Research Square

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

Single-cell and Spatial Transcriptomics Reveal Cholesterol Metabolism Heterogeneity in Pancreatic Ductal Adenocarcinoma and Identify NSDHL as a Key Regulatory Gene

Yuanfei Gao

gaoyf2024@lzu.edu.cn

Lanzhou University Second Clinical Medical School: Lanzhou University Second Hospital

Zhongqing Li

Lanzhou University Second Clinical Medical School: Lanzhou University Second Hospital

Yanxi Mu

Lanzhou University Second Clinical Medical School: Lanzhou University Second Hospital

Bo Ren

Lanzhou University Second Clinical Medical School: Lanzhou University Second Hospital

Wancheng Li

Lanzhou University Second Clinical Medical School: Lanzhou University Second Hospital

Wence Zhou

Lanzhou University Second Clinical Medical School: Lanzhou University Second Hospital

Research Article

Keywords: Pancreatic ductal adenocarcinoma, Cholesterol metabolism, NSDHL, Single-cell RNA sequencing, Spatial transcriptomics, Tumor heterogeneity

Posted Date: November 2nd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11313598/v1

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

Single-cell and Spatial Transcriptomics Reveal Cholesterol Metabolism Heterogeneity in Pancreatic Ductal Adenocarcinoma and Identify NSDHL as a Key Regulatory Gene

Yuanfei Gao[1], Zhongqing Li[1],Yanxi Mu [1], Bo Ren [1], Wancheng Li [1], Wence Zhou [*]

Abstract

Background Pancreatic ductal adenocarcinoma (PDAC) is one of the most lethal malignancies, and reprogramming of cholesterol metabolism plays a critical role in its malignant progression. However, the heterogeneity of cholesterol metabolism at the single-cell level and its key regulatory genes remain unclear.

Methods This study integrated single-cell RNA sequencing, spatial transcriptomics, and bulk transcriptomic data (TCGA, GEO, ICGC) to systematically dissect the cellular landscape of cholesterol metabolism in PDAC. Malignant ductal epithelial cells were scored using five independent algorithms and divided into low (LCM), intermediate (ICM), and high (HCM) cholesterol metabolic subgroups. Cell communication, pseudotime trajectory, hdWGCNA, and XGBoost feature importance ranking were performed. Functional experiments including CCK-8, colony formation, EdU, wound healing, Transwell, and nude mouse tumor models were used to validate the role of NSDHL.

Results Malignant ductal epithelial cells exhibited the highest cholesterol metabolic activity. The HCM subgroup showed enhanced communication with CAFs and endothelial cells, increased stemness, and coordinated activation of pro-tumor pathways (hypoxia, TNFα, NF-κB) and metabolic pathways (glycolysis, fatty acid synthesis). Pseudotime analysis revealed that cholesterol metabolism gradually increased along the LCM→ICM→HCM trajectory. Through hdWGCNA combined with XGBoost, NSDHL was identified as the most critical gene for the HCM phenotype. NSDHL was significantly overexpressed in PDAC tissues, enriched in high-malignancy regions, and correlated with poor prognosis. Knockdown of NSDHL significantly inhibited PDAC cell proliferation, migration, and invasion, and delayed tumor growth in nude mice.

Conclusion This study reveals the heterogeneity landscape of cholesterol metabolism in

[1]Department of General Surgery, The Second Hospital & Clinical Medical School, Lanzhou University, Lanzhou, China

[*] Corresponding author: Wence Zhou, Department of General Surgery, The Second Hospital & Clinical Medical School, Lanzhou University, Lanzhou, China. Email: zhouwc@lzu.edu.cn, zhouwc129@163.com

PDAC and identifies NSDHL as a key gene driving the HCM phenotype and malignant progression, providing a new theoretical basis for prognosis evaluation and targeted therapy of PDAC.

Keywords: Pancreatic ductal adenocarcinoma; Cholesterol metabolism; NSDHL; Single-cell RNA sequencing; Spatial transcriptomics; Tumor heterogeneity

Introduction

Pancreatic ductal adenocarcinoma (PDAC) is one of the most lethal malignancies, with a median overall survival of 4 months and a 5-year survival rate of approximately 13%, More than half of patients have already developed distant metastasis at the time of diagnosis, and this insidious nature of the disease leads to extremely poor prognosis, It is predicted to become the second leading cause of cancer-related death by 2030[1]. Despite significant advances in genomics and targeted therapy in recent years, the survival outcomes of PDAC patients have not been substantially improved. Therefore, a deeper understanding of the molecular mechanisms driving the aggressive growth and therapeutic resistance of PDAC is of great scientific and clinical significance for developing more effective diagnostic and therapeutic strategies.

Metabolic reprogramming has been widely recognized as a hallmark of cancer, and abnormal activation of lipid metabolism is particularly prominent in PDAC[2, 3]. Lipids such as fatty acids, cholesterol, phospholipids, and sphingolipids not only reshape oncogenic signaling but also regulate the microenvironment and promote metastasis and drug resistance[4]. Among various lipids, cholesterol metabolism has attracted considerable attention due to its central role in maintaining membrane homeostasis and integrating signals. Dysregulation of cholesterol metabolism in the tumor microenvironment differentially affects the functions of immune cells such as T cells, constructing an immunosuppressive microenvironment through metabolic reprogramming and facilitating immune evasion[5]. Notably, recent studies have revealed broader regulatory functions of cholesterol metabolism in the tumor microenvironment—it is not only a bioenergetic adaptation mechanism for tumor cells to cope with nutritional stress and oxidative pressure, but also a signaling regulatory hub that integrates oncogenic signals and stromal metabolic support. For example, USP20-driven cholesterol metabolism coordinates malignant progression and microenvironment remodeling in PDAC through the STAT3-USP20-HMGCR axis[6]; simultaneous targeting of ISG15 and HMGCR disrupts cancer stem cell self-renewal through interference with metabolic balance[7]. However, the cellular heterogeneity of cholesterol metabolism in PDAC and its dynamic changes across different cell states and tumor progression stages still lack systematic single-cell level analysis.

In recent years, the rapid development of single-cell RNA sequencing and spatial

transcriptomics technologies has provided unprecedented resolution for dissecting tumor heterogeneity. These technologies have been successfully applied in PDAC research, revealing lineage plasticity of tumor cells, subtype diversity of cancer-associated fibroblasts, and the complex interaction network of the tumor-immune microenvironment[8, 9]. However, in existing studies, there is still a lack of systematic characterization of PDAC cell heterogeneity from the perspective of cholesterol metabolism. In particular, as the recognized cell of origin for PDAC, whether ductal epithelial cells contain heterogeneous subpopulations with significantly different cholesterol metabolism levels, and the functional differences and clinical significance of these subpopulations in tumor progression, remain key scientific questions that urgently need to be addressed.

Based on the above research gaps, this study integrated single-cell transcriptomics, spatial transcriptomics, and bulk transcriptomic data to systematically depict the cellular landscape of cholesterol metabolism in PDAC. We focused on the malignant ductal epithelial cell population and analyzed the molecular characteristics of cholesterol metabolism heterogeneity subpopulations and their interactions with the tumor microenvironment through multidimensional bioinformatics analyses including cell communication networks, pseudotime differentiation trajectories, transcriptional regulatory networks, and machine learning models. This study demonstrates the role of high-cholesterol metabolism cells as signaling hubs in the tumor microenvironment and their close association with enhanced stemness, poor prognosis, and activation of pro-tumor pathways. Notably, through feature importance ranking, we identified NSDHL (NAD(P)-dependent steroid dehydrogenase-like protein), a key enzyme in the cholesterol biosynthesis pathway, as the most important gene. Previous studies have shown that inhibition of NSDHL leads to accumulation of upstream sterol intermediates, which induce cholesterol efflux through activation of the nuclear receptor LXRα, thereby inhibiting EGFR/KRAS-dependent tumor growth[10]. However, the specific role of NSDHL in cholesterol metabolism heterogeneity in PDAC and its potential as a biomarker and therapeutic target still lack systematic functional studies. Therefore, this study further elucidated the function of NSDHL in PDAC malignant progression and its potential as a biomarker and therapeutic target through bioinformatics and experiments, providing a new theoretical basis for precision diagnosis and treatment of PDAC.

Materials and methods

Data collection

The scRNA-seq datasets GSE154778[11] and GSE155698[12] related to pancreatic cancer were obtained from the GEO database. Spatial transcriptomics data were obtained from the GEO database, including three sections: GSE211895-GSM6505133, GSE211895-GSM6505134, and GSE211895-GSM6505135 [13]. Cholesterol metabolism-related genes

were collected from the Molecular Signatures Database (https://www.gsea-msigdb.org/gsea/msigdb/index.jsp), including five cholesterol metabolism-related gene sets: Hallmark Cholesterol Homeostasis genes, GOBP Regulation of Cholesterol genes, WP Cholesterol Metabolism genes, WP Cholesterol Biosynthesis genes, and Reactome Cholesterol Biosynthesis genes. Transcriptomic data were obtained from the GEO database, with three independent datasets: GSE28735[14], GSE62452[15] , and GSE71729[16]. Additionally, transcriptomic data of 29 candidate genes in pancreatic cancer samples were obtained from the TCGA database, and an independent validation dataset for pancreatic cancer was obtained from the ICGC database.

Data analysis

Seurat software package was used to process single-cell RNA sequencing data. Quality control filtering retained cells with mitochondrial gene expression ratio below 20% and detected gene count exceeding 200, and genes expressed in at least 3 cells with expression levels between 200-7000 were selected for subsequent analysis. To identify and remove potential doublets, we used the DoubletFinder software package, determining the optimal pK value by parameter scanning to maximize the mean-variance normalized bimodality coefficient (BCmvn), and removing doublets based on the pANN score threshold. The quality-controlled expression data were normalized and linearly scaled using the "LogNormalize" method. 3000 highly variable genes were identified through the "FindVariableFeatures" function. Based on highly variable genes, principal component analysis (PCA) was used for dimensionality reduction, and the number of significant principal components was determined by JackStraw analysis combined with the Elbow Plot (Supplementary Figure S1A-B). To correct systematic errors between different samples, the Harmony algorithm was used for batch effect correction. The corrected data were visualized in two dimensions using UMAP (Supplementary Figure S1C-D). Cell clustering was performed using graph-based clustering algorithms (FindNeighbors and FindClusters functions) with a resolution of 0.8. Cell subpopulation annotation was based on highly expressed specific genes in each cluster and known classical cell markers.

Single-cell Differential Gene Analysis and Functional Enrichment

The FindAllMarkers function in the Seurat package was used to identify specific marker genes for each cell cluster. The screening criteria for differentially expressed genes were: $|\log_2(\text{fold change})| > 0.3$ and adjusted P value < 0.05 (Bonferroni correction). Subsequently, the marker genes of each cell cluster were input into the clusterProfiler package (v4.0) in R software for Gene Ontology (GO) enrichment analysis. GO terms with adjusted P value < 0.05 were selected as significantly enriched.

Identification of Malignant Epithelial Cells

Based on single-cell gene expression profiles, the inferCNV software package in R software was used for copy number variation (CNV) analysis (https://github.com/broadinstitute/inferCNV). B cells were used as reference cells because of their genomic stability and lack of extensive CNVs. The parameter settings for inferCNV analysis were as follows: cutoff value of 0.1, clustering by groups = TRUE, denoising and hidden Markov model (HMM) enabled. To reduce false positives in CNV inference, the default Bayesian latent mixture model was used to evaluate the posterior probability of variation in each cell, with a threshold of 0.5. Subsequently, based on the CNV patterns of epithelial cells relative to the B cell reference group, CNV scores for all genes were calculated, and hierarchical clustering combined with k-means clustering was used to group epithelial cells and reference B cells. Epithelial cells with low CNV scores that clustered with B cells were identified as non-malignant epithelial cells, while epithelial cells with high CNV scores that formed independent clusters were identified as malignant epithelial cells.

Gene Set Scoring Algorithm in scRNA-seq

Five independent algorithms (AUCell, UCell, singscore, ssGSEA, and AddModuleScore) were used to quantitatively evaluate the activity of cholesterol metabolism-related gene sets in each single cell, to improve the robustness of inference results through multi-algorithm cross-validation. For each cell, the median of the scores obtained from the five algorithms was calculated as its final comprehensive cholesterol metabolism activity score. Based on the comprehensive scores of all malignant cells, cells were divided into three groups using tertiles: low cholesterol metabolism (LCM), intermediate cholesterol metabolism (ICM), and high cholesterol metabolism (HCM).

CytoTRACE Analysis

The CytoTRACE R package (default parameters) was used to infer the differentiation state of each malignant cell. This algorithm is based on the principle that transcriptional diversity gradually decreases during cell differentiation, and evaluates differentiation potential by calculating gene expression diversity[17]. Each cell receives a CytoTRACE score between 0 and 1, with higher scores representing higher transcriptional diversity, i.e., lower differentiation (more stem-like). By comparing CytoTRACE score differences among LCM, ICM, and HCM groups, we evaluated whether cholesterol metabolism activity is associated with specific differentiation states.

Cell Communication

The CellChat R package was used for cell communication network analysis, with ligand-receptor interaction identification based on its built-in CellChatDB database[18, 19]. Malignant cells were divided into three groups according to LCM, ICM, and HCM. For each group, based on the average expression levels of ligands and receptors, the probabilities of ligand-receptor interactions between malignant cells and other cell types in the tumor microenvironment (such as immune cells and stromal cells) were inferred separately. Through this method, we compared the differences in communication network strength and signaling pathway activity among LCM, ICM, and HCM groups. Permutation test was used to identify signaling pathways with significant changes between groups ($P < 0.05$). Finally, circle plots and bubble plots were used to visualize the cell-cell communication network, displaying the main signal sending and receiving patterns in each group.

PROGENy Pathway Activity Inference

The PROGENy (Pathway RespOnsive GENes) method was used for pathway activity inference [20]. We applied the progeny R package to the standardized expression matrix of all malignant cells and calculated activity scores for 14 signaling pathways including JAK-STAT, NF-κB, MAPK, PI3K-AKT, and TGF-β. Following the processing pipeline for single-cell data in existing studies, we integrated the obtained pathway scores as a new assay layer into the Seurat object and performed scaling. Finally, we extracted the pathway activity scores corresponding to LCM, ICM, and HCM groups for subsequent visualization and comparative analysis of signaling pathway activity in these three groups.

scMetabolism

The scMetabolism R package was used for metabolic pathway activity analysis. The sc.metabolism.Seurat function was applied to the standardized expression matrix of all malignant cells, using the KEGG metabolic gene set (85 pathways) as the reference database, with the VISION algorithm for scoring[21]. This analysis calculated an activity score for each KEGG metabolic pathway in each single cell, and the resulting metabolic pathway activity score matrix was integrated as a new assay layer (METABOLISM) into the Seurat object. Based on LCM, ICM, and HCM, we extracted and compared pathway activity scores among the three groups, with $P < 0.05$ considered statistically significant.

