective for neighboring spot interactions (max distance = 5), (iii) para perspective for broader tissue correlations (radius = 15), (iv) another intrinsic perspective for repeated local interactions, and (v) a para perspective with RCTD estimates (radius = 15). The last two perspectives were included to simulate the link between cell-type compositions in spots and pathway activities. It is important to note that cycling cells were excluded from earlier analyses. Predictors with an R^2 value below 10% for pathway activities were ignored for each specimen before aggregation.

2.13 Spatial trajectory analysis

The pseudo-time trajectory was identified using stLearn [20] and PAGA trajectory analysis on SME-normalized data, revealing links within subclusters. This algorithm mapped malignancy progression across sections, highlighting spatial and transcriptional connections among sub-clusters.

2.14 Machine learning algorithms

This integrated model was built using a variety of algorithms, including RSF, LASSO, Enet, Stepwise Cox, SuperPC, CoxBoost, Ridge, survival SVM, plsRcox, and GBM. AUC was calculated using the 'timeROC' package using various computing methods to build the risk score with the highest C-index across all validation datasets. Importantly, Cox regression analysis with the 'survival' package in R showed the risk score as an independent prognostic factor.

2.15 Immunotherapy efficacy prediction

Data for the single-cell cohort for immunotherapy (GSE123813, GSE145281) was obtained from GEO database. Bulk transcriptomic cohort immunotherapy prediction was completed using the submap algorithm.

2.16 Potential targeted drug prediction

Drug sensitivity data for cancer cell lines (CCLs) were obtained from the Cancer Therapeutics Response Portal (CTRP v.2.0, October 2015) and the PRISM dataset (19Q4, released December 2019, https://depmap.org/portal/prism/).

2.17 Cell culture and processing transfection

H520 and H460 cells were sourced from the Chinese Academy of Sciences Cell Bank and cultured at 37 °C with 5% CO_2 in RPMI 1640 medium with 10% fetal bovine serum. Using Lipofectamine 3000, specific siRNAs (GenePharma, Suzhou, China) were transfected into H520 and H460 cells to transiently suppress SLC7A5 expression. The siRNA sequences used were: negative control (NC) (5′- CGGGAG GGTTATGTGCCAATAGCTA-3′), siRNA-SLC7A5#1 (5′-CGGGAAGGGTGATGTGTCCAATCTA-3′), and siRNA-SLC7A5#2 (5′- GGGAACATTGTGCTGGCATTATAC A-3′).

2.18 Cell viability assay

Cells were collected after digestion and centrifuged. The counted cells were put into a density of 2000–3000 cells/well into 96-well plates. Moreover, we assessed cell viability via a Cell Counting Kit-8 (APExBIO, United States) at intervals of 24, 48, 72, and 96 h on basis of the manufacturer's instructions. For these treated cells, a drug concentration gradient of 0, 2, 4, 6, 8, 10, and 12 µg/mL was used

for analysis. Through this methodical approach, the proliferation condition of the cells could be monitored accurately over time.

2.19 Western blotting (WB)

We disrupted cells in an ice-cold lysis buffer containing protease and phosphatase inhibitors and quantitated protein concentrations using bicinchoninic acid. We separated protein samples by a 4–12% sodium dodecyl sulfate-polyacrylamide gel electrophoresis gradient and subsequently transferred onto polyvinylidene difluoride membranes. They were blocked and incubated with specific antibodies twice. The immunoreactive proteins were visualized through a chemiluminescent solution. The primary antibodies employed were anti-SLC7A5 (HUABIO, ER1803-41, 1:1000), anti-PARP (Abmart, T40050, 1:1000), anti-Cleaved-PARP (Abmart, T55035, 1:1000), anti-Caspase 3 (Abmart, M005851, 1:1000), anti-Cleaved-Caspase 3 (Abmart, MB0711, 1:1000), anti-BCL-2 (Solarbio, K003505P, 1:1000), anti-Bax (Solarbio, K001435M, 1:1000), anti-glyceraldehyde 3-phosphate dehydrogenase (GAPDH; Santa, sc-137179, 1:1000), and anti-beta-actin (Santa, sc-58673, 1:1000).

2.20 Transwell assays

Cell migration was assessed using an 8-μm pore Boyden chamber assay. H520 and H460 cells (1×10^5) were placed in the upper chamber with FBS-free medium, while the lower chamber contained medium with 10% FBS. After 24 h, cells were stained and counted in six random fields under a microscope.

2.21 The 5-ethynyl-2′-deoxyuridine (EdU) assay

Cells were transferred into 24-well plates (5×10^4 cells) and cultured overnight in a 37 °C incubator after transfection. EdU Cell Proliferation Kit with Alexa Fluor 594 was used to detect EdU as per the protocol provided by BeyoClick. EdU-positive cells were differentiated from non-proliferating cells after staining with Azide 594 and Hoechst 33,342. Images from three randomly selected fields were visualized and captured under a microscope to ensure a representative sampling. EdU-positive rate = EdU-positive cell count/(EdU-positive cell count + EdU negative cell count) × 100%.

2.22 Immunofluorescence staining

Cells were seeded onto 18-mm coverslips and stabilized overnight at 37 °C. Briefly, these slides were blocked using 5% bovine serum albumin, followed by incubation with primary antibodies and fluorescent-labeled secondary antibodies. The nuclei were stained using DAPI (Thermo). The samples were visualized through LSM 880 laser scanning (Zeiss). The primary antibodies used was anti-BCL-2 (Abmart, T40056, 1:200).

2.23 Colony formation assay

To make sure the colony formation capability of the cells, nearly $8 * 10^2$ to $1 * 10^3$ cells were seeded per well of 6-well plates and incubated at 37 °C for 10 to 14 days to provide plenty of time for colony process. After incubation, the colonies were fixed for 15 to 30 min using methanol for preservation. Then, these colonies were stained with 0.1% crystal violet for 15 min which were easier to count. The last procedure referred to meticulous cognizing the number of the colonies, to provide an explicit estimate of the ability of cell proliferation and form colonies under provided environment.

2.24 Apoptosis assay

Cells were cultured overnight in both complete and serum-depleted media. Adherent and floating cells were collected and analyzed for apoptosis using flow cytometry with an Annexin V-FITC/PI staining kit. After washing with cold PBS, cells were resuspended in binding buffer and stained with Annexin V-FITC/PI in the dark at room temperature for 30 min. Apoptotic cells were identified using PI and Annexin V markers on a FACS flow cytometer (BD Biosciences). All experiments were conducted in triplicate.

2.25 Statistical analysis

The chi-squared test identified differences in categorical variables between subgroups with a P-value of 0.05 indicating significance. Adjusted P-values for multiple tests were calculated using the Benjamini-Hochberg method. All analyses and graphs were generated using R software (version 4.1.3). All experimental results are based on data from at least three independent replicates to ensure the reliability and reproducibility of the findings.

3 Results

3.1 Identification of programmed cell death gene panel (PCD)

The workflow design of our study is illustrated in Fig. 1. We first performed differential analysis on five bulk cohorts

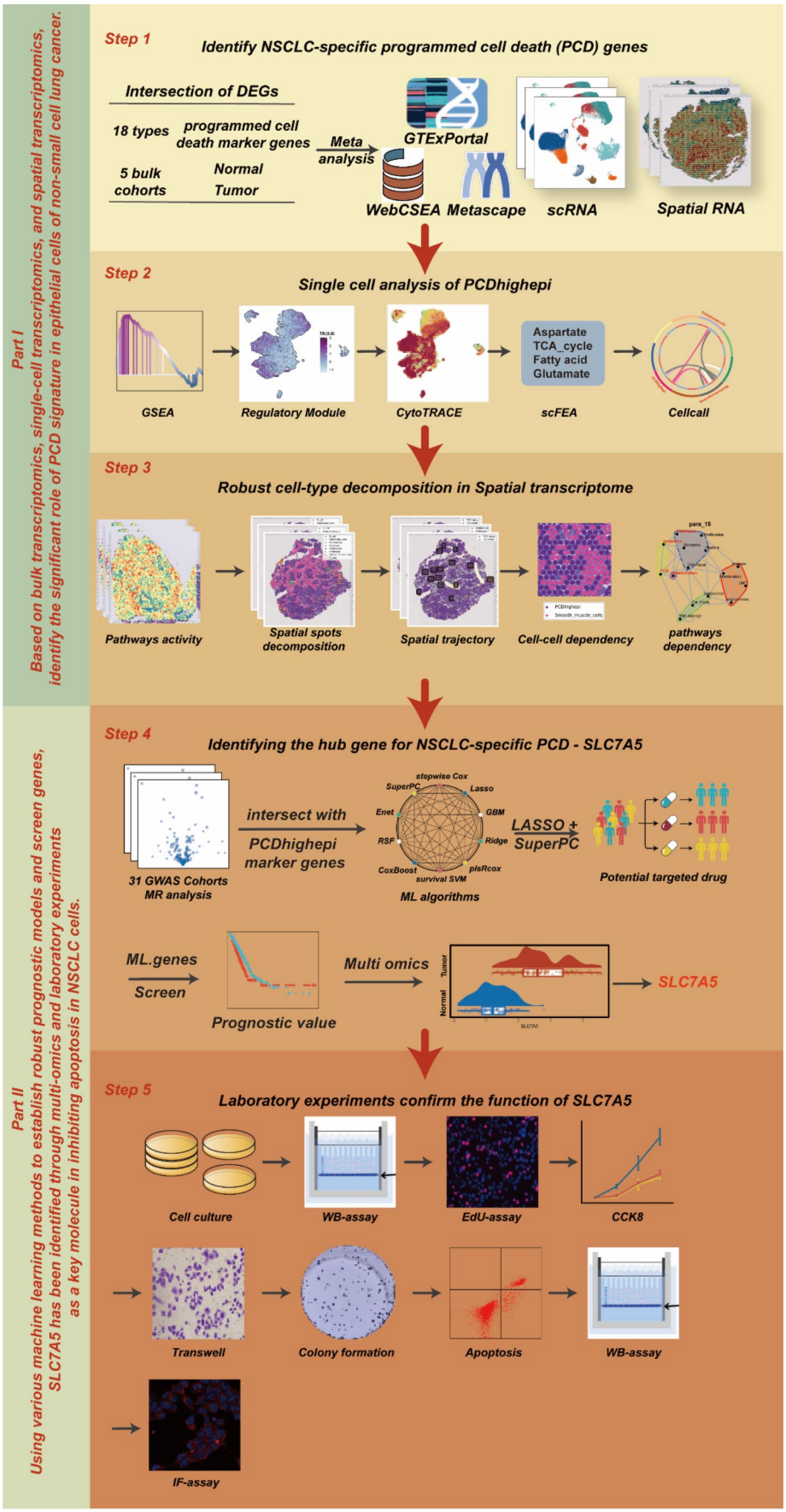

Fig. 1 The workflow of the study

(GSE18842, GSE101929, GSE134381, GSE33532, and GSE43458) (Supplementary Table 1) to identify genes commonly upregulated in tumor tissues (Fig. 2A). Next, we intersected these with 18 programmed cell death genes (Supplementary Table 2) and conducted a meta-analysis (Fig. 2B) to ultimately identify NSCLC-specific programmed cell death genes (Supplementary Table 3). Subsequent enrichment analysis using the Metascape database revealed that this genomic signature is primarily associated with apoptosis-related pathways (Fig. 2C, D), such as GO:0097190, GO:0010506, GO:2,001,233, hsa04210, and GO:0000278. Moreover, tissue-specific analysis using the GTEx database showed that this genomic signature is predominantly expressed in lung tissue (Fig. 2E). Furthermore, cell type-specific analysis using the WebCESA database indicated that this genomic signature is primarily expressed in T cells, dendritic cells, macrophages, epithelial cells, basal cells, mast cells, pneumocytes, and mesenchymal stem cells (Fig. 2F).

