Human Mutation

WILEY

SPECIAL ISSUE Identification of Immuno-Diagnostic and Therapeutic Markers in Tumors and Inflammatory Diseases Based on Single-Cell and Machine Learning Methods

RESEARCH ARTICLE OPEN ACCESS

Single-Cell RNA Sequencing–Integrated Fibroblast Markers and Candidate Drug–Target Associations in Pancreatic Ductal Adenocarcinoma

Tengfei Ji | Hongsheng Wu | Zongmin Zheng | Keqiang Ma

Department of Hepatobiliary Surgery, The GuangZhou Affiliated Hospital of Jinan University (Huadu District People's Hospital of Guangzhou, Huadu Institute of Medicine, Guangzhou), Guangzhou, China

Correspondence: Keqiang Ma (mkq9928@hotmail.com)

Received: 17 April 2026 | **Revised:** 14 August 2026 | **Accepted:** 14 September 2026

Guest Editor: Peipei Gong

Keywords: cancer-associated fibroblasts | pancreatic ductal adenocarcinoma | single-cell RNA sequencing | survival-associated genes | weighted gene co-expression network analysis

ABSTRACT

Background: Pancreatic ductal adenocarcinoma (PDAC) has dismal outcomes and is shaped by a fibroblast-rich tumor microenvironment (TME). A cell type–resolved, module-level understanding of cancer-associated fibroblasts (CAFs) and their therapeutic relevance remains incompletely characterized.

Methods: We integrated single-cell RNA-seq data from six treatment-naïve PDAC resection specimens in GSE212966; individual tumor-stage information was not available in the public metadata. Standardized preprocessing, clustering, and major-cell-type annotation were performed. Contact-mediated intercellular signaling toward fibroblasts was profiled. High-dimensional weighted gene co-expression network analysis (hdWGCNA) was applied to fibroblasts to identify co-expression modules and hub genes. Associations with overall survival were assessed by univariate Cox models and Kaplan–Meier analysis; immune infiltration was evaluated with CIBERSORT and ESTIMATE; pathway associations were examined by gene set enrichment analysis (GSEA). Candidate compounds were screened via Enrichr/DSigDB and docked to target proteins using a PyMOL–Chem3D–AutoDockTools workflow. SEMA3C was further examined in PDAC cell lines by siRNA-mediated knockdown followed by qRT–PCR, CCK-8, wound-healing, and Transwell invasion assays.

Results: We resolved 10 cell subpopulations and observed broad contact signals to fibroblasts, including the CD99–CD99 axis. hdWGCNA delineated fibroblast modules, from which four key modules were prioritized. Five genes—SEMA3C, PMP22, CPE, SNX9, and UACA—were significantly associated with overall survival. These biomarkers showed both positive and negative associations with specific immune-cell fractions but were positively correlated with the StromalScore, ImmuneScore, and ESTIMATEScore. High-expression groups were enriched for cancer-relevant programs such as epithelial–mesenchymal transition, inflammatory response, and TNFα–NF-κB signaling, whereas several proliferative and metabolic pathways were enriched in the CPE-low group. In vitro, siRNA-mediated SEMA3C knockdown in PANC-1 and Capan-2 cells reduced cell viability, wound closure, and invasion, supporting a cell-intrinsic protumorigenic role of SEMA3C in PDAC cells. Coumestrol emerged as a computationally prioritized compound, and molecular docking predicted a potentially favorable interaction with SNX9 (predicted binding energy −7.5 kcal/mol); this association requires experimental validation.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Copyright © 2026 Tengfei Ji et al. *Human Mutation* published by John Wiley & Sons Ltd.

Conclusion: An integrative single-cell and network framework identified fibroblast-related co-expression modules and five survival-associated genes in PDAC. The prognostic relevance of these genes and the predicted coumestrol–SNX9 association require further independent and experimental validation.

1 | Introduction

More than 90% of pancreatic cancers are pancreatic ductal adenocarcinoma (PDAC), which is among the most lethal solid tumors [1–3]. The 5-year survival rate remains in the single digits, and most patients present with unresectable or metastatic disease at diagnosis [4, 5]. Despite advances in perioperative management, chemotherapy regimens such as gemcitabine- or FOLFIRINOX-based combinations, and a limited number of targeted and immunotherapeutic strategies, population-level survival gains have been modest, with pervasive drug resistance and frequent relapse [6–8]. Beyond the intrinsic genomic instability and polyclonal evolution of malignant epithelial cells, the tumor microenvironment (TME)—characterized by profound immune suppression, a dense stromal barrier, and aberrant vasculature—has emerged as a major constraint on therapeutic efficacy [9–11]. Traditional tumor-cell–centric paradigms therefore fail to account for the pronounced interpatient variability in treatment response and the rapid development of acquired resistance [12]. Reframing PDAC through the lens of tumor–microenvironment co-evolution has become a central avenue for understanding disease biology and identifying new therapeutic entry points [13, 14]. In this context, systematically delineating the cellular composition, interactions, and molecular programs of the TME—and linking these features to clinical outcomes and treatment response—holds both scientific and translational value [15].

More than 90% of pancreatic cancers are pancreatic ductal adenocarcinoma (PDAC), which is among the most lethal solid tumors [1–3]. The 5-year survival rate remains in the single digits, and most patients present with unresectable or metastatic disease at diagnosis [4, 5]. Despite advances in perioperative management, chemotherapy regimens such as gemcitabine- or FOLFIRINOX-based combinations, and a limited number of targeted and immunotherapeutic strategies, population-level survival gains have been modest, with pervasive drug resistance and frequent relapse [6–8]. Beyond the intrinsic genomic instability and polyclonal evolution of malignant epithelial cells, the tumor microenvironment (TME)—characterized by profound immune suppression, a dense stromal barrier, and aberrant vasculature—has emerged as a major constraint on therapeutic efficacy [9–11]. Traditional tumor-cell–centric paradigms therefore fail to account for the pronounced interpatient variability in treatment response and the rapid development of acquired resistance [12]. Reframing PDAC through the lens of tumor–microenvironment co-evolution has become a central avenue for understanding disease biology and identifying new therapeutic entry points [13, 14]. In this context, systematically delineating the cellular composition, interactions, and molecular programs of the TME—and linking these features to clinical outcomes and treatment response—holds both scientific and translational value [15].

Within the PDAC TME, cancer-associated fibroblasts (CAFs) are abundant and functionally pivotal in shaping tumor ecology and determining therapeutic response [16–18]. By secreting and remodeling extracellular matrix (ECM), CAFs drive desmoplasia and tissue stiffening, alter interstitial pressure and fluid flow to create a drug-delivery barrier, and, through cytokines, chemokines, and metabolites, reprogram angiogenesis and immune infiltration to foster immune suppression and escape [19, 20]. CAFs also act as "signal amplifiers" engaging in extensive ligand–receptor interactions with malignant, endothelial, and immune cells to coordinate tumor growth, invasion, and metastasis [21]. Importantly, CAFs are not a homogeneous population but display marked molecular and functional heterogeneity: Single-cell and spatial omics studies have delineated myofibroblastic (myCAF), inflammatory (iCAF), and antigen-presenting (apCAF) phenotypes that differ in spatial localization, secretory programs, immune regulation, and metabolic crosstalk [22, 23]. Multiple studies link CAF heterogeneity to patient survival, chemotherapy sensitivity, and responsiveness to immunotherapy [24, 25]; however, discrepancies persist across studies regarding subpopulation definitions, marker sets, and functional interpretations [26]. However, most previous studies have focused on predefined CAF subtypes or individual markers, whereas cell-type-resolved co-expression modules linking fibroblast programs to prognosis and immune context in PDAC remain insufficiently characterized. Outside PDAC, CD99 has been detected in CAFs, and homophilic CD99 engagement has been shown to regulate β1-integrin/FAK-dependent cell adhesion [27, 28]. Nevertheless, whether CD99–CD99-mediated contact influences CAF activation in PDAC remains unknown.

To address these gaps, we integrated a publicly available PDAC single-cell dataset and applied high-dimensional weighted gene co-expression network analysis (hdWGCNA) to identify fibroblast-related co-expression modules at cellular resolution. We then prioritized module-associated genes and assessed their relationships with survival, immune infiltration, and functional pathways in the TCGA-PDAC cohort. In parallel, we combined drug-perturbation resources with molecular docking to generate hypothesis-level candidate compound–target associations. We hypothesized that fibroblast-specific co-expression analysis could identify clinically relevant stromal programs and generate experimentally testable hypotheses. Our contributions include a CAF-centered, module-level analytical framework for PDAC, five fibroblast-related survival-associated genes, and preliminary computational drug–target predictions requiring further validation.

2 | Materials and Methods

2.1 | Data Sources

Single-cell RNA-seq [29]. Public dataset GSE212966 was retrieved from GEO. The dataset contains six PDAC and six adjacent noncancerous resection specimens; only the six PDAC samples were included in the present analysis. All six PDAC

specimens were obtained before treatment, whereas individual tumor-stage information was not available in the public metadata. Sample-level batch effects were addressed during data integration. All data were publicly available and de-identified; no additional ethical approval was required.

TCGA-PDAC cohort (TCGA-PAAD). RNA-seq expression data [30] and matched clinical information were obtained from the TCGA-PAAD project through the GDC portal, extracting raw counts and FPKM values. We retained Primary Solid Tumor (01A) samples with overall survival (OS) $\geq$ 30 days, yielding 129 tumor cases for expression and prognostic analyses.

2.2 | Preprocessing, Dimensionality Reduction, and Clustering of Single-Cell RNA-seq Data

Seurat was used to import single-cell RNA-seq count matrices and performed quality control by discarding cells with fewer than 200 or more than 6000 detected genes, retaining 49,191 cells for downstream analyses [31]. Counts were normalized and variance-stabilized using SCTransform, followed by principal component analysis (PCA). To mitigate intersample batch effects, we applied Harmony to the PCA embeddings [32]. A k-nearest neighbor (KNN) graph was then constructed from Euclidean distances in the top 50 principal components using FindNeighbors, and unsupervised clustering was performed with FindClusters (Louvain community detection; resolution = 0.05). Two-dimensional embeddings were generated with UMAP for visualization. Cluster identities were assigned based on the expression of canonical marker genes curated from the CellMarker database, with manual curation using average-expression profiles and marker distributions.

2.3 | Differential Expression and Marker Identification Across Subpopulations

Differential expression was computed in a "one cluster versus all others" design using Seurat's FindAllMarkers (default Wilcoxon rank-sum test). Thresholds were set to logfc.threshold = 0.25 and min.pct = 0.30, with only.pos = TRUE to retain upregulated markers. Multiple testing was controlled using Seurat's default Bonferroni correction (p_val_adj), and genes with p_val_adj $<$ 0.05 were considered significant. For downstream analyses, markers were ranked by adjusted p value, average log fold change, and the proportion of expressing cells within each subcluster.

2.4 | Functional Annotation of Cell Subpopulations

For each subpopulation, the upregulated (highly expressed) marker genes identified by differential analysis were subjected to functional annotation/enrichment using clusterProfiler to summarize the principal biological features and putative functions of each cluster [33].

2.5 | Cell–Cell Communication Analysis

We inferred intercellular communication using CellChat and its curated ligand–receptor knowledgebase [34]. Treating fibroblasts as the receiver population, we extracted ligand–receptor interactions from other subpopulations to fibroblasts and restricted the analysis to the "Cell–Cell Contact" category. Interaction strength (communication probability) and statistical significance were estimated under CellChat's default framework, with significance defined as adjusted $p < 0.05$. Bubble plots were used to visualize the probability and significance of intersubpopulation interactions.