Pseudotime Trajectory Analysis

The Monocle3 R package (version 1.3.1) was used for pseudotime trajectory analysis[22]. The quality-controlled malignant ductal epithelial cell expression matrix was used as input. Preprocessing, UMAP dimensionality reduction, and cell clustering were performed sequentially, and the principal graph was learned to construct cell state transition trajectories

(Supplementary Figure S3A-B). The node corresponding to the LCM cell population was selected as the root node, and the pseudotime value of each cell was calculated. The graph_test() function was used to identify genes significantly changing along pseudotime ($q < 0.05$), and gene expression dynamics were visualized as heatmaps.

DrugReflector

The DrugReflector pre-trained model was used for virtual compound screening[23]. By comparing differentially expressed genes between HCM and ICM malignant cells, transcriptomic features characterizing the HCM state were defined. The FindMarkers function in the Seurat package was used to identify differentially expressed genes between the two groups, and the obtained gene expression features were formatted as a ranked list containing genes and their corresponding $\log_2$ fold change values as input queries for DrugReflector.

hdWGCNA

To identify gene co-expression modules associated with HCM and their core regulatory genes, high-dimensional weighted gene co-expression network analysis was performed on malignant cell transcriptomic data using the hdWGCNA R package[24, 25]. First, single cells were aggregated into metacells to reduce data noise, then a weighted co-expression network was constructed by calculating pairwise expression correlations between genes, and genes were clustered into different co-expression modules based on the topological overlap matrix. To ensure the network conformed to scale-free topology characteristics, we set the soft threshold to 3, which optimizes network connectivity while maintaining the scale-free topology model fit above the recommended threshold. To precisely locate modules associated with the target phenotype, we used LCM, ICM, and HCM as phenotypic features and performed module-trait correlation analysis. Modules significantly associated with HCM status were retained as key modules. We identified the blue module as an important module and intersected it with HCM differentially expressed genes, ultimately screening 29 core Hub genes for subsequent research.

Construction and Validation of Multi-gene Prognostic Risk Model

RNA-seq expression data of 29 candidate genes and clinical information for four survival endpoints (OS, DSS, PFI, DFI) were obtained from The Cancer Genome Atlas (TCGA) database. The TCGA-PAAD dataset was used as the training set. Fifteen Cox regression-based algorithms were used to construct risk scoring models, including Lasso (alpha = 1), Ridge (alpha = 0), Elastic Net (alpha from 0.1 to 0.9, 9 types), stepwise Cox regression (bidirectional, backward, forward), and CoxBoost. The optimal regularization parameters for each algorithm were determined through cross-validation,

outputting gene coefficient tables and risk scores. In the external validation phase, independent validation sets for PAAD were obtained from GEO and ICGC databases. Intersection genes between each validation set and the training set were taken, and expression values were converted by exp() and weighted with coefficients to calculate risk scores. Valid models needed to satisfy univariate Cox regression $P < 0.05$ and $HR > 1$ for each endpoint. Random-effects Meta-analysis was used to combine the HR values of the optimal model in each validation set. Patients were divided into high/low risk groups according to the optimal cutpoint of risk scores, and survival differences between groups were evaluated by Kaplan-Meier curves and Log-rank tests. R software packages (ComplexHeatmap, survminer, meta) were used for visualization of heatmaps, KM curves, and forest plots.

Feature Importance Ranking of Candidate Genes

The XGBoost algorithm was used for feature importance ranking[26]. The expression levels of the 29 screened genes in malignant cells were used as input features, and HCM was used as the prediction target to construct an XGBoost classification model. The model was trained based on the gradient boosting tree framework, and the importance of each feature was evaluated by the Gain metric—this metric measures the average contribution of each feature to the reduction of the model loss function at all split nodes. Higher Gain values indicate greater importance of the feature in distinguishing HCM status.

Spatial Differential Expression Analysis

This study analyzed 10x Visium spatial transcriptomics data from three PAAD sections, processed using the Seurat package. The Cottrazm pipeline was used for spatial deconvolution, which integrates spatial transcriptomics data, H&E staining images, and single-cell reference data to predict the cellular composition of each spatial spot and reconstruct cell type-specific gene expression profiles at sub-spot resolution[27]. Based on the proportion of tumor cells (TumorCells) in the deconvolution results, spatial spots were divided into three categories: hMal (high malignancy, $TumorCells > 0.5$), iMal (borderline, $0 < TumorCells \leq 0.5$), and nMal (normal, $TumorCells = 0$). Wilcoxon rank-sum test was used to evaluate the expression differences of NSDHL among the three groups, reporting the mean, standard error, and sample size for each group.

Immunohistochemistry Analysis Based on HPA Database

This study used the Human Protein Atlas (HPA) database (https://www.proteinatlas.org). The HPA database integrates transcriptomic and proteomic data from various normal and tumor tissues, including RNA-seq and immunohistochemistry (IHC) protein expression

profiles. The expression levels of genes in normal and tumor tissues were analyzed using HPA immunohistochemistry data.

Differential Expression Analysis

Three independent GEO datasets were analyzed. Gene expression values of each dataset were z-score standardized to eliminate dimensional differences between platforms, and extreme outlier samples with |z-score| > 3 were removed. Non-parametric tests were used for between-group comparisons: Wilcoxon rank-sum test for two-group comparisons and Kruskal-Wallis test for multi-group comparisons, with two-sided $P < 0.05$ as the significance threshold. Additionally, TCGA-PAAD and CPTAC datasets were simultaneously used for differential expression evaluation to further verify the robustness of the results.

Survival Prognosis and ROC Diagnostic Analysis

Based on NSDHL expression levels in the TCGA-PAAD cohort, patients were divided into NSDHL high-expression and low-expression groups. Kaplan-Meier survival curves were used to compare overall survival differences between the two groups, and two-sided Log-rank test was used to evaluate statistical significance. Additionally, timeROC() function was used for time-dependent ROC analysis to calculate the area under the curve (AUC) at different time points to quantify predictive accuracy.

BioGRID Protein-Protein Interaction Analysis

This study used the BioGRID database (https://thebiogrid.org) for protein-protein interaction (PPI) analysis. Using "NSDHL" as the search term, with species limited to Homo sapiens, entries with clear experimental evidence supporting physical interactions were retained, and a PPI network was constructed and visualized.

Immune-related Feature Analysis

Based on RNA-seq data from the TCGA-PAAD cohort, NSDHL gene expression levels were extracted from tumor samples. The immune feature set referenced the immune-related features used by Thorsson et al. [28] in pan-cancer immune subtyping analysis, including 18 features: SNV Neoantigens, Indel Neoantigens, Wound Healing, Proliferation, Macrophage Regulation, TGF-β Response, TCR Shannon diversity, Lymphocyte Infiltration Score, Intratumor Heterogeneity, TCR Richness, Leukocyte Fraction, TIL Regional Fraction, IFN-γ Response, BCR Evenness, TCR Evenness, Stromal Fraction, BCR Richness, and BCR Shannon diversity. For immune features with available data in the current cancer type and with ≥10 samples having available data, Spearman rank correlation test (R stats::cor.test, method = "spearman") was used to evaluate the direction and strength of correlation. Benjamini-Hochberg method (R stats::p.adjust, method = "BH") was used for multiple testing

correction of all immune feature correlation tests included in the final analysis. Confidence intervals were constructed using the Fisher z-transform method (R atanh() function). Visualization was performed using the ggplot2 R package to draw dumbbell plots.

Virtual Knockout and Gene Enrichment Analysis

The scTenifoldKnk (version 1.0.2) was used for virtual knockout analysis[29]. The scTenifoldKnk function was executed to virtually knock out NSDHL, with the parameter qc set to FALSE (since quality control had already been completed). This tool constructs a wild-type single-cell gene regulatory network and forces all outgoing edge weights of the NSDHL node to zero, simulating the perturbed network after NSDHL knockout. By comparing the wild-type network with the pseudo-knockout network, the perturbation distance and adjusted P value of each gene were calculated. Genes with adjusted $P < 0.05$ were selected as significantly perturbed downstream genes after NSDHL virtual knockout. To quantify the effect size of each perturbed gene, the perturbation distance output by scTenifoldKnk was $\log_2$ transformed as the effect size for ranking and visualization. Subsequently, the clusterProfiler R package (version 4.10.1) [30] was used for KEGG pathway enrichment analysis of significantly perturbed genes, with adjusted $P < 0.05$ as the significance threshold.

Cell Source and Culture

The normal pancreatic cell line hTERT-HPNE was purchased from Beina Biotechnology, and human pancreatic cancer cell lines SW 1990, MIA PaCa-2, AsPC-1, BxPC-3, Capan-1, and PANC-1 were purchased from the Cell Bank of the Chinese Academy of Sciences. All cells were maintained in our laboratory and cultured strictly according to the recommended conditions of the cell bank or manufacturer. The specific culture conditions were as follows: Capan-1 cells were cultured in IMDM medium supplemented with 20% fetal bovine serum (FBS) and 1% penicillin-streptomycin (P/S); SW 1990, MIA PaCa-2, PANC-1, and hTERT-HPNE cells were cultured in DMEM medium supplemented with 10% FBS and 1% P/S; AsPC-1 and BxPC-3 cells were cultured in RPMI-1640 medium supplemented with 10% FBS and 1% P/S. All cells were incubated at 37°C, 5% CO_2, and saturated humidity. The medium was changed every 2-3 days, and when cell confluence reached 80%-90%, cells were digested with 0.25% trypsin-EDTA and passaged at appropriate ratios.

Western Blot

Cells were lysed with RIPA lysis buffer (containing protease and phosphatase inhibitors), and protein concentration was determined by BCA method. Equal amounts of protein (10 μL per well) were separated by SDS-PAGE and transferred to PVDF membranes. Membranes were blocked with 5% skim milk at room temperature for 1 hour, then incubated with anti-NSDHL antibody (1:1000, Proteintech, China) and anti-β-Tubulin antibody (1:10000,

Proteintech, China) at 4°C overnight. After washing, membranes were incubated with HRP-conjugated secondary antibodies (1:5000, Proteintech, China) at room temperature for 2 hours. Enhanced chemiluminescence (ECL) substrate was used for color development, and protein bands were detected by a chemiluminescence imaging system. ImageJ software was used to quantify band gray values, and NSDHL expression levels were normalized to β-Tubulin as an internal reference.

Construction of Stable Transfected Cell Lines

Ready-to-use lentiviruses carrying shRNA targeting NSDHL and non-targeting control shRNA (shNC) were purchased from Hanbio Biotechnology (Shanghai). shNC vector sequence: TTCTCCGAACGTGTCACGTAA; sh-NSDHL-1 vector sequence: GGCAAGAGGATATGCTGTCATAGTA; sh-NSDHL-2 vector sequence: CATTGGCACCAAGAATGTCATTGAA; sh-NSDHL-3 vector sequence: CCTATGCCATGAAACCCATTGACTA. The lentiviral vector carried green fluorescent protein (GFP) reporter gene and puromycin resistance gene. Target cells were seeded in 6-well plates (density of 2×10^5 cells/well) and cultured overnight. The next day, according to the formula: virus volume per well (μL) = (MOI × cell number) / virus titer (TU/mL) × 1000, 10 μL of lentivirus was added per well to infect cells. After 4 hours of infection, 1 mL of complete medium was added to each well. The next day, the virus-containing medium was replaced with fresh complete medium. After 48 hours of infection, GFP expression was observed under a fluorescence microscope to confirm transduction efficiency, and puromycin at 4 μg/mL was added for selection. After 3 passages, stably transduced cell lines were obtained.

CCK8

CCK-8 kit was used to evaluate cell proliferation ability. Stably NSDHL-knockdown cells and control cells in logarithmic growth phase were collected and seeded in 96-well plates at 2000 cells per well in 100 μL complete medium. Five replicate wells were set for each group. Cells were cultured at 37°C, 5% CO_2 incubator. At 0, 24, 48, 72, and 96 hours after inoculation, 10 μL CCK-8 solution (Servicebio, China) was added to each well and incubated at 37°C for 2-4 hours. Absorbance (OD value) at 450 nm was measured using a microplate reader (BioTek, USA). Cell proliferation curves were plotted with OD values against time points. Each experiment was independently repeated 3 times, and data were expressed as mean ± standard deviation.

Colony Formation

Stably NSDHL-knockdown cells and control cells in logarithmic growth phase were collected and seeded in 6-well plates at 1000 cells per well, with 3 replicate wells per group, gently shaken to distribute cells evenly. The 6-well plates were cultured at 37°C, 5% CO_2,

saturated humidity incubator, with fresh complete medium changed every 3 days, continuously cultured for 10-14 days until visible colonies appeared. The medium was discarded, cells were gently washed twice with PBS, and fixed with 4% paraformaldehyde at room temperature for 20 min. After discarding the fixative and washing with PBS, 0.1% crystal violet staining solution was added for 20 min at room temperature. Excess stain was washed off with PBS, air-dried naturally, and photographed under an inverted microscope or gel imaging system

EdU Assay

EdU cell proliferation detection kit (Beyotime, Shanghai) was used. Stably NSDHL-knockdown cells and control cells in logarithmic growth phase were collected and seeded in 24-well plates at 2×10^4 cells per well, cultured overnight. EdU labeling working solution was prepared by diluting EdU stock solution (10 mM) with complete medium at 1:500 ratio to a final concentration of 20 μM. First, 500 μL of cell culture medium was added, followed by an equal volume of EdU working solution, and incubated at 37°C for 2 hours. After EdU labeling, cells were fixed with 4% paraformaldehyde at room temperature for 15 min, permeabilized with 0.3% Triton X-100 for 10 min, washed with PBS, and incubated with prepared Click reaction solution at room temperature in the dark for 30 min for detection. Cell nuclei were counterstained with Hoechst 33342 (1 μg/mL) for 10 min, washed 3 times with PBS, and images were collected using a fluorescence microscope.