3.2 Programmed cell death (PCD) gene panel in epithelial cells of NSCLC patients

By integrating five single-cell NSCLC cohorts (Supplementary Fig. 1), we explored the cell-specific expression of PCD. After eliminating batch effects, we performed dimensionality reduction, clustering, and cell annotation (Fig. 3A). Subsequently, we evaluated the expression of PCD in different cell types using 13 single-cell scoring methods and compared the average scores (Fig. 3B-C). The results showed that PCD scores were highest in NSCLC epithelial cells. To examine the specificity of PCD in tumor tissues, we used the R package Cottrazm to identify and characterize benign and malignant regions in spatial transcriptomics (Supplementary Fig. 2A-D). Based on spatial pathway scoring, PCD scored higher in malignant tissues than in benign tissues (Fig. 3D-G). According to the scoring, we divided the epithelial cells into two groups (PCD groups): high PCD group (PCDhighepi) and low PCD group (PCDlowepi). Differential enrichment analysis using Hallmark, KEGG, and Reactome pathways showed that cell cycle-related pathways, including g2m_checkpoint, e2f_targets, myc_targets, and p53_signaling_pathway, were upregulated in PCDhighepi (Fig. 4A). Moreover, enrichment analysis in 12 tumor states of CancerSEA indicated that cell cycle, metastasis, and hypoxia were enhanced in PCDhighepi (Fig. 4B). Studies have shown that hypoxia, by activating hypoxia-inducible factors (HIF), regulates the expression of multiple genes, which can promote cell survival in harsh environments or, in some cases, induce apoptosis [21]. During tumor metastasis, cells often evade apoptosis by altering cell cycle regulation and enhancing survival under hypoxic conditions, thereby promoting metastasis [22]. To further explore the intrinsic mechanisms of pathway activation, we performed transcriptional regulatory analysis in NSCLC epithelial cells. Six transcriptional modules were identified through the analysis of connectivity specificity index (Fig. 4C), and it was found that PCDhighepi scored higher in modules 1 and 2 (Fig. 4D-E). Transcription factor activity analysis showed enhanced activity of representative genes such as KLF5, HIF1A, MYC, E2F4, and BACH1 in modules 1 and 2 in PCDhighepi (Fig. 4F). Studies have shown that KLF5 promotes cell cycle progression by facilitating the expression of cell cycle-related genes, particularly during the G1/S phase transition [23]. Additionally, KLF5 is involved in apoptosis regulation, affecting cell survival through interactions with apoptosis-related genes such as the BCL-2 family [24]. MYC is an oncogene that is extensively involved in the regulation of cell proliferation and apoptosis. MYC can activate multiple cell cycle-related genes to promote cell proliferation. However, overexpression of MYC may also induce apoptosis, primarily through the upregulation of p53 and its related apoptotic pathways [25].

3.3 PCDhighepi as the origin of epithelial cells in NSCLC patients

Further, using the Monocle algorithm, we performed pseudotime analysis to explore the developmental hierarchy within the PCD group. Initially, we analyzed the tumor stemness characteristics of the PCD group using the CytoTRACE algorithm (Fig. 5A-B), and the results showed that PCDhighepi had higher tumor stemness scores (Fig. 5C). Cells inferred to have high stemness scores by CytoTRACE analysis are considered the starting point of cell development. The developmental trajectory predicted by the Monocle algorithm showed that the proportion of PCDhighepi cells gradually decreased during development (Fig. 5D-E). Genes associated with cell development were grouped into four clusters (Fig. 5F), with GPX2 and KRT5 being primarily expressed during early development, while IGFBP7 and CLDN18 were mainly expressed during late development. Furthermore, pseudotime analysis evaluated tumor-related characteristic pathways within the PCD group, such as epithelial-mesenchymal transition (EMT), angiogenesis (ANG), DNA damage repair (DNA), phosphatidylinositol 3-kinase (PI3K) pathway, apoptosis (APO), and the programmed cell death gene panel (PCD) (Fig. 5G). Tumor-related features such as apoptosis (APO) and epithelial-mesenchymal transition (EMT) gradually decreased during pseudotime, with PCDhighepi exhibiting higher tumor-related characteristics throughout pseudotime compared to PCDlowepi. Subsequently, using the RCTD deconvolution method, we mapped the cell types of single-cell

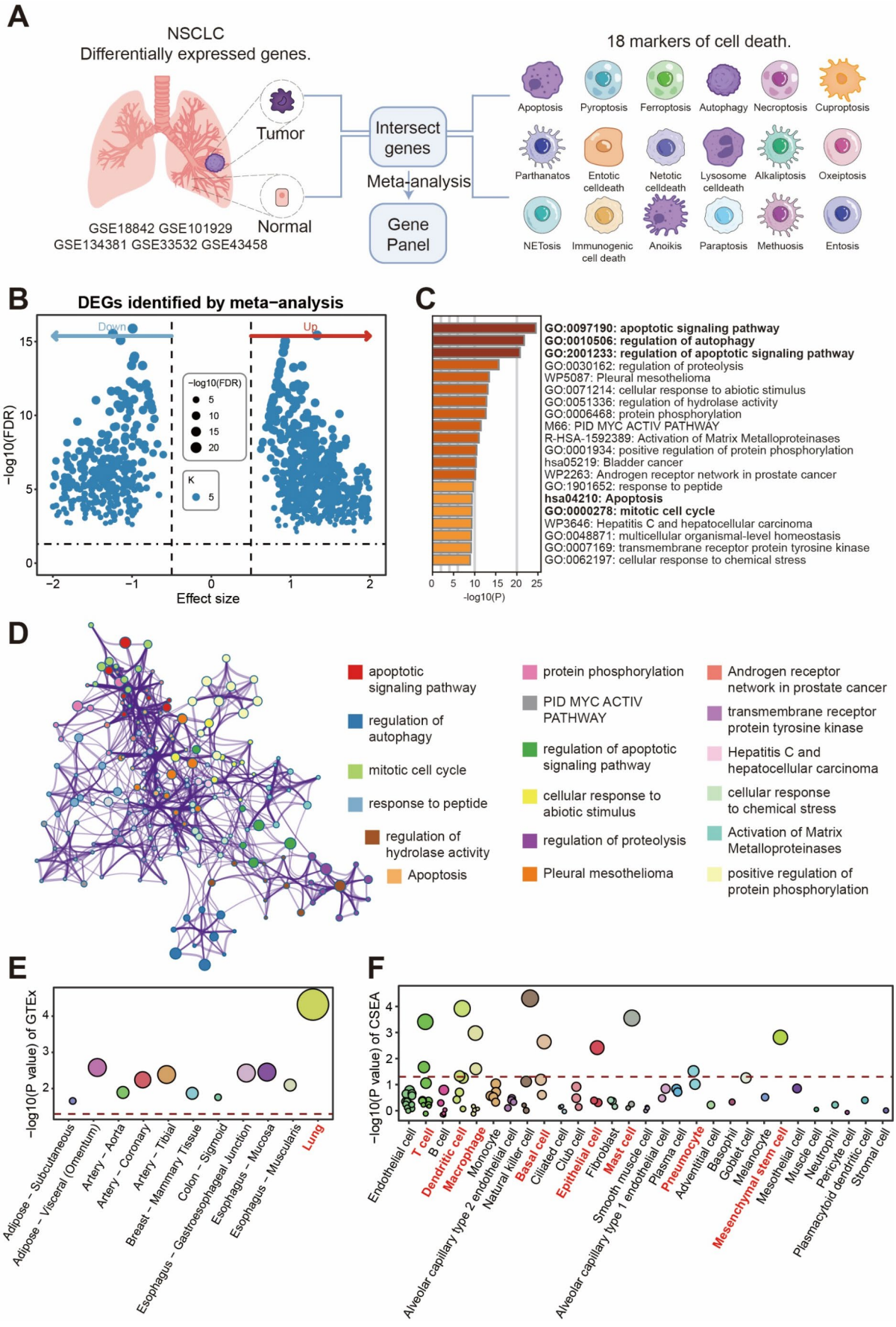
A
NSCLC
Differentially expressed genes.
Tumor
Normal
GSE18842 GSE101929
GSE134381 GSE33532 GSE43458
Intersect genes
Meta-analysis
Gene Panel
18 markers of cell death.
Apoptosis
Pyroptosis
Ferroptosis
Autophagy
Necroptosis
Cuproptosis
Parthanatos
Entotic celldeath
Netotic celldeath
Lysosome celldeath
Alkaliptosis
Oxeiptosis
NETosis
Immunogenic cell death
Anoikis
Paraptosis
Methuosis
Entosis
B
DEGs identified by meta−analysis
Down
Up
−log10(FDR)
5
10
15
20
K
5
−log10(FDR)
Effect size
C
GO:0097190: apoptotic signaling pathway
GO:0010506: regulation of autophagy
GO:2001233: regulation of apoptotic signaling pathway
GO:0030162: regulation of proteolysis
WP5087: Pleural mesothelioma
GO:0071214: cellular response to abiotic stimulus
GO:0051336: regulation of hydrolase activity
GO:0006468: protein phosphorylation
M66: PID MYC ACTIV PATHWAY
R-HSA-1592389: Activation of Matrix Metalloproteinases
GO:0001934: positive regulation of protein phosphorylation
hsa05219: Bladder cancer
WP2263: Androgen receptor network in prostate cancer
GO:1901652: response to peptide
hsa04210: Apoptosis
GO:0000278: mitotic cell cycle
WP3646: Hepatitis C and hepatocellular carcinoma
GO:0048871: multicellular organismal-level homeostasis
GO:0007169: transmembrane receptor protein tyrosine kinase
GO:0062197: cellular response to chemical stress
-log10(P)
D
apoptotic signaling pathway
regulation of autophagy
mitotic cell cycle
response to peptide
regulation of hydrolase activity
Apoptosis
protein phosphorylation
PID MYC ACTIV PATHWAY
regulation of apoptotic signaling pathway
cellular response to abiotic stimulus
regulation of proteolysis
Pleural mesothelioma
Androgen receptor network in prostate cancer
transmembrane receptor protein tyrosine kinase
Hepatitis C and hepatocellular carcinoma
cellular response to chemical stress
Activation of Matrix Metalloproteinases
positive regulation of protein phosphorylation
E
−log10(P value) of GTEx
Adipose − Subcutaneous
Adipose − Visceral (Omentum)
Artery − Aorta
Artery − Coronary
Artery − Tibial
Breast − Mammary Tissue
Colon − Sigmoid
Esophagus − Gastroesophageal Junction
Esophagus − Mucosa
Esophagus − Muscularis
Lung
F
−log10(P value) of CSEA
Endothelial cell
T cell
B cell
Dendritic cell
Macrophage
Monocyte
Alveolar capillary type 2 endothelial cell
Natural killer cell
Basal cell
Ciliated cell
Club cell
Epithelial cell
Fibroblast
Mast cell
Smooth muscle cell
Alveolar capillary type 1 endothelial cell
Plasma cell
Pneumocyte
Adventitial cell
Basophil
Goblet cell
Melanocyte
Mesenchymal stem cell
Mesothelial cell
Muscle cell
Neutrophil
Pericyte cell
Plasmacytoid dendritic cell
Stromal cell

Fig. 2 Identification and function of the non-small cell lung cancer-specific programmed cell death gene panel (PCD). **A** Schematic diagram of the NSCLC PCD origin. **B** Volcano plot showing differentially expressed genes of NSCLC-specific programmed cell death genes after meta-analysis of five bulk cohorts. The volcano plot shows the meta effect sizes on the x-axis while the y-axis indicates the −log10-transformed meta P values. **C** Bar chart displaying pathways related to PCD from enrichment analysis. **D** Network diagram showing the interaction relationships of pathways related to PCD from enrichment analysis. **E** Tissue specificity analysis of PCD. **F** Cell type specificity analysis of PCD

cohorts onto spatial transcriptomics data (Fig. 6A-C, Supplementary Fig. 3A). Developmental trajectories of the PCD group were inferred using the stlearn method. The developmental trajectory in spatial transcriptomics showed a transition from PCDhighepi to PCDlowepi, and the trajectory tree revealed detailed evolutionary relationships between cell clusters (Fig. 6D-I, Supplementary Fig. 3B-C). Additionally, a positive correlation was observed between the PCD score and the enrichment score of trajectory evolution genes (Fig. 6J-L, Supplementary Fig. 3D). Thus, our findings suggest that PCDhighepi represents the developmental origin of cells within the PCD group.

