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.

Spatial epithelial and stromal remodeling in isolated human pancreatic intraepithelial neoplasia before invasion

Atsuhiro Masuda

atmasuda@med.kobe-u.ac.jp

Kobe University Graduate School of Medicine https://orcid.org/0000-0002-2138-1207

Masahiro Tsujimae

Division of Gastroenterology, Department of Internal Medicine, Kobe University Graduate School of Medicine

Tetsuhisa Ko

Division of Gastroenterology, Department of Internal Medicine, Kobe University Graduate School of Medicine

Mika Miki

Division of Gastroenterology, Department of Internal Medicine, Kobe University Graduate School of Medicine

Kohei Okamoto

Division of Gastroenterology, Department of Internal Medicine, Kobe University Graduate School of Medicine

Kento Ogawa

Division of Gastroenterology, Department of Internal Medicine, Kobe University Graduate School of Medicine

Tatsuya Kageyama

Division of Gastroenterology, Department of Internal Medicine, Kobe University Graduate School of Medicine

Takafumi Tokunaga

Division of Gastroenterology, Department of Internal Medicine, Kobe University Graduate School of Medicine

Yuta Kawase

Division of Gastroenterology, Department of Internal Medicine, Kobe University Graduate School of Medicine

Kaoruko Kanamaru

Division of Gastroenterology, Department of Internal Medicine, Kobe University Graduate School of Medicine

Kenta Yamamoto
Division of Gastroenterology, Department of Internal Medicine, Kobe University Graduate School of Medicine
Akira Shirohata
Division of Gastroenterology, Department of Internal Medicine, Kobe University Graduate School of Medicine
Yusuke Yokotani
Division of Gastroenterology, Department of Internal Medicine, Kobe University Graduate School of Medicine
Yuki Oka
Division of Gastroenterology, Department of Internal Medicine, Kobe University Graduate School of Medicine
Noriko Juri
Division of Gastroenterology, Department of Internal Medicine, Kobe University Graduate School of Medicine
Yosuke Irie
Division of Gastroenterology, Department of Internal Medicine, Kobe University Graduate School of Medicine
Yoshiyuki Harada
Division of Gastroenterology, Department of Internal Medicine, Kobe University Graduate School of Medicine
Kae Nagao
Division of Gastroenterology, Department of Internal Medicine, Kobe University Graduate School of Medicine
Shinya Kohashi
Division of Gastroenterology, Department of Internal Medicine, Kobe University Graduate School of Medicine
Noriko Inomata
Division of Gastroenterology, Department of Internal Medicine, Kobe University Graduate School of Medicine
Hisahiro Uemura
Division of Gastroenterology, Department of Internal Medicine, Kobe University Graduate School of Medicine
Shigeto Masuda
Division of Gastroenterology, Department of Internal Medicine, Kobe University Graduate School of Medicine
Masanori Gonda
Division of Gastroenterology, Department of Internal Medicine, Kobe University Graduate School of Medicine
Shohei Abe

Division of Gastroenterology, Department of Internal Medicine, Kobe University Graduate School of Medicine

Arata Sakai
Division of Gastroenterology, Department of Internal Medicine, Kobe University Graduate School of Medicine

Takashi Kobayashi
Division of Gastroenterology, Department of Internal Medicine, Kobe University Graduate School of Medicine https://orcid.org/0000-0002-9202-7643

Koki Kamizaki
Section of Molecular and Pathological Medicine, Department of Gastroenterology, Kobe University Graduate School of Medicine

Sara Horie
Division of Molecular Oncology, National Cancer Center Research Institute

Yasunori Kogure
National Cancer Center Research Institute https://orcid.org/0000-0003-1158-5131

Keisuke Kataoka
Keio University School of Medicine

Keitaro Fujita
Advanced Therapeutic Target Discovery, Department of Gastroenterology, Kobe University Graduate School of Medicine

Akihito Machinaga
Eisai Co., Ltd.

Toshio Imai
Advanced Therapeutic Target Discovery, Department of Gastroenterology, Kobe University Graduate School of Medicine

Yasuaki Yanagimoto
Division of Hepato-Biliary-Pancreatic Surgery, Department of Surgery, Kobe University Graduate School of Medicine

Yoshihide Nanno
Division of Hepato-Biliary-Pancreatic Surgery, Department of Surgery, Kobe University Graduate School of Medicine

Keitaro Sofue
Department of Radiology, Kobe University Graduate School of Medicine

Takayuki Kodama
Division of Molecular and Genomic Pathology, Division of Pathology, Kobe University Graduate School of Medicine; Department of Diagnostic Pathology, Kobe University Hospital

Norimitsu Uza
Division of Gastroenterology, Department of Internal Medicine, Kobe University Graduate School of Medicine

Tomoo Itoh
Department of Diagnostic Pathology, Kobe University Hospital
Yashuhiro Minami
Kobe University, Graduate School of Medicine https://orcid.org/0000-0003-3514-4285
Yuzo Kodama
Department of Gastroenteroogy, Kobe University Graduate School of Medicine, Kobe

Article

Keywords:

Posted Date: October 4th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11014749/v1

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

Additional Declarations: Yes there is potential Competing Interest. A. Machinaga is an employee of Eisai Co., Ltd. K. Fujita and T. Imai are affiliated with the Advanced Therapeutic Target Discovery Department of Gastroenterology, Kobe University Graduate School of Medicine, an endowed chair funded by Eisai Co., Ltd. The remaining authors declare no potential conflicts of interest.

Abstract

Pancreatic ductal adenocarcinoma develops through pancreatic intraepithelial neoplasia, but when epithelial–stromal remodeling emerges before invasion remains unclear. Here we integrate spatial transcriptomics, histological validation and patient-derived organoids to characterize isolated human low- and high-grade precursor lesions from pancreata without coexisting invasive carcinoma. High-grade lesions show 3.4-fold periductal stromal expansion, enrichment of myofibroblastic fibroblast programs and elevated stromal interleukin-6. Epithelial basal-like features emerge focally before invasion and expand further in invasive carcinoma organoids. Spatial ligand–receptor analyses identify candidate growth-factor and inflammatory interactions between epithelial and stromal compartments. Conditioned medium from one high-grade precursor-derived organoid line induces both myofibroblastic and inflammatory programs in pancreatic stellate cells, providing proof-of-principle functional support. These findings identify high-grade pancreatic intraepithelial neoplasia as a biologically active preinvasive epithelial–stromal ecosystem and motivate validation of candidate mediators in larger cohorts and perturbation studies.

Introduction

Pancreatic ductal adenocarcinoma (PDAC) remains highly lethal, largely because most tumors are diagnosed when curative surgery is no longer feasible.[1–2] In contrast, Japanese early detection efforts, including the Japan Study Group on the Early Detection of Pancreatic Cancer, have shown that patients diagnosed at stage 0 (carcinoma in situ, corresponding to HG-PanIN) can achieve excellent survival, highlighting the importance of understanding precursor biology.[3–5] PDAC develops through pancreatic intraepithelial neoplasia (PanIN),[6] progressing from low-grade (LG-PanIN) to high-grade (HG-PanIN) before invasion over years to decades.[6–9] However, the cellular events that convert a microscopic precursor lesion into an invasive, stroma-rich cancer remain unclear.