Wound Healing Assay

Stably NSDHL-knockdown cells and control cells in logarithmic growth phase were collected and seeded in 6-well plates at 5×10^5 cells per well, cultured in complete medium until cell confluence reached 90-95%. A sterile 200 μL pipette tip was used to create a straight scratch in the center of each well. After scratching, cells were gently washed with phosphate-buffered saline (PBS) to remove detached cells, and the remaining cells were cultured in serum-free medium. Images were observed and photographed under an inverted microscope at 4× magnification at 0h and 24h after scratching.

Transwell Migration and Invasion

24-well Transwell chambers (Corning, USA) with polycarbonate membrane pore size of 8.0 μm were used. For invasion assay, Matrigel matrix gel (Corning, USA) was diluted with serum-free medium at 1:9 ratio, and 50 μL was coated on the bottom surface of the upper chamber membrane, incubated at 37°C for 1 hour to solidify; for migration assay, the upper chamber was not coated. Stably NSDHL-knockdown cells and control cells in logarithmic growth phase were collected and resuspended in serum-free medium. For migration assay, cells were seeded at 5×10^4 cells/100 μL; for invasion assay, cells were seeded at 2.5×10^5

cells/100 μL. 200 μL of cell suspension was added to the upper chamber, and 700 μL of complete medium containing 10% FBS was added to the lower chamber. Transwell chambers were incubated at 37°C, 5% CO_2 incubator (migration assay: 24 h; invasion assay: 48 h). After incubation, non-migrated cells on the upper membrane surface were gently wiped off with a cotton swab, and cells on the lower membrane surface were fixed with 4% paraformaldehyde for 15 min and stained with 0.1% crystal violet at room temperature for 20 min. After PBS washing, images were taken under an inverted microscope.

Subcutaneous Tumor Formation in Nude Mice

BALB/c nude mice (male, 4-6 weeks old, body weight 18-22 g) purchased from Cavens Company were used for subcutaneous tumor formation experiments. All animal experiments were conducted in accordance with the guidelines approved by the Experimental Animal Management and Use Committee (IACUC) of Lanzhou University Second Hospital (approval number: D2026517). Mice were housed under specific pathogen-free (SPF) conditions with 12 h light/12 h dark cycles, with free access to water and food. They were randomly divided into two groups (n = 6). Stably NSDHL-knockdown cells and control cells in logarithmic growth phase were collected, washed twice with PBS, and resuspended in PBS to a density of 1×10^7 cells/mL. Cells were subcutaneously injected into the right axillary region of each mouse. Tumor length (L) and width (W) were measured with vernier calipers every 4 days, and tumor volume (V) was calculated according to the formula $V = (L \times W^2) / 2$. When tumors grew to a certain length, mice were euthanized, tumor tissues were excised, photographed, and weighed.

HE Staining

Subcutaneous tumor tissues from nude mice were fixed with 4% paraformaldehyde for 24 h, then dehydrated through graded ethanol, cleared with xylene, and embedded in paraffin to prepare 4 μm thick continuous sections. Sections were deparaffinized in xylene I and II for 10 min each, rehydrated through graded ethanol (100%, 95%, 85%, 75%) for 5 min each, and rinsed with distilled water. Subsequently, sections were immersed in hematoxylin staining solution at room temperature for 10 min, rinsed under running water, differentiated in 1% hydrochloric acid ethanol for several seconds, blued in 0.6% ammonia water, and rinsed under running water. Then immersed in eosin staining solution at room temperature for 3 min, dehydrated through graded ethanol, cleared with xylene, and mounted with neutral balsam. Images were observed and collected under a light microscope to evaluate tumor tissue morphology, nuclear atypia, and necrosis.

Immunohistochemistry Staining

Paraffin sections were deparaffinized in xylene, rehydrated through graded ethanol, and antigen retrieval was performed using citrate buffer (10 mM, pH 6.0): microwave medium heat for 8 min, pause for 8 min, then medium-low heat for 8 min, naturally cooled to room temperature. PBS washing 3 times, 5 min each. Sections were immersed in 3% H_2O_2 at room temperature for 15 min to inactivate endogenous peroxidase, followed by PBS washing. Blocking was performed with 5% bovine serum albumin (BSA) at room temperature for 30 min. After discarding the blocking solution, anti-Ki67 antibody (1:2000, Proteintech, China) was added and incubated at 4°C in a humid chamber overnight. The next day, after PBS washing, HRP-labeled secondary antibody was added and incubated at room temperature for 30-60 min. After PBS washing, DAB color development solution was applied in the dark for 5 min, color development was controlled under a microscope, and the reaction was terminated by running water. Hematoxylin counterstaining for 2 min, hydrochloric acid ethanol differentiation, ammonia water bluing, graded ethanol dehydration, xylene clearing, and neutral balsam mounting. Images were observed and photographed under a light microscope. ImageJ software was used for image analysis, and all sections were blindly scored by two independent observers.

Statistical Analysis

Statistical analysis and plotting were performed using R (version 4.3.1) and GraphPad Prism (version 10). Data were expressed as mean ± standard deviation (SD) or median (interquartile range, IQR). For two-group comparisons: normally distributed data were analyzed by unpaired two-sided t-test, and non-normally distributed data were analyzed by Wilcoxon rank-sum test. For multi-group comparisons: normally distributed data were analyzed by one-way ANOVA combined with Tukey post-hoc test, and non-normally distributed data were analyzed by Kruskal-Wallis test combined with Dunn post-hoc test. Correlation analysis used Pearson or Spearman correlation coefficients according to data distribution. Kaplan-Meier survival curves were compared using two-sided Log-rank test. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$.

Results

Through the Seurat pipeline for single-cell sequencing analysis, 19 transcriptionally distinct cell clusters were identified in pancreatic cancer samples (Figure 1A). Samples were divided into normal and tumor samples according to tissue type (Figure 1B). Based on typical marker genes, these 19 cell clusters were annotated into 10 categories, including ductal cells, T cells, macrophages, neutrophils, cancer-associated fibroblasts (CAFs), mast cells, plasma cells, acinar cells, B cells, and endothelial cells (Figure 1C). Stratification by tissue origin showed that these cells included 43,060 tumor tissue-derived cells and 5,593 normal tissue-derived cells (Figure 1D). Marker gene expression supported accurate cell type classification

(Figure 1E, Supplementary Figure S1E). For example, CD3D and CD3E were enriched in T cells, MS4A1 and CD79A marked B cells, and MZB1 was specifically expressed in plasma cells. The unique expression patterns of these typical lineage-specific marker genes confirmed the reliability of our cell type annotation. Cell type-specific differentially expressed genes (DEGs) further supported the accuracy of cell type annotation (Supplementary Figure S1F). We compared the proportion differences of various cell types between normal and tumor samples (Figure 1F). In normal samples, T cells were the most abundant, followed by macrophages and acinar cells; while in tumor samples, ductal cells significantly increased in proportion and became the most abundant cell population. Given this proportion reversal phenomenon, we performed GO functional enrichment analysis on each major cell type (Figure 1G). The results showed that the functional characteristics of ductal cells were concentrated in cell-cell junction organization, innate immune response regulation, and amino acid metabolism modification. In contrast, immune cells such as T cells, B cells, and macrophages were mainly enriched in immune effector functions (such as lymphocyte differentiation and phagocytosis), while CAFs and endothelial cells were mainly involved in maintaining microenvironment structure.

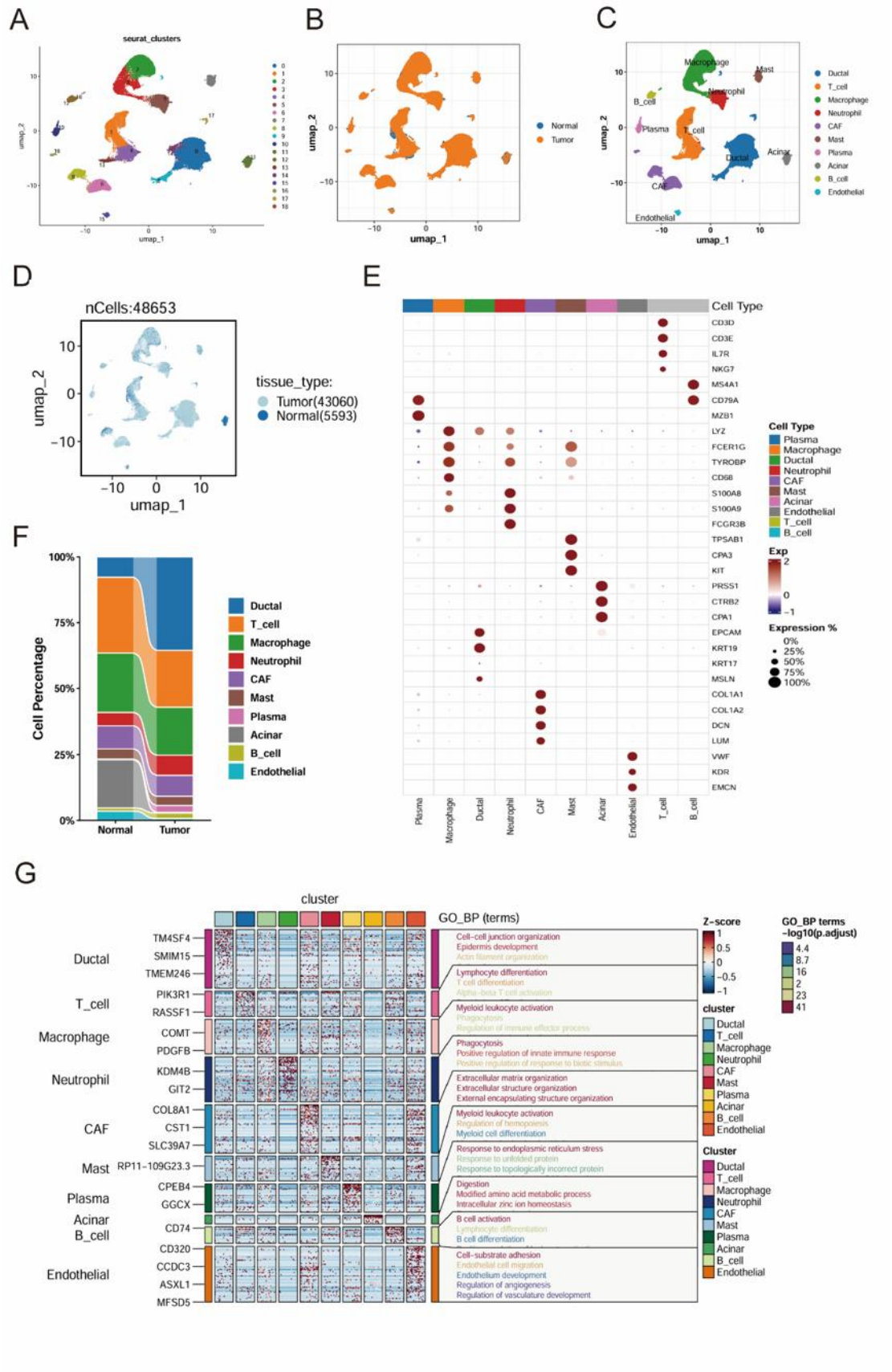


Figure 1. Single-cell transcriptomic atlas of pancreatic cancer. (A) UMAP plot showing 19 transcriptionally distinct cell clusters identified in single-cell RNA sequencing data. (B)

UMAP plot colored by tissue origin (normal and tumor samples). (C) UMAP plot showing 10 major cell types annotated based on typical marker genes. (D) Bar chart showing cell numbers from tumor and normal tissues. (E) Dot plot showing representative marker gene expression of each cell type. (F) Proportion differences of major cell types between normal and tumor samples. (G) GO enrichment analysis of marker genes for each major cell type.

To systematically dissect the heterogeneity of the pancreatic cancer epithelial compartment, we first extracted all epithelial cells based on typical epithelial marker genes and performed unsupervised clustering, resolving them into 12 transcriptionally distinct cell clusters (Figure 2A), and divided them into normal and malignant populations according to tissue origin (Figure 2B). Given that ductal cells are the cell of origin for PDAC and play a central role in tumorigenesis, we further extracted the ductal epithelial cell population for downstream analysis (Figure 2C). To evaluate genomic aberrations and characterize intrapopulation heterogeneity, we performed copy number variation (CNV) analysis using B cells as diploid reference. The heatmap showed the genome-wide CNV landscape of individual ductal epithelial cells (Figure 2D), based on which these cells could be divided into 5 CNV subpopulations. UMAP spatial analysis showed that these 5 subpopulations were clearly separated (Figure 2E), supporting the reliability of CNV-based classification. We further quantified the CNV score of each subpopulation (Figure 2F). Compared with the other four subpopulations, cluster 1 had significantly lower CNV scores (Figure 2G), indicating lack of large-scale chromosomal changes, and was therefore annotated as normal ductal epithelial cells, while the remaining subpopulations were identified as malignant ductal epithelial cells. To verify the accuracy of this refined annotation strategy, we analyzed the tissue origin composition of normal and malignant ductal epithelial cells (Figure 2H-I). The clear separation in UMAP space further validated the reliability of this strategy, confirming that we could accurately distinguish the two cell types for subsequent downstream analysis.