3.4 Metabolic differences in single-cell metabolomics analysis of the programmed cell death gene panel (PCD) group

Our previous findings indicated that the glutathione metabolism and pyrimidine metabolism pathways were activated in PCDhighepi (Fig. 3B). To further explore the metabolic characteristics of the PCD group, we employed the scFEA algorithm. Among the 169 metabolic pathways studied, the majority were upregulated in PCDhighepi, including pathways related to the tricarboxylic acid cycle, glycogen synthesis, and fatty acid uptake (Supplementary Fig. 4A). Conversely, pathways upregulated in PCDlowepi included those related to glucose uptake and tyrosine excretion (Supplementary Fig. 4A). Further analysis of PCDhighepi revealed activation of several pathways related to Glycolysis_TCA_cycle (Supplementary Fig. 4B), as well as pathways associated with pyrimidine synthesis (Supplementary Fig. 4C), N-linked glycan synthesis (Supplementary Fig. 4D), O-linked glycan synthesis (Supplementary Fig. 4E), sialic acid synthesis (Supplementary Fig. 4F), and purine synthesis (Supplementary Fig. 4G). These findings indicate that PCDhighepi exhibits stronger metabolic characteristics. In summary, the results highlight the significant activation of metabolic pathways in PCDhighepi and suggest that its metabolic capabilities are enhanced compared to PCDlowepi.

3.5 Interaction between PCDhighepi and smooth muscle cells

Research has shown that programmed cell death supports the cancer immunity cycle and reshapes the inflammatory tumor microenvironment (TME) [26]. To further explore the interaction between the PCD group and other cell types at the single-cell level, we analyzed cell-cell communication. Our analysis identified strong communication between PCDhighepi and smooth muscle cells (Fig. 7A). These interactions primarily activate pathways such as the Hippo signaling pathway, GnRH signaling pathway, FoxO signaling pathway, Focal adhesion, and EGFR tyrosine kinase inhibitor resistance (Fig. 7B). Ligand-receptor pairs are direct pathways for intercellular communication. Specifically, our analysis revealed key ligand-receptor pairs from PCDhighepi to smooth muscle cells, including WNT7B-FZD1, GAS6-AXL, LIF-IL6ST, BMP2-BMPR2, and BMP2-AVCR1 (Fig. 7C, Supplementary Fig. 5). Conversely, key ligand-receptor pairs from smooth muscle cells to PCDhighepi included WNT2-LRP6, WNT2-LRP5, WNT2-FZD5, SFRP2-FZD5, SFRP1-FZD6, and IGF1-IGF1R (Fig. 7C). This suggests multiple communication pathways between PCDhighepi and smooth muscle cells.

To confirm the interaction between PCDhighepi and smooth muscle cells, we further analyzed spatial transcriptomics deconvolution results, which revealed colocalization of PCDhighepi and smooth muscle cells (Fig. 7D-G). Spatial dependency analysis using the mistyR algorithm revealed a strong dependency between PCDhighepi and smooth muscle cells in colocalized, neighboring regions, and extended neighboring regions (15 points) (Fig. 7H). These findings were further validated in multiple spatial transcriptomic samples (Supplementary Fig. 6A-B). Overall, our comprehensive analysis suggests that PCDhighepi interacts with smooth muscle cells through multiple pathways, highlighting the important communication between these cell types.

3.6 Dependency of the programmed cell death gene panel (PCD) on apoptosis and other tumor states

Our previous study reported that programmed cell death gene panel (PCD) is primarily enriched in apoptosis-related pathways, a finding confirmed by spatial transcriptomic analysis. Using the CancerSEA database, which includes 12 tumor cell states, we scored PCD in various tumor states through spatial transcriptomic enrichment analysis. Initially, we observed an enrichment of apoptosis in the spatial transcriptomic data, with higher scores in malignant regions compared to benign regions (Fig. 8A). Subsequent

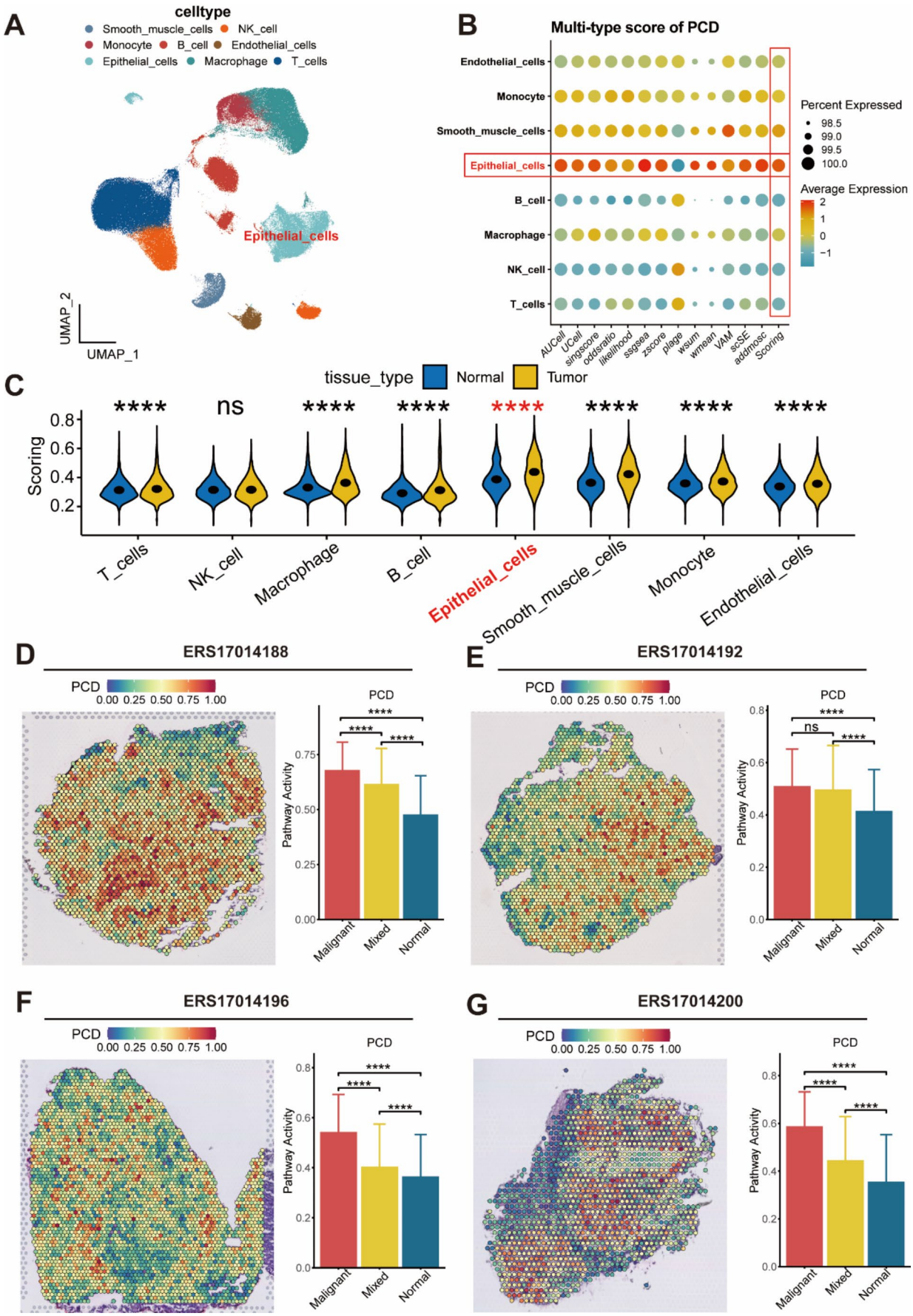
A
celltype
Smooth_muscle_cells
NK_cell
Monocyte
B_cell
Endothelial_cells
Epithelial_cells
Macrophage
T_cells
Epithelial_cells
UMAP_2
UMAP_1
B
Multi-type score of PCD
Endothelial_cells
Monocyte
Smooth_muscle_cells
Epithelial_cells
B_cell
Macrophage
NK_cell
T_cells
AUCell
UCell
singscore
oddsratio
likelihood
ssgsea
zscore
plage
wsum
wmean
VAM
scSE
addmosc
Scoring
Percent Expressed
98.5
99.0
99.5
100.0
Average Expression
2
1
0
−1
C
tissue_type
Normal
Tumor
Scoring
0.8
0.6
0.4
0.2
****
ns
****
****
****
****
****
****
T_cells
NK_cell
Macrophage
B_cell
Epithelial_cells
Smooth_muscle_cells
Monocyte
Endothelial_cells
D
ERS17014188
PCD
0.00 0.25 0.50 0.75 1.00
PCD
Pathway Activity
Malignant
Mixed
Normal
E
ERS17014192
F
ERS17014196
G
ERS17014200

Fig. 3 Elevated expression of the programmed cell death gene panel (PCD) in epithelial cells. **A** UMAP showing cell types after batch correction and dimensionality reduction clustering. **B** Bubble plot displaying the PCD scores across various cell types using multiple enrichment analysis methods. **C** Violin plot showing the average enrichment scores across different cell types. **D-G** Enrichment scores and regional difference analysis of PCD in spatial transcriptomics. The size and color of the bubbles represent the PCD score levels. ****$P < 0.0001$, ***$P < 0.001$, **$P < 0.01$, *$P < 0.05$, ns $P > 0.05$

pathway dependency analysis using the mistyR algorithm revealed that PCD primarily depends on apoptosis, hypoxia, and metastasis, with this dependency present in colocalized, neighboring regions, and extended neighboring regions (15 points) (Fig. 8B-E). This finding was further validated across multiple spatial transcriptomic samples (Supplementary Fig.7A-C). PCDhighepi primarily activates apoptosis-related pathways. To confirm this association, we combined spatial transcriptomic deconvolution results with tumor state enrichment analysis and found that PCDhighepi exhibited higher activity in pathways or tumor states related to metastasis (Fig. 8F), hypoxia (Fig. 8G), apoptosis (Fig. 8H), and PCD (Fig. 8H). Moreover, mistyR analysis showed a strong spatial dependency of PCDhighepi on PCD, apoptosis, hypoxia, and metastasis pathways or tumor states in both spatial and neighboring regions (Fig. 8I-J). Overall, our comprehensive analysis indicates that PCD depends on pathways or tumor states related to apoptosis, hypoxia, and metastasis.

3.7 Mendelian randomization (MR) analysis of the causal relationship between the programmed cell death gene panel (PCD) and lung cancer

We collected 31 genome-wide association study (GWAS) cohorts related to lung cancer and conducted MR analysis to investigate the causal relationship between PCD SNPs and lung cancer GWAS. In the ieu-a-966 cohort, the genes BIRC5, CADM1, CCT5, and CD74 demonstrated a causal relationship with lung cancer (Supplementary Fig. 8A, C). Similarly, in the ebi-a-GCST90013922 cohort, the genes AGER, CDKN2A, CLIC3, MIF, and SLC7A5 also showed a causal relationship with lung cancer (Supplementary Fig. 8B, D). In total, 58 genes were found to have a causal relationship with lung cancer across the 31 GWAS cohorts (Supplementary Table 4).

3.8 Robust machine learning model predicts patient survival and guides treatment

We identified characteristic genes of PCDhighepi using the Findmarker function ($|\log FC| > 0.5$) (Supplementary Table 5) and intersected these genes with those identified through MR analysis to establish a prognostic model (Supplementary Fig. 9, Supplementary Table 6). Multiple machine learning algorithms were employed to predict patient survival time, with Lasso + SuperPC ranking among the top five in average C-index (Supplementary Fig. 10A). Since Lasso + SuperPC retains gene weights (unlike the top four methods) (Supplementary Fig. 11), we used this method to develop a prognostic model named NSCLCPCD. Using the median NSCLCPCD score as a boundary, patients were divided into high-risk and low-risk groups. The TCGA training dataset and seven validation datasets consistently showed that the high-risk group had significantly worse survival outcomes than the low-risk group (Supplementary Fig. 10B-G).

ROC analysis indicated that NSCLCPCD has strong discriminatory power. The AUC values for 1–5 year survival rates in the TCGA-NSCLC dataset were 0.64, 0.62, 0.60, 0.59, and 0.54, respectively. The corresponding AUC values for GSE11969 were 0.78, 0.68, 0.60, 0.55, and 0.57; for GSE12428 were 0.40, 0.58, 0.63, 0.62, and 0.61; for GSE31210 were 0.81, 0.76, 0.71, 0.72, and 0.74; for GSE41271 were 0.68, 0.69, 0.62, 0.61, and 0.61; for GSE42127 were 0.73, 0.72, 0.62, 0.63, and 0.65; and for GSE72094 were 0.71, 0.70, 0.67, 0.72, and 0.76 (Supplementary Fig. 10H-L). These results underscore the robust performance of NSCLCPCD in predicting outcomes across multiple independent cohorts.