2.6 | High-Dimensional WGCNA for Fibroblast-Related Co-Expression Modules and Hub Genes

We imported the single cell.rds object with hdWGCNA and subsetted the fibroblast population based on prior annotations [35]. A weighted co-expression network was constructed on this subset. The soft-thresholding power was selected at $\beta = 8$ using the scale-free topology criterion, after which a signed adjacency matrix and topological overlap matrix (TOM) were computed. Hierarchical clustering with dynamic tree cutting was applied to detect co-expression modules, yielding nine modules (M1–M9). For each module, we computed the module eigengene (ME) and intramodular connectivity/module membership and designated hub genes based on these connectivity metrics for downstream analyses.

2.7 | Survival Association Analysis

Candidate genes were taken from the fibroblast-related set identified by hdWGCNA. Tumor samples with complete survival information and a follow-up duration (days) of $\geq$ 30 were included. OS was used as the endpoint (time unit: days; event = death; survivors were censored). For each gene, univariate Cox proportional hazards models were fitted with survival::coxph, treating normalized gene expression as a continuous covariate. We report the hazard ratio (HR), 95% confidence interval, and Wald test p value, and define $p < 0.05$ as statistically significant to designate survival-associated genes. Kaplan–Meier curves were generated using survfit after dichotomizing patients into high and low expression groups using the optimal cutpoint determined with survminer::surv_cutpoint (minimum proportion per group = 0.20), and differences were assessed using the log-rank test. The candidate genes were evaluated individually and were not combined into a multigene risk score or prognostic signature. The analyses were intended to identify fibroblast-related genes associated with survival rather than to establish independent prognostic factors.

2.8 | Immune Infiltration Assessment and Correlation Analyses

Immune-cell composition was inferred using the CIBERSORT R implementation with the LM22 signature [36]. After obtaining the relative fractions of the 22 immune cell types for each sample, we evaluated associations between study biomarkers (gene expression or signature scores) and inferred infiltration levels using Spearman rank correlation, reporting the correlation coefficient (ρ) and p value; Benjamini–Hochberg adjustment was applied where multiple comparisons were made.

In addition, the estimate package was applied to the TCGA-PDAC cohort to compute StromalScore, ImmuneScore, and ESTIMATEScore [37]. Associations between these scores and the biomarkers were assessed with Spearman correlation, with multiple-testing correction applied as appropriate.

2.9 | Gene Set Enrichment Analysis (GSEA)

Samples were dichotomized into high- and low-expression groups based on the median expression of the target gene, and enrichment testing was performed with GSEA v4.4.0 under a binary phenotype design. The HALLMARK collection (MSigDB) was used as the gene set resource. Both positive and negative enrichment scores were examined to identify pathways enriched in the high- and low-expression groups, respectively. An FDR q value < 0.05 was considered statistically significant.

2.10 | CELL Culture and siRNA Transfection

The human pancreatic duct epithelial cell line HPDE6-C7 (SCC442, Sigma-Aldrich, Darmstadt, Germany) and the PDAC cell lines PANC-1 (CRL-1469, ATCC, Manassas, VA, United States) and Capan-2 (HTB-80, ATCC) were used in this study. Cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum and 1% penicillin–streptomycin at 37°C in a humidified atmosphere with 5% CO_2, maintained as adherent monolayers and passaged with trypsin–EDTA when reaching 80%–90% confluence. Experiments were performed with cells in the logarithmic growth phase.

SEMA3C expression was silenced using small interfering RNAs targeting *SEMA3C* (si-SEMA3C) and a non-targeting control siRNA (si-NC). The sequences were si-NC, 5′-UUCUCCGAACGUGUCACGUTT-3′; si-SEMA3C#1, 5′-GCGCUACUAAUUGGGAAGATT-3′; and si-SEMA3C#2, 5′-GGGCUGAGGACCUUGCAGAAGATT-3′. The siRNA oligonucleotides, including the nontargeting control, were synthesized by Shanghai GeneChem Co. Ltd. (Shanghai, China). PANC-1 and Capan-2 cells were seeded 24 h before transfection to reach 50%–60% confluence and transfected with 50-nM siRNA using Lipofectamine 3000 (Thermo Fisher Scientific, United States) following the manufacturer's instructions. Cells were collected at the indicated time points for RNA extraction and functional assays.

2.11 | Quantitative Real-Time PCR (qRT–PCR)

Total RNA was extracted from the cultured cells using TRIzol reagent (15596026CN, Invitrogen, Carlsbad, CA, United States). The amount and purity of the RNA were then assessed using a NanoDrop spectrophotometer (Thermo Fisher Scientific, United States). Next, first-strand cDNA was created from the total RNA using a reverse transcription kit (D7178S, Beyotime, Shanghai, China). Finally, quantitative PCR was performed using SYBR Green qPCR Mix (D7260, Beyotime, China) on an Applied Biosystems real-time PCR platform (Applied Biosystems, Foster City, CA, United States).

The amplification process began with a 10-min denaturation step at 95°C, succeeded by 30 cycles of 15 s at 95°C and 30 s at 60°C, culminating in a 5-min extension at 72°C. Relative mRNA levels were determined employing the $2^{-\Delta\Delta Ct}$ method, utilizing GAPDH as the reference gene. Primer sequences are presented in Table 1. Each sample underwent analysis in technical triplicate, and no-template controls were incorporated into each experimental run.

2.12 | Cell Viability Assay

Cell viability was examined by plating cells in 96-well plates at a density of 2×10^3 cells per well. These cells were further cultured in DMEM medium containing 10% FBS and 1% penicillin–streptomycin. Cells were incubated at 37°C with 5% CO_2, and cell viability was measured at 0, 24, 48, and 72 h after transfection. Viability was evaluated by Cell Counting Kit-8 (CCK-8; C0037, Beyotime). In each well, 10 μL of CCK-8 solution was added, and it was incubated for 4 h. Absorbance at 450 nm was then read with a microplate reader (Thermo Fisher Scientific). Relative viability was determined using the following equation: $(OD_{450}(\text{sample}) - OD_{450}(\text{blank}))/(OD_{450}(\text{si-NC}) - OD_{450}(\text{blank})) \times 100\%$, normalized to si-NC control. Three replicate measurements were analyzed for each experimental condition ($n = 3$).

2.13 | Wound-Healing Assay

Cells were plated in six-well plates at approximately 5×10^5 cells per well in 2 mL of complete medium and incubated at 37°C with 5% CO_2 until approximately 90% confluence

TABLE 1 | The primers utilized in this study.

Gene	Accession no.	Primers (5′-3′)	
		Forward	Reverse
SEMA3C	NM_006379	ACCCACTGACTCAATGCAGAGG	CAGCCACTTGATAGATGCCTGC
PMP22	NM_000304	GGCAATGGACACGCAACTGATC	TGATCGACAGGATCATGGTGGC
CPE	NM_001873	TTGCTCCTGAGACCAAGGCTGT	GAGGAGCTGTATTCGTGAGCAC
SNX9	NM_016224	GGTCAGCCATTCCAATCCCTTC	ATGGCGACACATCCTGGTCATC
UACA	NM_018003	GGACACCACTTGTTCTGGCTAC	GCAACCTAGCATGAGGGCAGTT
GAPDH	NM_002046	GTCTCCTCTGACTTCAACAGCG	ACCACCCTGTTGCTGTAGCCAA

was reached. A straight scratch was then introduced using a sterile 200-μL pipette tip, and cell debris was removed by rinsing the wells twice with PBS. Then, serum-free DMEM was added, and phase-contrast images were acquired at 0 and 48 h using an inverted microscope. The wound-healing rate was ascertained by calculating the percentage reduction in mean wound width at 48 h compared to the initial measurement. Each experimental condition was tested in triplicate before repeating the assay independently at least three times [38].

2.14 | Transwell Invasion Assay

Cell invasion was evaluated using Transwell inserts with 8-μm pores. The upper surface of each insert was coated with Matrigel and incubated at 37°C for 30 min. A total of 3×10^4 transfected cells suspended in serum-free medium were added to the upper chamber, whereas medium containing serum was added to the lower chamber as a chemoattractant. After 48 h, noninvading cells remaining on the upper surface were gently removed. Cells that had invaded through the membrane were fixed with 4% paraformaldehyde for 15 min and stained with crystal violet for 10 min. Invaded cells were counted in at least five randomly selected microscopic fields per insert. Three replicate measurements were analyzed per condition.

2.15 | Targeted Drug Prediction and Molecular Docking

The fibroblast-related candidate genes were queried against DSigDB using the enrichR R package [39]. Hits with adjusted $p < 0.05$ were retained, entries explicitly restricted to specific cell lines were excluded, and nondrug chemical entities were removed by manual literature curation. Candidate small molecules were prioritized in ascending order of p value, and 3D ligand structures were downloaded from PubChem in SDF/3D format. The three-dimensional structure of coumestrol was obtained from PubChem (CID: 5281707).

For receptor structures, target proteins were identified on UniProt, and cross-referenced crystallographic entries were retrieved, preferentially selecting X-ray structures with lower resolution values and greater sequence coverage. Molecular docking was used as a representative structural assessment rather than as a systematic druggability ranking of all candidate genes. Representative compound–target pairs were selected according to the availability and completeness of experimentally resolved receptor structures and the availability of compatible candidate ligands.

Structure preparation and docking proceeded as follows: crystallographic waters and cocrystallized small ligands were removed, and hydrogens were added in PyMOL; ligands were energy-minimized in ChemBioOffice (Chem3D). Receptors and ligands were then prepared and docked using the AutoDockTools suite. For SNX9–coumestrol docking, the grid box was centered at $x = 11.115$, $y = 43.771$, and $z = 48.921$, with $126 \times 126 \times 126$ grid points and a spacing of 1.0 Å. Docking poses were retained if the predicted binding energy was ≤ -5.0 kcal/mol and if at least one key ligand–receptor contact (e.g., hydrogen bond or hydrophobic interaction) lay within 3.5 Å. Resulting complexes were visually inspected and measured in PyMOL.

2.16 | Statistical Analysis

Bioinformatic analyses were performed using R Version 4.3.0, and the Benjamini–Hochberg procedure was applied for multiple testing where appropriate, with an FDR q value < 0.05 considered statistically significant. In vitro data were analyzed using GraphPad Prism Version 10.6.1 and are presented as mean ± SD or SEM, as indicated in the figure legends ($n = 3$). Differences were assessed using two-tailed unpaired Welch's t-tests or one- or two-way ANOVA followed by appropriate multiple-comparisons tests. $p < 0.05$ was considered statistically significant.

3 | Results

3.1 | Single-Cell Atlas of PDAC

After quality control, normalization, batch correction, and dimensionality reduction/clustering, we resolved 10 major cell subpopulations in the UMAP embedding (Figure 1A–B). Using marker information curated from CellMarker 2.0 and pertinent literature, clusters were annotated as follows: T cells (*IL7R*, *CCL5*, *CD2*, *CD3D*); fibroblasts (*DCN*, *LUM*, *COL1A1*, *COL1A2*); macrophages (*C1QA*, *C1QB*, *HLA-DRA*); epithelial cells (*KRT19*, *MUC1*, *KRT8*); B cells (*IGHA1*, *CD79A*, *MS4A1*); smooth muscle cells (*MYH11*, *RGS5*, *ACTA2*, *MYL9*); endothelial cells (*VWF*, *EPAS1*, *COL4A1*); mast cells (*CPA3*, *CTSG*, *KIT*); Schwann cells (*S100B*, *CDH19*, *PLP1*); and MKI67$^+$ progenitor-like cells (*TOP2A*, *MKI67*). Marker expression patterns across clusters are summarized in a bubble plot (Figure 1C). Functional annotation of fibroblast-specific upregulated genes revealed significant enrichment of ECM and matrix-remodeling pathways (Figure 1D).