Early clinical detection often relies on indirect imaging rather than direct tumor visualization. Focal pancreatic parenchymal atrophy and main pancreatic duct (MPD) dilatation can precede overt tumor formation.[3,10] In our recent large-scale multicenter CT study, focal parenchymal atrophy appeared approximately 3 years before diagnosis, followed by MPD dilatation at approximately 1 year, suggesting that microscopic neoplastic lesions can be associated with tissue-scale remodeling long before a radiologically visible mass forms.[10] These observations raise the basic question of how microscopic HG-PanIN lesions are related to widespread stromal remodeling before invasion. Because dense desmoplasia and activated fibroblast states are defining features of invasive PDAC,[11–18] defining when CAF-like fibroblast programs first emerge is important for pancreatic cancer interception.

Diagnosing HG-PanIN before invasion is technically and clinically challenging because these lesions are microscopic, usually non-mass-forming, and often require endoscopic pancreatic duct sampling approaches such as pancreatic juice cytology, endoscopic nasopancreatic drainage, serial pancreatic

juice aspiration cytologic examination, or pancreatic duct brushing cytology.[3–5] This diagnostic pathway has been developed most actively in Japan, where HG-PanIN/carcinoma in situ is managed as stage 0 pancreatic cancer and can therefore lead to surgical resection before invasion.[3–5] Resected HG-PanIN tissues without any coexisting invasive carcinoma are exceptionally uncommon, and opportunities to establish organoid models of such lesions are inherently limited.

Epithelial plasticity is another feature of early PDAC development. In invasive PDAC, classical and basal-like molecular subtypes have distinct clinical associations, with basal-like tumors being linked to worse prognosis and therapeutic resistance.[19–23] Whether this subtype heterogeneity is already present in precursor lesions and how it relates to stromal remodeling remains unclear. Spatial transcriptomics enables the analysis of these epithelial–stromal relationships while preserving tissue architecture.[24–25] Spatial studies of the human pancreas have largely profiled invasive PDAC,[26–27] and recent analyses of PanIN have examined incidental lesions in organ-donor or grossly normal pancreata[28–29] or PanIN within pancreata resected for invasive carcinoma,[30] a setting in which paracrine signals, immune cells, and stromal programs from invasive tumors may confound the identification of intrinsic precursor biology.

Here, we focused on truly isolated PanIN lesions from resected specimens in which no invasive carcinoma was identified anywhere in the pancreas. This rare Japanese cohort, enabled by active stage 0 detection and surgical resection,[3–5] allowed us to study HG-PanIN as the most advanced lesion rather than as a bystander lesion coexisting with invasive PDAC. Therefore, these lesions offer a view of precursor pathology with minimal confounding from invasive tumors. We integrated standard Visium and Visium HD spatial transcriptomics with patient-derived organoids from HG-PanINs and invasive PDAC, including conditioned medium (CM) experiments in primary pancreatic stellate cells. This design allowed us to define a spatially resolved human precursor atlas and to test whether the precursor epithelium could functionally engage stromal cells before invasion.

Results

Stromal expansion precedes pancreatic cancer invasion

First, we quantified stromal remodeling across PanIN stages (Fig. 1a) using Masson's trichrome staining with automated image segmentation (Fig. 1b). Stromal proportion, defined as the collagen-positive area within 200 μm of the ductal basement membrane, was intermediate in LG-PanIN and markedly increased in HG-PanIN as acinar tissue decreased (Fig. 1c and 1d). HG-PanIN showed a 3.4-fold stromal expansion compared with the normal pancreas (normal: median 17.2%; LG-PanIN: 29.6%; HG-PanIN: 57.6%; Kruskal–Wallis $P < 0.001$ with Dunn's post hoc normal vs. HG-PanIN $P < 0.001$ and LG-PanIN vs. HG-PanIN $P < 0.001$; Supplementary Table 1), approaching levels typical of invasive PDAC, whereas the LG-PanIN increase over normal tissue was not statistically significant. Immunohistochemistry further confirmed stage-associated upregulation of αSMA and IL-6 in the periductal stroma ($P < 0.001$ for both; see CAF-like fibroblast section below).

Spatial transcriptomic profiling using Visium (normal, n = 2; LG-PanIN, n = 5; HG-PanIN, n = 5) and Visium HD (LG-PanIN, n = 2; HG-PanIN, n = 2) confirmed this stage-dependent compositional shift (Fig. 1e–g; Extended Data Fig. 1). As all PanIN samples were isolated precursor lesions without concurrent invasive PDAC in the resected pancreas, these changes are unlikely to reflect the secondary effects of nearby carcinomas. Module score-based enrichment showed a significantly increasing stromal trend from normal to LG-PanIN to HG-PanIN (Jonckheere–Terpstra, P = 0.036), and Visium HD recapitulated the same pattern at higher spatial resolution.

CAF-like programs expand in high-grade precursor lesions

Next, we characterized fibroblasts expressing CAF-associated transcriptional programs in precursor microenvironments. Because these fibroblasts occur before invasive carcinoma, we refer to them as CAF-like fibroblasts and their signatures as myCAF-like and iCAF-like programs. Using established myCAF-associated (ACTA2, TAGLN, MYL9) and iCAF-associated (IL6, CXCL12, PDGFRA) signatures, Visium spatial mapping revealed marked expansion of CAF-like programs at the HG-PanIN stage. These programs were minimal in the normal pancreas, sparse in the LG-PanIN, and expanded in the HG-PanIN (Fig. 2a–c). The myCAF-like proportion increased significantly from normal to HG-PanIN (Jonckheere–Terpstra, P = 0.025), whereas the iCAF-like proportion did not reach significance (P = 0.233), producing a relative shift toward myCAF-like dominance in HG-PanIN.

Protein-level validation supported these findings. αSMA expression increased across PanIN stages, with strong staining in 78.4% of the HG-PanIN periductal stroma (Fig. 2e and 2f; Supplementary Table 2). IL-6 was also elevated in the HG-PanIN stroma, and spatial transcriptomics localized IL6 primarily to stromal spots with high iCAF signature scores rather than to epithelial or immune compartments (Fig. 2g and 2h; Supplementary Table 3).

Analysis of periductal CAF-like fibroblasts further showed that CAF-like program-positive regions extended farther from the epithelium in HG-PanINs than in LG-PanINs. Within each stage, iCAF-like spots tended to lie farther from the epithelium than myCAF-like spots in most samples (3 of 5 in both LG-PanIN and HG-PanIN), but this spatial separation did not reach significance at the biological-replicate level (per-sample medians, paired Wilcoxon test, n = 5 per stage, P = 0.31 in both stages; Cliff's $\delta < 0.13$; Fig. 2d). Apparent spot-level differences (nominal P = 0.003 in LG-PanIN and $P = 2.2 \times 10^{-17}$ in HG-PanIN) were driven by the large number of spots (pseudoreplication) and are reported for transparency only; they do not constitute evidence of reproducible between-sample separation, for which the per-sample test above is the appropriate inference (see Methods). These findings indicate that HG-PanIN is accompanied by expansion of a CAF-like stromal compartment before invasion, whereas the relative spatial positioning of the iCAF-like programs and myCAF-like programs remains exploratory.

Focal basal-like drift accompanies PanIN progression

Next, we evaluated whether classical and basal-like epithelial programs were detectable in precursor lesions. Visium mapping showed maintained but spatially heterogeneous classical scores across stages (Fig. 3a), whereas basal-like scores were minimal in normal ducts and increased from LG-PanIN to HG-PanIN (Fig. 3b; Extended Data Figs. 2–4).