Comparison with other prognostic models showed that NSCLCPCD consistently ranked among the top in C-index across independent cohorts (Supplementary Fig. 12A), except for a lower rank in the GSE12428 cohort. Using the median NSCLCPCD score as a threshold, we associated the two risk groups with clinical indicators, finding significant differences in survival status, tumor grade, tumor stage, tumor T stage, and tumor N stage (Supplementary Fig. 12B). A higher proportion of patients with grade 2–4 tumors was observed in the high-risk group compared to the low-risk group (Supplementary Fig. 12C), and within these grades, NSCLCPCD effectively predicted shorter survival times in high-risk group patients (Supplementary Fig. 12D). To assess the independent prognostic value of NSCLCPCD, we performed univariate and multivariate Cox regression analyses on overall survival (OS), progression-free interval (PFI), and disease-specific survival (DSS) in the TCGA-NSCLC dataset (Supplementary Fig. 12E-J). Our findings indicate that NSCLCPCD is a significant risk factor in univariate analysis ($HR > 1$, $p < 0.001$) (Supplementary Fig. 12E-G) and remains a significant risk factor in multivariate analysis for OS (HR: 1.476, $p = 0.03$) and DSS (HR: 1.780, $p < 0.001$) (Supplementary Fig. 12H-J), demonstrating its strong prognostic ability in NSCLC patients. However, it is worth noting that NSCLCPCD's performance in

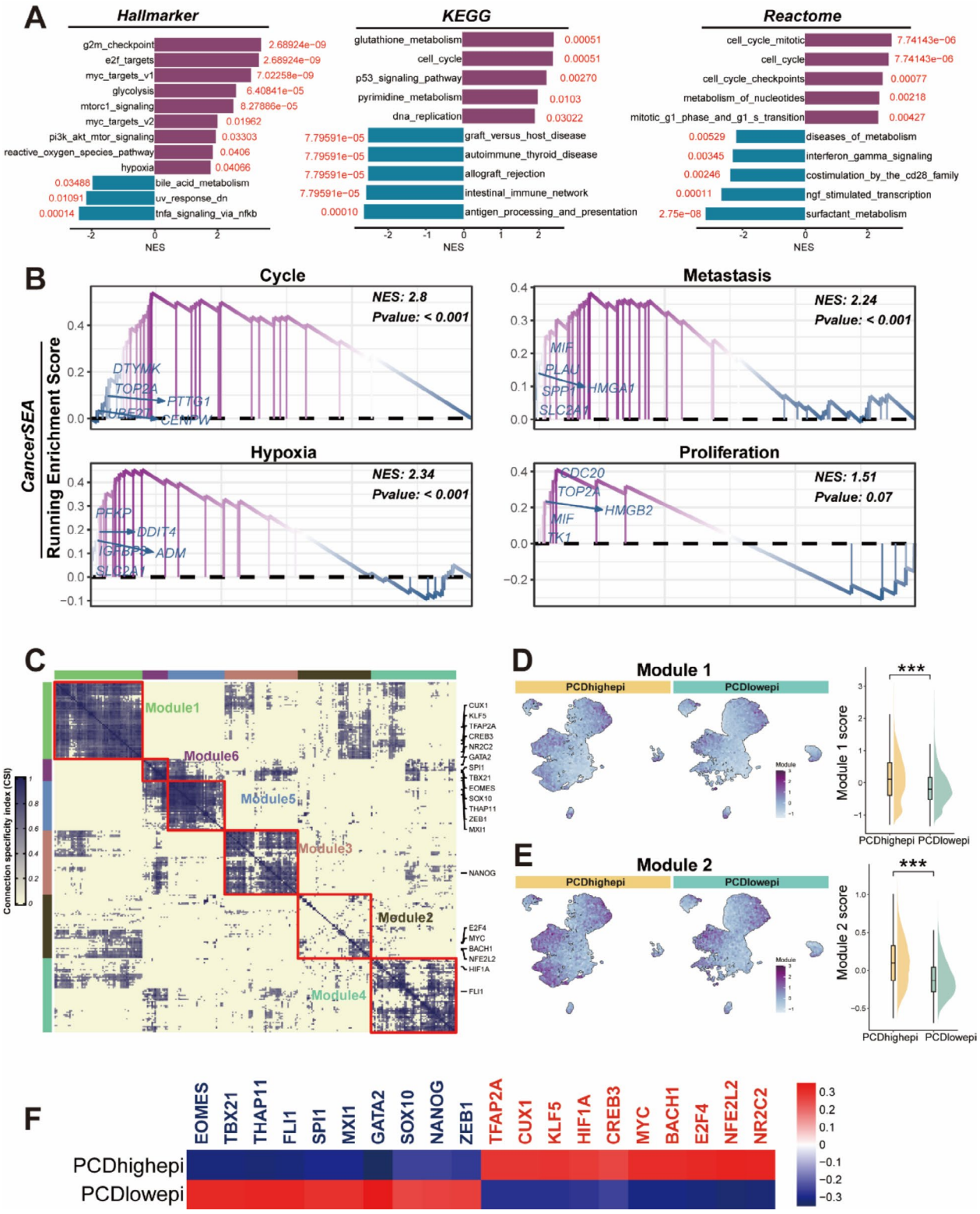

Fig. 4 Identification and functional analysis of the PCDhighepi cell population. **A** Schematic diagram of the identification of PCD high-score (PCDhighepi) cell clusters. **B** Bar chart showing the enrichment analysis of Reactome, Hallmark, and KEGG pathways. **C** Line graph showing the tumor state enrichment analysis in CancerSEA. **D** Regulatory module analysis of epithelial cells. **E-F** Comparative analysis of regulatory submodules within the programmed cell death gene panel (PCD). **G** Heatmap showing the activity of different transcription factors within the PCD group (Red text: higher transcription factor activity in PCDhighepi; Blue text: higher TF activity in PCDlowepi). $^{****}P < 0.0001$, $^{***}P < 0.001$, $^{**}P < 0.01$, $^{*}P < 0.05$, ns $P > 0.05$

PFI multivariate analysis was less satisfactory (HR: 1.197, $p = 0.111$).

Using the Submap algorithm, we further explored differences in immune therapy responses between the risk groups to predict immune therapy responses in the TCGA and GSE42127 cohorts. Results indicated that low-risk patients showed relatively better responses to immune therapy (Supplementary Fig. 13A-B). Additionally, TIDE scores indicated that low-risk patients might have better responses to immune therapy (Supplementary Fig. 13C). Moreover, in publicly available single-cell immune therapy cohorts, cells responding to immune therapy had lower risk scores (Supplementary Fig. 13D-E). Given the poor immune therapy responses in high-risk patients, analysis of the CTRP and PRISM databases showed that the high-risk group was more sensitive to clinical drugs, including paclitaxel, docetaxel, erlotinib, and gefitinib (Supplementary Fig. 13F-I). Additionally, we were able to identify suitable drugs for high-risk patients, such as leptomycin B, SB-743,921, epothilone-b, and cabazitaxel (Supplementary Fig. 13J-K).

3.9 Knockdown of SLC7A5 inhibits cell proliferation and increases apoptosis

Prognostic analysis of genes retained in the NSCLCPCD model across pan-cancer showed that only SLC7A5 was a risk factor for both LUAD and LUSC (Supplementary Fig. 14A). Further analysis of SLC7A5 expression indicated that SLC7A5 was highly expressed in tumor tissues in the bulk cohorts of LUAD and LUSC (GSE31547, GSE40791, GSE21933, GSE33479) (Supplementary Fig. 14B-E). Moreover, in the proteomics cohorts (LUAD_CPTAC and LUSC_CPTAC), SLC7A5 expression levels were higher in tumor tissues than in normal tissues (Supplementary Fig. 14F-G). Additionally, spatial transcriptomics revealed that SLC7A5 expression was higher in tumor regions compared to normal regions (Supplementary Fig. 14H-I). Collectively, these findings indicate that SLC7A5 is a risk factor for NSCLC.

In vitro experiments were conducted to investigate the biological function of SLC7A5. We knocked down SLC7A5 expression using a pair of siRNAs in the H520 and H460 cell lines. After 48 h of transfection, WB assay confirmed the significant knockdown of SLC7A5 (Fig. 9A-B). Following SLC7A5 knockdown, the EdU assay showed a decrease in proliferation in both cell lines (Fig. 9C-D). The CCK8 assay also demonstrated a reduction in proliferation after SLC7A5 knockdown (Fig. 9E-F). Furthermore, invasion assays indicated that the migration ability of H520 and H460 cells was significantly reduced after SLC7A5 knockdown (Fig. 9G-I). The two-dimensional colony formation assays revealed a significant decrease in colony-forming ability in H520 and H460 cells after SLC7A5 knockdown (Fig. 9J-L). These results suggest that SLC7A5 promotes the proliferation and migration of NSCLC cells.

We further explored the relationship between SLC7A5 and apoptosis. Following SLC7A5 knockdown, the apoptosis rate of cells significantly increased (Fig. 10A-B). Western blot analysis further demonstrated that SLC7A5 knockdown inhibited the expression of PARP, BCL-2, and Caspase3, while the expression of Cleaved-PARP, Bax, and Cleaved-Caspase3 increased (Fig. 10C-F). Immunofluorescence analysis showed reduced BCL-2 intensity in the SLC7A5 knockdown group (Fig. 10G-H). Overall, these results indicate that SLC7A5 inhibits apoptosis in NSCLC cells.

4 Discussion

Programmed cell death processes, particularly apoptosis, autophagy, and necrosis, play crucial roles in the onset, progression, and therapeutic response of non-small cell lung cancer (NSCLC). Among these, apoptosis is a primary form of programmed cell death that is tightly regulated by genes and plays a vital role in maintaining tissue homeostasis and eliminating damaged or abnormal cells. In NSCLC, dysregulation of the apoptotic pathway is a key factor in tumorigenesis and progression. Our study found that an NSCLC-specific programmed cell death gene panel is enriched in tumor epithelial cells and is associated with apoptotic pathways. Moreover, we discovered that tumor epithelial cells with a high enrichment score for the programmed cell death gene panel (PCDhighepi) exhibit upregulated pathways related to the cell cycle, tumor metastasis, and tumor hypoxia. Studies indicate that the relationship between apoptosis and the cell cycle is complex and closely intertwined [27]. Typically, cell cycle arrest or abnormalities can trigger apoptosis. For example, at the G1 and G2 checkpoints, cells usually assess the extent of DNA damage, and if the damage is too severe to be repaired, they will activate apoptotic pathways to prevent the transmission of harmful genomic information [28]. The p53 protein, a critical regulator linking the cell cycle (e.g. CDK4) and apoptosis, plays a significant role in the DNA damage response [29]. p53 can promote apoptosis by inducing the expression of apoptotic genes such as Bax and PUMA, while also inhibiting cell cycle progression, allowing time for DNA damage repair.