3.2 | Cell–Cell Communication Analysis

In multicellular systems, cellular functions depend on intercellular interactions, often mediated by multisubunit protein complexes. Our communication analysis revealed a widespread network of cross-cluster signaling within the PDAC microenvironment. After extracting ligand–receptor pairs for each cluster and restricting to the Cell–Cell Contact category, predicted CD99–CD99 interactions were observed from T cells, macrophages, and epithelial cells to fibroblasts (Figure 2).

3.3 | hdWGCNA Identifies Fibroblast-Related Modules

The annotated single-cell.rds file was imported into hdWGCNA, retaining genes expressed in ≥ 5% of cells to initialize

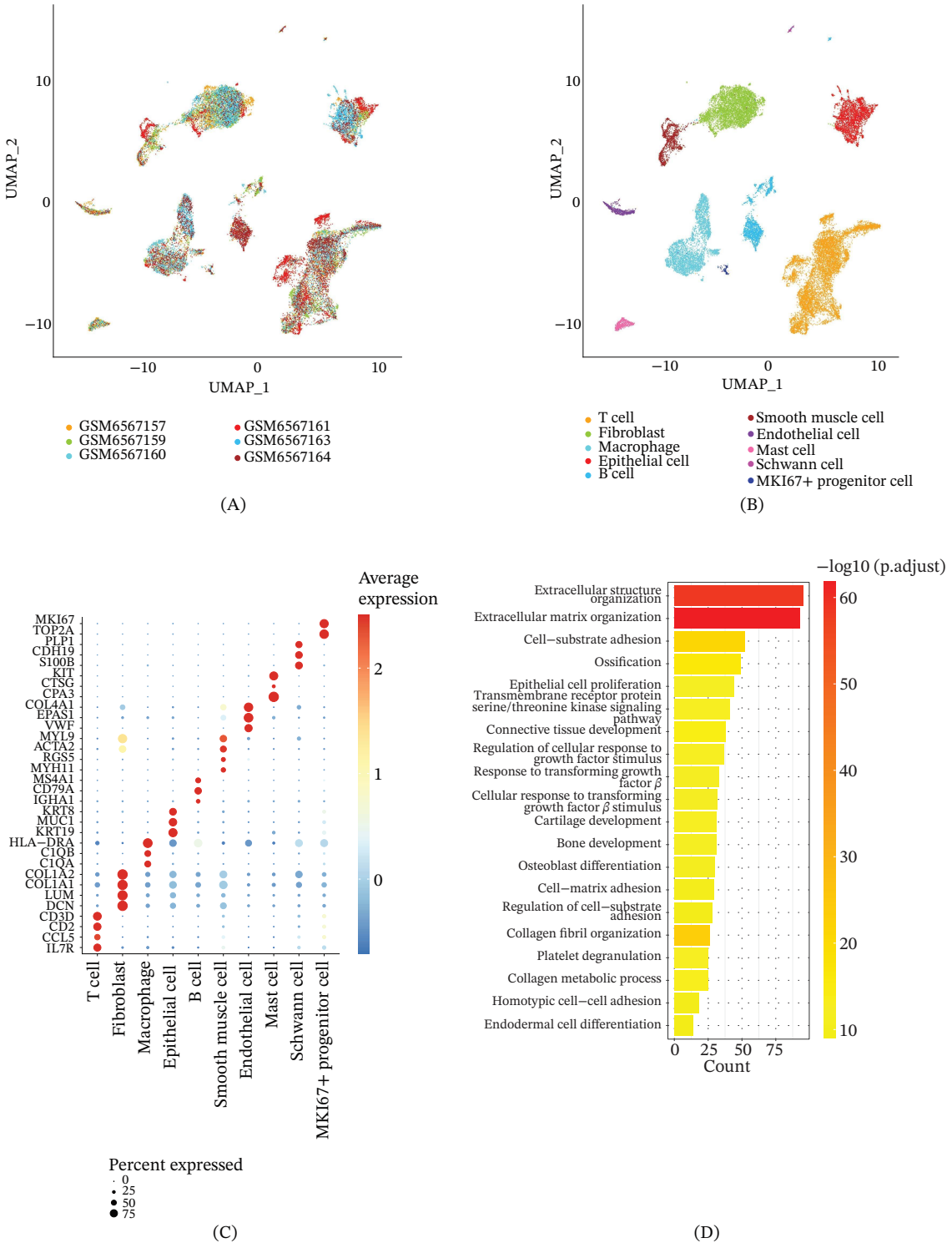


FIGURE 1 | Single-cell atlas of PDAC. (A) UMAP embedding colored by sample after batch correction. (B) UMAP embedding colored by cell type (10 subpopulations). (C) Bubble plot of canonical marker expression across clusters. (D) Functional annotation of fibroblast-enriched genes.

the analysis object. To mitigate sparsity, metacells were generated using a KNN approach ($k = 25$) with the maximum number of shared cells per metacell pair set to 10. Weighted co-expression analysis was then performed within the fibroblast subset: The soft-thresholding power was automatically selected at $\beta = 8$ based on the scale-free topology criterion (Figure 3A), from which a signed network and topological overlap were constructed (Figure 3B). Using MEs and intramodular connectivity, we prioritized hub genes and delineated nine robust modules (Figure 3C).

We quantified, for each module, the proportion of expressing cells and the average expression across cell subpopulations and visualized intermodule correlations (Figure 4A). M1, M3, M4, and M8 exhibited higher expression prevalence and levels in fibroblasts (Figure 4B) and were therefore designated as key

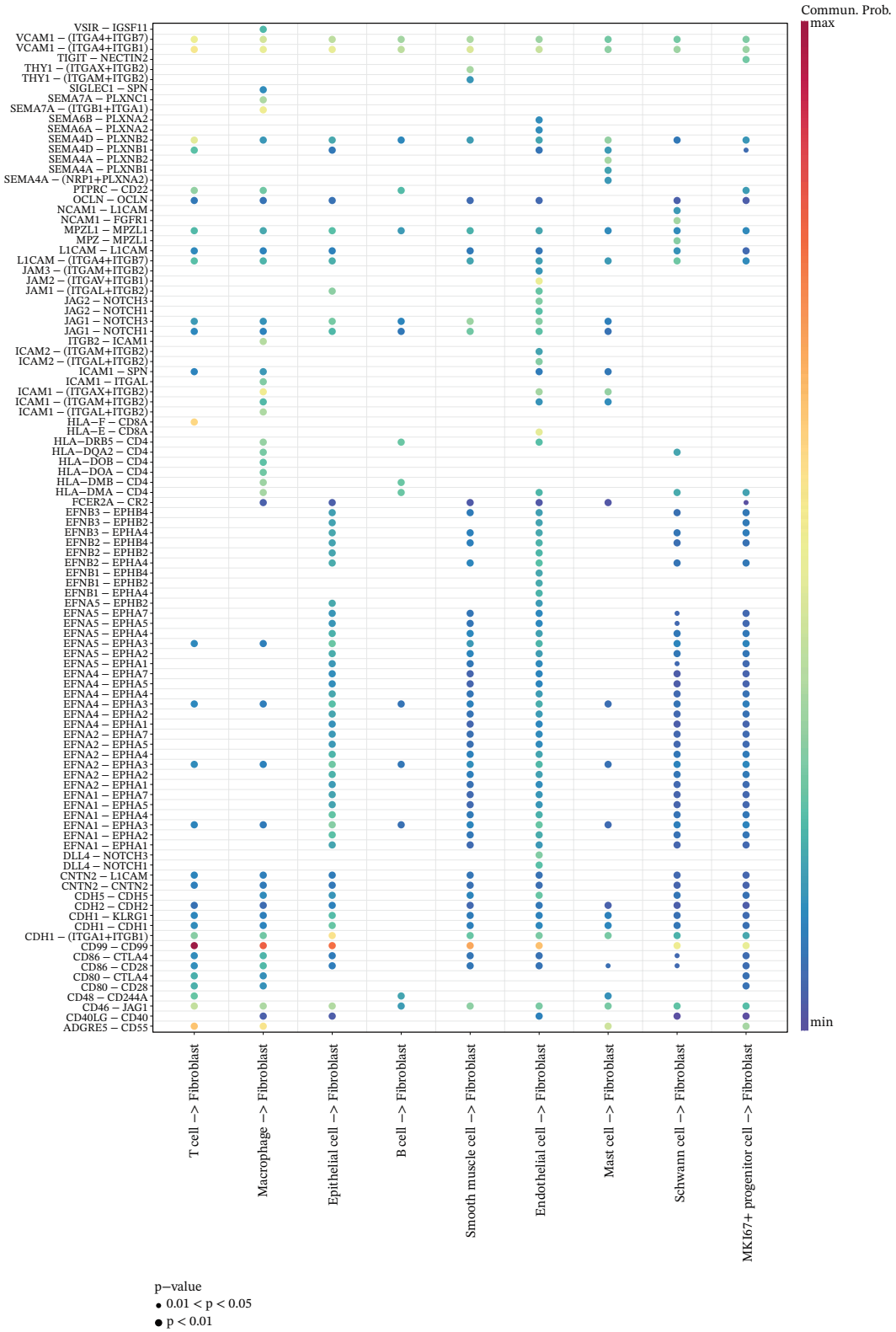


FIGURE 2 | Ligand–receptor interactions targeting fibroblasts. Bubble plot of ligand–receptor interactions from all cell subpopulations to fibroblasts.

modules. Hub-gene interaction networks were then extracted for these modules (Figure 4C–F): The inner ring (10 genes) denotes the primary hubs, whereas the outer ring (15 genes) represents secondary hubs, together delineating the core regulatory architecture within each module.

3.4 | Survival Association Analysis

To assess the survival associations of fibroblast-related genes derived from hdWGCNA, we fitted univariate Cox models in the TCGA-PDAC cohort. Five genes—*SEMA3C*, *PMP22*, *CPE, SNX9*,