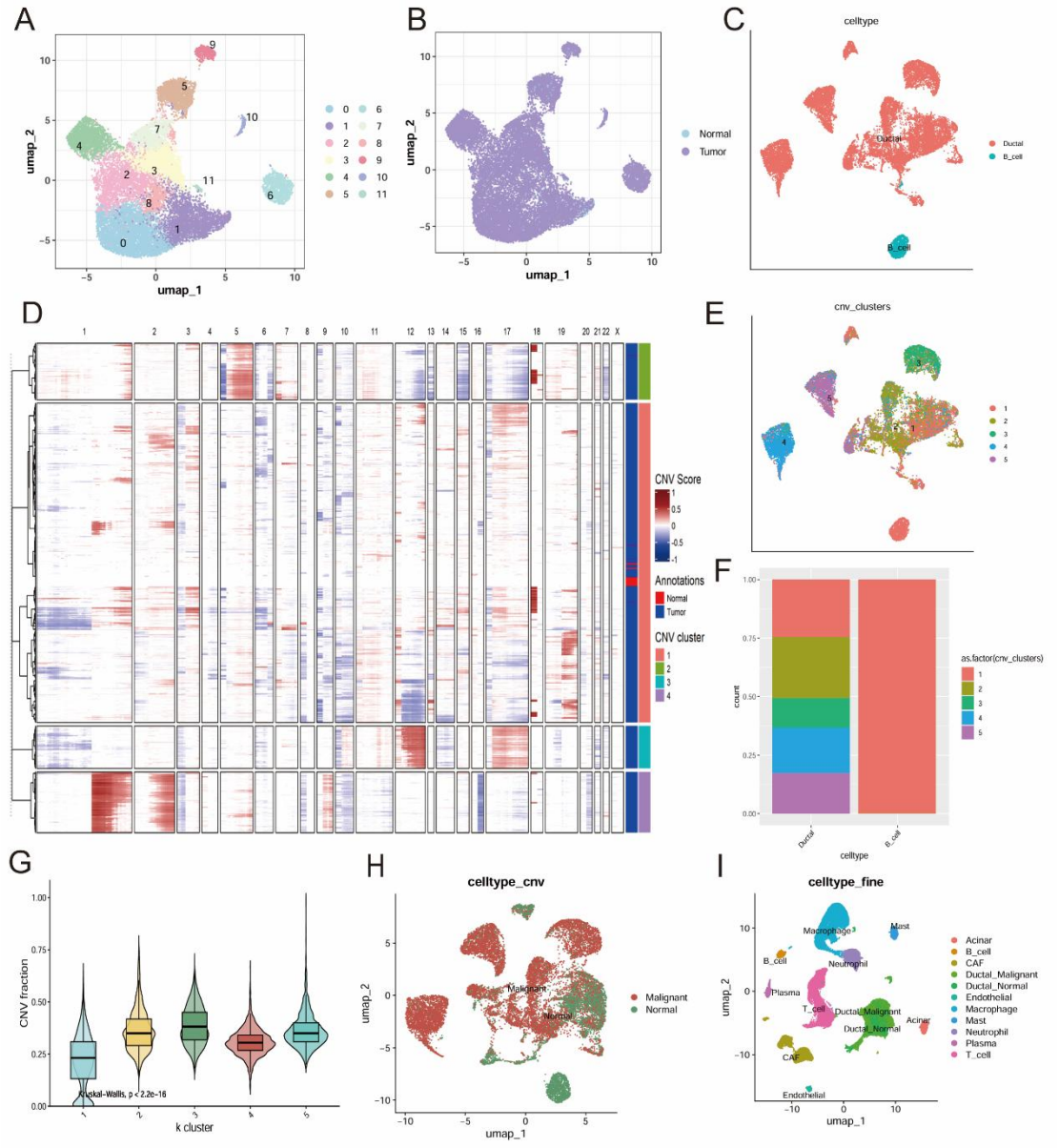


Figure 2. Heterogeneity of the pancreatic cancer epithelial compartment. (A) UMAP plot showing 12 transcriptionally distinct cell clusters of epithelial cells. (B) UMAP plot colored by normal and malignant epithelial cells. (C) UMAP plot of ductal epithelial cells. (D) CNV heatmap showing the genome-wide CNV landscape of individual ductal epithelial cells. (E) UMAP plot showing 5 subpopulations based on CNV classification. (F) Quantitative comparison of CNV scores among subpopulations.

To investigate the cholesterol metabolism landscape of different cell types in the pancreatic tumor microenvironment, we scored the cholesterol activity of each cell population using five cholesterol metabolism-related gene sets. The results showed that ductal epithelial cells had the highest cholesterol metabolism activity among all cell types (Figure 3A-B). Notably, when ductal epithelial cells were stratified into normal and malignant

subpopulations, malignant ductal epithelial cells consistently scored higher than normal ductal epithelial cells in all methods, suggesting that malignant transformation may further upregulate cholesterol metabolism activity (Figure 3C). UMAP plots clearly showed the distribution pattern of cholesterol metabolism activity in malignant ductal epithelial cells (Figure 3D). To further characterize the metabolic heterogeneity within malignant ductal epithelial cells, we performed density distribution analysis of their cholesterol metabolism activity scores. The results showed that malignant ductal epithelial cells exhibited a trimodal distribution, based on which they could be divided into three subpopulations with different cholesterol metabolism levels: low cholesterol metabolism group (LCM), intermediate cholesterol metabolism group (ICM), and high cholesterol metabolism group (HCM) (Figure 3E). UMAP visualization showed that these three subpopulations were clearly separated in space (Figure 3F), indicating substantial differences in their metabolic states. Notably, HCM group cells had significantly higher CytoTRACE stemness scores than the other two groups, and there was a strong positive correlation between cholesterol metabolism scores and CytoTRACE scores (Figure 3I), suggesting that enhanced cholesterol metabolism is closely associated with upregulated stemness features.

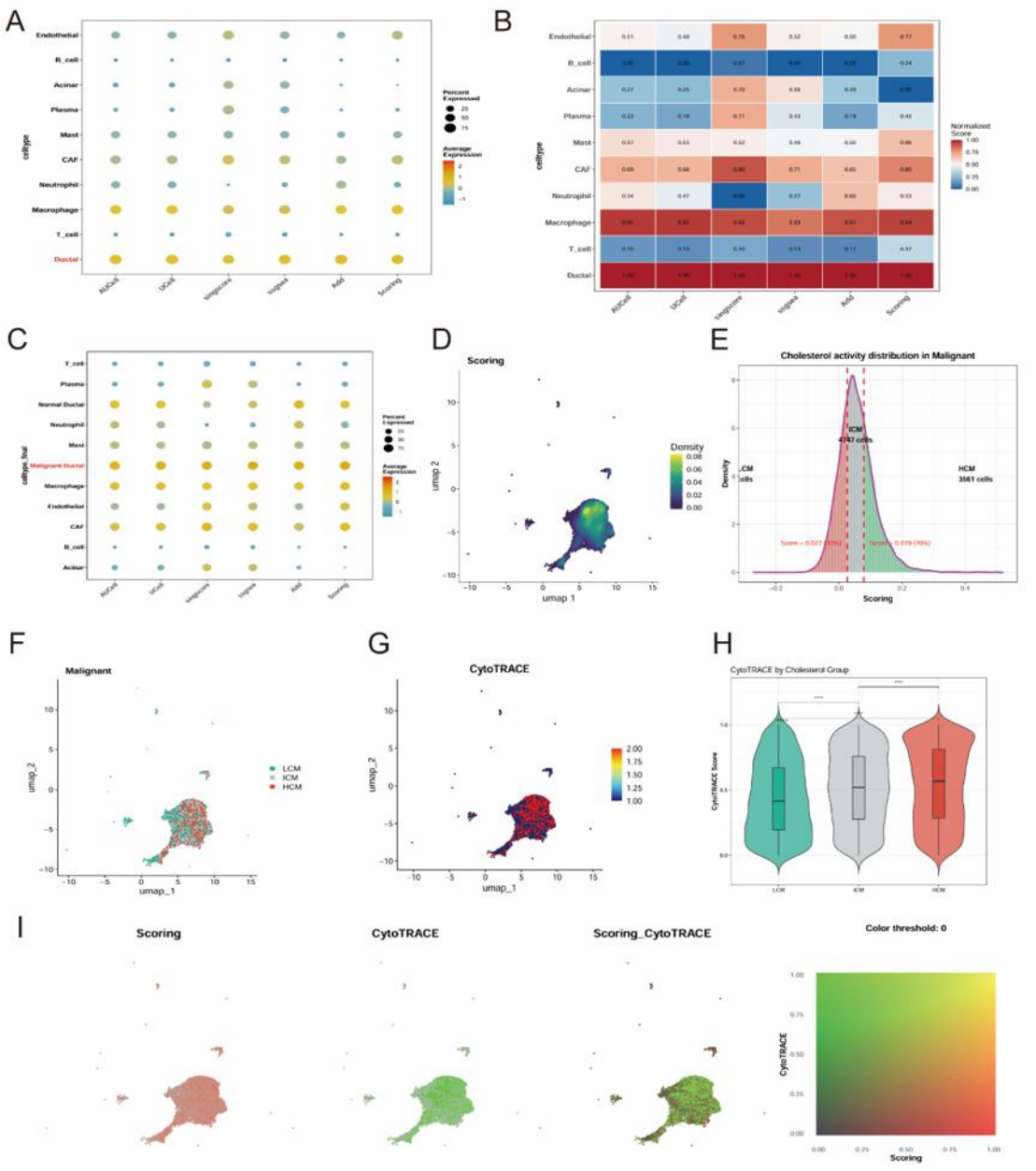


Figure 3. Cholesterol metabolism landscape of pancreatic cancer. (A-B) Cholesterol metabolism activity scores of each major cell type. (C) Comparison of cholesterol metabolism activity between normal and malignant ductal epithelial cells. (D) UMAP plot showing the distribution of cholesterol metabolism activity in malignant ductal epithelial cells. (E) Density distribution of cholesterol metabolism scores showing trimodal distribution, based on which LCM, ICM, and HCM subpopulations were defined. (F) UMAP visualization of LCM, ICM, and HCM subpopulations. (G-H) CytoTRACE scores and group comparisons of LCM, ICM, and HCM subpopulations. (I) Correlation between cholesterol metabolism scores and CytoTRACE scores.

To quantitatively evaluate the cell-cell communication ability of different cell types in the pancreatic tumor microenvironment, we used CellChat to integrate single-cell data and calculated the total number of interactions and interaction weights of each cell type in the global signaling network. Among all malignant ductal subpopulations, the HCM group had

significantly higher interaction frequency and strength with multiple cell types than the LCM and ICM groups, with particularly prominent communication with CAFs and endothelial cells (Figure 4A-B, Supplementary Figure S2A). Directed signal strength analysis further showed that the HCM group had the highest output and input signal strength (Figure 4C), indicating that cholesterol metabolism levels are positively correlated with both signal sending and receiving abilities, suggesting that HCM cells are core hubs in the cell-cell communication network. At the ligand-receptor pair level, from LCM to ICM to HCM, communication between malignant ductal cells and T cells and neutrophils gradually increased, mainly mediated by HLA family molecules; the HCM group also involved significantly more ligand-receptor pairs than the other two groups (Figure 4D). Further detection of integrin and adhesion-related gene expression levels revealed that CD44, ITGAV, and ITGB1 had the highest expression in the three subpopulations, showing a gradient increasing trend of LCM < ICM < HCM (Figure 4E). These genes mediate hyaluronic acid binding, TGF-β activation, and extracellular matrix adhesion, respectively, suggesting that HCM cells may have stronger migration, signal transduction, and matrix interaction abilities. In multiple signaling networks, HCM cells exhibited the strongest signal sending and receiving abilities (Figure 4F). At the signaling pathway level, HCM cells showed the highest activity scores in multiple pathways including Hypoxia, TNFα, and NF-κB (Figure 4G, Supplementary Figure S2B), suggesting close association with hypoxia adaptation, inflammatory signaling, immune regulation, and anti-apoptotic survival signals. Synergistic activation of multiple pro-tumor pathways may jointly drive the invasive phenotype of HCM cells. Additionally, metabolic pathway enrichment analysis showed that HCM cells had significantly higher activity in pathways including glycolysis/gluconeogenesis, fatty acid biosynthesis, and amino acid metabolism (including glycine, serine, and threonine metabolism) than LCM cells (Figure 4H), indicating that HCM cells have synergistic advantages in metabolic reprogramming and signaling hub functions.

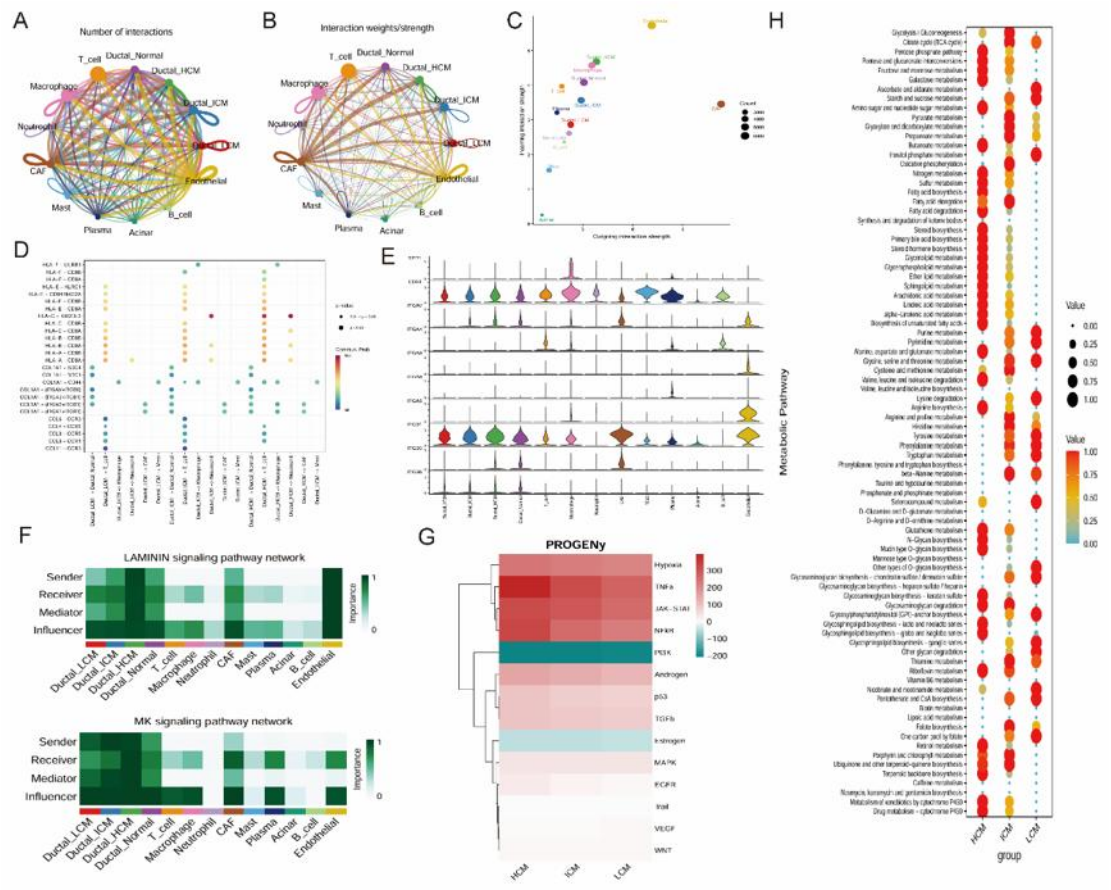


Figure 4. Cell communication and signaling networks of malignant ductal epithelial cells. (A-B) CellChat analysis showing the number and strength of interactions between HCM cells and other cell types. (C) Scatter plot of signal sending and receiving strength of cell-cell communication. (D) Bubble plot of ligand-receptor pair communication among different cell subpopulations. (E) Violin plot of adhesion and integrin-related gene expression in different cell types. (F) Signaling network with the strongest signal sending/receiving ability in HCM cells. (G) Heatmap of PROGENy signaling pathway activity in different cholesterol metabolism subpopulations. (H) Metabolic pathway enrichment analysis.