Autophagy, another key process, plays a more complex role in cell metastasis, as it can either suppress or promote tumor cell metastasis [30, 31]. In the early stages of tumors, autophagy often acts as a tumor-suppressive mechanism by degrading damaged organelles and proteins, thus maintaining normal cellular function [32]. However, during tumor

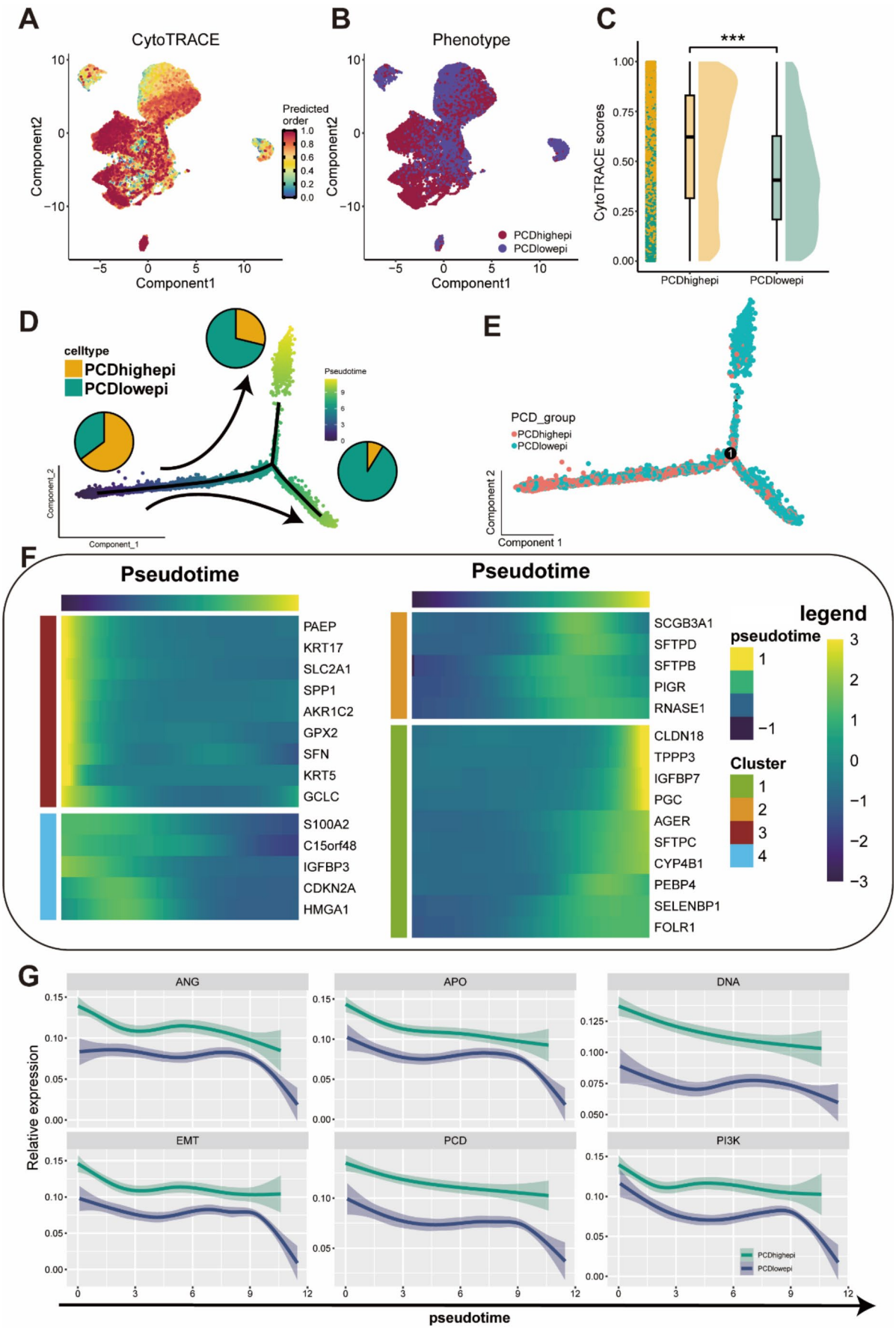
A
CytoTRACE
Predicted order
Component2
Component1
B
Phenotype
PCDhighepi
PCDlowepi
C
***
CytoTRACE scores
PCDhighepi
PCDlowepi
D
celltype
PCDhighepi
PCDlowepi
Pseudotime
Component_2
Component_1
E
PCD_group
PCDhighepi
PCDlowepi
Component 2
Component 1
F
Pseudotime
Pseudotime
PAEP
KRT17
SLC2A1
SPP1
AKR1C2
GPX2
SFN
KRT5
GCLC
S100A2
C15orf48
IGFBP3
CDKN2A
HMGA1
SCGB3A1
SFTPD
SFTPB
PIGR
RNASE1
CLDN18
TPPP3
IGFBP7
PGC
AGER
SFTPC
CYP4B1
PEBP4
SELENBP1
FOLR1
legend
pseudotime
Cluster
G
ANG
APO
DNA
EMT
PCD
PI3K
Relative expression
PCDhighepi
PCDlowepi
pseudotime

Fig. 5 Exploring the cellular origin of the programmed cell death gene panel (PCD) at the single-cell level. **A** CytoTRACE analysis of the differentiation potential of PCD group cells. **B** Grouping of the PCD group. **C** Raincloud plot showing the comparative differences in CytoTRACE scores. **D**-**E** Analysis and display of cell type proportions along the developmental trajectory of PCD group cells. **F** Pseudotime analysis of genes related to cell developmental pathways. **G** Comparative pseudotime analysis of pathways related to tumor activity. ****$P < 0.0001$, ***$P < 0.001$, **$P < 0.01$, *$P < 0.05$, ns $P > 0.05$

progression to metastasis, the role of autophagy may reverse, especially in harsh microenvironments such as hypoxia and nutrient deprivation, where autophagy helps tumor cells survive and adapt to new environments [33].

Previous studies, such as those referenced [34, 35], primarily rely on single algorithms like Non-negative Matrix Factorization (NMF) to distinguish cellular subpopulations and their functions. While these approaches provide useful insights, they may have limitations when identifying gene panels specific to certain conditions or capturing the full complexity of cell functions. In contrast, our study represents a significant advancement by first identifying an apoptosis gene panel specific to non-small cell lung cancer (NSCLC). We further differentiate cell subpopulations by integrating 13 scoring algorithms, including AUCell, UCell, singscore, and ssgsea. This comprehensive scoring strategy offers a more robust and nuanced way of characterizing the functional heterogeneity among cell subpopulations. Additionally, to evaluate the function of these subpopulations, we employed multiple state-of-the-art methods such as Cytotrace, monocle, and spatial trajectory analysis. These approaches allowed us to assess both the differentiation potential and spatial distribution of cells, providing deeper insights into the functional landscape that were not addressed by prior studies.

Our study found that tumor epithelial cells with high enrichment scores for the programmed cell death gene panel (PCDhighepi) originate from tumor epithelial cells and evolve into cells with low enrichment scores for the panel (PCDlowepi). Additionally, our research suggests that autophagy helps maintain metabolic balance and survival in tumor stem cells under adverse conditions like hypoxia and nutrient deprivation by degrading damaged mitochondria and proteins [36]. For instance, in breast cancer, studies have shown that autophagy promotes tumor stem cell survival by maintaining mitochondrial function and reactive oxygen species (ROS) balance [37]. Moreover, autophagy also plays a role in regulating stem cell-related signaling pathways such as Wnt/β-catenin, Notch, and Hedgehog in the self-renewal of tumor stem cells [38]. Autophagy is also important in the differentiation of tumor stem cells. For example, in glioblastoma, inhibiting autophagy can promote the conversion of tumor stem cells into terminally differentiated cells, reducing their tumor-forming ability [39]. Conversely, enhancing autophagy may maintain the undifferentiated state of tumor stem cells, increasing their malignant potential [40]. Thus, the regulation of autophagy could be a potential strategy to induce tumor stem cell differentiation, reducing tumor heterogeneity and invasiveness. Based on the critical role of autophagy in tumor stem cells, targeting autophagy is emerging as a new direction in cancer therapy. For example, chloroquine and its derivatives can inhibit autophagy by preventing autophagosome-lysosome fusion, thereby enhancing the cytotoxic effects of chemotherapy on tumor stem cells [41].

Autophagy plays a role not only in the homeostasis of smooth muscle cells but also in the regulatory interactions between tumor cells and smooth muscle cells. These interactions mainly influence tumor angiogenesis [42], tumor cell invasion and metastasis, and remodeling of the tumor microenvironment. Our research also found that tumor epithelial cells with high enrichment scores for the programmed cell death gene panel (PCDhighepi) exhibit strong communication with smooth muscle cells in the tumor microenvironment. Importantly, autophagy plays a crucial role in regulating the response of smooth muscle cells to angiogenesis. Studies have shown that hypoxic conditions in the tumor microenvironment can activate autophagy through the HIF-1α pathway, thereby enhancing the survival of smooth muscle cells and supporting neovascularization [43]. During tumor cell invasion and metastasis, autophagy regulates the abnormal proliferation of smooth muscle cells and the remodeling of the extracellular matrix, affecting the invasive capacity of tumor cells [44]. The remodeling of the tumor microenvironment is a key factor in tumor progression. Autophagy plays an important role in smooth muscle cells' regulation of the tumor microenvironment. Specifically, it has been found that smooth muscle cells can regulate the secretion of matrix metalloproteinases through autophagy, thus affecting the remodeling of the tumor stroma and the invasiveness of tumor cells [45, 46]. Given the important role of autophagy in the interaction between tumor cells and smooth muscle cells, targeting the autophagy pathway may provide new insights for cancer treatment. For example, using autophagy inhibitors might reduce tumor-associated angiogenesis, thereby inhibiting tumor growth. Additionally, regulating autophagy could influence the migration and secretion properties of smooth muscle cells, reducing the invasive and metastatic capabilities of tumor cells.

In addition, we developed the NSCLCPCD model to predict patient survival and personalize treatment. However, the NSCLCPCD model showed lower performance in the GSE12428 cohort. We first considered the differences in clinical and demographic characteristics between the GSE12428 cohort and other datasets. The GSE12428

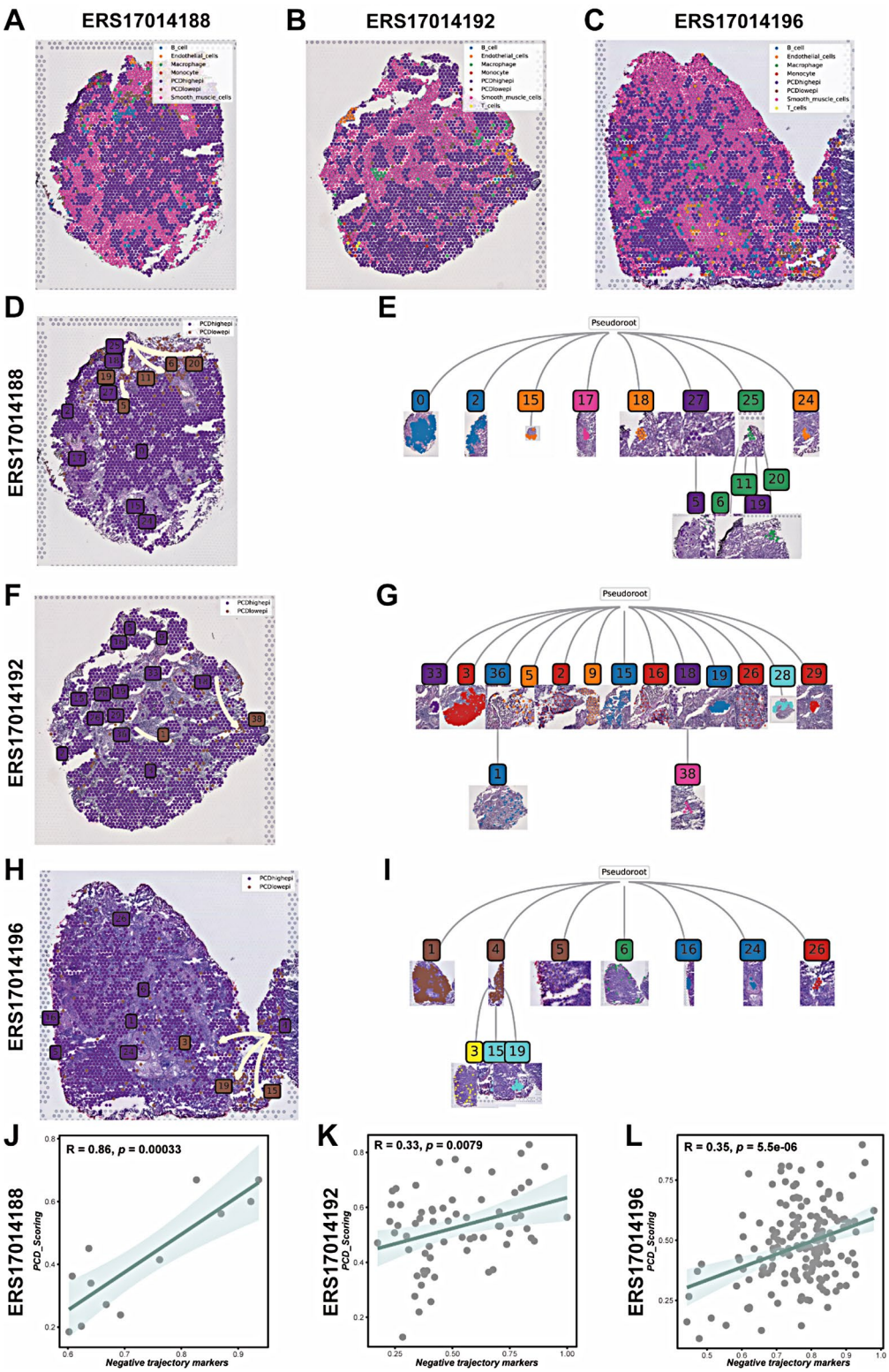

Fig. 6 Exploring the cellular origin of the programmed cell death gene panel (PCD) using spatial transcriptomics (ST). **A**-**C** Cell types after ST deconvolution. **D**-**E** Cell developmental trajectories and developmental trajectory trees in ST ERS17014188. **F**-**G** Cell developmental trajectories and developmental trajectory trees in ST ERS17014192. **H**-**I** Cell developmental trajectories and developmental trajectory trees in ST ERS17014196. **J**-**L** Scatter plots showing the correlation between PCD gene panel scores and developmental trajectory genes