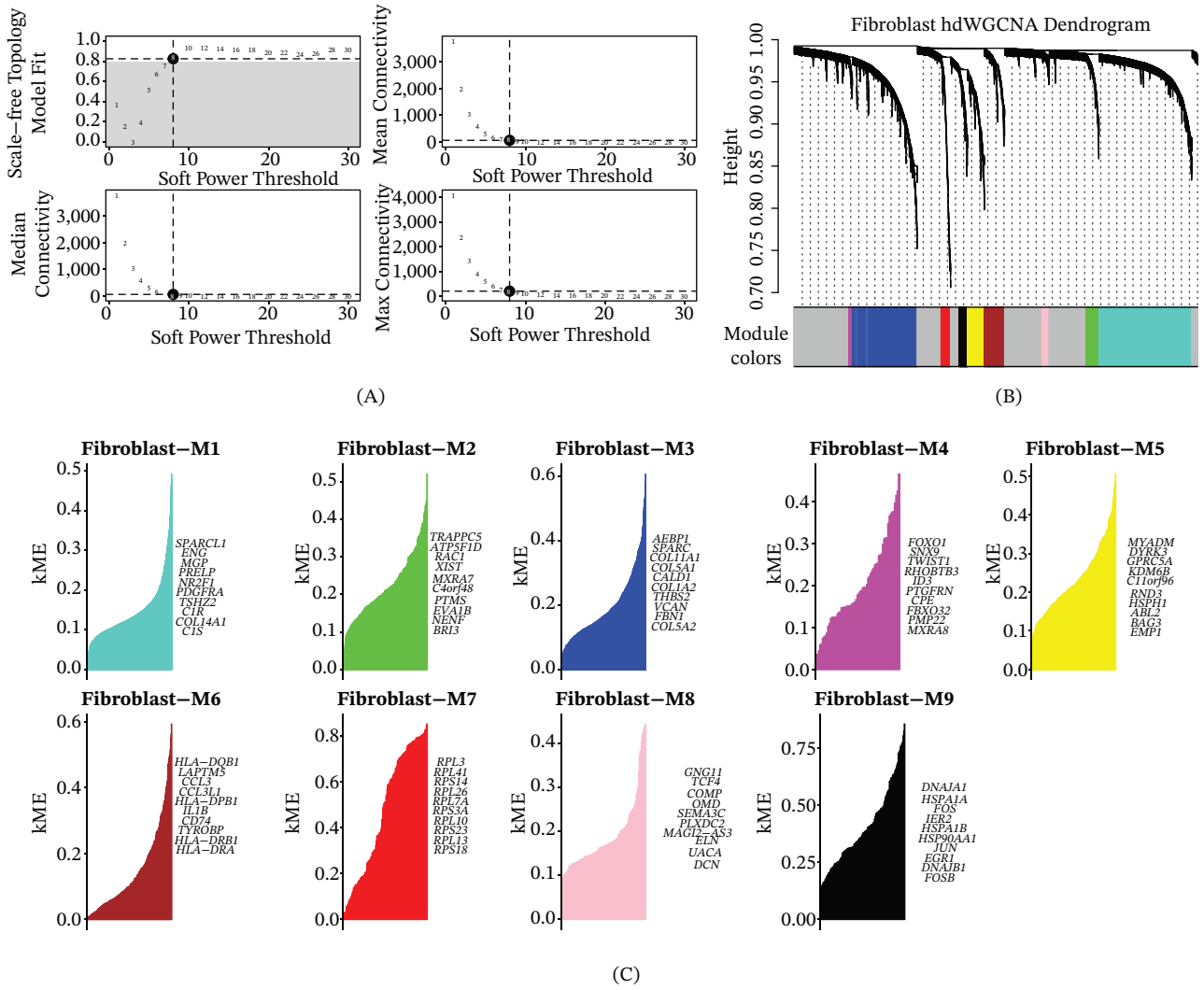


FIGURE 3 | hdWGCNA-based fibroblast module identification. (A) Scale-free topology fit and connectivity across candidate soft-thresholding powers. $\beta = 8$ was selected as the smallest power reaching a scale-free topology fit of approximately $R^2 = 0.80$. (B) Co-expression network dendrogram: upper panel, gene clustering tree; lower panel, module assignments. (C) Module delineation and hub genes: *y*-axis shows kME (module membership) indicating connectivity to the module eigengene; right panel lists module hub genes.

and *UACA*—were significantly associated with OS (Figure 5A). Patients were then dichotomized into high- and low-expression groups using survminer with optimal cutpoints (function surv_cutpoint; minimum proportion per group = 0.20) and analyzed by Kaplan–Meier methods. All five genes showed significant survival stratification (log-rank $p < 0.05$; Figure 5B–F). These genes were evaluated as individual survival-associated candidates and were not combined into a prognostic signature or risk-score model.

3.5 | Correlations Between Biomarkers and Immune Infiltration

Using CIBERSORT–inferred proportions of 22 immune cell types, we computed Spearman correlations between the five biomarkers and immune infiltration levels and summarized the results in a correlation plot (Figure 6A). Multiple associations were statistically significant (adjusted $p < 0.05$), indicating associations with both innate and adaptive immune compartments. Consistently, ESTIMATE analysis revealed significant positive correlations of all five biomarkers with the StromalScore, ImmuneScore, and ESTIMATEScore (Figure 6B; Spearman $\rho > 0$, FDR < 0.05), supporting their association with a more stroma- and immune-enriched microenvironment.

3.6 | GSEA

To investigate the functional roles of the five biomarkers in tumors, samples were dichotomized at the median expression of each biomarker and analyzed using GSEA v4.4.0 under a binary phenotype design (FDR < 0.05). High-expression groups showed significant enrichment of cancer-related HALLMARK pathways, including epithelial–mesenchymal transition (EMT), inflammatory response, and TNFα signaling via NF-κB, among others (Figure 7A–E). Conversely, the CPE-low group showed enrichment of E2F targets, MYC targets V1/V2, and oxidative phosphorylation, demonstrating that the analysis captured pathways associated with both high- and low-expression phenotypes (Figure 7A).

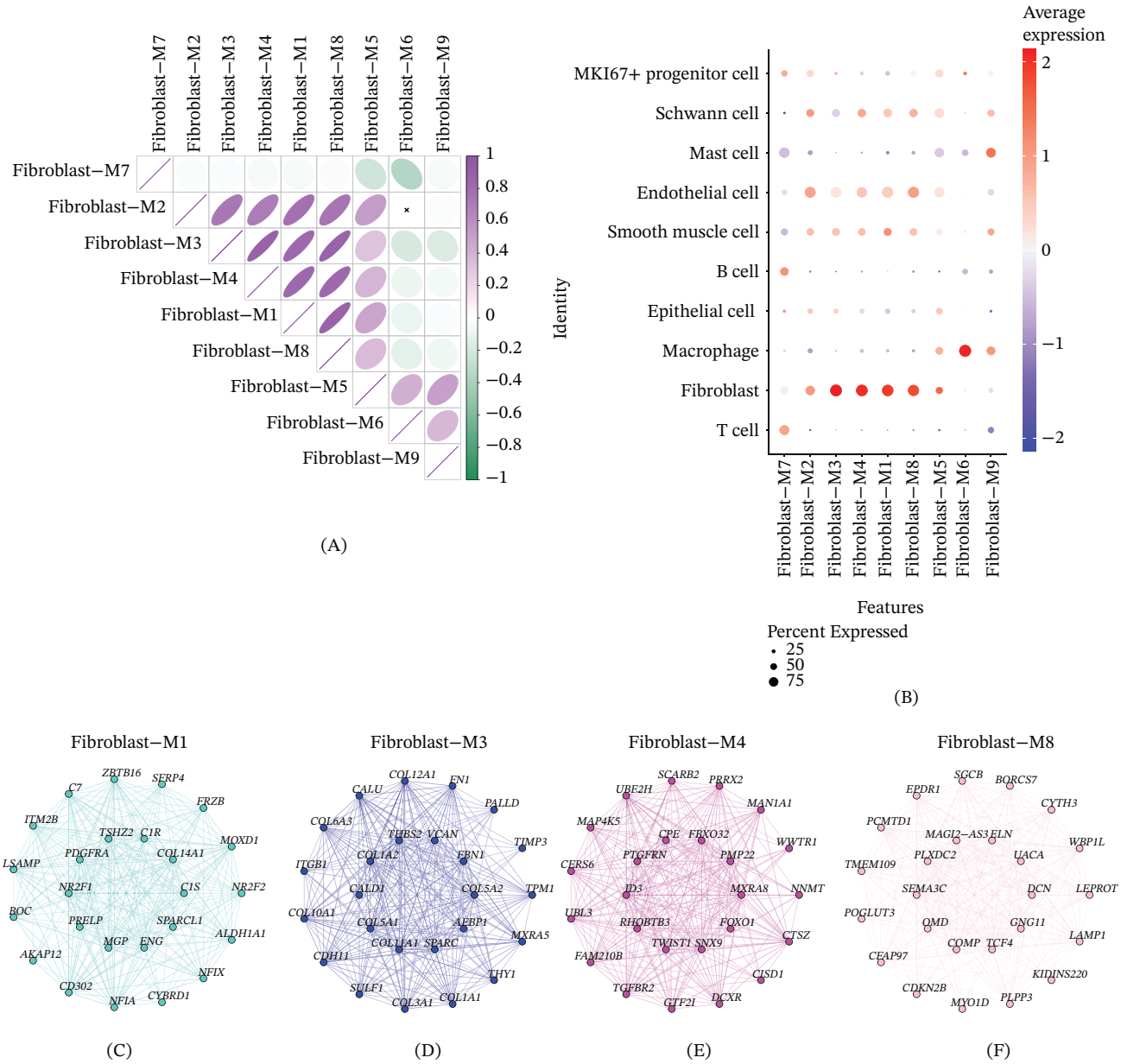


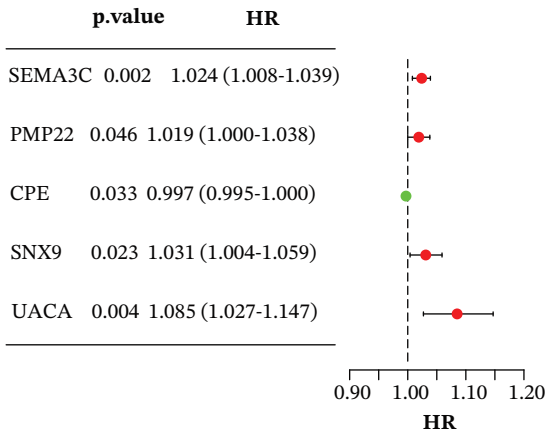
FIGURE 4 | Key modules and hub-gene networks. (A) Intermodule correlation matrix. (B) Expression of module genes across cell types; red = high, blue = low; dot size indicates the proportion of expressing cells. (C–F) Co-expression networks of module hub genes; inner ring, primary hubs; outer ring, secondary hubs.

3.7 | Experimental Validation of SEMA3C in PDAC Cell Lines

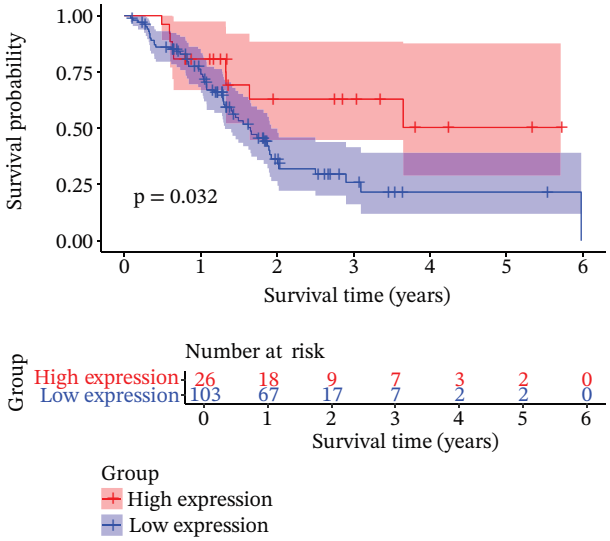
The basal expression of the five fibroblast-related candidate genes was first examined in a panel of pancreatic cell lines. qRT–PCR showed that SEMA3C, PMP22, CPE, SNX9, and UACA were expressed at higher levels in the PDAC cell lines PANC-1 and Capan-2 than in the nonmalignant pancreatic ductal epithelial cell line HPDE6-C7, with SEMA3C showing one of the largest increases (Figure 8A). Two independent SEMA3C-targeting siRNAs significantly reduced SEMA3C mRNA expression in PANC-1 and Capan-2 cells compared with si-NC. Because si-SEMA3C#1 produced the greater knockdown efficiency, it was selected for subsequent functional assays (Figure 8B,C). SEMA3C knockdown significantly reduced cell viability at 48 and 72 h (Figure 8D,E), decreased wound closure at 48 h (Figure 8F,H), and reduced the number of invading cells in the Matrigel-coated Transwell assay (Figure 8G,I). These findings support a cell-intrinsic role of SEMA3C in promoting PDAC-cell proliferation, migration, and invasion; however, they do not establish a CAF-specific function of SEMA3C.