Visium HD confirmed distinct but overlapping classical and basal marker domains in HG-PanINs (Fig. 3c; Extended Data Figs. 5 and 6). Most epithelial cells occupied intermediate transcriptional states in both LG-PanINs (66.5%) and HG-PanINs (68.6%), with similar threshold-based basal-like fractions (7.4% and 7.6%, respectively; Fig. 3d). However, continuous Subtype Index analysis revealed a small but significant leftward shift toward the basal-like direction in HG-PanIN (Kolmogorov–Smirnov $D = 0.028$, $P = 0.011$; Wilcoxon $P = 0.004$; Fig. 3e), supporting gradual redistribution along a classical-to-basal spectrum rather than abrupt subtype switching.

Immunohistochemistry results corroborated these findings. Basal markers p63 and KRT5 were rare in normal ducts and LG-PanIN but were significantly elevated in HG-PanIN ($P < 0.001$ for both; Fig. 3f and 3g; Supplementary Tables 4 and 5). Conversely, GATA6 expression scores were reduced in HG-PanIN (Kruskal–Wallis $P < 0.001$; Supplementary Table 6). The magnitude of the transcriptomic shift was small, and categorical basal-like fractions were similar between LG-PanIN and HG-PanIN. Therefore, spatial transcriptomics and IHC together supported focal basal-like drift rather than overt subtype conversion before invasion.

Spatial profiling nominates epithelial–stromal interactions

The stromal expansion observed during PanIN progression prompted us to investigate the epithelial–stromal crosstalk by ligand-receptor (L-R) analysis using curated pathway databases.

Epithelial-to-stromal L-R analysis revealed stage-associated enrichment of candidate fibroblast-engaging pathways in HG-PanIN, including TGFβ, PDGF, FGF, and WNT signaling (Fig. 4a). Stromal-to-epithelial interactions showed increased ECM–integrin, matricellular, and IGF signaling, whereas CXCL12–CXCR4 activity was decreased (Fig. 4b and 4c). This decrease was primarily driven by reduced epithelial CXCR4 expression rather than loss of stromal CXCL12, indicating diminished epithelial responsiveness to an otherwise active iCAF-associated signal.

A focused comparison of the nine key epithelial-to-stromal L-R pairs showed stage-associated increases in interaction scores (Fig. 4d–f). WNT7B-FZD7 was significantly increased in HG-PanIN ($P = 0.0079$), followed by TGFB1–TGFBR1 ($P = 0.032$); FGF2–FGFR1 and WNT7A–FZD1 also showed an upward trend.

Visium HD analysis showed similar patterns at higher spatial resolutions (Fig. 4g–j). Disease stage was the dominant axis of L-R variation, whereas basal-like epithelium showed modestly higher activity than its classical counterparts within each stage. Given the limited number of Visium HD samples, these subtype-stratified L-R findings should be interpreted as exploratory and hypothesis-generating. In the

two-sample-per-group Visium HD comparison, basal-like epithelial cells showed increased HGF–MET, WNT7A–FZD1, WNT7B–FZD7, and FGF2–FGFR1 activities in HG-PanINs (Fig. 4h), although these effect size estimates require validation in larger cohorts. Subtype-stratified profiling further suggested that growth factor signaling was most prominent in the basal-like-to-myCAF-like axis, whereas inflammatory signaling, including IL6-IL6R and IL1B-IL1R1, was preferentially active in the basal-like-to-iCAF-like axis (Fig. 4i and 4j).

Across standard Visium, Visium HD, and CellChat analyses, candidate growth-factor and morphogen interactions, including PDGF, FGF, WNT, TGFβ, and HGF/MET-related interactions, were enriched in HG-PanIN, with Hedgehog detectable but more modestly represented. Because L-R analysis infers potential communication from spatially resolved expression patterns, these pathways should be interpreted as candidate axes that require functional validation. Thus, disease stage shaped the broad L-R remodeling pattern, whereas basal-like epithelial populations contributed subtype-specific growth factors and inflammatory signaling.

Organoids reveal subtype-associated ligand programs

To test the biological consequences of epithelial subtype heterogeneity, we established patient-derived organoids from HG-PanIN and invasive PDAC specimens (Fig. 5a; Extended Data Fig. 7).[31–34] HG-PanIN cells were obtained by brushing the pancreatic duct during ERCP.

We generated organoids from two HG-PanIN cases (KKp048 and KKp055) and two invasive PDAC cases (KKp011 and KKp060) and performed single-cell RNA sequencing (Fig. 5b). HG-PanIN organoids remained morphologically stable across passages (Extended Data Fig. 8) and harbored KRAS mutations without TP53 or SMAD4 alterations, whereas invasive PDAC organoids carried additional driver alterations, including TP53 and CDKN2A, with SMAD4 alterations observed in other invasive PDAC organoid lines in the same panel (Extended Data Fig. 9d). These profiles support their precursor and invasive characteristics.

In xenograft experiments, invasive PDAC organoids formed progressively growing adenocarcinomas, whereas HG-PanIN organoids failed to establish tumors after 84 days (Extended Data Fig. 9a–c), supporting their preinvasive phenotype.

Integrated single-cell analysis classified the organoid epithelial cells into classical-like, intermediate, and basal-like states (Fig. 5c). HG-PanIN organoids were predominantly classical-like (~95%), with minimal basal-like representation (~1%), whereas invasive PDAC organoids showed substantial basal-like representation (cell-weighted average ~52%; Fig. 5d). This transition was accompanied by a continuous leftward shift in the classical-minus-basal score distributions rather than discrete clustering (Fig. 5e and 5f).

We investigated whether the epithelial subtypes expressed distinct fibroblast-engaging ligands. In HG-PanIN organoids, basal-like cells were rare (KKp048: 0.2%; KKp055: 1.9%) but showed a distinct ligand profile, with intermediate cells displaying transitional expression (Fig. 5g). Across HG-PanIN and invasive PDAC organoids, basal-like cells were enriched for inflammatory chemokines (CXCL5, CXCL8, CXCL3, CCL2, CCL20), TGF-β family members, ECM modifiers, and Hedgehog pathway components, whereas classical-like cells showed WNT- and growth-factor-associated profiles (Fig. 5h). These subtype-specific ligand profiles indicate that classical- and basal-like epithelial cells engage distinct stromal programs.

Precursor organoid medium activates stellate cells

Next, we tested whether these ligand programs translated into stromal activation in pancreatic stellate cells (PSCs) using conditioned medium (CM), as PSCs are an established source and experimental model of activated fibroblasts in pancreatic cancer.[35–36]

Primary human PSCs exposed to CM from the HG-PanIN organoid KKp048 acquired a transcriptional profile distinct from that of controls (PC1, 94% variance; Fig. 6a and 6b), with 8,738 differentially expressed genes. Both myCAF-like and iCAF-like signature scores increased significantly after KKp048 CM exposure (two-sample t-test on per-sample module scores, $P = 0.0028$ and $P = 0.0014$; $n = 3$ independent culture replicates per group; Fig. 6c), including upregulation of inflammatory mediators such as IL6, CXCL8, CXCL1, CXCL2, IL1B, and CCL2, and myCAF-associated genes, including ACTA2, CSRP2, CNN1, and TPM1 (Fig. 6d). Because this experiment used KKp048 as the representative HG-PanIN, these data provide proof-of-principle functional support rather than evidence that all HG-PanIN organoids share the same secretome activity. KKp048 CM shifted PSCs toward a hybrid myCAF-like/iCAF-like activation state.