To investigate the differentiation trajectory of malignant ductal subpopulations with different cholesterol metabolism states, we performed pseudotime analysis based on UMAP dimensionality reduction. LCM, ICM, and HCM cells occupied different regions in UMAP space (Figure 5A). Pseudotime trajectory analysis showed that malignant ductal cells gradually evolved along a continuous differentiation axis from LCM-like state to ICM-like state to HCM-like state, suggesting that cholesterol metabolism levels gradually increase during differentiation, and HCM represents the terminal differentiation state with the highest malignant potential (Figure 5B-C, Supplementary Figure S3C). These findings position cholesterol metabolism not only as a static biomarker but also as a dynamic feature during tumor progression, providing a theoretical basis for targeting cholesterol metabolism in advanced pancreatic cancer. To identify dynamic gene expression changes along the

differentiation trajectory, we clustered the expression patterns of differentially expressed genes according to pseudotime, dividing them into Cluster 1, 2, and 3, corresponding to late, intermediate, and early cell function modules, respectively (Figure 5D, Supplementary Figure S3D-E). The expression trend plots of genes in each cluster along pseudotime showed that NSDHL, as a key enzyme in cholesterol biosynthesis, maintained relatively constant activity throughout the differentiation trajectory (Figure 5E, Supplementary Figure S3F), suggesting it may serve as a fundamental driver of cholesterol metabolism reprogramming. To further explore potential intervention strategies targeting HCM cells, we performed differentially expressed gene analysis between HCM and LCM subpopulations using DrugReflector (Figure 5F), screening the top 10 differentially expressed genes in the two subpopulations (Figure 5G-H). Based on these gene features for drug prediction, the results showed that BRD-A23067620 was the most effective candidate compound against the HCM group (Figure 5I-K), providing a potential direction for drug development targeting high-cholesterol metabolism malignant ductal cells.

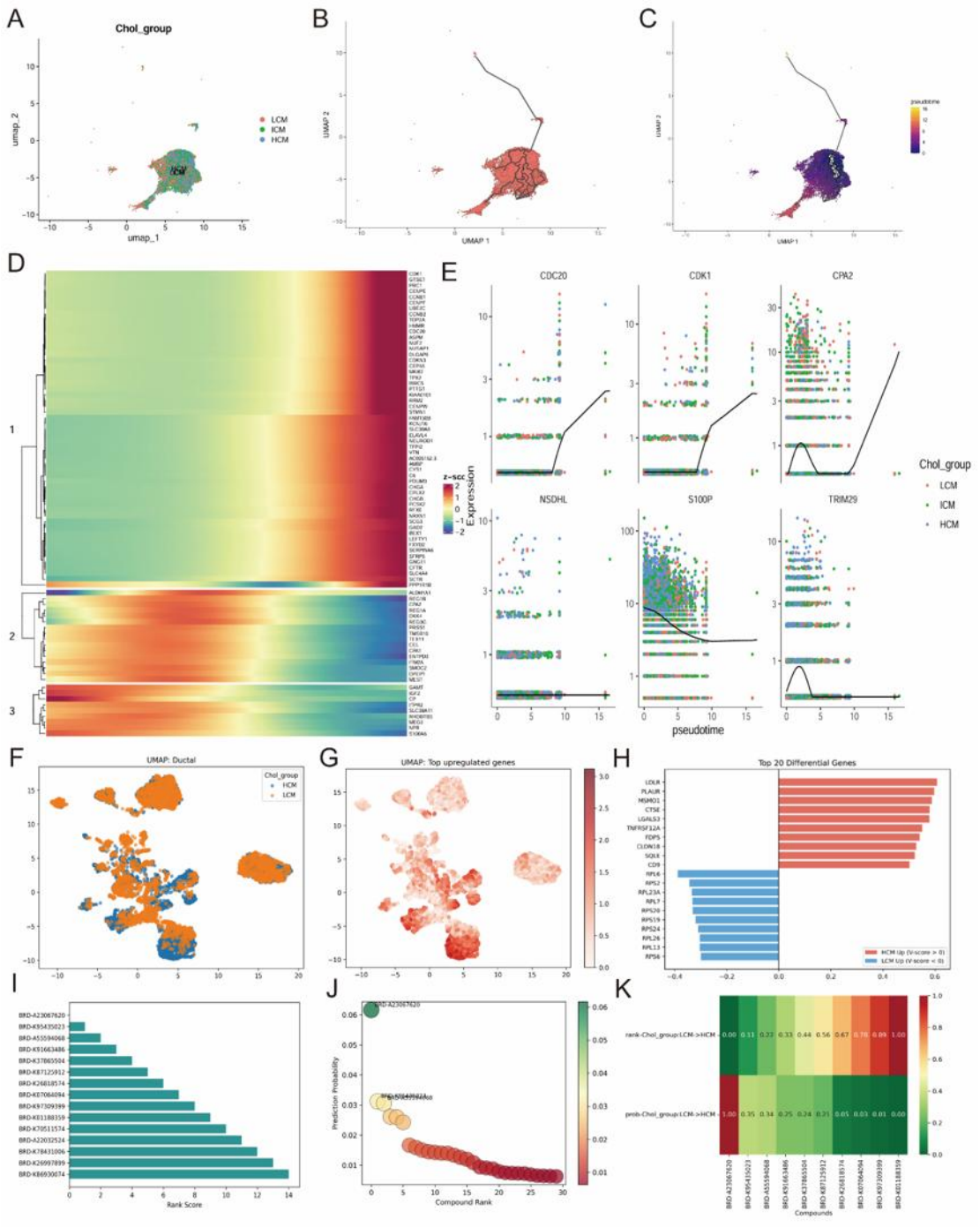


Figure 5. Pseudotime trajectory and drug prediction. (A) UMAP distribution of cholesterol metabolism subpopulations in malignant ductal epithelial cells. (B-C) Pseudotime trajectory display. (D) Differentially expressed genes along pseudotime clustered into Cluster 1, 2, and 3. (E) Expression dynamics of representative genes along pseudotime. (F) DrugReflector analysis of differentially expressed genes between HCM and LCM subpopulations. (G-H) Top 10 differentially expressed genes in each subpopulation. (I-K) Drug prediction results.

To reveal the gene expression programs associated with the HCM group, we performed high-dimensional weighted gene co-expression network analysis (hdWGCNA). Based on scale-free topology criteria, soft threshold power 3 was selected to optimize network

connectivity (Figure 6A). Hierarchical clustering identified 8 co-expression modules, represented by different colors (Figure 6B). Module eigengene (ME) analysis revealed key contributing genes of each module (Figure 6C, Supplementary Figure S3H), and the module correlation heatmap showed good independence among modules (Figure 6D). Bubble plots further showed that blue, green, red, and brown modules were specifically enriched in HCM cells (Figure 6E). After projecting module genes into UMAP space, the blue module showed the most significant enrichment (Figure 6F, Supplementary Figure S3G). Therefore, we extracted blue module genes (Figure 6G) and intersected them with HCM differentially expressed genes, ultimately screening 29 core Hub genes for subsequent research (Figure 6H).

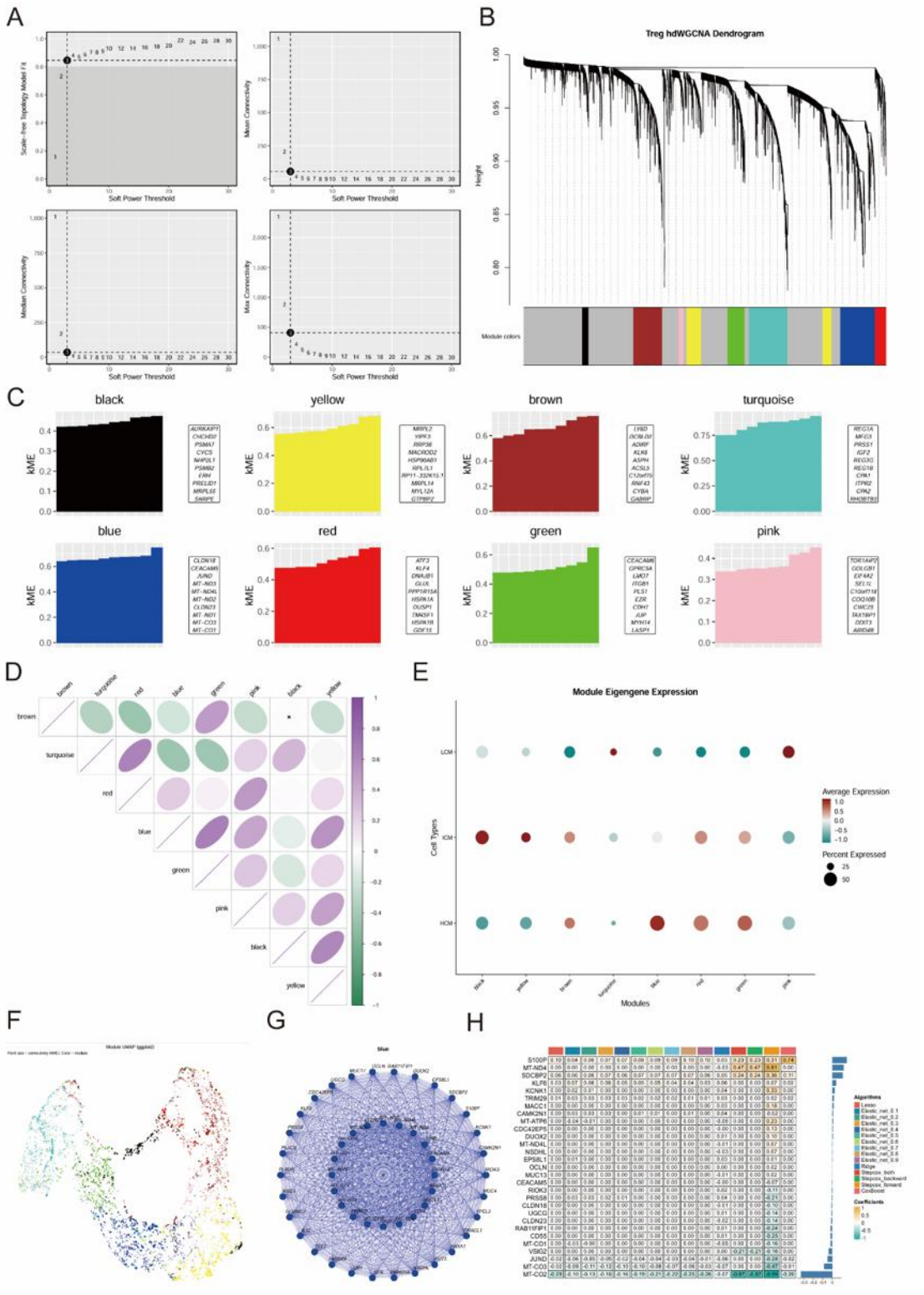


Figure 6. hdWGCNA identifies HCM-related co-expression modules. (A) Soft threshold power selection. (B) Hierarchical clustering dendrogram showing 8 co-expression modules. (C) Module eigengene (ME) analysis. (D) Module correlation heatmap. (E) Bubble plot showing modules specifically enriched in HCM cells. (F) UMAP projection of module genes, with the blue module showing the most significant enrichment. (G) Blue module hub gene co-expression network. (H) Intersection of blue module genes with HCM differentially expressed genes,

ultimately obtaining 29 core Hub genes.

Based on 29 candidate genes in the TCGA-PAAD training set, we constructed 15 Cox regression prognostic risk scoring models and independently validated them in multiple external validation datasets. Among all models, the Ridge model achieved the highest average AUC across validation cohorts (average AUC = 0.644, training set AUC = 0.663) (Figure 7A). Random-effects Meta-analysis further confirmed that this risk score has consistent prognostic predictive ability across multiple cohorts, with high-risk group patients having significantly worse overall survival than low-risk group patients (Figure 7B). Multi-group KM survival analysis showed that in the training set, the high-risk group had significantly lower OS, RFS, DSS, DFI, and PFI than the low-risk group, and this result was consistently validated in validation sets and the total dataset (Figure 7C-K).

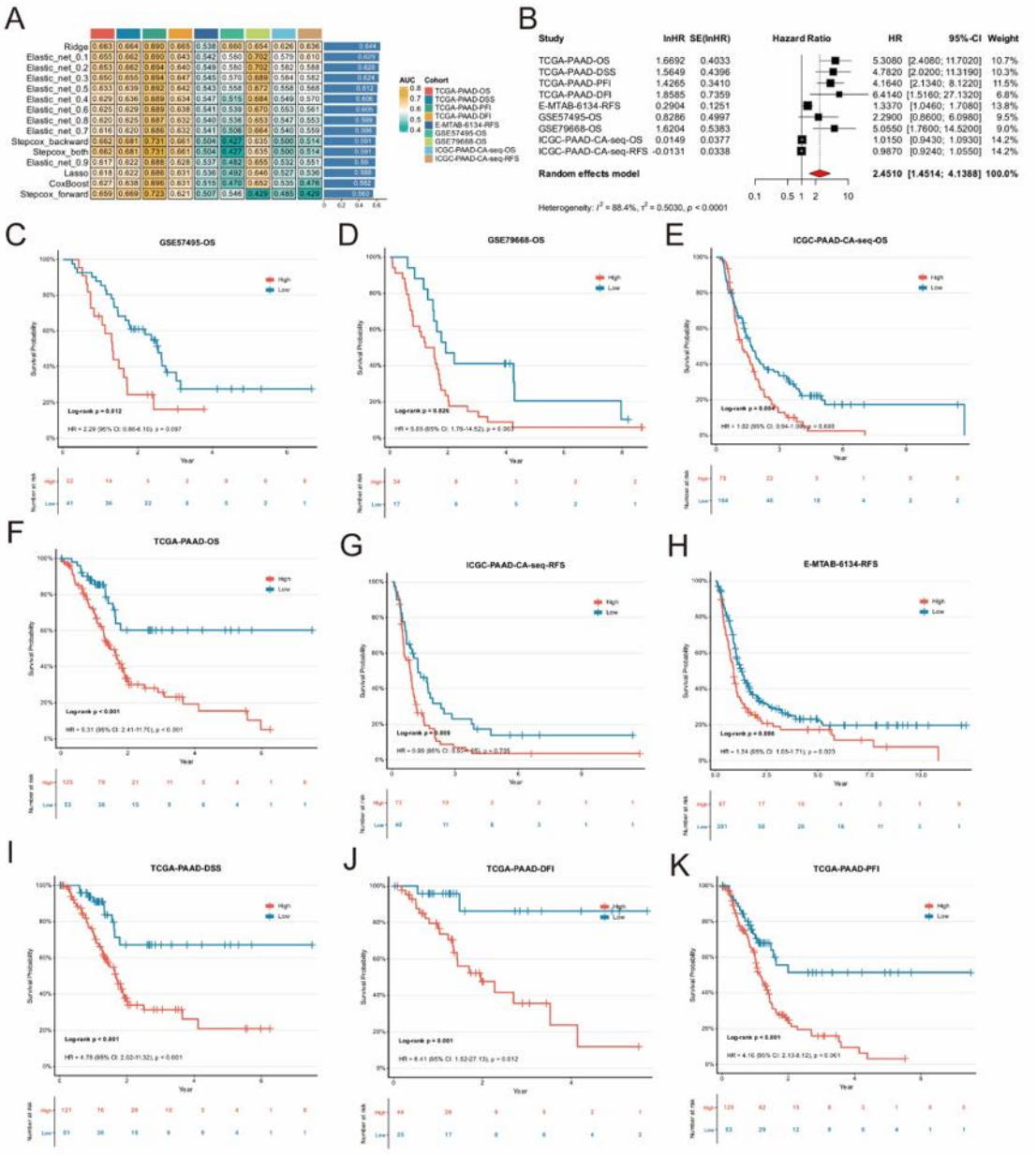


Figure 7. Construction and validation of multi-gene prognostic risk model. (A) Time-dependent AUC heatmap of 15 Cox regression models across multiple cohorts and survival endpoints. (B) Random-effects Meta-analysis forest plot. (C-K) Kaplan-Meier survival curves of OS, RFS, DSS, DFI, and PFI between high-risk and low-risk groups in each cohort.