3.8 | Drug Prioritization and Molecular Docking

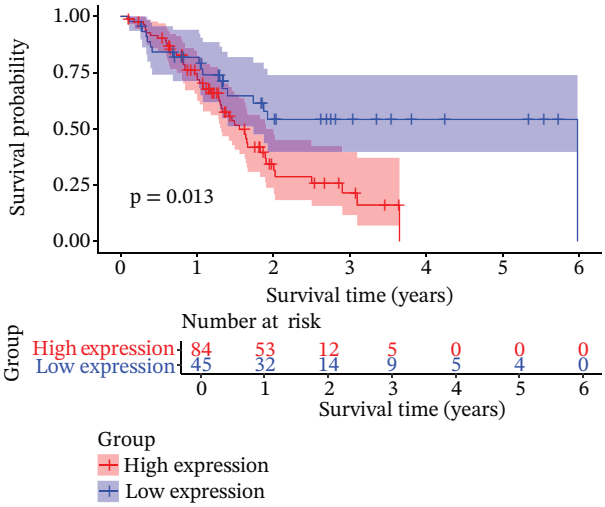
Following DSigDB screening of the complete candidate-gene set, the SNX9–coumestrol pair was selected for representative structural assessment because an experimentally resolved SNX9 structure (PDB: 3DYT) and a compatible candidate ligand were available. The predicted binding energy for the coumestrol–SNX9



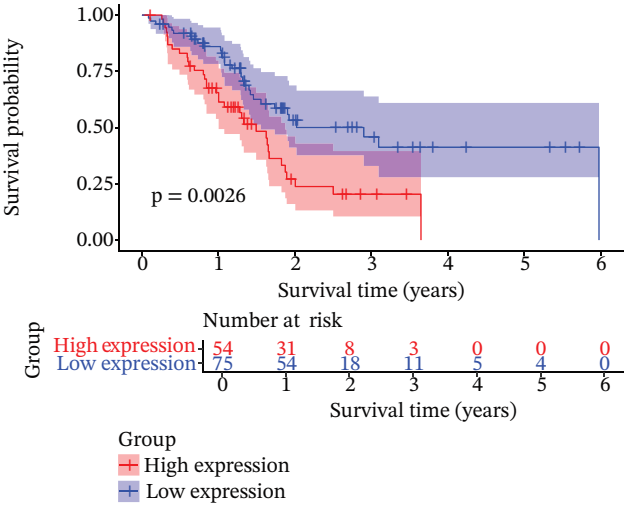
(A)



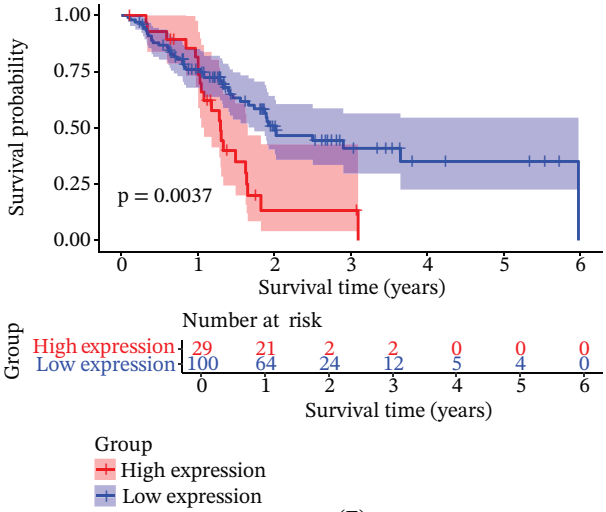
(B)



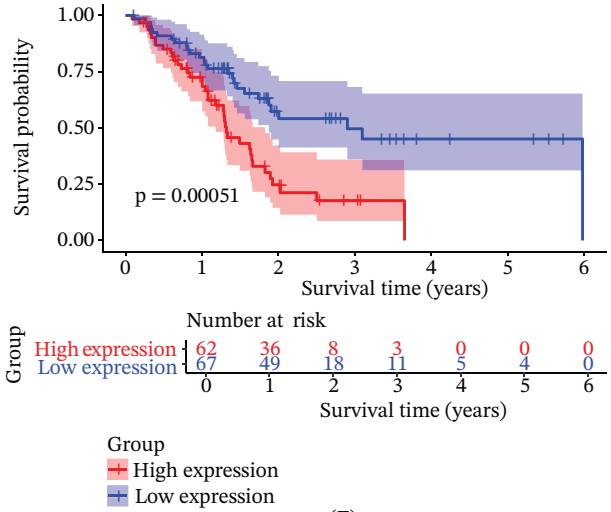
(C)



(D)



(E)



(F)

FIGURE 5 | Survival association analysis of candidate genes. (A) Forest plot of univariate Cox models for survival-associated genes. (B–F) Kaplan–Meier survival curves stratified by gene expression: (B) *CPE*, (C) *PMP22*, (D) *SEMA3C*, (E) *SNX9*, and (F) *UACA*.

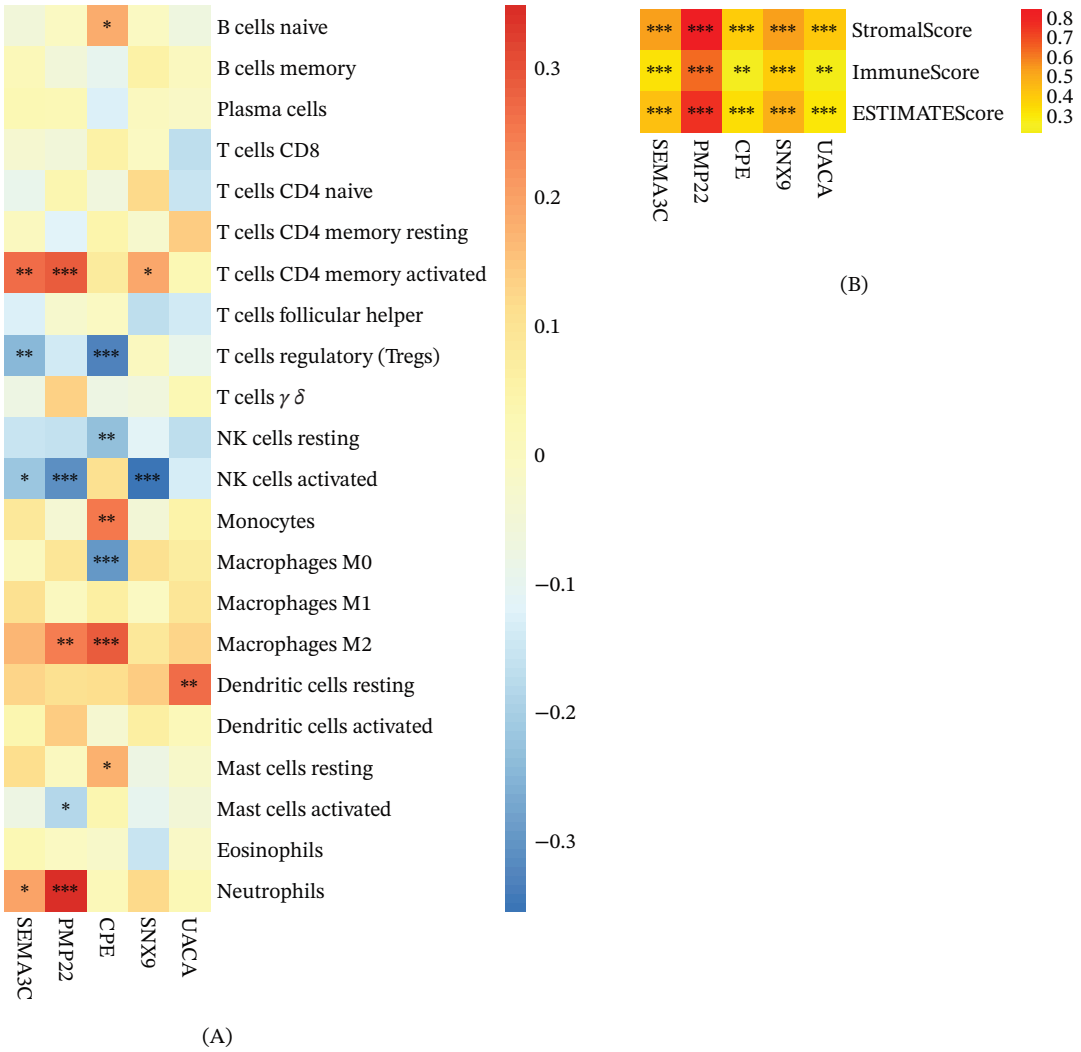


FIGURE 6 | Correlations between biomarkers and immune infiltration. (A) Heatmap of correlations between the five biomarkers and the 22 immune cell fractions inferred by CIBERSORT. (B) Heatmap of correlations between the five biomarkers and ESTIMATE-derived StromalScore, ImmuneScore, and ESTIMATEScore. *Statistical significance is indicated as follows: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

complex was −7.5 kcal/mol, meeting the prespecified computational threshold of ≤ −5.0 kcal/mol and yielding a computationally favorable docking pose; predicted ligand–receptor contacts were within 3.5 Å (Figure 9). This selection was not intended as a systematic druggability ranking and does not indicate that the remaining four genes lack therapeutic potential. These docking results indicate structural compatibility in silico but do not establish direct molecular binding or pharmacological activity.

4 | Discussion

PDAC is characterized by an exceedingly dense stroma and a profoundly immunosuppressive TME, in which CAFs are widely regarded as central drivers of fibrosis, immune evasion, and therapeutic resistance [40]. In this study, we profiled PDAC at the single-cell and downstream network levels, showing that CAF-upregulated genes are enriched in ECM and matrix-remodeling pathways; intercellular communication analysis further suggested that immune/epithelial cells modulate CAF states through a CD99–CD99 contact axis. Within the fibroblast subset, hdWGCNA identified nine co-expression modules and five fibroblast-related genes associated with OS (*SEMA3C*, *PMP22*, *CPE*, *SNX9*, *UACA*). Collectively, these findings delineate a CAF-centric framework that couples stromal remodeling with immune regulation and provide a roadmap for mechanistic studies and translational interventions.

CD99 is a homophilic cell-surface adhesion molecule, and previous studies outside PDAC have detected CD99 expression in CAFs and linked CD99 engagement to the regulation of β1-integrin/FAK-dependent adhesion. In the present analysis, CD99–CD99 interactions were predicted between T cells, macrophages, epithelial cells, and fibroblasts. Such homophilic contacts could plausibly alter the duration or strength of immune/epithelial–fibroblast contact and thereby affect adhesion, cytoskeletal organization, and matrix-remodeling programs associated with CAF states. Nevertheless, CellChat infers ligand–receptor compatibility from transcriptomic data and cannot determine signaling direction, CAF activation status, or causality. The proposed CD99-dependent mechanism should therefore be regarded as hypothesis-generating.