We extended the CM experiments to invasive PDAC organoids (KKp011 and KKp060) along with KKp048 (Fig. 6e). All CM-treated PSCs were separated from controls, and PC2 distinguished the CM sources (Fig. 6f). MyCAF-like scores increased across all CM-treated groups, whereas iCAF-like scores increased from the control group to KKp048 and invasive PDAC organoids (Fig. 6g). This stage-associated iCAF-like intensification reflected inflammatory mediators including IL6, CXCL1, CXCL8, CCL2, CCL5, LIF, and CXCL2, whereas myCAF markers were more uniformly increased (Fig. 6h). Differentially expressed genes and pathway analyses revealed that HG-PanIN and invasive PDAC CM were associated with distinct PSC programs. The extended four-condition experiment (Fig. 6e–h) was performed independently of the initial KKp048-only experiment (Fig. 6a–d), both of which demonstrated significant myCAF-like and iCAF-like activation by KKp048 CM, with absolute P-values differing because of distinct independent culture replicates and reference distributions.

Finally, 2D cancer cell line CM experiments showed that basal-type BxPC-3 cells preferentially increased an iCAF-like program compared with classical-type PANC-1 cells (Extended Data Fig. 10). These functional data are consistent with a model in which HG-PanIN-derived epithelium engages stromal

programs, whereas invasive and basal-like epithelial states intensify inflammatory CAF-like polarization during progression (Fig. 6i).

Discussion

Our integrated spatial transcriptomic and functional analyses showed that true isolated HG-PanINs were accompanied by stromal expansion, CAF-like program diversification, epithelial basal-like drift, and candidate epithelial–stromal signaling before invasion. Specifically, the stromal area expanded 3.4-fold at the HG-PanIN stage, basal-like epithelial features appeared gradually, and HG-PanIN organoid CM shifted PSCs toward CAF-like transcriptional states, with stronger iCAF-like polarization observed after exposure to invasive PDAC organoid CM. These data place HG-PanIN as a preinvasive epithelial–stromal ecosystem rather than simply a histologic transition state.

The stromal increase from 17% in the normal pancreas to 58% in HG-PanINs was one of the largest quantitative changes observed in this study. This finding extends the traditional view that dense desmoplasia is primarily a feature of established invasive cancer. As all PanIN samples lacked concurrent invasive carcinoma anywhere in the resected pancreas, these microenvironmental changes were unlikely to reflect paracrine effects from nearby PDAC. This distinction is central: our HG-PanIN samples were not PanIN lesions incidentally found in a pancreas already harboring invasive carcinoma, but truly isolated HG-PanIN lesions representing precursor pathology. Therefore, the data provide a biological framework for our clinical observation that focal pancreatic atrophy and MPD dilatation can precede overt tumor detection by years.[10] These findings complement recent profiling of incidental PanIN in organ-donor and grossly normal pancreata[28–29] and of PanIN within PDAC resection specimens,[30] which established that neoplastic and stromal programs begin early. The present cohort differs in examining clinically diagnosed HG-PanIN as the most advanced lesion in pancreata proven free of invasive carcinoma, and in pairing these lesions with precursor-derived organoids for functional testing.

CAF-like program analysis suggests functional diversification within the precursor stroma. Standard Visium indicated program-associated spatial patterns, but Visium HD did not reproduce fixed compartmentalization, suggesting that spot-level inference was limited in this setting. Nonetheless, co-localization of αSMA with TAGLN, MYL9, and CNN1 in the periductal stroma supports a CAF-like rather than vascular identity, while IL-6/iCAF-associated programs indicate inflammatory remodeling before invasion.

Our data indicate that epithelial subtype heterogeneity begins before invasion but remains subtle at the PanIN stage. Basal-like fractions by categorical Visium HD assignment were similar in LG-PanIN and HG-PanIN (7.4% and 7.6%, respectively), whereas continuous Subtype Index analysis and IHC for p63/KRT5 revealed a coordinated shift toward basal-like features in HG-PanIN. The predominance of intermediate states (~67%–69%) argues against a binary subtype switch. Instead, PanIN progression appears to involve gradual redistribution along a classical-to-basal continuum, with marked basal-like expansion

becoming clearer in invasive PDAC organoids. The discrepancy between the modest distributional shift in transcriptomic Subtype Index (Kolmogorov–Smirnov D = 0.028) and the clear increase in $p63^+/KRT5^+$ cells by IHC likely reflects the emergence of rare, focally distributed basal-like cells whose transcriptional contribution is diluted in spot-level averaged spatial data, while remaining detectable as discrete protein-positive nuclei by IHC.

Whether the p63+/KRT5+ cells in HG-PanINs represent bona fide precursor cells or nascent invasive cells is an important consideration. A precursor interpretation is supported by several observations: all HG-PanIN samples lacked invasive carcinoma, HG-PanIN organoids did not engraft in xenografts, and genomic profiling showed KRAS mutations without TP53 or SMAD4 alterations typical of invasive diseases. However, the molecular drivers of TP63 induction and basal-like drift in PanIN remain unresolved and may involve hypoxia-adapted programs, MYC activation, EMT, and GATA6 downregulation.[37–38]

Functional data link the epithelial state to stromal remodeling. Spatial L-R profiling suggested that basal-like epithelium engages growth-factor pathways with myCAF-like programs and inflammatory pathways with iCAF-like programs, whereas CM assays showed that HG-PanIN-derived soluble factors can shift PSCs toward both states. Invasive PDAC organoid CM was associated with stronger iCAF-like polarization, indicating that the inflammatory arm of epithelial–stromal crosstalk may intensify with disease progression. Although our experiments primarily support epithelial-to-stromal signaling, reciprocal stromal-to-epithelial interactions identified by L-R analysis point to bidirectional epithelial–stromal co-evolution that warrants direct functional perturbation experiments.

MyCAF induction observed with organoid CM should be interpreted with caution. Previous studies have shown that complete myCAF differentiation often requires direct cell-cell contact, whereas soluble factors preferentially induce iCAF programs.[14] Thus, the myCAF signature increase in our CM experiments may reflect partial contractile activation or residual Matrigel-derived factors, despite the Matrigel-only controls. Consistent iCAF intensification across invasive PDAC organoids and the basal-type cell line CM is more directly aligned with established IL-1/JAK-STAT-driven iCAF biology.[16]

Duct obstruction and pancreatitis-like remodeling may also contribute to periductal fibrosis in isolated HG-PanINs, particularly in cases of MPD stricture or upstream ductal dilation. However, the consistent enrichment of CAF-like programs and IL-6-positive periductal stroma in isolated HG-PanIN in the absence of invasive carcinoma supports the interpretation that precursor-associated epithelial–stromal remodeling is already present before histologically detectable invasion.

Establishing organoids from ERCP brush-derived HG-PanINs was a technical strength of this study. These lesions are microscopic and rarely available as fresh experimental material; however, the precursor identity of the resulting organoids was supported by the clinical context, serial passage stability, genomic features consistent with precursor disease, and the absence of xenograft growth. This

platform brings human precursor biology within the scope of functional studies that are otherwise difficult to access.