To identify key genes driving the prognostic risk score, we performed feature importance ranking on these 29 candidate genes, and the results showed that NSDHL ranked first (Figure 8A). As a key enzyme in the cholesterol biosynthesis pathway, this ranking of NSDHL further supports its role as a core gene of the HCM phenotype. To investigate the spatial distribution of NSDHL expression associated with malignant progression, we performed spatial transcriptomics analysis and divided tissue regions into hMAL (high malignancy), iMAL (borderline), and nMAL (normal) groups according to pathological features. The results showed that NSDHL had the highest expression in the hMAL region (Figure 8B-D), and this spatial gradient was highly consistent with the pseudotime trajectory observed in single-cell data (cholesterol metabolism gradually increasing from LCM to ICM to HCM). To further analyze the cellular distribution characteristics of NSDHL, we projected its expression level onto the UMAP dimensionality reduction space of the single-cell transcriptomic atlas. The results showed that NSDHL was significantly enriched in pancreatic ductal epithelial cell populations (Figure 8E). Based on this finding, we further split its expression into normal and tumor parts for comparison, and found that NSDHL expression in tumor ductal epithelial cells was significantly higher than in normal ductal epithelial cells ($P < 0.05$, Figure 8F), suggesting its upregulation is closely related to the occurrence and development of pancreatic cancer. Immunohistochemistry staining analysis through the HPA database also confirmed that NSDHL is highly expressed in pancreatic tumor tissues (Figure 8G). To systematically evaluate the expression level of NSDHL in pancreatic cancer, we further performed differential expression validation analysis based on 3 GEO datasets, CPTAC dataset, and TCGA dataset, and the results consistently showed that NSDHL expression in tumor tissues was significantly elevated compared to normal tissues (Figure 8H-M). Subsequently, Western blot results showed that NSDHL expression in pancreatic cancer cell lines was significantly higher than in normal pancreatic cells (Figure 8N). Survival prognosis analysis showed that patients with high NSDHL expression had worse prognosis (Figure 8O), and diagnostic ROC curve analysis showed that NSDHL had moderate to good diagnostic accuracy (Figure 8P).

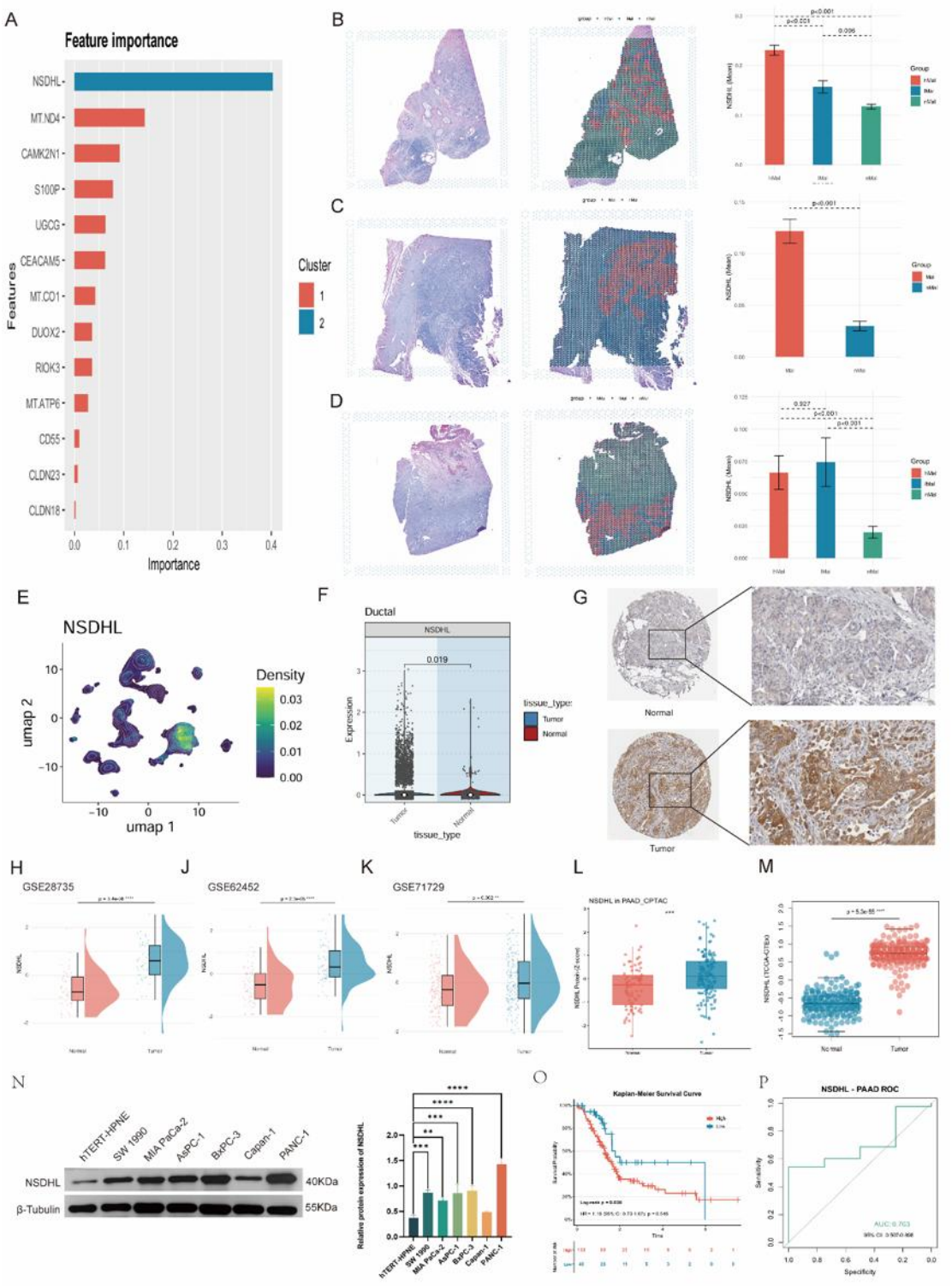


Figure 8. Expression and clinical significance validation of NSDHL in PDAC. (A) XGBoost feature importance ranking. (B-D) Spatial transcriptomics analysis showing highest NSDHL expression in hMAL region. (E) UMAP projection showing NSDHL enrichment in ductal epithelial cells. (F) Comparison of NSDHL expression between normal and tumor ductal epithelial cells. (G) HPA database NSDHL immunohistochemistry staining. (H-M) NSDHL differential expression validation in 3 GEO datasets, CPTAC, and TCGA datasets. (N) Western blot detection of NSDHL expression in pancreatic cancer cell lines. (O) Kaplan-Meier survival analysis of

patients with high and low NSDHL expression. (P) Time-dependent ROC curve of NSDHL.

To investigate the interacting protein network of NSDHL, we performed protein-protein interaction analysis using the BioGRID database. The results showed that NSDHL interacts with multiple proteins including ORC4 and SERBP1 (Figure 9A). To investigate the association between NSDHL expression and the pancreatic cancer immune landscape, we analyzed the correlation between NSDHL expression and 18 immune-related features. The results showed that NSDHL reached statistically significant levels with 10 immune features (BH-adjusted $P < 0.05$), with the strongest positive correlation being SNV Neoantigens (rho = 0.496). Additionally, NSDHL expression was positively correlated with wound healing, proliferation, and intratumor heterogeneity features (Figure 9B), suggesting that tumors with high NSDHL expression have higher mutational load and proliferative activity. To identify downstream genes regulated by NSDHL, we performed virtual knockout analysis by perturbing NSDHL in the single-cell regulatory network. After virtual knockout, the perturbed genes TFF1, REG4, and MMP1 were the most significant (Figure 9C). To deeply understand the functions of these perturbed genes, we performed KEGG pathway enrichment analysis on significantly upregulated genes, and the results showed that enriched pathways were mainly concentrated in immune regulation and epithelial homeostasis-related pathways (Figure 9D). This enrichment pattern suggests that NSDHL may normally inhibit immune activation and maintain epithelial integrity, and its loss triggers a coordinated response of inflammatory reactions, antigen presentation, and adhesion remodeling.

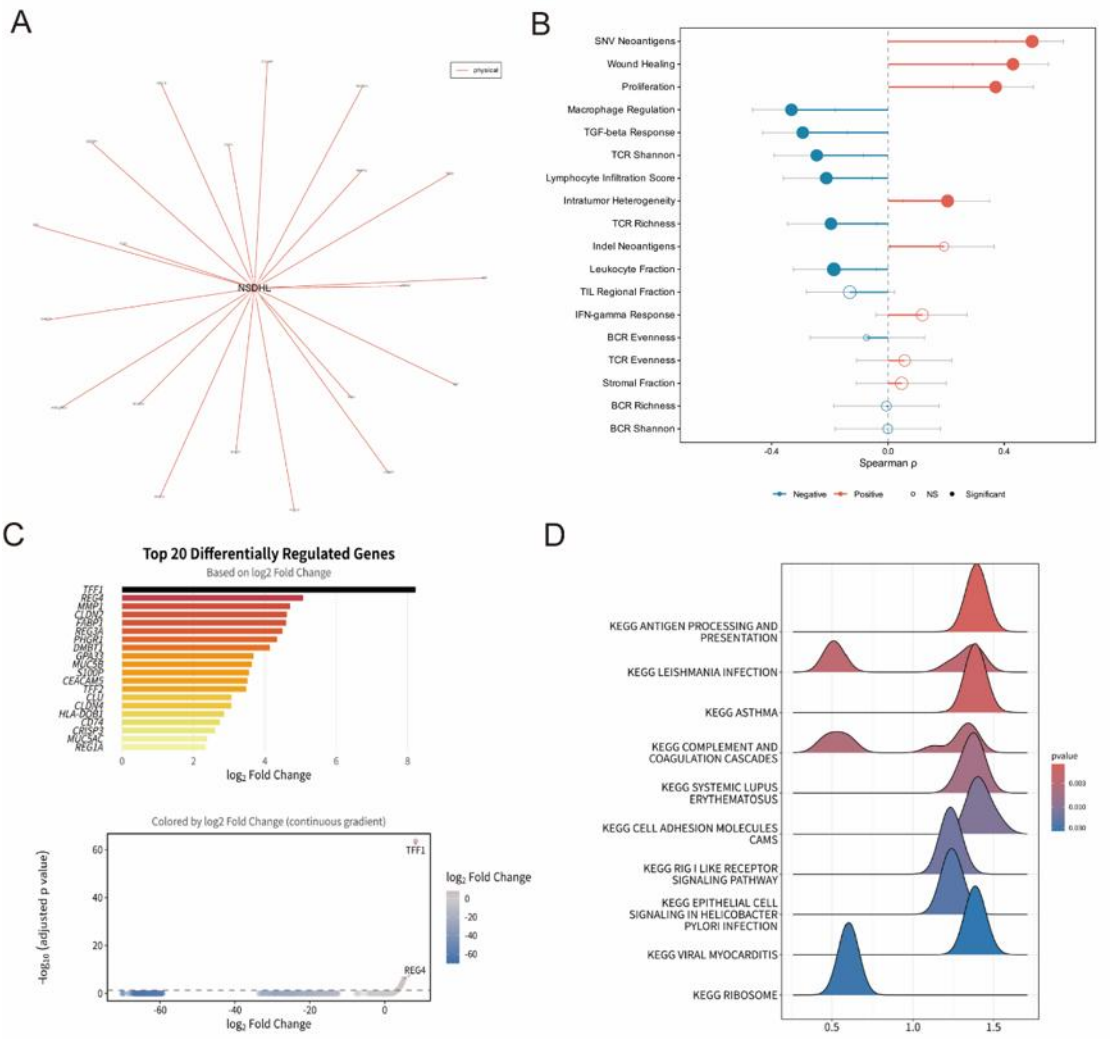


Figure 9. NSDHL-related network and functional prediction. (A) Protein-protein interaction network. (B) Forest plot showing the correlation between NSDHL expression and 18 immune-related features. (C) Top 20 differentially regulated genes after virtual knockout of NSDHL. (D) KEGG enrichment analysis of significantly perturbed genes.