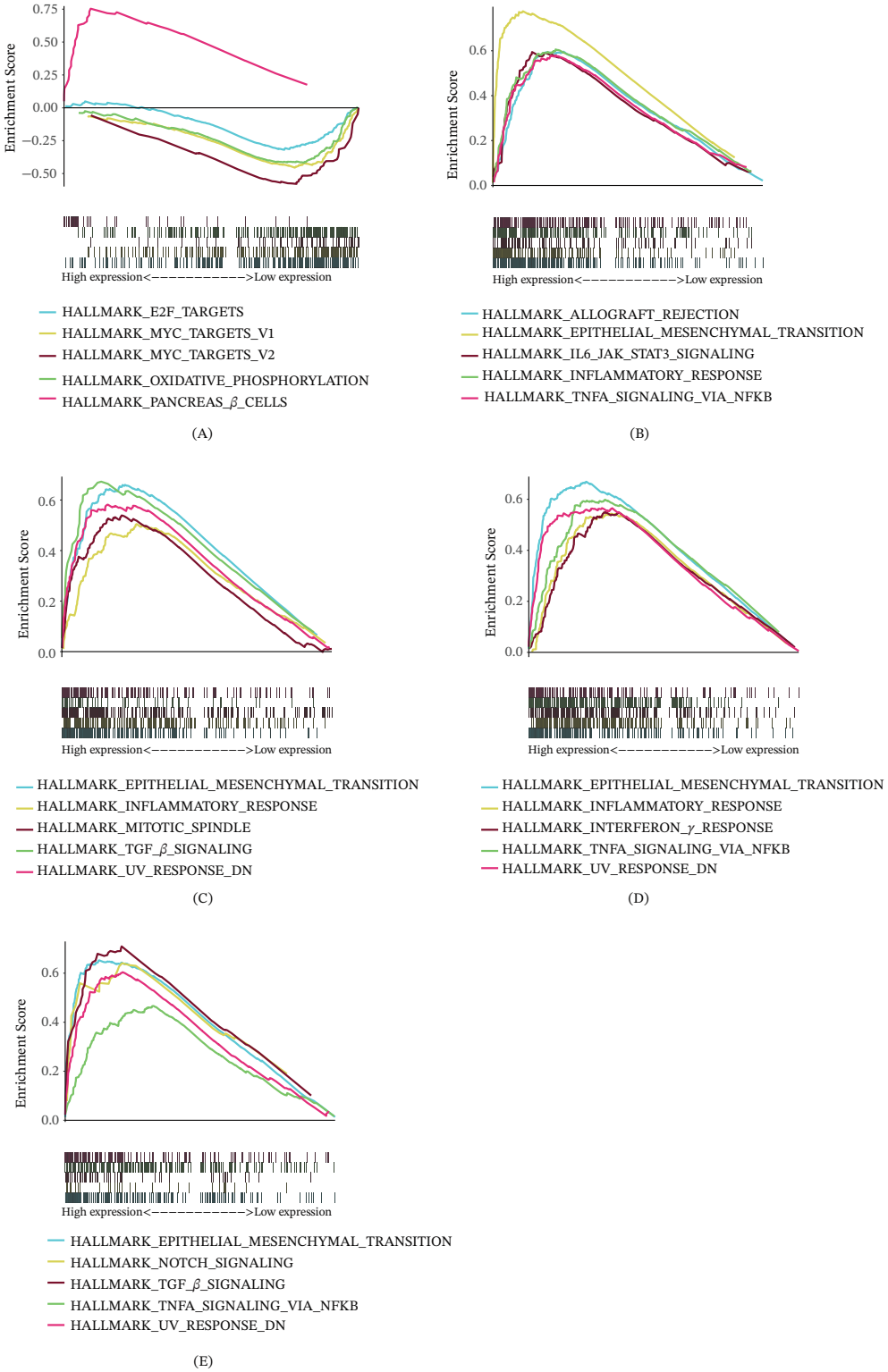


FIGURE 7 | GSEA by biomarker expression. (A–E) GSEA plots for *CPE*, *PMP22*, *SEMA3C*, *SNX9*, and *UACA* (high vs. low expression). Upper panel shows the running enrichment score (ES); lower panels indicate the positions of gene-set members along the ranked list. Positive enrichment scores indicate enrichment toward the high-expression phenotype, whereas negative enrichment scores indicate enrichment toward the low-expression phenotype.

SEMA3C is upregulated in PDAC, where it promotes tumor-cell survival and reshapes the immune milieu (including modulation of PD-L1 and associations with M2-like macrophage markers), and its high expression correlates with poorer survival [41]. Additional work links *SEMA3C* to tumor cell stemness/drug resistance and peritoneal metastasis, underscoring pleiotropic promalignant roles [42]. Consistent with these observations, our siRNA-mediated knockdown

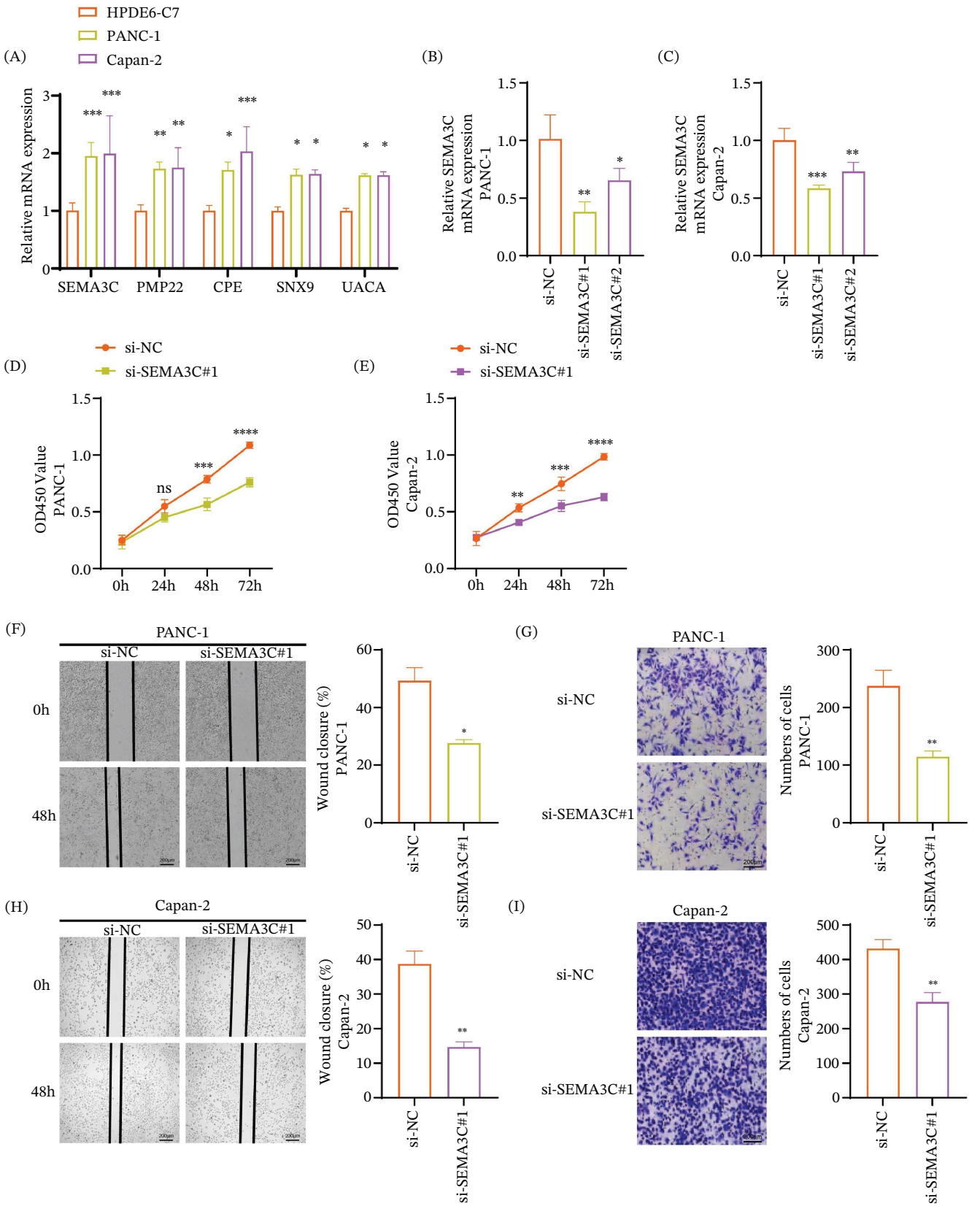


FIGURE 8 | Experimental validation of SEMA3C in PDAC cells. (A) qRT–PCR analysis of *SEMA3C*, *PMP22*, *CPE*, *SNX9*, and *UACA* expression in HPDE6-C7, PANC-1, and Capan-2 cells. (B, C) Knockdown efficiency of si-SEMA3C#1 and si-SEMA3C#2 compared with si-NC in (B) PANC-1 and (C) Capan-2 cells. (D, E) CCK-8 assays showing the effects of SEMA3C knockdown on cell viability in (D) PANC-1 and (E) Capan-2. (F, H) Wound-healing assays in (F) PANC-1 and (H) Capan-2 cells transfected with si-NC or si-SEMA3C#1. (G, I) Matrigel-coated Transwell invasion assays in (G) PANC-1 and (I) Capan-2 cells after si-NC or si-SEMA3C#1 transfection. Data are shown as mean ± SD. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, $^{****}p < 0.0001$.

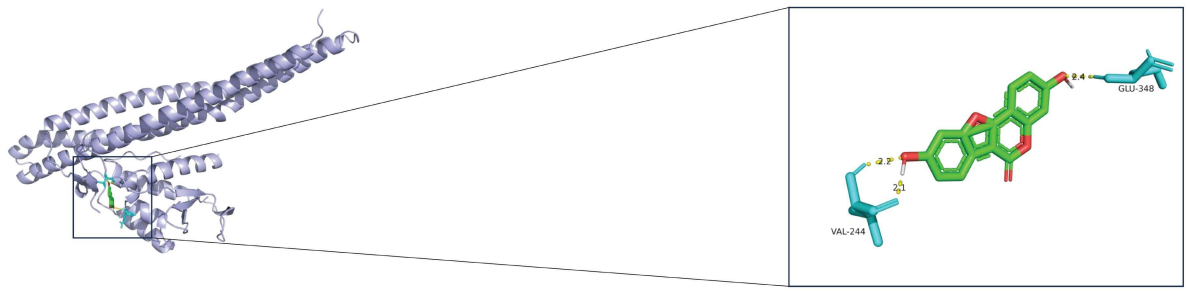


FIGURE 9 | Predicted molecular docking pose of coumestrol with *SNX9*. Docking visualization: receptor protein in blue, ligand (coumestrol) in green, interacting residues in cyan; hydrogen bonds shown as yellow dashed lines with annotated distances in Å.

experiments in PANC-1 and Capan-2 cells showed that SEMA3C depletion reduced cell viability, migration, and invasion in vitro, providing direct functional support for its protumorigenic role in PDAC. In parallel, CAF-derived *SEMA3C* has been shown in colorectal cancer to drive liver metastasis via an NRP2–MAPK axis, supporting a stroma-origin, paracrine mechanism and providing external support for its function in CAF contexts [43]. For *PMP22*, recent studies focusing on m6A regulation indicate that, in PDAC settings, *METTL14* can influence p53-effector transcripts related to *PMP22* to promote proliferation and invasion, although direct, reproducible in vivo PDAC evidence remains limited [44]. *CPE* has been summarized across multiple cancers as a promoter of growth and metastasis; in PDAC cells, *CPE* knockdown suppresses proliferation, suggesting a protumor role, but direct data on its expression and function in CAFs are still lacking [45]. For *SNX9*, recent work demonstrates that deleting *SNX9* in T cells alleviates $CD8^+$ T-cell exhaustion and enhances CAR-T efficacy, highlighting a key regulatory role on the immune side and implying that *SNX9* may affect PDAC outcomes via immune pathways [46]. *UACA* (Nucling) has been associated with inferior survival in recent transcriptome-wide and TWAS analyses in breast and other solid tumors [47]; however, direct evidence in PDAC/CAF contexts remains insufficient and will require spatial and higher resolution validation. Importantly, PMP22, CPE, and UACA should not be interpreted as established PDAC CAF drivers. Their inclusion in the present study was based primarily on their membership in fibroblast-derived co-expression modules and their associations with survival and stromal scores. Thus, our results position these genes as module-associated candidate markers rather than demonstrating fibroblast-specific causal functions. Spatial localization and perturbation in defined CAF subsets will be necessary to determine whether they participate directly in PDAC stromal biology.

We observed mixed positive and negative correlations between the five genes and specific immune-cell fractions, together with increased StromalScore, ImmuneScore, and ESTIMATEScore. For example, PMP22 and CPE were positively associated with M2 macrophages, whereas SEMA3C, PMP22, and SNX9 were negatively associated with activated NK cells. These findings suggest an immune-enriched but potentially dysfunctional microenvironment rather than uniformly increased antitumor immunity. Dense CAF-derived matrix may contribute to spatial T-cell exclusion even when immune-cell transcripts are detectable in bulk tissue. However, CIBERSORT estimates cell fractions but does not provide spatial information; therefore, the present correlations are not direct evidence of CAF-mediated T-cell exclusion. Mechanistically, *SEMA3C*'s regulation of PD-L1 and association with M2-like macrophage markers implicate a potential CAF/tumor-cell–M2 macrophage–T-cell suppression axis [48]. *SNX9*'s control of T-cell exhaustion offers an immunologic lever that can help explain the concordance between high-expression strata and worse outcomes [49]. At the microenvironmental level, CAFs can suppress T-cell function via cytokine signaling, extracellular vesicles, and metabolic coupling, which aligns with our association patterns [50]. Moreover, the frequent activation of EMT and TNFα–NF-κB pathways in PDAC supports a self-reinforcing loop linking matrix remodeling, inflammatory amplification, and immune suppression [51].