Our finding that Hedgehog is detectable but less dominant than HGF, PDGF, FGF2, and WNT may help contextualize prior Hedgehog inhibitor failures in PDAC.[39–41] Stromal activation appears multi-axis and subtype-dependent, implying that future interception strategies may need to account for CAF-like heterogeneity rather than relying on indiscriminate stromal suppression. The identified features, including iCAF-like enrichment, basal marker expression, and signaling pathway activation, may also inform the development of biomarkers in high-risk surveillance cohorts.[42–43]

This study had several limitations. The numbers of biological replicates were modest (standard Visium: normal n = 2, LG-PanIN n = 5, HG-PanIN n = 5; Visium HD: n = 2 per stage; HG-PanIN organoids: n = 2 lines; invasive PDAC organoids: n = 2 lines), and the extended CM experiment used KKp048 as the HG-PanIN representative. This sample size limitation reflects the rarity of clinically diagnosed isolated HG-PanIN rather than routine cohort availability. HG-PanIN is usually microscopic and non-mass-forming, often requires specialized endoscopic cytology for diagnosis, and is resected before invasion, mainly in early detection settings where carcinoma in situ is treated as stage 0 pancreatic cancer.[3–5] Therefore, establishing organoids from such lesions is technically and clinically difficult, and the HG-PanIN CM data should be viewed as proof-of-principle functional support that requires validation in additional precursor-derived organoid lines.

Some effects, including the basal-like distributional shift (Kolmogorov–Smirnov D = 0.028) and Visium HD subtype-stratified L-R fold changes, were statistically detectable but small or based on few biological replicates. In addition, the 10 standard Visium PanIN sections were derived from four unique patients, with two patients contributing sections to both the LG-PanIN and HG-PanIN groups; therefore, the across-stage trend statistics should be regarded as exploratory rather than as tests over independent patients, and confirmation in larger, patient-independent cohorts is needed.

Likewise, conditioned medium comparisons were based on only three culture replicates per group (compared by two-sample t-test). These functional results are therefore presented as proof of principle, supported by consistent effect direction and independent experimental replication, rather than by large-sample inference (see Methods). The cross-sectional design also precludes direct lineage or causal inference. Organoid-derived CM may contain Matrigel-related confounders, and extended passaging may alter secretome profiles, although driver mutation profiles remain stable. Finally, spatial profiling was not performed on invasive PDAC tissues in this study. Formal spot-level deconvolution was not applied to standard Visium data, and immune cell dynamics remain to be comprehensively characterized. Larger cohorts, additional HG-PanIN organoid lines, and mechanistic perturbation experiments are required to establish the generalizability and identify actionable mediators.

In summary, isolated human HG-PanIN lesions were accompanied by marked periductal stromal expansion, CAF-like fibroblast program enrichment, focal basal-like epithelial drift, and candidate

epithelial–stromal signaling before histologically detectable invasion. Together with functional conditioned medium experiments using HG-PanIN-derived organoids, these findings support the concept that HG-PanIN represents a biologically active, preinvasive epithelial–stromal ecosystem. Larger cohorts and perturbation studies are required to establish generalizability and identify the functional mediators of this precursor microenvironment.

Methods

Study Design and Patient Samples

This study was approved by the Institutional Review Board of Kobe University Hospital (approval number: B200341), and written informed consent was obtained from all participants in accordance with the Declaration of Helsinki. Surgically resected pancreatic tissues containing normal ducts, low-grade pancreatic intraepithelial neoplasia (LG-PanIN), and high-grade PanIN (HG-PanIN) were obtained from patients who underwent pancreatectomy at Kobe University Hospital between 2021 and 2025. For spatial transcriptomic analysis, we analyzed 12 sections using standard Visium (normal n = 2, LG-PanIN n = 5, HG-PanIN n = 5) and four additional cases using Visium HD (LG-PanIN n = 2, HG-PanIN n = 2). Visium HD analysis focused on LG-PanIN and HG-PanIN to enable direct comparison of epithelial–stromal interactions at the LG-to-HG transition at near-single-cell resolution, as standard Visium had already established a normal baseline. The Visium HD cohort consisted of four patients, independent of the standard Visium cohort (see Extended Data Fig. 1 for detailed sample allocation). The 10 standard Visium PanIN sections were obtained from four unique patients (Supplementary Table 8): two patients each contributed two LG-PanIN and two HG-PanIN sections (together 8 of the 10 PanIN sections), and the remaining two patients each contributed a single section (one LG-PanIN and one HG-PanIN, respectively), which were co-mounted within a single CytAssist capture area (see Spatial Transcriptomics). Because the standard Visium biological replicates were partially nested within patients, with the same patients represented in both stage groups, across-stage trend comparisons were interpreted as exploratory rather than as inferences over fully independent patients (see Statistical Analysis). For immunohistochemical validation, we evaluated 177–186 individual duct lesions per marker (the exact count varying by antibody): normal (n = 58–77 ducts), LG-PanIN (n = 31–40 ducts), and HG-PanIN (n = 68–89 ducts). For stromal fibrosis quantification using Masson’s trichrome staining, 65 cases were analyzed (normal, n = 23; LG-PanIN, n = 15; HG-PanIN, n = 27). To investigate the intrinsic tumor microenvironment of precursor lesions without contamination from adjacent invasive carcinomas, all PanIN samples (both LG-PanIN and HG-PanIN) were selected from resection specimens in which no concurrent invasive PDAC was identified anywhere in the entire resected pancreas. Specifically, “isolated” PanIN was defined as lesions from patients whose final pathological diagnosis was either LG-PanIN or HG-PanIN (carcinoma in situ) without any invasive carcinoma component in the resected specimen, thereby minimizing field effects from coexisting invasive tumors and enriching for precursor pathology. The complete absence of invasive carcinoma was confirmed by exhaustive histological evaluation of the entire resected specimen, including serial sections and surgical margins. Pathological

diagnoses were independently confirmed by two board-certified pathologists (T.I. and T.K.) according to the WHO 2019 classification criteria.

Histological Analysis

Tissue sections (5 μm) were stained with hematoxylin and eosin (H&E) and independently evaluated by two experienced pathologists using a multi-headed microscope for consensus diagnosis. PanINs were classified according to the WHO 2019 classification: LG-PanIN (PanIN-1A, -1B, and -2), characterized by flat-to-papillary epithelial lesions with basally oriented nuclei and minimal cytological atypia; and HG-PanIN (PanIN-3), defined by papillary or micropapillary architecture with significant nuclear atypia, loss of polarity, and frequent mitotic figures. Masson's trichrome staining was performed on consecutive FFPE sections using the Masson's Trichrome Stain Kit (Sigma-Aldrich). Briefly, the sections were deparaffinized, rehydrated, and stained sequentially with Weigert's iron hematoxylin, Biebrich scarlet–acid fuchsin, phosphomolybdic–phosphotungstic acid, and aniline blue. Whole-slide digital images were acquired using a NanoZoomer slide scanner (Hamamatsu Photonics) at 20× magnification and analyzed using QuPath[44] and ImageJ. Periductal regions were defined as areas within 200 μm of the ductal epithelial basement membrane. Stromal and collagen areas were quantified using the ImageJ software (NIH) with fixed threshold values for color segmentation. Unstained (white) areas corresponding to adipose tissue and tissue voids were excluded from quantification. For each case, 3–5 representative regions were analyzed by an investigator blinded to the clinical information.