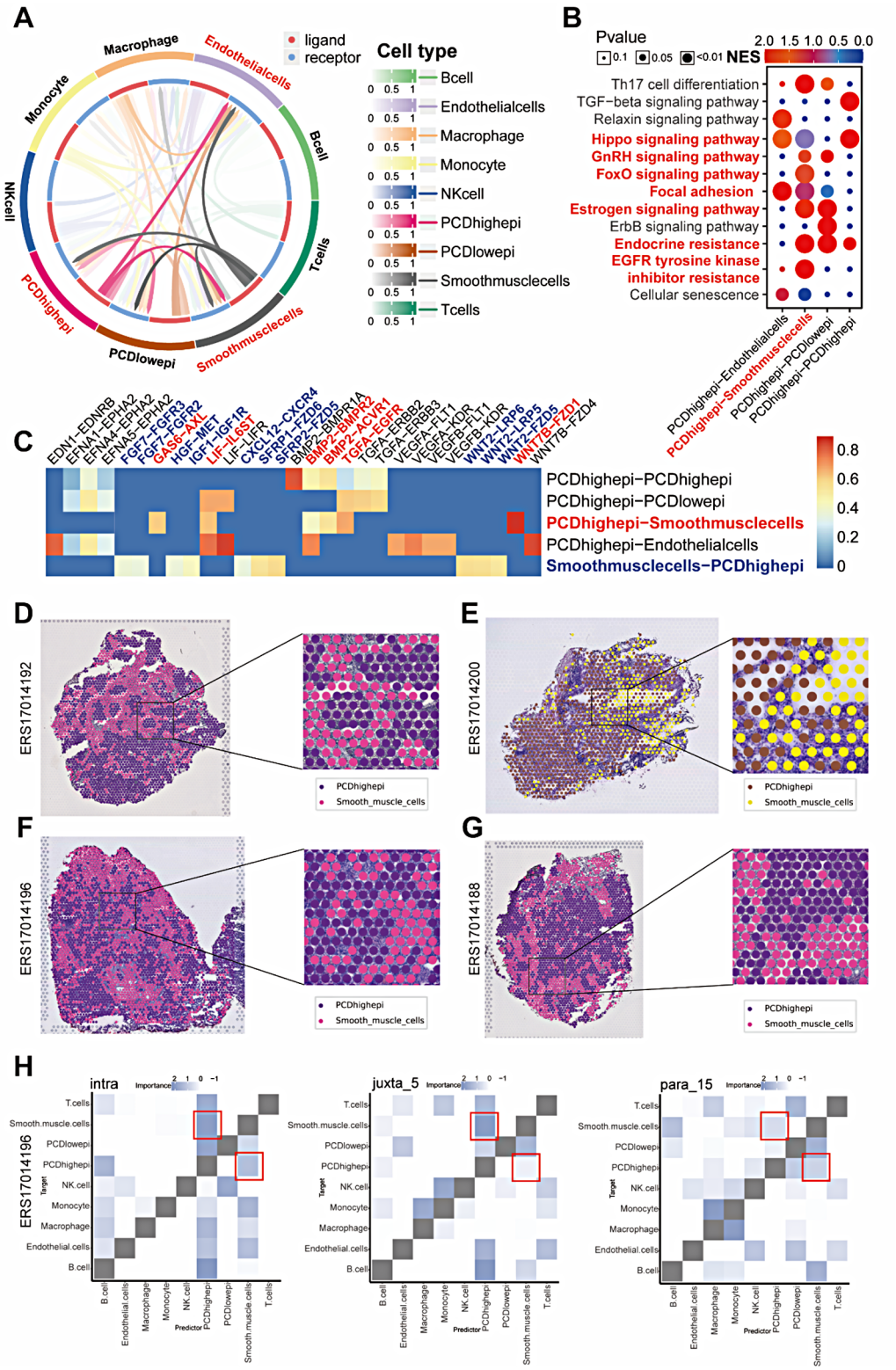

Fig. 7 Interaction between PCD high-score (PCDhighepi) cell clusters and smooth muscle cells. **A** Analysis of interaction strength between PCDhighepi and various cell types. **B** Analysis of activated pathways in various cell communications. **C** Analysis of activated ligand-receptor pairs. **D**-**G** Spatial transcriptomics deconvolution results show that PCDhighepi and smooth muscle cells are in close proximity. **H** Heatmap showing the intercellular dependency analysis in co-localized, neighboring, and extended neighboring (15-point) regions within the spatial transcriptomics context

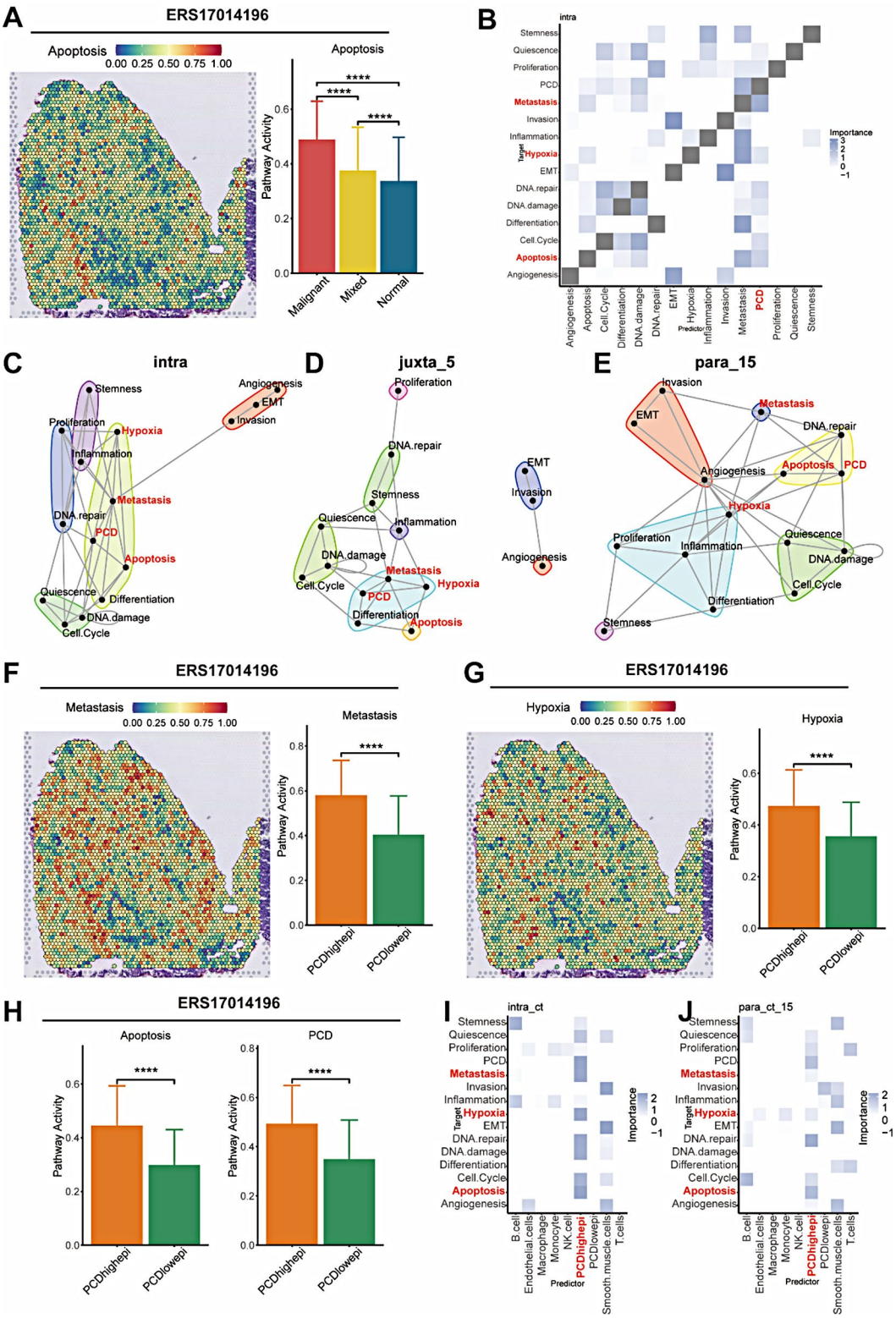

Fig. 8 Exploration of programmed cell death gene panel (PCD)-related pathways in spatial transcriptomics (ST). **A** Enrichment results of the ST cell cycle pathway and comparison of differences between regions. **B** Heatmap showing the pathways dependent on PCD within the spatial context in different regions. **C**-**E** Network diagrams showing the pathways dependent on DDP in the intra (**C**), juxta_5 (**D**), and para_15 (**E**) regions within the spatial context. ****$P < 0.0001$, ***$P < 0.001$, **$P < 0.01$, *$P < 0.05$, ns $P > 0.05$