Recent in vitro and in vivo studies suggest that coumestrol exhibits antitumor activity across multiple malignancies, for example by inducing apoptosis and limiting invasion in colorectal cancer, with involvement of ferroptosis-related pathways [52]. As a natural haspin inhibitor, a sprout-enriched diet reduced intestinal polyps and improved cachexia-related systemic phenotypes in mouse models, pointing to translational pharmacologic potential [53]. Epidemiologic analyses from PLCO further indicate that higher intake of total isoflavones and selected components (including coumestrol) associates with a lower risk of pancreatic cancer, predominantly among former smokers [54]. Overall, coumestrol remains at a preclinical/early-evidence stage in oncology; logical next steps are to test it in PDAC models, particularly CAF coculture and immunocompetent systems, evaluate combination efficacy with standard regimens, and assess pharmacokinetics and safety while attending to potential endocrine and transporter-related interactions. Importantly, epidemiological associations with dietary coumestrol intake do not demonstrate that physiologically achievable plasma or tumor-tissue concentrations are sufficient to engage SNX9 or exert antitumor effects. Absorption, metabolism, protein binding, endocrine activity, and intratumoral exposure may differ substantially from concentrations used in pharmacological experiments. Therefore, our docking result should be interpreted only as a computational hypothesis rather than evidence of therapeutic efficacy.

This study nonetheless has limitations. First, the single-cell analysis included only six treatment-naïve PDAC resection specimens, and individual tumor-stage information was unavailable. This limited sample size may not capture the full interpatient, stage-related, and treatment-related heterogeneity of CAF states. The survival analyses used gene-specific Kaplan–Meier optimal cutpoints and univariable Cox models to screen individual survival-associated candidates; no multigene risk score or clinical prediction model was constructed. Multivariable adjustment and independent external-cohort validation were not performed. Therefore, the identified genes should not be interpreted as independent prognostic factors, and their survival associations require validation in future cohorts. The immune-infiltration landscape was inferred by transcriptomic deconvolution and lacks spatial, imaging, or protein-level confirmation. Future multiplex immunofluorescence or spatial-transcriptomic studies should examine the colocalization of the five genes and CD99 with CAF markers such as FAP, ACTA2/α-SMA, and PDGFRA, together with CD8-positive T cells. The linkage between the five genes and CAFs is primarily based on transcriptomic and network associations; lineage specificity and causal functions require validation in defined myCAF, iCAF, and apCAF populations. The SEMA3C-knockdown experiments in PANC-1 and Capan-2 cells demonstrate only a cancer-cell-intrinsic phenotype, and the absence of CAF–tumor-cell coculture experiments prevents conclusions regarding paracrine or CAF-specific SEMA3C functions. Finally, molecular docking provides predictive evidence only; direct binding assays, molecular-dynamics simulations, cellular target-engagement experiments, pharmacokinetic analyses, and in vivo efficacy and safety studies remain necessary.

5 | Conclusion

This single-cell–anchored study characterized the PDAC microenvironment and identified nine fibroblast co-expression modules and five fibroblast-related survival-associated genes—SEMA3C, PMP22, CPE, SNX9, and UACA. Intercellular communication analysis identified a putative CD99–CD99 contact axis, whereas immune and pathway analyses linked the candidate genes to stromal enrichment and tumor-related biological programs. SEMA3C knockdown reduced PDAC-cell viability, migration, and invasion in vitro. Molecular docking further generated a preliminary coumestrol–SNX9 interaction hypothesis. Collectively, these findings provide candidate biomarkers and experimentally testable mechanisms, but their prognostic independence, CAF specificity, and pharmacological relevance require further validation.

Nomenclature

BH	Benjamini–Hochberg
CAF	cancer-associated fibroblast
CD99	cluster of differentiation 99
CIBERSORT	Cell type Identification By Estimating Relative Subsets Of RNA Transcripts
CI	confidence interval
ECM	extracellular matrix
EMT	epithelial–mesenchymal transition
ESTIMATE	Estimation of Stromal and Immune cells in Malignant Tumor tissues using Expression
FDR	false discovery rate
FPKM	fragments per kilobase of transcript per million mapped reads
GSEA	gene set enrichment analysis
HALLMARK	MSigDB Hallmark gene-set collection
hdWGCNA	high-dimensional weighted gene co-expression network analysis
HR	hazard ratio
KNN	k-nearest neighbor
kME	module membership
LM22	leukocyte gene-signature matrix of 22 immune cell types
MSigDB	Molecular Signatures Database
NF-κB	nuclear factor kappa-B
OS	overall survival
PCA	principal component analysis
PDB	Protein Data Bank
PDAC	pancreatic ductal adenocarcinoma
QC	quality control
scRNA-seq	single-cell RNA sequencing
TCGA	The Cancer Genome Atlas
TME	tumor microenvironment
TOM	topological overlap matrix
UMAP	uniform manifold approximation and projection
WGCNA	weighted gene co-expression network analysis

Author Contributions

Tengfei Ji designed the study, performed data analysis, curated the dataset, drafted the original manuscript, and prepared visualizations. Hongsheng Wu and Zongmin Zheng assisted in data processing, validated results, and revised the manuscript. Keqiang Ma supervised the research, provided resources, and critically revised the manuscript.

Acknowledgments

The authors have nothing to report.

Funding

This project was supported by the Huadu District Basic and Applied Basic Research Joint Funding Project (District-Hospital Collaboration), Guangzhou (No. 24HDQYLH010, Tengfei Ji) and the Key Medical Disciplines and Specialities Program of Guangzhou [2025-2027].

Ethics Statement

Ethical approval was not required for this study because it did not involve any human or animal experiments.

Consent

The authors have nothing to report.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The datasets generated and/or analyzed during the current study are available in the [GSE212966] repository, [https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE212966].

References

1. D. Zhang, S. Fu, J. Xu, and X. Sun, "Clinical Significance of CA-199 and LINC01197 in Pancreatic Cancer," *Cancer* 22, no. 2 (2020): 95–105, https://doi.org/10.32604/oncologie.2020.012439.

2. L. Periyasamy, B. Murugantham, R. Sundaraj, S. Krishnamoorthi, and S. Muthusami, "Screening of MMP-2 Inhibiting Phytoconstituents for the Development of Newer Pancreatic Cancer Treatment Modalities," *Protein and Peptide Letters* 30, no. 4 (2023): 304–313, https://doi.org/10.2174/0929866530666230213113835.

3. S. Zhan, and P. Feng, "Cost-Effectiveness Analysis of Irinotecan Hydrochloride Liposome in Combination With 5-Fluorouracil and Leucovorin in Locally Advanced or Metastatic Pancreatic Ductal Adenocarcinoma," *Frontiers in Pharmacology* 16 (2025): 1596658, https://doi.org/10.3389/fphar.2025.1596658.

4. E. T. Pavlidis, I. N. Galanis, and T. E. Pavlidis, "Updates in the Diagnosis and Management of Ductal Adenocarcinoma of the Pancreas," *World Journal of Clinical Oncology* 16, no. 6 (2025): 105601, https://doi.org/10.5306/wjco.v16.i6.105601.

5. A. Mukund, M. A. Afridi, A. Karolak, M. A. Park, J. B. Permuth, and G. Rasool, "Pancreatic Ductal Adenocarcinoma (PDAC): A Review of Recent Advancements Enabled by Artificial Intelligence," *Cancers* 16, no. 12 (2024): 2240, https://doi.org/10.3390/cancers16122240.

6. T. Conroy, F. Castan, A. Lopez, et al., "Five-Year Outcomes of FOLFIRINOX vs Gemcitabine as Adjuvant Therapy for Pancreatic Cancer: A Randomized Clinical Trial," *JAMA Oncology* 8, no. 11 (2022): 1571–1578, https://doi.org/10.1001/jamaoncol.2022.3829.

7. P. K. Chintamaneni, S. Pindiprolu, S. S. Swain, et al., "Conquering Chemoresistance in Pancreatic Cancer: Exploring Novel Drug Therapies and Delivery Approaches Amidst Desmoplasia and Hypoxia," *Cancer Letters* 588 (2024): 216782, https://doi.org/10.1016/j.canlet.2024.216782.

8. Y. Ju, D. Xu, M. M. Liao, et al., "Barriers and Opportunities in Pancreatic Cancer Immunotherapy," *NPJ Precision Oncology* 8, no. 1 (2024): 199, https://doi.org/10.1038/s41698-024-00681-z.

9. Y. Masugi, "The Desmoplastic Stroma of Pancreatic Cancer: Multilayered Levels of Heterogeneity, Clinical Significance, and Therapeutic Opportunities," *Cancers* 14, no. 13 (2022): 3293, https://doi.org/10.3390/cancers14133293.

10. F. H. Miller, C. Lopes Vendrami, N. A. Hammond, P. K. Mittal, P. Nikolaidis, and A. Jawahar, "Pancreatic Cancer and Its Mimics," *Radiographics* 43, no. 11 (2023): e230054, https://doi.org/10.1148/rg.230054.

11. B. Zhang, C. Zhang, C. Chen, et al., "Hybrid Membranes-Mediated Biomimetic-Nanoparticle Carrying miR-665 for Effective Tumor Treatment by Remodeling Tumor Microenvironment," *International Journal of Pharmaceutics* 675 (2025): 125479, https://doi.org/10.1016/j.ijpharm.2025.125479.

12. M. Shakiba, and D. A. Tuveson, "Macrophages and Fibroblasts as Regulators of the Immune Response in Pancreatic Cancer," *Nature Immunology* 26, no. 5 (2025): 678–691, https://doi.org/10.1038/s41590-025-02134-6.

13. M. M. Chen, Q. Gao, H. Ning, et al., "Integrated Single-Cell and Spatial Transcriptomics Uncover Distinct Cellular Subtypes Involved in Neural Invasion in Pancreatic Cancer," *Cancer Cell* 43, no. 9 (2025): 1656–1676.e10, https://doi.org/10.1016/j.ccell.2025.06.020.

14. Z. Song, J. Yu, M. Wang, et al., "CHDTEPDB: Transcriptome Expression Profile Database and Interactive Analysis Platform for Congenital Heart Disease," *Congenital Heart Disease* 18, no. 6 (2023): 693–701, https://doi.org/10.32604/chd.2024.048081.

15. W. Croft, H. Pearce, S. Margielewska-Davies, et al., "Spatial Determination and Prognostic Impact of the Fibroblast Transcriptome in Pancreatic Ductal Adenocarcinoma," *eLife* 12 (2023): 12, https://doi.org/10.7554/eLife.86125.

16. S. Menezes, M. H. Okail, S. M. A. Jalil, H. M. Kocher, and A. J. M. Cameron, "Cancer-Associated Fibroblasts in Pancreatic Cancer: New Subtypes, New markers, new targets," *Journal of pathology.* 257, no. 4 (2022): 526–544, https://doi.org/10.1002/path.5926.

17. F. Poyia, C. M. Neophytou, M. I. Christodoulou, and P. Papageorgis, "The Role of Tumor Microenvironment in Pancreatic Cancer Immunotherapy: Current Status and Future Perspectives," *International Journal of Molecular Sciences* 25, no. 17 (2024): 9555, https://doi.org/10.3390/ijms25179555.