Immunohistochemistry

Immunohistochemistry was performed on FFPE tissue sections (5 μm). Primary antibodies used were: anti-αSMA (clone 1A4, Dako, 1:200), anti-IL-6 (clone EPR21711, Abcam, 1:100), anti-p63 (clone 4A4, Roche Ventana, prediluted), anti-KRT5 (clone EP1601Y, Abcam, 1:200), and anti-GATA6 (clone D61E4, Cell Signaling, 1:400). Heat-induced antigen retrieval was performed using CC1 solution at 95 °C for 64 min, followed by primary antibody incubation for 32 min at 37 °C. Signal detection was performed using an OptiView DAB IHC Detection Kit. Immunohistochemical expression was evaluated using a 4-point scoring system (0–3) based on staining intensity and the percentage of positive cells. For αSMA and IL-6, the periductal stromal region (within 200 μm of the ductal epithelium) was assessed. For p63 and KRT5, positive cell counts per high-power field (HPF, 40× magnification) were recorded. For GATA6, GATA6-negative cells per HPF were counted, and a GATA6 Expression Score was calculated as max(GATA6-negative cells across all samples) minus the number of GATA6-negative cells per HPF. This inverted scoring was used so that higher numerical values consistently corresponded to retained classical-marker expression, paralleling the directional convention used for the basal markers p63 and KRT5 (where higher counts indicate stronger basal-like phenotype), thereby enabling comparable directional interpretation across all subtype markers (Fig. 3g). αSMA and IL-6 were scored semiquantitatively on a 4-point scale: 0 (negative), 1 (weak), 2 (moderate), and 3 (strong). Sample sizes:

αSMA: Normal (n = 77), LG-PanIN (n = 31), HG-PanIN (n = 74); IL-6: Normal (n = 58), LG-PanIN (n = 35), HG-PanIN (n = 89); TP63: Normal (n = 76), LG-PanIN (n = 36), HG-PanIN (n = 74); KRT5: Normal (n = 68), LG-PanIN (n = 37), HG-PanIN (n = 78); GATA6: Normal (n = 69), LG-PanIN (n = 40), HG-PanIN (n = 68). Scoring was performed independently by two pathologists who were blinded to the clinical information of the patients.

Spatial Transcriptomics

Spatial gene expression profiling was performed using standard Visium and Visium HD platforms (10x Genomics). For the standard Visium, FFPE tissue sections (5 μm) were placed on Visium Spatial Gene Expression slides. After deparaffinization and H&E staining, the tissue sections were imaged using a Keyence BZ-X810 microscope. Regions of interest (6 mm × 6 mm) encompassing ductal epithelium, PanIN lesions, and surrounding stroma were selected by a pathologist (T.I.) blinded to prior molecular data; ROIs were chosen to include the PanIN lesion with at least 200 μm of surrounding periductal stroma on all sides. Tissue permeabilization was optimized (18 min at 37 °C), and spatial gene expression libraries were prepared following the manufacturer's protocol for FFPE. For four of the 12 standard Visium FFPE samples, two sections were co-mounted within a single CytAssist capture area as a sample-multiplexing strategy (Slide V44A10-307 region A1: PAT01 section 9 + PAT02 section 11; region D1: PAT02 section 17 + PAT04 section 3), and the two co-mounted sections in each capture area were post hoc segmented based on histology in the Loupe Browser to export per-section feature-barcode matrices. For Visium HD, FFPE tissue sections (5 μm) were mounted on Visium HD slides and processed for spatial gene expression, following the Visium HD FFPE Tissue Prep Handbook (CG000684 RevA), including hematoxylin and eosin staining, decrosslinking, probe hybridization, and library construction. Standard Visium libraries were sequenced using an MGI DNBSEQ-G400 (DNBSEQ-G400RS PE100 flow cell). Visium HD libraries were sequenced on an MGI DNBSEQ-G400 (LG-PanIN samples) or an Illumina NextSeq 2000 (HG-PanIN samples), targeting at least 25,000 read pairs per spot for standard Visium and approximately 400 reads per 8 μm bin or higher for Visium HD. Standard Visium raw data were processed using Space Ranger (v2.0.1) and Visium HD data were processed using Space Ranger (v3.0.1), both aligned to GRCh38-2020-A. For standard Visium, quality control excluded spots with <500 UMI counts, <200 detected genes, or >20% mitochondrial reads, retaining 39,693 high-quality spots across 12 samples (mean: 3,308 spots/sample). For Visium HD, quality filtering retained 8-μm bins with 50–2,000 unique genes and <25% mitochondrial content, yielding 542,481 high-quality bins across 4 samples (86.9% retention rate). Data were analyzed using Seurat (v4.3.0–5.0.1)[45–46] in R (v4.2.0–4.4.1). Raw counts were normalized using SCTransform[47] with variance stabilization. Principal component analysis was performed on the top 3,000 variable features, and the first 30 principal components were used for UMAP visualization and clustering. Batch effects were corrected using Harmony (v1.0).[48]

Cell type annotation and module-score enrichment

Cell type annotation was performed using marker gene expression. Ductal and epithelial cells were identified using KRT7, KRT8, KRT18, KRT19, EPCAM, and CDH1. Acinar cells were labeled with PRSS1, CPA1, CTRB1, AMY2A, and REG1A. Stromal cells were identified using COL1A1, COL1A2, COL3A1, VIM, DCN, LUM, ACTA2, and FAP. The module scores were calculated using the AddModuleScore function in Seurat. The complete gene lists for all cell types and subtype signatures are provided in Supplementary Table 7. Standard Visium spots (55 μm) typically encompass multiple cells; accordingly, module-score-based "iCAF-like," "myCAF-like," and epithelial-subtype designations applied to standard Visium denote the dominant spot-level transcriptional program of a multicellular spot rather than single-cell identities, and no reference-based spot deconvolution (e.g., RCTD or SPOTlight) was applied. Visium HD bins (8 μm) approach single-cell resolution, but small residual mixing cannot be fully excluded. Throughout, "spots" or "bins" denote spatial measurement units, not individual cells.

CAF-like program classification and spatial analysis

CAF-like fibroblast programs were classified based on established signatures: myCAF-like (key markers: ACTA2, TAGLN, MYL9, TPM1, TPM2, MYH11, CNN1, LMOD1, and PDGFRB) and iCAF-like (key markers: IL6, CXCL12, CXCL1, CXCL2, CCL2, LIF, IL1A, PDGFRA, and HAS1); the complete gene signatures used for module score calculation are provided in Supplementary Table 7. Module scores were calculated for each program, and the cells were classified based on score thresholds. For spatial distance analysis, the Euclidean distance from each CAF-like spot to the nearest epithelial spot was calculated, and periductal CAF-like fibroblast analysis was restricted to spots within 300 μm of the nearest epithelial spot.

Epithelial Molecular Subtype Analysis

The epithelial molecular subtypes were classified using established gene signatures. Classical key markers include GATA6, GATA4, HNF4A, HNF1A, HNF1B, CDH1, CLDN4, MUC1, MUC5AC, TFF1, TFF2, and AGR2. The key basal markers were TP63, KRT5, KRT6A, KRT14, KRT17, S100A2, S100A8, S100A9, LAMC2, and MMP7. The complete classical and basal-like signatures used for the module score calculation, including additional EMT-associated genes in the basal-like signature (e.g., VIM, ZEB1, SNAI1/2, and TWIST1), are provided in Supplementary Table 7. Spots were classified as classical-like (score difference >0.1), intermediate (−0.1 to 0.1), or basal-like (<−0.1) based on the difference between classical and basal module scores. Because the basal-like signature includes mesenchymal/EMT genes (e.g., VIM), which are also expressed by stromal cells, subtype scores were interpreted within the histopathology-annotated epithelial compartments to limit stromal spillover in mixed multicellular spots.