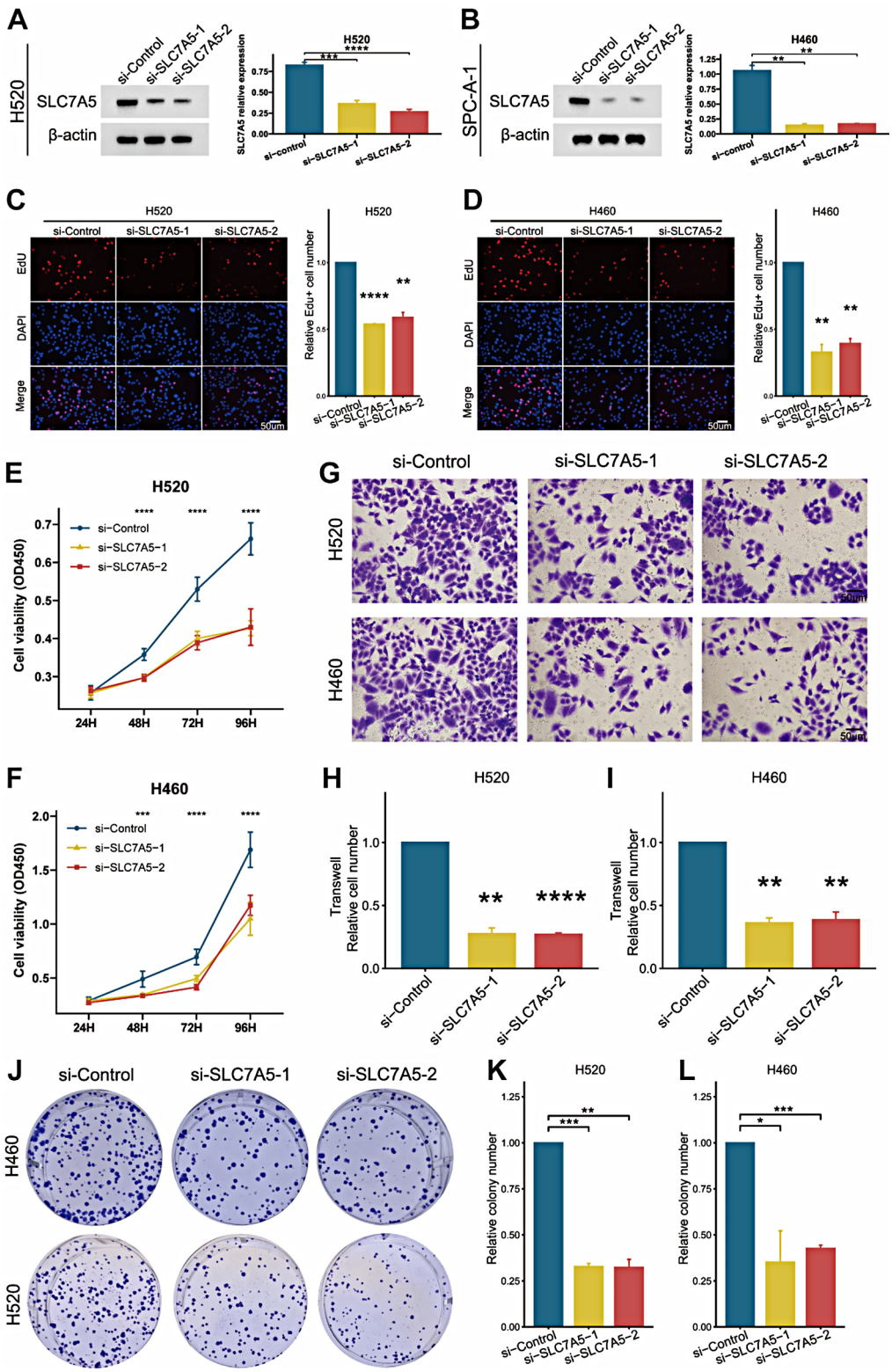

Fig. 9 Knockdown of SLC7A5 inhibits NSCLC cell proliferation. **A-B** Western blot analysis in H520 and H460 cells shows the efficiency of SLC7A5 knockdown and the statistical analysis in bar charts. **C-D** The effect of SLC7A5 knockdown on DNA synthesis in cells is shown by EdU assays. **E-F** The effect of SLC7A5 knockdown on cell viability is shown by Cell Counting Kit-8 assays. **G-I** The effect of SLC7A5 knockdown on cell invasion is shown by Transwell analysis, with statistical analysis presented in bar charts. **J-L** The effect of SLC7A5 knockdown on cell colony formation is shown by 2D colony formation assays, with statistical analysis presented in bar charts. $^{****}P < 0.0001$, $^{***}P < 0.001$, $^{**}P < 0.01$, $^{*}P < 0.05$, ns $P > 0.05$

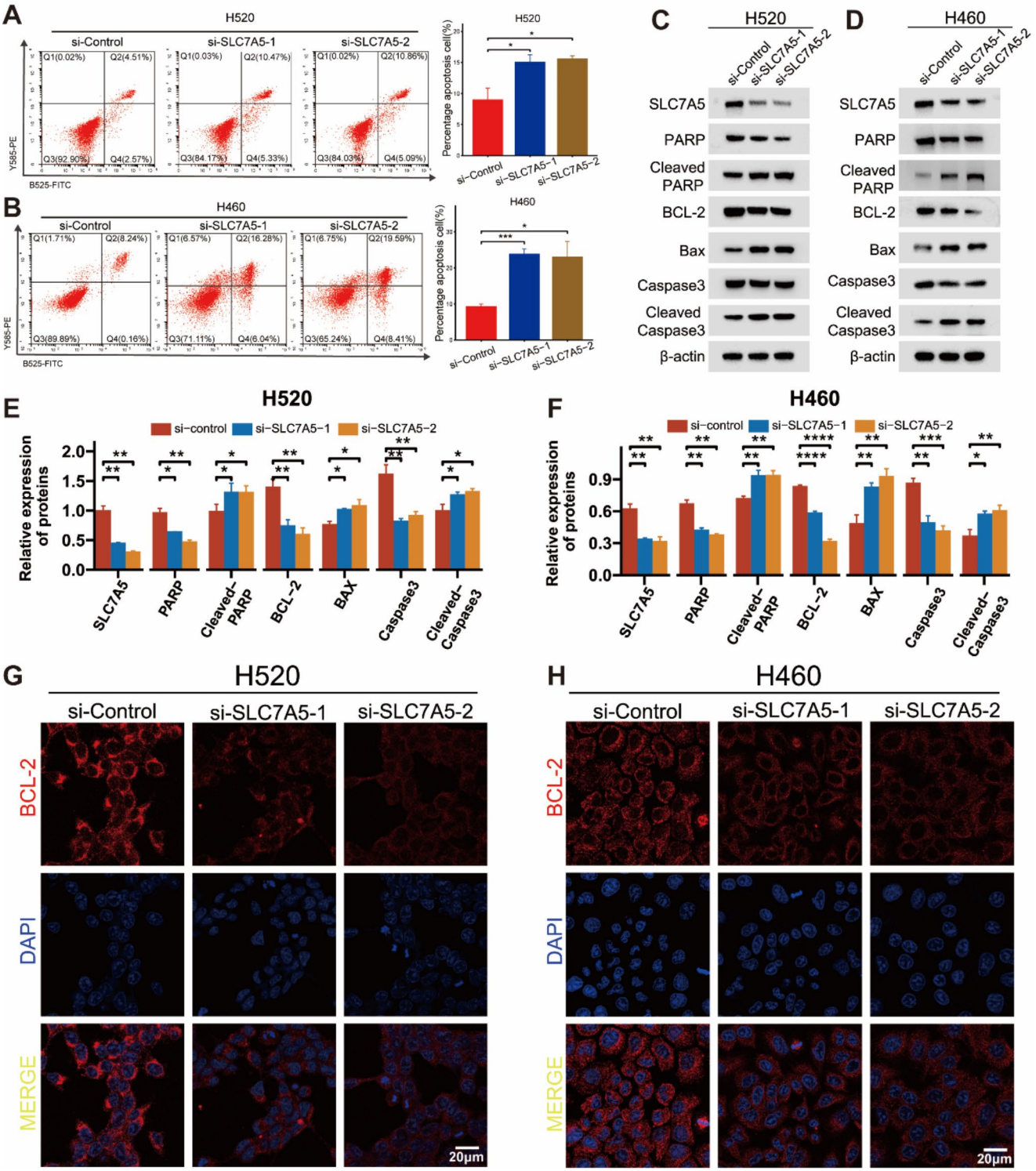

Fig. 10 Knockdown of SLC7A5 significantly increases apoptosis in NSCLC cells. **A**-**B** Flow cytometry experiments show that knockdown of SLC7A5 increases the proportion of apoptosis in H520 and H460 cells. **C**-**F** Western blot analysis shows that knockdown of SLC7A5 inhibits the expression of PARP, BCL-2, and Caspase 3, while increasing the expression of Cleaved-PARP, Bax, and Cleaved-Caspase 3, with the statistical analysis presented in bar charts. **G**-**H** Immunofluorescence experiments show that knockdown of SLC7A5 inhibits BCL-2 expression. $^{****}P < 0.0001$, $^{***}P < 0.001$, $^{**}P < 0.01$, $^{*}P < 0.05$, ns $P > 0.05$

cohort had a higher proportion of early-stage patients and greater heterogeneity within the patient population, which may have contributed to the decreased predictive performance of the model. The relatively small sample size in the GSE12428 cohort, along with differences in technical platforms, batch effects, and data preprocessing, may also have affected the model's performance. Although NSCLCPCD demonstrated significance in overall survival (OS) and disease-specific survival (DSS) in multivariate analysis (Supplementary Fig. 12E-J), its performance in progression-free interval (PFI) was less consistent. We believe that PFI might be influenced by different biological mechanisms and could differ from OS and DSS. For example, PFI may be more sensitive to short-term disease dynamics, treatment responses, or recurrence patterns—factors that may not be fully captured by the current model. Additionally, differences in treatment regimens between cohorts may impact PFI. While different treatments might control disease progression in the short term, they may not have a significant impact on long-term survival metrics such as OS and DSS.

Finally, our study discovered that the NSCLC-specific programmed cell death gene SLC7A5 promotes cell proliferation and inhibits apoptosis. Research shows that SLC7A5 overexpression is significantly associated with the histopathological grading of breast cancer patients, and SLC7A5 mRNA expression positively correlates with the expression of the proliferation marker Ki-67 and hypoxia-inducible factor 1-alpha [47]. Further, clinical pathological studies have shown that SLC7A5 protein is localized in the plasma membrane of tumors and metastases originating from various tissues, including the lungs [48–50]. Squamous cell carcinomas tend to express SLC7A5 more than adenocarcinomas. For example, 91% of surgically removed NSCLC showed high levels of SLC7A5 [51]. In a clinicopathological correlation study of completely resected pathological stage I-III NSCLC, the 5-year survival rate of SLC7A5-positive patients (51.8%) was significantly lower than that of SLC7A5-negative patients (87.8%; $P < 0.005$) [51]. Multivariate analysis showed that SLC7A5-positive expression is an independent predictor of poor prognosis [6]. In order to treat cancer, high-affinity SLC7A5-specific inhibitors have been developed, such as JPH203. Due to the novel mechanism of SLC7A5 inhibitors, they may help treat cancers resistant to current therapies, either alone or in combination with other anticancer drugs [52]. In summary, we suggest that SLC7A5 could serve as a therapeutic target for NSCLC, and combining it with other chemotherapeutic agents may enhance chemotherapy response. However, while the original sequencing data we collected came from two disease types, lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC), we used H520 and H460 cell lines for experimental validation. We acknowledge that this approach has certain limitations in terms of consistency between data sources and experimental models. The use of the H460 cell line (derived from large cell lung cancer) may limit the broad applicability of the results to LUAD. However, we believe that the use of this model still provides valuable insights into the regulatory mechanisms of SLC7A5 across different types of lung cancer. Although H460 is not a direct match for LUAD, its molecular characteristics are highly relevant to the SLC7A5-related mechanisms studied, making it valuable for our research. For LUSC, H520 is a well-established model for studying lung squamous cell carcinoma and aligns with the LUSC data in our study. In future research, we plan to use additional LUAD-specific cell lines (such as A549) to further validate the role of SLC7A5 in lung adenocarcinoma.

5 Conclusion

Our data indicate that programmed cell death is mainly associated with pathways related to apoptosis, tumor metastasis, and hypoxia. Additionally, it suggests that SLC7A5 is an obvious risk indicator for the prognosis of non-small cell lung cancer (NSCLC) and an effective target for enhancing apoptosis in NSCLC tumor cells.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s13402-024-01025-6.

Acknowledgements None.

Author contributions Huimin Li is the first author of this article. The Corresponding author is Shuixian Huang and Junzhi Liu. (I) Conception and design of the experiments: HML and JZL; (II) Experimentation: HML, YHJ and JZL (III) Administrative support: JZL and SXH; (IV) Data collection and assembly: JZL, HML and YZ; (V) Data analysis and writing: JZL and HML. All authors have read and approved the final version of the manuscript.

Funding This study was supported by the Pudong New Area health system leading talent training program (Grant No. PWRI2021-07) and Pudong New Area Clinical Characteristic Discipline (Grant No. PWYts2021-15).

Data availability No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication None.

Competing interests The authors declare no competing interests.