Given that NSDHL had the highest expression levels in PANC-1 and BxPC-3 pancreatic cancer cell lines, we constructed NSDHL stable knockdown cell lines in these two cell lines through lentiviral transfection and verified the knockdown efficiency by Western blot (Figure 10A). Functional experiments showed that after NSDHL knockdown, CCK-8 assay showed significantly inhibited cell proliferation ability (Figure 10B), colony formation assay also showed significantly inhibited cell proliferation ability (Figure 10C), EdU assay showed significantly reduced proportion of EdU-positive cells (Figure 10D), wound healing assay showed significantly weakened cell migration ability (Figure 10E), and Transwell migration and invasion assays further confirmed that NSDHL knockdown significantly reduced the migration and invasion abilities of pancreatic cancer cells (Figure 11A). To evaluate the pro-tumor role of NSDHL in vivo, we constructed a nude mouse subcutaneous tumor formation

model. The results showed that compared with the control group, the NSDHL knockdown group had significantly slowed tumor growth rate and significantly reduced tumor weight (Figure 11B-D). HE staining suggested that the NSDHL knockdown group had lower nuclear atypia, and immunohistochemistry staining showed that the NSDHL knockdown group had lower Ki67 expression (Figure 11E). The above results indicate that NSDHL promotes the proliferation, migration, and invasion of pancreatic cancer both in vitro and in vivo.

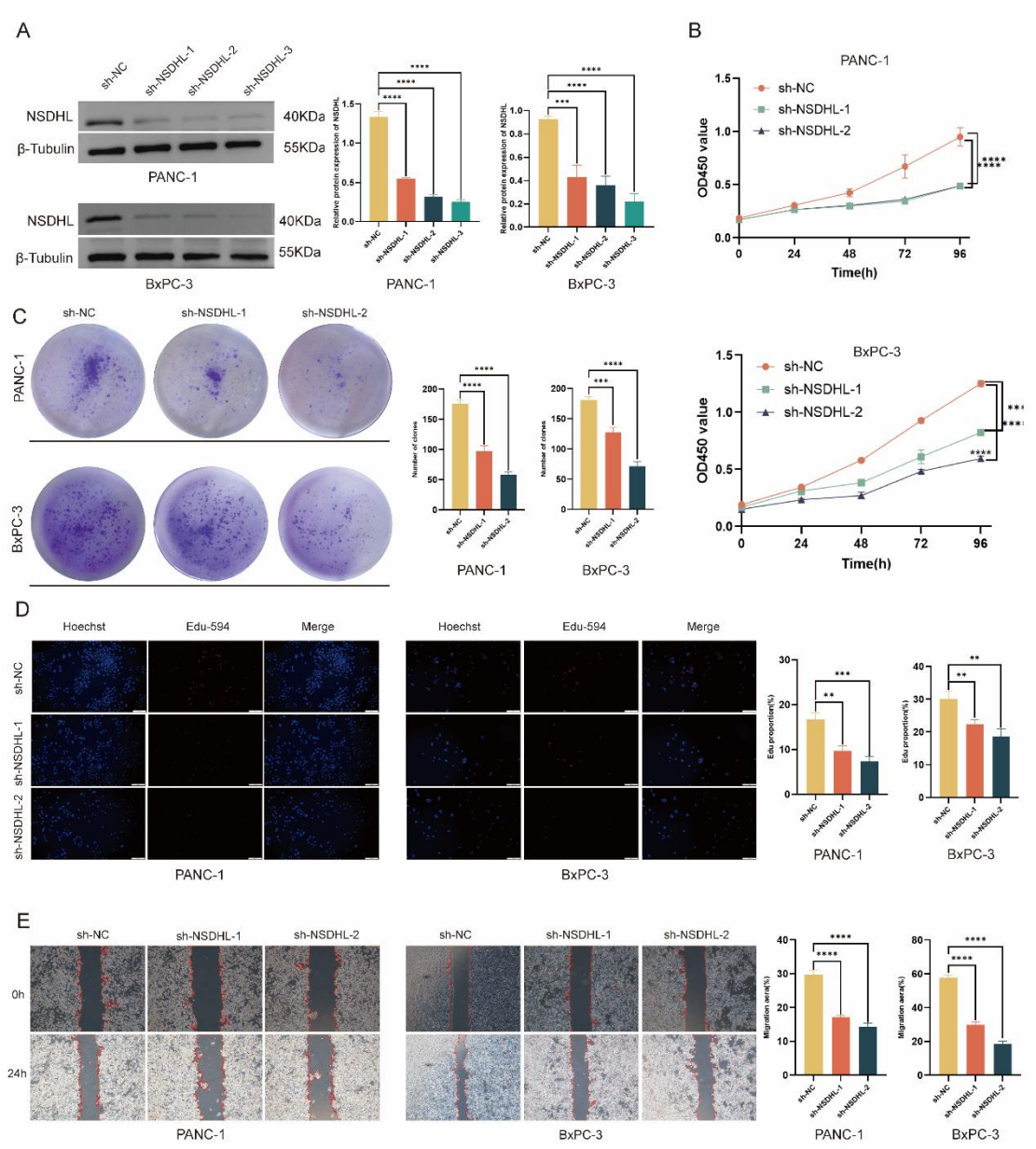


Figure 10. In vitro functional validation of NSDHL knockdown. (A) Western blot validation of NSDHL knockdown efficiency. (B) CCK-8 proliferation assay. (C) Colony formation assay. (D) EdU assay. (E) Wound healing assay. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, $****P < 0.0001$.

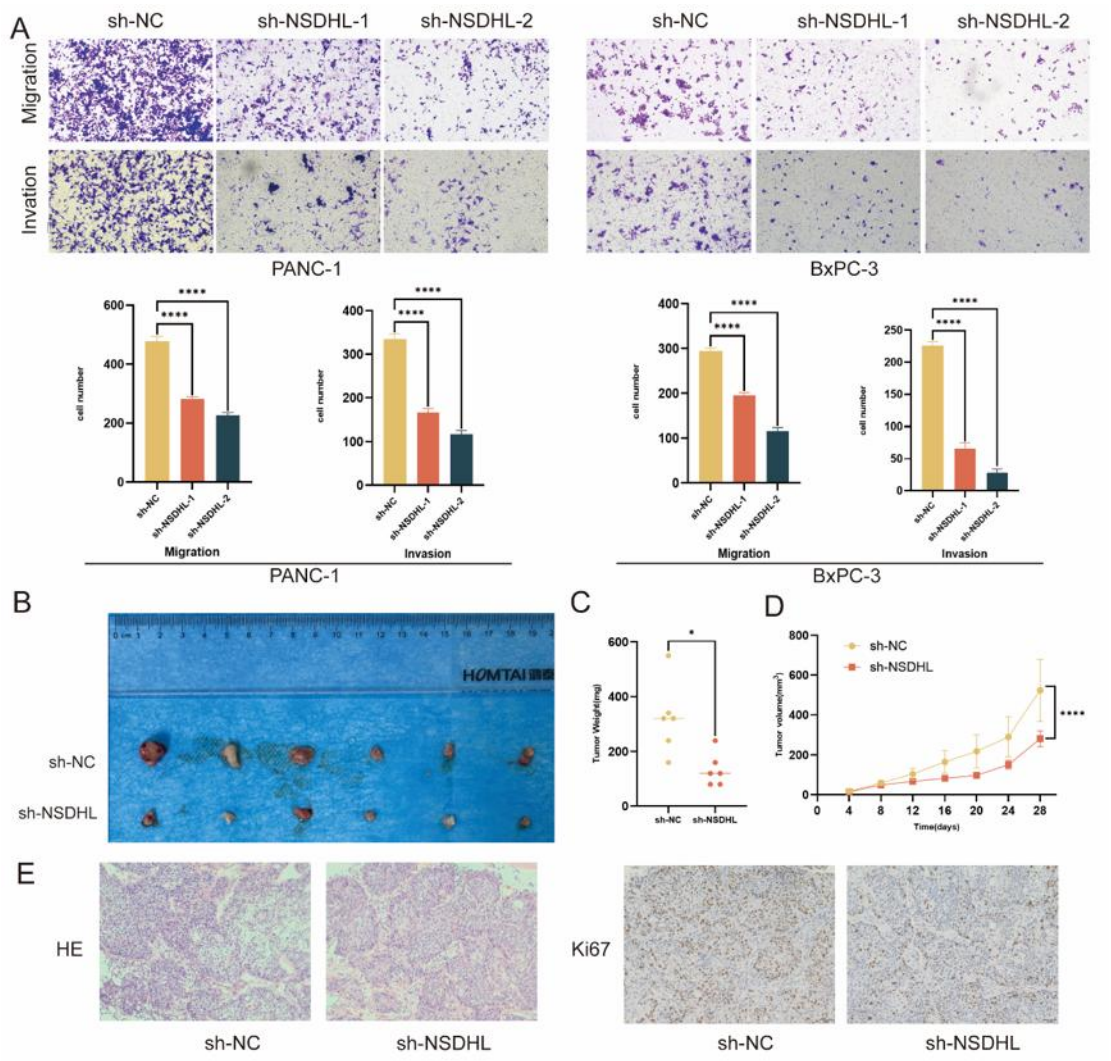


Figure 11. In vitro and in vivo experimental validation of NSDHL knockdown. (A) Transwell migration and invasion assay. (B) Excised tumors from nude mouse subcutaneous tumor formation model. (C) Tumor weight. (D) Tumor growth curve. (E) HE staining and immunohistochemistry staining. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$.

Discussion

The extremely high lethality of PDAC is closely related to its high heterogeneity, early metastasis, and intrinsic resistance to existing therapeutic regimens[31, 32]. Among the many factors driving malignant progression of PDAC, metabolic reprogramming, especially abnormal activation of lipid metabolism, has become one of the recognized core events in the pancreatic cancer field [4]. Although previous studies have revealed the key roles of cholesterol metabolism in maintaining tumor cell membrane homeostasis, integrating oncogenic signals, and remodeling the microenvironment[33, 34], the heterogeneity of this process at the single-cell level, its dynamic evolution characteristics, and its specific impact on tumor cell functional phenotypes remain largely a "black box." Currently, most studies on

cholesterol metabolism in PDAC are based on bulk samples, which cannot resolve functional differences among different cell subpopulations, nor can they distinguish metabolic interactions between tumor cells and other cell types in the microenvironment. This limitation in resolution directly leads to a lack of systematic understanding of cholesterol metabolism heterogeneity and its clinical significance, and also limits the precise discovery of related therapeutic targets.

This study specifically addresses the above gaps by using single-cell and spatial transcriptomics technologies to systematically depict the heterogeneity landscape of cholesterol metabolism in PDAC at the single-cell level. The relationship between high cholesterol and pancreatic cancer is extremely complex. Blood cholesterol levels are negatively correlated with pancreatic cancer risk[35], but cholesterol metabolism reprogramming within tumors is key to driving their malignant progression[36, 37]. This seemingly contradictory phenomenon suggests that the negative correlation at the epidemiological level is not a "protective effect" of cholesterol, but rather a "reverse causation" result of preclinical pancreatic cancer actively consuming blood cholesterol to meet its own malignant proliferation needs. The highly active cholesterol metabolism reprogramming within tumors, through mechanisms such as driving membrane biogenesis, integrating oncogenic signals, promoting metastatic organ selection, and inducing therapeutic resistance, becomes the true core driver of PDAC malignant progression. Our study clearly reveals that malignant ductal epithelial cells are the population with the highest cholesterol metabolism activity in the tumor microenvironment, and can be further divided into three functionally distinct subpopulations (LCM, ICM, HCM) based on metabolic activity levels. Among them, the HCM subpopulation not only shows extremely active cholesterol metabolism but also has significantly enhanced communication strength with cancer-associated fibroblasts (CAFs) and endothelial cells, constituting the core hub of the microenvironment signaling network.

The malignant progression of tumors depends not only on genetic mutations in cancer cells themselves but is also deeply constrained by complex interactions with the microenvironment. In solid tumors rich in fibrotic stroma such as pancreatic cancer, hypoxia caused by abnormal vascular networks is one of the core stressors driving tumor invasiveness [38]. On one hand, the hypoxic environment directly induces upregulation of glycolytic flux through transcription factors such as HIF-1α to meet the energy demands of rapid proliferation; on the other hand, it can activate the NF-κB signaling axis, synergizing with inflammatory factors such as TNFα to jointly construct a pro-tumor inflammatory microenvironment that supports immune evasion and chemotherapy resistance [39-41]. Meanwhile, to adapt to nutritional deprivation and oxidative stress, cancer cells must undergo drastic metabolic reprogramming, not only relying on aerobic glycolysis (Warburg effect) but

also enhancing de novo fatty acid synthesis to supply cell membrane construction and signaling molecule generation, thereby maintaining high proliferation rate and stemness characteristics [42-44]. Through PROGENy pathway analysis, HCM cells showed synergistic activation of pro-tumor inflammatory pathways such as hypoxia, TNFα, and NF-κB, and metabolic profiling results showed synchronous upregulation of multiple metabolic pathways including glycolysis and fatty acid synthesis. This suggests that the enhancement of cholesterol metabolism is not an isolated event but is closely coupled with pro-tumor signal transduction and global metabolic reprogramming, jointly shaping the invasive phenotype of HCM cells. This "metabolism-signal-communication" trinity synergistic model provides a new mechanistic understanding for explaining the rapid progression of PDAC.

Based on pseudotime trajectory analysis, we further revealed that cholesterol metabolism levels dynamically increase with malignant transformation progression, namely from LCM through ICM to HCM, suggesting that HCM represents the terminal differentiation state with the highest malignant potential. This finding elevates cholesterol metabolism from a static biomarker to a dynamic driver in tumor evolution, representing an important conceptual breakthrough. Its clinical significance lies in that therapeutic strategies targeting cholesterol metabolism may show better efficacy in later-line or advanced patients, because tumor cells at these stages may be more inclined toward the HCM phenotype. Through hdWGCNA and machine learning feature importance ranking, we further identified NSDHL from numerous candidate genes as the most representative core gene of the HCM phenotype. Previous studies have found that NSDHL is a key enzyme in the C4-demethylation step of the cholesterol synthesis pathway, and its functional execution depends on precise subcellular localization through Golgi apparatus transport to the endoplasmic reticulum membrane and lipid droplet surface[45]. NSDHL acts as a oncogene in multiple malignant tumors including breast cancer, ovarian cancer, and gastric cancer, and is an important risk factor for tumor development[46-48]. Our study focuses NSDHL from the previous pan-cancer level to PDAC-specific heterogeneity regulation.