Ligand-Receptor Interaction Analysis

Ligand-receptor interactions were analyzed using two complementary approaches. First, 44 curated L-R pairs spanning 17 signaling pathways (BMP, Chemokine, FGF, Hedgehog, HGF, IGF, IL1, IL6, Matricellular,

PDGF, TGFβ, VEGF, WNT, and additional growth factor / ECM-integrin / morphogen pathways) were drawn from CellPhoneDB (v4.0)[49] and analyzed with a geometric mean-based scoring approach. Analysis focused on candidate epithelial-to-stromal and stromal-to-epithelial interactions, with key pathways including HGF, PDGF, FGF, TGF-β, Hedgehog, and WNT. Interaction scores were calculated as the geometric mean of ligand expression in the source compartment cells and receptor expression in the target compartment cells. For Visium HD, L-R scores were further stratified by epithelial subtype (Classical, Basal-like) and CAF-like fibroblast programs (iCAF-like and myCAF-like) to enable subtype-specific interaction profiling. Interactions were retained if both ligand and receptor were expressed in ≥5% of respective cell populations. Statistical significance was assessed using the Wilcoxon rank-sum test with Benjamini-Hochberg correction. Second, CellChat (v2.2)[50] was applied to Visium HD data using subtype-level identities (Classical, Basal-like, Intermediate, iCAF-like, myCAF-like, Mixed) for unbiased, database-wide statistical inference of cell-cell communication. The Secreted Signaling subset of CellChatDB.human was used, and the communication probabilities were computed using the truncated mean method.

Patient-Derived Organoid Establishment and Culture

Patient-derived organoids were established from HG-PanINs and invasive PDAC specimens. For HG-PanIN, samples were obtained by ERCP brush cytology. Organoids were established from the limited cellular material obtained by pancreatic duct brushing. Invasive PDAC organoids were established from EUS-guided tissue acquisition or surgical specimens. Tissue was digested with collagenase II (2 mg/mL) and dispase II (1.25 mg/mL) at 37 °C for 30–60 min, and cell clusters were filtered through a 70 μm strainer, washed with Advanced DMEM/F12, and embedded in Matrigel (Corning). Organoids were cultured in Advanced DMEM/F12 supplemented with B27, N2, HEPES, GlutaMAX, N-acetylcysteine (1.25 mM), EGF (50 ng/mL), Noggin (100 ng/mL), FGF10 (100 ng/mL), gastrin (10 nM), A83-01 (500 nM), and Y-27632 (10 μM). The medium was changed every 3–4 days and organoids were passaged every 7–10 days.

Whole-Genome Sequencing of Patient-Derived Organoids

Genomic DNA was extracted from the organoids using a DNeasy Blood & Tissue Kit (Qiagen). WGS libraries were prepared using the TruSeq Nano DNA Library Prep Kit and sequenced on an Illumina NovaSeq 6000, targeting 40× coverage for tumor organoids and 30× coverage for normal samples. Genomon pipeline (https://github.com/Genomon-Project) version 2.6.3 was used for sequence alignment and mutation calling.[51] Briefly, the reads were aligned to the hs37d5 human reference genome. Putative somatic mutations were selected based on: (i) Fisher's exact $P < 0.1$; (ii) EBCall[52] $P < 1 \times 10^{-4}$; (iii) ≥ 4 variant reads in tumor; (iv) tumor variant allele frequency (VAF) ≥ 0.1; and (v) germline VAF < 0.05. These were filtered by removing (i) synonymous single nucleotide variants, (ii) variants in repetitive regions, and (iii) variants in unidirectional reads. Candidate mutations were then filtered further,

unless found ≥ 10 times in COSMIC version 70, by excluding known germline variants from NCBI dbSNP build 131, or those with frequency ≥ 0.0001 in 1000 Genomes (released in October 2014), National Heart, Lung, and Blood Institute Exome Sequencing 6500, the Human Genome Variation Database version 2.00, Tohoku Medical Megabank Organization 3.5KJPNv2, or gnomAD genome collection version 2.1.1. Somatic mutations in the established pancreatic cancer driver genes (Extended Data Fig. 9d)[53] were considered driver mutations.

To obtain allele-specific copy number (CN) data, we adopted ascatNgs version 4.2.1, provided in "dockstore-cgpwgs" version 2.1.1, executed using the Apptainer container platform version 1.2.4, with the reference files obtained from the ASCAT GitHub repository (https://github.com/VanLoo-lab/ascat). Focal segments were defined as segments ≤ 0.5× arm length. Focal CN segments were annotated using RefGene, and those affecting known driver genes (Extended Data Fig. 9d)[22,53] were selected, including CN gains or amplifications affecting oncogenes and deletions affecting tumor suppressor genes.

Xenograft Experiments

Female SCID mice (6–8 weeks old) were used for xenograft experiments using IACUC-approved protocols at Kobe University (approval no. P220306). Organoids were dissociated and approximately 1×10^6 viable cells in Matrigel:PBS (1:1) were subcutaneously injected into the right flank (n = 5 mice per organoid line). Tumor formation was monitored twice weekly for up to 84 days using calipers. Tumor growth was assessed using the longest diameter.

Single-Cell RNA Sequencing of Patient-Derived Organoids

Organoids were dissociated into single cells using TrypLE Express supplemented with DNase I. Dead cells were removed using the Dead Cell Removal Kit (Miltenyi Biotec), and cell viability (>85%) and cell concentration were confirmed using a Countess II automated cell counter (Thermo Fisher Scientific) before fixation. Single-cell suspensions from each of the four organoid lines were fixed using the 10x Genomics Fixation of Cells protocol (formaldehyde-based) and hybridized to the Chromium Human Transcriptome Probe Set v1.0.1 carrying sample-specific 4-plex probe barcodes (KKp060 = BC001, KKp048 = BC002, KKp011 = BC003, and KKp055 = BC004). Probe-hybridized cells from all four samples were pooled and loaded onto a single Chromium Next GEM Chip using the Chromium Next GEM Single Cell Fixed RNA Sample Kit (10x Genomics). The pooled library was sequenced using an Illumina NovaSeq X (paired-end). Raw data were processed using Cell Ranger v8.0.0, with the cellranger multi workflow against the GRCh38-2020-A reference and Human Transcriptome Probe Set v1.0.1 reference, generating per-sample demultiplexed feature-barcode matrices. Quality filtering retained cells with 500–6,000 genes, 1,000–50,000 UMI, and <20% mitochondrial content. Doublets were removed using DoubletFinder. After filtering, 63,517 high-quality cells were retained (30,557 from the HG-PanIN organoids and 32,960 from the PDAC organoids). The data were normalized using SCTransform with cell

cycle regression. Batch effects were corrected using the Harmony (v1.0). Clustering was performed using the Louvain algorithm (resolution = 0.8) and cell types were annotated using canonical markers and SingleR with the Human Primary Cell Atlas reference. Epithelial subtype classification used the same gene signatures as spatial transcriptomics.

Pancreatic Stellate Cell Isolation and Culture

Primary human pancreatic stellate cells (PSCs) were commercially obtained (ScienCell Research Laboratories). PSCs were characterized by αSMA expression and stellate morphology, and maintained in DMEM supplemented with 10% FBS. Cells at passages 3–8 were used for the experiments. PSCs were verified to be mycoplasma negative by PCR-based testing prior to use.