18. H. Masuda, "Cancer-Associated Fibroblasts in Cancer Drug Resistance and Cancer Progression: A Review," *Cell Death Discovery* 11, no. 1 (2025): 341, https://doi.org/10.1038/s41420-025-02566-x.

19. M. Chakkera, J. B. Foote, B. Farran, and G. P. Nagaraju, "Breaking the Stromal Barrier in Pancreatic Cancer: Advances and Challenges," *Biochimica Et Biophysica Acta Reviews On Cancer* 1879, no. 1 (2024): 189065, https://doi.org/10.1016/j.bbcan.2023.189065.

20. O. Olajubutu, O. D. Ogundipe, A. Adebayo, and S. K. Adesina, "Drug Delivery Strategies for the Treatment of Pancreatic Cancer," *Pharmaceutics* 15, no. 5 (2023): 1318, https://doi.org/10.3390/pharmaceutics15051318.

21. Y. Liu, A. Sinjab, J. Min, et al., "Conserved Spatial Subtypes and Cellular Neighborhoods of Cancer-Associated Fibroblasts Revealed by Single-Cell Spatial Multi-Omics," *Cancer Cell* 43, no. 5 (2025): 905–924, https://doi.org/10.1016/j.ccell.2025.03.004.

22. T. Zhang, Y. Ren, P. Yang, J. Wang, and H. Zhou, "Cancer-Associated Fibroblasts in Pancreatic Ductal Adenocarcinoma," *Cell Death & Disease* 13, no. 10 (2022): 897.

23. H. Huang, Z. Wang, Y. Zhang, et al., "Mesothelial Cell-Derived Antigen-Presenting Cancer-Associated Fibroblasts Induce Expansion of Regulatory T Cells in Pancreatic Cancer," *Cancer Cell* 40, no. 6 (2022): 656–673, https://doi.org/10.1016/j.ccell.2022.04.011.

24. F. Ge, C. Zeng, J. Wang, et al., "Cancer-Associated Fibroblasts Drive Early Pancreatic Cancer Cell Invasion via the SOX4/MMP11 Signalling Axis," *Biochimica Et Biophysica Acta Molecular Basis Of Disease* 1870, no. 1 (2024): 166852, https://doi.org/10.1016/j.bbadis.2023.166852.

25. K. Boelaars, E. Rodriguez, Z. R. Huinen, et al., "Pancreatic Cancer-Associated Fibroblasts Modulate Macrophage Differentiation via Sialic Acid-Siglec Interactions," *Communications Biology* 7, no. 1 (2024): 430, https://doi.org/10.1038/s42003-024-06087-8.

26. A. Brichkina, P. Polo, S. D. Sharma, N. Visestamkul, and M. Lauth, "A Quick Guide to CAF Subtypes in Pancreatic Cancer," *Cancers* 15, no. 9 (2023): 2614, https://doi.org/10.3390/cancers15092614.

27. K. J. Lee, Y. Kim, Y. H. Yoo, et al., "CD99-Derived Agonist Ligands Inhibit Fibronectin-Induced Activation of β1 Integrin Through the Protein Kinase A/SHP2/Extracellular Signal-Regulated Kinase/PTPN12/Focal Adhesion Kinase Signaling Pathway," *Molecular And Cellular Biology* 37, no. 14 (2017), https://doi.org/10.1128/MCB.00675-16.

28. K. Edlund, C. Lindskog, A. Saito, et al., "CD99 Is a Novel Prognostic Stromal Marker in Non-Small Cell Lung Cancer," *International Journal Of Cancer* 131, no. 10 (2012): 2264–2273, https://doi.org/10.1002/ijc.27518.

29. L. Ye, X. Tang, J. Zhong, et al., "Unraveling the Complex Pathophysiology of White Matter Hemorrhage in Intracerebral Stroke: A Single-Cell RNA Sequencing Approach," *CNS Neuroscience & Therapeutics* 30, no. 3 (2024): e14652, https://doi.org/10.1111/cns.14652.

30. J. Wu, W. Wu, J. Qin, et al., "Exploring the Impact of Cuproptosis on Prostate Cancer Prognosis via RNA Methylation Regulation Based on Single Cell and Bulk RNA Sequencing Data," *Frontiers In Pharmacology* 16 (2025): 1573611, https://doi.org/10.3389/fphar.2025.1573611.

31. L. Merotto, G. Sturm, A. Dietrich, M. List, and F. Finotello, "Making Mouse Transcriptomics Deconvolution Accessible With Immunedeconv," *Bioinformatics Advances* 4, no. 1 (2024): , https://doi.org/10.1093/bioadv/vbae032.

32. I. Korsunsky, N. Millard, J. Fan, et al., "Fast, Sensitive and Accurate Integration of Single-Cell Data With Harmony," *Nature Methods* 16, no. 12 (2019): 1289–1296, https://doi.org/10.1038/s41592-019-0619-0.

33. N. Rusk, "Expanded CIBERSORTx," *Nature Methods* 16, no. 7 (2019): 577, https://doi.org/10.1038/s41592-019-0486-8.

34. S. Jin, M. V. Plikus, and Q. Nie, "Cellchat For Systematic Analysis Of Cell–Cell Communication From Single-Cell Transcriptomics," *Nature Protocols* 20, no. 1 (2025): 180–219, https://doi.org/10.1038/s41596-024-01045-4.

35. B. S. White, A. de Reyniès, A. M. Newman, et al., "Community Assessment of Methods to Deconvolve Cellular Composition From Bulk Gene Expression," *Nature Communications* 15, no. 1 (2024): 7362, https://doi.org/10.1038/s41467-024-50618-0.

36. C. B. Steen, C. L. Liu, A. A. Alizadeh, and A. M. Newman, "Profiling Cell Type Abundance and Expression in Bulk Tissues With CIBERSORTx," *Methods in Molecular Biology* 2117 (2020): 135–157, https://doi.org/10.1007/978-1-0716-0301-7_7.

37. W. Bai, "Protocol to Detect Immune Levels, Abnormal Metabolism, and Signaling Pathways in Tumor Tissue Based on scRNA-seq Obtained From Patient Databases," *STAR Protocols* 5, no. 2 (2024): 103065, https://doi.org/10.1016/j.xpro.2024.103065.

38. Q. Zeng, K. Feng, Y. Yu, and Y. Lv, "Hsa_Circ_0000021 Sponges miR-3940-3p/KPNA2 Expression to Promote Cervical Cancer Progression," *Current Molecular Pharmacology* 17 (2023): e170223213775, https://doi.org/10.2174/1874467216666230217151946.

39. M. V. Kuleshov, M. R. Jones, A. D. Rouillard, et al., "Enrichr: A Comprehensive Gene Set Enrichment Analysis Web Server 2016 Update," *Nucleic Acids Research* 44, no. W1 (2016): W90–W97, https://doi.org/10.1093/nar/gkw377.

40. C. J. Halbrook, C. A. Lyssiotis, M. Pasca di Magliano, and A. Maitra, "Pancreatic cancer: Advances and challenges," *Cell* 186, no. 8 (2023): 1729–1754, https://doi.org/10.1016/j.cell.2023.02.014.

41. D. Zhang, A. Lindstrom, E. J. Kim, et al., "SEMA3C Supports Pancreatic Cancer Progression by Regulating the Autophagy Process and Tumor Immune Microenvironment," *Frontiers Ii Oncology* 12 (2022): 890154, https://doi.org/10.3389/fonc.2022.890154.

42. S. Tomizawa, S. Takano, R. Eto, T. Takayashiki, S. Kuboki, and M. Ohtsuka, "Semaphorin 3 C Enhances Putative Cancer Stemness and Accelerates Peritoneal Dissemination in Pancreatic Cancer," *Cancer Cell International* 23, no. 1 (2023): 155, https://doi.org/10.1186/s12935-023-03008-3.

43. Y. Zhang, A. Zuo, Y. Ba, et al., "Cancer-Associated Fibroblast-Derived SEMA3C Facilitates Colorectal Cancer Liver Metastasis via NRP2-Mediated MAPK Activation," *Proceedings of the National Academy of Sciences of the United States of America* 122, no. 21 (2025): e2423077122, https://doi.org/10.1073/pnas.2423077122.

44. S. Meng, and C. Wang, "Analysis of METTL14 Expression in Pancreatic Cancer and Adjacent Tissues and Its Prognostic Value for Patient Outcomes," *Clinical And Experimental Medicine* 25, no. 1 (2025): 3, https://doi.org/10.1007/s10238-024-01506-w.

45. S. Hareendran, X. Yang, V. K. Sharma, and Y. P. Loh, "Carboxypeptidase E and Its Splice Variants: Key Regulators of Growth and Metastasis in Multiple Cancer Types," *Cancer Letters* 548 (2022): 215882, https://doi.org/10.1016/j.canlet.2022.215882.

46. M. P. Trefny, N. Kirchhammer, P. Auf der Maur, et al., "Deletion of SNX9 Alleviates CD8 T Cell Exhaustion for Effective Cellular Cancer Immunotherapy," *Nature Communications* 14, no. 1 (2023), https://doi.org/10.1038/s41467-022-35583-w.

47. Q. Zhu, E. Schultz, J. Long, et al., "UACA Locus Is Associated With Breast Cancer Chemoresistance and Survival," *NPJ Breast Cancer* 8, no. 1 (2022): 39, https://doi.org/10.1038/s41523-022-00401-5.

48. M. Li, B. Zhou, and C. Zheng, "An Integrated Bioinformatic Analysis of Bulk and Single-Cell Sequencing Clarifies Immune Microenvironment And Metabolic Profiles Of Lung Adenocarcinoma To Predict Immunotherapy Efficacy," *Frontiers In Cell And Developmental Biology* 11 (2023): 1163314, https://doi.org/10.3389/fcell.2023.1163314.

49. E. Jenkins, T. Whitehead, M. Fellermeyer, S. J. Davis, and S. Sharma, "The Current State and Future of T-Cell Exhaustion Research," *Oxford Open Immunology* 4, no. 1 (2023), https://doi.org/10.1093/oxfimm/iqad006.

50. V. Milosevic, and A. Östman, "Interactions Between Cancer-Associated Fibroblasts and T-Cells: Functional Crosstalk With Targeting and Biomarker Potential," *Upsala Journal of Medical Sciences* 129 (2024): e10710, https://doi.org/10.48101/ujms.v129.10710.

51. W. Luo, T. Wen, and X. Qu, "Tumor Immune Microenvironment-Based Therapies in Pancreatic Ductal Adenocarcinoma: Time to Update the Concept," *Journal Of Experimental & Clinical Cancer Research : CR* 43, no. 1 (2024): 8, https://doi.org/10.1186/s13046-023-02935-3.

52. J. Geng, Y. Wang, F. Lv, et al., "Coumestrol Facilitates Apoptosis in Colorectal Cancer Cells by Interacting With ZIP8 Protein via the Ferroptosis Pathway," *Journal of Cancer* 15, no. 14 (2024): 4656–4667, https://doi.org/10.7150/jca.94628.

53. H. Tanaka, S. Matsuyama, T. Ohta, et al., "The Natural HASPIN Inhibitor Coumestrol Suppresses Intestinal Polyp Development, Cachexia, and Hypogonadism in a Mouse Model of Familial Adenomatous Polyposis (Apc(Min/+))," *Biology* 13, no. 9 (2024): 736, https://doi.org/10.3390/biology13090736.

54. C. Liu, M. Reger, H. Fan, J. Wang, and J. Zhang, "Dietary Intake of Isoflavones and Coumestrol and Risk of Pancreatic Cancer in the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial," *British Journal Of Cancer* 132, no. 3 (2025): 266–275, https://doi.org/10.1038/s41416-024-02929-8.