References

1. C. Gridelli, A. Rossi, D.P. Carbone, J. Guarize, N. Karachaliou, T. Mok, F. Petrella, L. Spaggiari, R. Rosell, Non-small-cell lung cancer. Nat. Rev. Dis. Primers. **1**, 15009 (2015)
2. L. Ouyang, Z. Shi, S. Zhao, F.T. Wang, T.T. Zhou, B. Liu, J.K. Bao, Programmed cell death pathways in cancer: a review of apoptosis, autophagy and programmed necrosis. Cell. Prolif. **45**(6), 487–498 (2012)
3. J.C. Reed, Dysregulation of apoptosis in cancer. J. Clin. Oncol. **17**(9), 2941–2953 (1999)
4. X. Li, S. He, B. Ma, Autophagy and autophagy-related proteins in cancer. Mol. Cancer. **19**(1), 12 (2020)
5. H. Xia, D.R. Green, W. Zou, Autophagy in tumour immunity and therapy. Nat. Rev. Cancer. **21**(5), 281–297 (2021)
6. H. Imai, K. Kaira, N. Oriuchi, K. Shimizu, H. Tominaga, N. Yanagitani, N. Sunaga, T. Ishizuka, S. Nagamori, K. Promchan et al., Inhibition of L-type amino acid transporter 1 has antitumor activity in non-small cell lung cancer. Anticancer Res. **30**(12), 4819–4828 (2010)
7. El R. Ansari, M.L. Craze, I. Miligy, M. Diez-Rodriguez, C.C. Nolan, I.O. Ellis, E.A. Rakha, A.R. Green, The amino acid transporter SLC7A5 confers a poor prognosis in the highly proliferative breast cancer subtypes and is a key therapeutic target in luminal B tumours. Breast Cancer Res. **20**(1), 21 (2018)
8. Y. Hao, S. Hao, E. Andersen-Nissen, W.M. 3 Mauck rd, S. Zheng, A. Butler, M.J. Lee, A.J. Wilk, C. Darby, M. Zager et al., Integrated analysis of multimodal single-cell data. Cell. **184**(13), 3573–3587e3529 (2021)
9. C. Hafemeister, R. Satija, Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. Genome Biol. **20**(1), 296 (2019)
10. I. Korsunsky, N. Millard, J. Fan, K. Slowikowski, F. Zhang, K. Wei, Y. Baglaenko, M. Brenner, P.R. Loh, S. Raychaudhuri, Fast, sensitive and accurate integration of single-cell data with Harmony. Nat. Methods. **16**(12), 1289–1296 (2019)
11. C. Hu, T. Li, Y. Xu, X. Zhang, F. Li, J. Bai, J. Chen, W. Jiang, K. Yang, Q. Ou et al., CellMarker 2.0: an updated database of manually curated cell markers in human/mouse and web tools based on scRNA-seq data. Nucleic Acids Res. **51**(D1), D870–D876 (2023)
12. R. Vento-Tormo, M. Efremova, R.A. Botting, M.Y. Turco, M. Vento-Tormo, K.B. Meyer, J.E. Park, E. Stephenson, K. Polanski, A. Goncalves et al., Single-cell reconstruction of the early maternal-fetal interface in humans. Nature. **563**(7731), 347–353 (2018)
13. C. Trapnell, D. Cacchiarelli, J. Grimsby, P. Pokharel, S. Li, M. Morse, N.J. Lennon, K.J. Livak, T.S. Mikkelsen, J.L. Rinn, The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat. Biotechnol. **32**(4), 381–386 (2014)
14. S. Aibar, C.B. Gonzalez-Blas, T. Moerman, V.A. Huynh-Thu, H. Imrichova, G. Hulselmans, F. Rambow, J.C. Marine, P. Geurts, J. Aerts et al., SCENIC: single-cell regulatory network inference and clustering. Nat. Methods. **14**(11), 1083–1086 (2017)
15. G.S. Gulati, S.S. Sikandar, D.J. Wesche, A. Manjunath, A. Bharadwaj, M.J. Berger, F. Ilagan, A.H. Kuo, R.W. Hsieh, S. Cai et al., Single-cell transcriptional diversity is a hallmark of developmental potential. Science. **367**(6476), 405–411 (2020)
16. D.M. Cable, E. Murray, L.S. Zou, A. Goeva, E.Z. Macosko, F. Chen, R.A. Irizarry, Robust decomposition of cell type mixtures in spatial transcriptomics. Nat. Biotechnol. **40**(4), 517–526 (2022)
17. M. Schubert, B. Klinger, M. Klunemann, A. Sieber, F. Uhlitz, S. Sauer, M.J. Garnett, N. Bluthgen, J. Saez-Rodriguez, Perturbation-response genes reveal signaling footprints in cancer gene expression. Nat. Commun. **9**(1), 20 (2018)
18. C.H. Holland, B. Szalai, J. Saez-Rodriguez, Transfer of regulatory knowledge from human to mouse for functional genomics analysis. Biochim. Biophys. Acta Gene Regul. Mech. **1863**(6), 194431 (2020)
19. J. Tanevski, R.O.R. Flores, A. Gabor, D. Schapiro, J. Saez-Rodriguez, Explainable multiview framework for dissecting spatial relationships from highly multiplexed data. Genome Biol. **23**(1), 97 (2022)
20. D. Pham, X. Tan, J. Xu, L.F. Grice, P.Y. Lam, A. Raghubar, J. Vukovic, M.J. Ruitenberg, Q. Nguyen, stLearn: integrating spatial location, tissue morphology and gene expression to find cell types, cell-cell interactions and spatial trajectories within undissociated tissues (2020)
21. M. Li, J. Tan, Y. Miao, P. Lei, Q. Zhang, The dual role of autophagy under hypoxia-involvement of interaction between autophagy and apoptosis. Apoptosis. **20**(6), 769–777 (2015)
22. S.U. Khan, K. Fatima, F. Malik, H. Kalkavan, A. Wani, Cancer metastasis: molecular mechanisms and clinical perspectives. Pharmacol. Ther. **250**, 108522 (2023)
23. C. Chen, M.S. Benjamin, X. Sun, K.B. Otto, P. Guo, X.Y. Dong, Y. Bao, Z. Zhou, X. Cheng, J.W. Simons et al., KLF5 promotes cell proliferation and tumorigenesis through gene regulation and the TSU-Pr1 human bladder cancer cell line. Int. J. Cancer. **118**(6), 1346–1355 (2006)
24. C. Rogerson, S. Ogden, E. Britton, O. Consortium, Y. Ang, A.D. Sharrocks, Repurposing of KLF5 activates a cell cycle signature during the progression from a precursor state to oesophageal adenocarcinoma. Elife **9** (2020)
25. N. Meyer, L.Z. Penn, Reflecting on 25 years with MYC. Nat. Rev. Cancer. **8**(12), 976–990 (2008)
26. T. Liang, L. Gu, X. Kang, J. Li, Y. Song, Y. Wang, W. Ma, Programmed cell death disrupts inflammatory tumor microenvironment (TME) and promotes glioblastoma evolution. Cell. Commun. Signal. **22**(1), 333 (2024)
27. B. Pucci, M. Kasten, A. Giordano, Cell cycle and apoptosis. Neoplasia. **2**(4), 291–299 (2000)
28. T.K. MacLachlan, N. Sang, A. Giordano, Cyclins, cyclin-dependent kinases and cdk inhibitors: implications in cell cycle control and cancer. Crit. Rev. Eukaryot. Gene Expr. **5**(2), 127–156 (1995)
29. X. Wang, E.R. Simpson, K.A. Brown, p53: Protection against Tumor Growth beyond effects on Cell cycle and apoptosis. Cancer Res. **75**(23), 5001–5007 (2015)
30. L. Galluzzi, F. Pietrocola, B. Levine, G. Kroemer, Metabolic control of autophagy. Cell. **159**(6), 1263–1276 (2014)
31. E.E. Mowers, M.N. Sharifi, K.F. Macleod, Autophagy in cancer metastasis. Oncogene. **36**(12), 1619–1630 (2017)
32. L. Galluzzi, F. Pietrocola, J.M. Bravo-San Pedro, R.K. Amaravadi, E.H. Baehrecke, F. Cecconi, P. Codogno, J. Debnath, D.A. Gewirtz, V. Karantza et al., Autophagy in malignant transformation and cancer progression. Embo J. **34**(7), 856–880 (2015)
33. R. Lock, C.M. Kenific, A.M. Leidal, E. Salas, J. Debnath, Autophagy-dependent production of secreted factors facilitates oncogenic RAS-driven invasion. Cancer Discov. **4**(4), 466–479 (2014)
34. X. Sun, F. Meng, M. Nong, H. Fang, C. Lu, Y. Wang, P. Zhang, Single-cell dissection reveals the role of aggrephagy patterns in tumor microenvironment components aiding predicting prognosis and immunotherapy on lung adenocarcinoma. Aging (Albany NY). **15**(23), 14333–14371 (2023)
35. W. Wang, T. Li, Z. Xie, J. Zhao, Y. Zhang, Y. Ruan, B. Han, Integrating single-cell and bulk RNA sequencing data unveils antigen presentation and process-related CAFS and establishes a predictive signature in prostate cancer. J. Transl Med. **22**(1), 57 (2024)

36. V. Espina, B.D. Mariani, R.I. Gallagher, K. Tran, S. Banks, J. Wiedemann, H. Huryk, C. Mueller, L. Adamo, J. Deng et al., Malignant precursor cells pre-exist in human breast DCIS and require autophagy for survival. PloS One. **5**(4), e10240 (2010)
37. L. Vera-Ramirez, S.K. Vodnala, R. Nini, K.W. Hunter, J.E. Green, Autophagy promotes the survival of dormant breast cancer cells and metastatic tumour recurrence. Nat. Commun. **9**(1), 1944 (2018)
38. X. Chen, Y. He, F. Lu, Autophagy in stem cell biology: a perspective on stem cell self-renewal and differentiation. Stem Cells Int. **2018**, 9131397 (2018)
39. A.J. Manea, S.K. Ray, Regulation of autophagy as a therapeutic option in glioblastoma. Apoptosis. **26**(11–12), 574–599 (2021)
40. P. Boya, P. Codogno, N. Rodriguez-Muela, Autophagy in stem cells: repair, remodelling and metabolic reprogramming. Development 2018, **145**(4)
41. P. Maycotte, S. Aryal, C.T. Cummings, J. Thorburn, M.J. Morgan, A. Thorburn, Chloroquine sensitizes breast cancer cells to chemotherapy independent of autophagy. Autophagy. **8**(2), 200–212 (2012)
42. S. Roy, S. Kumaravel, A. Sharma, C.L. Duran, K.J. Bayless, S. Chakraborty, Hypoxic tumor microenvironment: implications for cancer therapy. Exp. Biol. Med. (Maywood). **245**(13), 1073–1086 (2020)
43. A. Emami Nejad, S. Najafgholian, A. Rostami, A. Sistani, S. Shojaeifar, M. Esparvarinha, R. Nedaeinia, S. Haghjooy Javanmard, M. Taherian, M. Ahmadlou et al., The role of hypoxia in the tumor microenvironment and development of cancer stem cell: a novel approach to developing treatment. Cancer Cell. Int. **21**(1), 62 (2021)
44. C.G. Chen, R.V. Iozzo, Extracellular matrix guidance of autophagy: a mechanism regulating cancer growth. Open. Biol. **12**(1), 210304 (2022)
45. J.L. Johnson, Matrix metalloproteinases: influence on smooth muscle cells and atherosclerotic plaque stability. Expert Rev. Cardiovasc. Ther. **5**(2), 265–282 (2007)
46. G.A. Cabral-Pacheco, I. Garza-Veloz, C. Castruita-De la Rosa, J.M. Ramirez-Acuna, B.A. Perez-Romero, J.F. Guerrero-Rodriguez, N. Martinez-Avila, M.L. Martinez-Fierro, The Roles of Matrix Metalloproteinases and Their Inhibitors in Human Diseases. Int. J. Mol. Sci. **21**(24) (2020)
47. R. Tornroos, E. Tina, A. Gothlin Eremo, SLC7A5 is linked to increased expression of genes related to proliferation and hypoxia in estrogen–receptor–positive breast cancer. Oncol. Rep. **47**(1) (2022)
48. K. Kaira, N. Oriuchi, H. Imai, K. Shimizu, N. Yanagitani, N. Sunaga, T. Hisada, S. Tanaka, T. Ishizuka, Y. Kanai et al., l-type amino acid transporter 1 and CD98 expression in primary and metastatic sites of human neoplasms. Cancer Sci. **99**(12), 2380–2386 (2008)
49. M. Ichinoe, T. Mikami, T. Yoshida, I. Igawa, T. Tsuruta, N. Nakada, N. Anzai, Y. Suzuki, H. Endou, I. Okayasu, High expression of L-type amino-acid transporter 1 (LAT1) in gastric carcinomas: comparison with non-cancerous lesions. Pathol. Int. **61**(5), 281–289 (2011)
50. M. Furuya, J. Horiguchi, H. Nakajima, Y. Kanai, T. Oyama, Correlation of L-type amino acid transporter 1 and CD98 expression with triple negative breast cancer prognosis. Cancer Sci. **103**(2), 382–389 (2012)
51. K. Kaira, N. Oriuchi, H. Imai, K. Shimizu, N. Yanagitani, N. Sunaga, T. Hisada, S. Tanaka, T. Ishizuka, Y. Kanai et al., Prognostic significance of L-type amino acid transporter 1 expression in resectable stage I-III nonsmall cell lung cancer. Br. J. Cancer. **98**(4), 742–748 (2008)
52. Y. Kanai, Amino acid transporter LAT1 (SLC7A5) as a molecular target for cancer diagnosis and therapeutics. Pharmacol. Ther. **230**, 107964 (2022)