To systematically validate the clinical significance and function of NSDHL in PDAC, we integrated multiple independent cohorts and spatial transcriptomics data, confirming the high expression of NSDHL in PDAC tumor tissues and its significant correlation with poor prognosis from multiple dimensions including transcription, protein, and tissue spatial localization. Its expression level is significantly enriched in the "high malignancy region" of spatial transcriptomics, highly consistent with the single-cell pseudotime trajectory, further corroborating the positive association between NSDHL and tumor malignancy. Notably, virtual knockout perturbation experiments suggest that NSDHL exerts its pro-tumor function by disrupting epithelial homeostasis and inhibiting immune activation-related pathways. Immune analysis suggests that NSDHL expression levels are positively correlated with

neoantigen load and proliferation features, suggesting that while enhancing the malignant potential of tumor cells themselves, it may also participate in shaping an immunosuppressive microenvironment. Functional experiments confirmed that NSDHL knockdown significantly inhibits the proliferation, migration, and invasion abilities of pancreatic cancer cells in vitro and delays tumor growth in vivo models. These results indicate that NSDHL acts as a "malignant metabolic switch" in PDAC and has potential as a therapeutic target.

In summary, this study through multidimensional bioinformatics and experimental validation not only reveals the heterogeneity landscape of cholesterol metabolism in PDAC but also clarifies the core node position of NSDHL as a link between high cholesterol metabolism status and tumor malignant progression. However, this study also has certain limitations. First, although we used multi-algorithm cross-validation, the classification of cholesterol metabolism status is still based on transcriptomic data, and future studies need to combine lipidomics or spatial metabolomics technologies for direct validation. Second, the downstream effector pathways regulated by NSDHL and its specific interaction mechanisms with the immune microenvironment still need to be deeply analyzed in more comprehensive in vivo models. Finally, the value of NSDHL as a biomarker in predicting treatment response, especially for cholesterol metabolism or targeted therapy, still needs to be tested in prospective clinical cohorts. Future research should focus on developing NSDHL-specific inhibitors and exploring their potential for combined application with existing chemotherapy or immunotherapy, with the aim of translating the theoretical findings of this study into clinical practice that improves the prognosis of PDAC patients.

Conclusion

In summary, this study through integrating single-cell transcriptomics, spatial transcriptomics, and bulk transcriptomic data, systematically depicts the cellular heterogeneity landscape of cholesterol metabolism in pancreatic ductal adenocarcinoma, revealing the HCM subpopulation with abnormally active cholesterol metabolism in malignant ductal epithelial cells and its multidimensional driving role in tumor evolution. Importantly, we identified and validated NSDHL as a key gene driving the HCM phenotype, which is significantly overexpressed in PDAC tissues, significantly correlated with poor patient prognosis, and specifically enriched in high-malignancy regions in spatial transcriptomics. Functional experiments confirmed that knockdown of NSDHL significantly inhibits the proliferation, migration, and invasion abilities of PDAC cells and delays tumor growth in vivo. This study elucidates the core position of cholesterol metabolism heterogeneity in PDAC malignant progression, providing experimental evidence for NSDHL as a prognostic biomarker and potential therapeutic target for clinical application. Future research should further explore the precise molecular mechanisms by which NSDHL

regulates cholesterol metabolism and downstream signaling pathways, and validate its value as a biomarker to guide individualized treatment through prospective clinical cohorts, thereby providing a more systematic biological basis for precision therapeutic strategies targeting cholesterol metabolism in pancreatic cancer.

Abbreviations

AUC: Area Under the Curve
CAF: Cancer-Associated Fibroblast
CCK-8: Cell Counting Kit-8
CNV: Copy Number Variation
DFI: Disease-Free Interval
DSS: Disease-Specific Survival
EdU: 5-Ethynyl-2'-deoxyuridine
GEO: Gene Expression Omnibus
GO: Gene Ontology
HCM: High Cholesterol Metabolism
hdWGCNA: High-Dimensional Weighted Gene Co-expression Network Analysis
HE: Hematoxylin and Eosin
ICGC: International Cancer Genome Consortium
ICM: Intermediate Cholesterol Metabolism
IHC: Immunohistochemistry
KEGG: Kyoto Encyclopedia of Genes and Genomes
LCM: Low Cholesterol Metabolism
OS: Overall Survival
PDAC: Pancreatic Ductal Adenocarcinoma
PFI: Progression-Free Interval
PROGENy: Pathway RespOnsive GENes
ROC: Receiver Operating Characteristic
scRNA-seq: Single-Cell RNA Sequencing
TCGA: The Cancer Genome Atlas
UMAP: Uniform Manifold Approximation and Projection
XGBoost: eXtreme Gradient Boosting

Acknowledgements
We sincerely thank all colleagues in our laboratory for all of the kindly advices and support.

Author contributions
Y.G. and W.Z. conceived and designed the study. Y.G. performed the bioinformatics analysis, performed the experiments, and wrote the manuscript. Z.L. and Y.M. assisted with experiments

and data analysis. B.R. and W.L. contributed to data analysis and interpretation. W.Z. supervised the study and revised the manuscript. All authors read and approved the final manuscript.

Funding

This work was supported by the National Natural Science Foundation of China (82260555); The Major Science and Technology Projects of Gansu Province (22ZD6FA021-4); The Joint Research Fund General Projects of Gansu Province (23JRRA1508).

Data availability

Publicly available datasets were analyzed in this study.

Declarations

Ethics approval and consent to participate

All animal experiments were conducted in accordance with the guidelines approved by the Experimental Animal Management and Use Committee (IACUC) of Lanzhou University Second Hospital (approval number: D2026517).

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests

References

[1] Stoop T F, Javed A A, Oba A, et al. Pancreatic cancer. Lancet, 2025, 405(10485): 1182-1202.
[2] Ma J, Bhardwaj V, Lobie P E, et al. Metabolic plasticity in pancreatic ductal adenocarcinoma progression and response to treatment. Mol Cancer, 2026, 25(1).
[3] Zhang R, Peng X, Du J X, et al. Oncogenic KRASG12D Reprograms Lipid Metabolism by Upregulating SLC25A1 to Drive Pancreatic Tumorigenesis. Cancer Res, 2023, 83(22): 3739-3752.
[4] Costamagna A, Milan G, De Santis M C, et al. Lipid signaling networks in pancreatic cancer progression and therapeutic perspectives. Trends Endocrinol Metab, 2026.
[5] Zhao L, Zheng R, Liu W, et al. Cholesterol metabolism: a new checkpoint in cancer immunity. Trends Mol Med, 2026, 32(8): 739-760.
[6] Jiang X, Zhao B, Wang T, et al. USP20-Driven Cholesterol Metabolism Links Inflammatory Signaling to Malignancy and Stromal Coevolution in Pancreatic Cancer. Cancer Res, 2026, 86(3): 712-729.
[7] Sun J, Wang J M, Zhang Q, et al. The simultaneous targeted Inhibition of ISG15 and HMGCR disrupts cancer stemness through metabolic collapse and induces synthetic lethality in pancreatic ductal adenocarcinoma. J Exp Clin Cancer Res, 2025, 44(1): 317.
[8] Qi H, Lv G, Zhang S, et al. Unraveling pancreatic ductal adenocarcinoma at single-cell resolution with spatial insights: From mechanisms to clinical translation. Cancer Lett, 2026,

645: 218391.
[9] Li Y, Du Y, Li R, et al. Spatial transcriptomics in pancreatic cancer: Advances, prospects and challenges. Crit Rev Oncol Hematol, 2024, 203: 104430.
[10] Gabitova L, Restifo D, Gorin A, et al. Endogenous Sterol Metabolites Regulate Growth of EGFR/KRAS-Dependent Tumors via LXR. Cell Rep, 2015, 12(11): 1927-1938.
[11] Lin W, Noel P, Borazanci E H, et al. Single-cell transcriptome analysis of tumor and stromal compartments of pancreatic ductal adenocarcinoma primary tumors and metastatic lesions. Genome Med, 2020, 12(1): 80.
[12] Steele N G, Carpenter E S, Kemp S B, et al. Multimodal Mapping of the Tumor and Peripheral Blood Immune Landscape in Human Pancreatic Cancer. Nat Cancer, 2020, 1(11): 1097-1112.
[13] Lyubetskaya A, Rabe B, Fisher A, et al. Assessment of spatial transcriptomics for oncology discovery. Cell Rep Methods, 2022, 2(11): 100340.
[14] Zhang G, He P, Tan H, et al. Integration of metabolomics and transcriptomics revealed a fatty acid network exerting growth inhibitory effects in human pancreatic cancer. Clin Cancer Res, 2013, 19(18): 4983-4993.
[15] Yang S, He P, Wang J, et al. A Novel MIF Signaling Pathway Drives the Malignant Character of Pancreatic Cancer by Targeting NR3C2. Cancer Res, 2016, 76(13): 3838-3850.
[16] Moffitt R A, Marayati R, Flate E L, et al. Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma. Nat Genet, 2015, 47(10): 1168-1178.
[17] Gulati G S, Sikandar S S, Wesche D J, et al. Single-cell transcriptional diversity is a hallmark of developmental potential. Science, 2020, 367(6476): 405-411.
[18] Jin S, Guerrero-Juarez C F, Zhang L, et al. Inference and analysis of cell-cell communication using CellChat. Nat Commun, 2021, 12(1): 1088.
[19] Jin S, Plikus M V, Nie Q. CellChat for systematic analysis of cell-cell communication from single-cell transcriptomics. Nat Protoc, 2025, 20(1): 180-219.
[20] Schubert M, Klinger B, Klünemann M, et al. Perturbation-response genes reveal signaling footprints in cancer gene expression. Nat Commun, 2018, 9(1): 20.
[21] Wu Y, Yang S, Ma J, et al. Spatiotemporal Immune Landscape of Colorectal Cancer Liver Metastasis at Single-Cell Level. Cancer Discov, 2022, 12(1): 134-153.
[22] Cao J, Spielmann M, Qiu X, et al. The single-cell transcriptional landscape of mammalian organogenesis. Nature, 2019, 566(7745): 496-502.
[23] Demeo B, Nesbitt C, Miller S A, et al. Active learning framework leveraging transcriptomics identifies modulators of disease phenotypes. Science, 2025, 390(6776): eadi8577.
[24] Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics, 2008, 9: 559.
[25] Morabito S, Reese F, Rahimzadeh N, et al. hdWGCNA identifies co-expression networks in high-dimensional transcriptomics data. Cell Rep Methods, 2023, 3(6): 100498.
[26] Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Association for Computing Machinery. 2016: 785–794
[27] Xun Z, Ding X, Zhang Y, et al. Reconstruction of the tumor spatial microenvironment along the malignant-boundary-nonmalignant axis. Nat Commun, 2023, 14(1): 933.

[28] Thorsson V, Gibbs D L, Brown S D, et al. The Immune Landscape of Cancer. Immunity, 2018, 48(4): 812-830.e814.
[29] Osorio D, Zhong Y, Li G, et al. scTenifoldKnk: An efficient virtual knockout tool for gene function predictions via single-cell gene regulatory network perturbation. Patterns (N Y), 2022, 3(3): 100434.
[30] Wu T, Hu E, Xu S, et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innovation (Camb), 2021, 2(3): 100141.
[31] Zhang J, Li H, Ru S, et al. Tumor heterogeneity: development, mechanisms, and therapeutic implications. Signal Transduct Target Ther, 2026, 11(1).
[32] Maddipati R. Metastatic heterogeneity in pancreatic cancer: mechanisms and opportunities for targeted intervention. J Clin Invest, 2025, 135(14).
[33] Daya T, Breytenbach A, Gu L, et al. Cholesterol metabolism in pancreatic cancer and associated therapeutic strategies. Biochim Biophys Acta Mol Cell Biol Lipids, 2025, 1870(2): 159578.
[34] Mordzińska-Rak A, Verdeil G, Hamon Y, et al. Dysregulation of cholesterol homeostasis in cancer pathogenesis. Cell Mol Life Sci, 2025, 82(1): 168.
[35] Chen W C, Boursi B, Mamtani R, et al. Total Serum Cholesterol and Pancreatic Cancer: A Nested Case-Control Study. Cancer Epidemiol Biomarkers Prev, 2019, 28(2): 363-369.
[36] Duan Z, Yang M, Yang J, et al. AGFG1 increases cholesterol biosynthesis by disrupting intracellular cholesterol homeostasis to promote PDAC progression. Cancer Lett, 2024, 598: 217130.
[37] Rademaker G, Hernandez G A, Seo Y, et al. PCSK9 drives sterol-dependent metastatic organ choice in pancreatic cancer. Nature, 2025, 643(8074): 1381-1390.
[38] Zhang H, Cao K, Xiang J, et al. Hypoxia induces immunosuppression, metastasis and drug resistance in pancreatic cancers. Cancer Lett, 2023, 571: 216345.
[39] Liu X, Feng M, Hao X, et al. m6A methylation regulates hypoxia-induced pancreatic cancer glycolytic metabolism through ALKBH5-HDAC4-HIF1α positive feedback loop. Oncogene, 2023, 42(25): 2047-2060.
[40] Godet I, Oza H H, Shi Y, et al. Hypoxia induces ROS-resistant memory upon reoxygenation in vivo promoting metastasis in part via MUC1-C. Nat Commun, 2024, 15(1): 8416.
[41] Mao H, Zhao X, Sun S C. NF-κB in inflammation and cancer. Cell Mol Immunol, 2025, 22(8): 811-839.
[42] Wu H, Fu M, Wu M, et al. Emerging mechanisms and promising approaches in pancreatic cancer metabolism. Cell Death Dis, 2024, 15(8): 553.
[43] Li Y, Li Z, Li Q, et al. Adaptation to cystine limitation stress confers a targetable lipid metabolism vulnerability in pancreatic ductal adenocarcinoma. Nat Commun, 2025, 17(1): 1343.
[44] Cheng C, Hu J, Mannan R, et al. Targeting PIKfyve-driven lipid metabolism in pancreatic cancer. Nature, 2025, 642(8068): 776-784.
[45] Caldas H, Herman G E. NSDHL, an enzyme involved in cholesterol biosynthesis, traffics through the Golgi and accumulates on ER membranes and on the surface of lipid droplets. Hum Mol Genet, 2003, 12(22): 2981-2991.
[46] Yoon S H, Lee S, Kim H S, et al. NSDHL contributes to breast cancer stem-like cell maintenance and tumor-initiating capacity through TGF-β/Smad signaling pathway in MCF-7

tumor spheroid. BMC Cancer, 2024, 24(1): 1370.
[47] Shanshan X, Jinshi R, Wanlin C, et al. NSDHL as a regulator of lipid peroxidation and ER stress-mediated apoptosis in ovarian cancer. Biol Direct, 2026, 21(1).
[48] Kang W, Yang Y, Chen C, et al. CPSF1 positively regulates NSDHL by alternative polyadenylation and promotes gastric cancer progression. Am J Cancer Res, 2022, 12(10): 4566-4583.

Supplementary Files

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

- renamede5df5.docx