Conditioned Medium Experiments

For conditioned medium (CM) experiments, cancer cell lines (BxPC-3 [ATCC CRL-1687] representing the basal-like subtype and PANC-1 [ATCC CRL-1469] representing the classical-like subtype) or patient-derived organoids were used as CM sources. BxPC-3 and PANC-1 cell lines were authenticated by short tandem repeat (STR) profiling, and both lines were confirmed to be mycoplasma-negative by PCR-based testing prior to use. The cancer cell lines were cultured in RPMI-1640 supplemented with 10% FBS until 80% confluence, washed twice with PBS, and incubated in serum-free DMEM for 48 h to generate the CM. For organoid-derived CM, patient-derived pancreatic cancer organoids KKp048 (HG-PanIN-derived), KKp011 (invasive PDAC-derived), and KKp060 (invasive PDAC-derived) were utilized. Organoids at passages 8–12 were cultured in Matrigel domes with complete organoid medium for 48 h. The CM was then collected, and Matrigel-only control CM was prepared by incubating cell-free Matrigel domes under identical conditions to account for Matrigel-derived growth factors. All CM was centrifuged (2,000 × g, 10 min) to remove debris and filtered through a 0.22 μm filter. The initial organoid CM experiment compared the control with the KKp048 CM (two groups × three independent culture replicates = six samples). The extended experiment included all four conditions: Control, KKp048 (HG-PanIN), KKp011 (invasive PDAC), and KKp060 (invasive PDAC) (four groups × three independent culture replicates = 12 samples). In both organoid CM experiments, n = 3 denotes independent culture replicates and KKp048 is the HG-PanIN representative of the extended experiment. PSCs (2×10^4 cells/well) were seeded in 6-well plates in DMEM with 10% FBS, allowed to adhere overnight, then switched to a 1:1 mixture of CM and fresh PSC medium for 24 h. Control PSCs were cultured in a 1:1 mixture of unconditioned base medium (or Matrigel-only CM for organoid experiments) and fresh PSC medium. PSCs were harvested for RNA extraction.

Bulk RNA Sequencing

Total RNA was extracted using TRIzol reagent after DNase I treatment. The RNA quality was assessed using an Agilent 2100 Bioanalyzer (RIN > 8.0). Libraries were prepared using the TruSeq Stranded mRNA LT Sample Prep Kit following poly (a) selection. Sequencing was performed on an Illumina NovaSeq X (paired-end 101 bp). Raw reads were quality-trimmed with Trimmomatic (v0.38), aligned to GRCh38 (NCBI annotation 109.20200522) using HISAT2 (v2.1.0), and transcript and gene abundances were estimated using StringTie (v2.1.3b). Gene-level read counts from StringTie were rounded to integers for downstream differential expression analyses. Gene expression quantification and differential expression analysis were performed using DESeq2 (v1.36)[54] in R (v4.4.2). Raw read counts were filtered to retain genes with ≥ 10 counts in ≥ 3 samples. Variance-stabilizing transformation (VST) was applied for visualization and module score calculations. Genes with adjusted $P < 0.05$ and $|\log2 \text{ fold change}| > 0.5$ were considered differentially expressed. For the 2D cell line experiments, three pairwise comparisons were performed (control vs. PANC-1, Control vs. BxPC-3, and PANC-1 vs. BxPC-3). For the extended organoid experiment, five comparisons were made (control vs. KKp048, Control vs. KKp011, Control vs. KKp060, KKp048 vs. KKp011, and KKp048 vs. KKp060). CAF-like program module scores were calculated as the mean VST-normalized expression of the respective marker genes across samples.

Statistical Analysis

Statistical analyses were performed using R (v4.2.0–4.4.1) and GraphPad Prism (v9.5.0). For categorical data, the chi-squared test or Fisher's exact test was used. For continuous variables, the unpaired two-tailed Student's t-test or Wilcoxon rank-sum test was applied based on the data distribution assessed using the Shapiro-Wilk test. For multiple group comparisons, a one-way analysis of variance (ANOVA) with Tukey's post hoc test or the Kruskal–Wallis test with Dunn's post hoc test was used. Multiple testing corrections were performed using the Benjamini-Hochberg method. Correlation analyses were performed using Spearman's rank correlation coefficients. Effect sizes were calculated using Cliff's δ, where appropriate. All tests were two-sided, with $P < 0.05$ considered statistically significant. For spatial transcriptomic data with small sample sizes at the biological replicate level (n = 2–5 per group), cell-level statistical tests should be interpreted with caution as they reflect cell-level differences rather than sample-level reproducibility. Because several patients contributed more than one section to the standard-Visium cohort, and two patients contributed sections to both the LG-PanIN and HG-PanIN groups (Supplementary Table 8), the across-stage trend tests (Jonckheere–Terpstra) were reported as exploratory rather than as inferences over fully independent patients. For the conditioned medium experiments, CAF-like program activation was quantified per sample as the mean VST-normalized expression of the marker genes of the program (Supplementary Table 7), and conditions were compared using an unpaired two-sample Student's t-test (n = 3 independent culture replicates per group). Given the small number of replicates, these comparisons relied on parametric assumptions and were interpreted as proof of principle, supported by the consistent direction of program score changes across replicates (shown as individual data points in Fig. 6c, 6g, and Extended Data Fig. 10d) and by independent reproduction of the KKp048 CM effect in two separate experiments (Fig. 6a–d and Fig. 6e–h). Data visualization was performed using ggplot2, ComplexHeatmap, and R software packages.

Declarations

Data availability

Further information and requests for resources and reagents should be directed to and fulfilled by the corresponding author, A. Masuda (atmasuda@med.kobe-u.ac.jp). The spatial transcriptomic and single-cell RNA sequencing data reported in this article have been deposited in the NCBI Gene Expression Omnibus as SuperSeries GSE332553, which comprises three SubSeries: Visium FFPE spatial transcriptomics of 12 pancreatic specimens (GSE331451), Visium HD spatial transcriptomics of four pancreatic specimens (GSE331450), and 10x Chromium single-cell RNA sequencing of four patient-derived pancreatic organoid lines (GSE331447). During peer review, editors will receive the private reviewer access details for distribution to reviewers, and the data will be made publicly available upon manuscript acceptance. The patient-derived organoid lines generated in this study are available from the corresponding author upon reasonable request and are subject to a material transfer agreement.

Code availability

The analysis scripts used to generate the reported figures and tables (spatial transcriptomic processing, module-score and subtype classification, ligand–receptor scoring and differential expression analyses) are maintained in a dedicated GitHub repository (https://github.com/masahiro-tsujimae/kobe-hgpanin-spatial). The repository will be made publicly available and archived with a citable DOI through Zenodo upon manuscript acceptance. Requests for access during peer review should be directed to the corresponding author (atmasuda@med.kobe-u.ac.jp).

Competing interests

A. Machinaga is an employee of Eisai Co., Ltd. K. Fujita and T. Imai are affiliated with the Advanced Therapeutic Target Discovery Department of Gastroenterology, Kobe University Graduate School of Medicine, an endowed chair funded by Eisai Co., Ltd. The remaining authors declare no potential conflicts of interest.

Funding

This work was supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI (grant numbers 24K18916, 22K20920, 24K11070, and 19K08444) and Japan Science and Technology Agency (JST) Moonshot R&D (grant no. JPMJMS2022).

Author contributions

M. Tsujimae: Conceptualization, data curation, formal analysis, investigation, methodology, visualization, writing–original draft, writing–review and editing. A. Masuda: Conceptualization, supervision, writing–original draft, writing–review and editing. T. Ko: Investigation, methodology. M. Miki: Investigation, methodology. K. Okamoto, K. Ogawa, T. Kageyama, T. Tokunaga, Y. Kawase, K. Kanamaru, K. Yamamoto, A. Shirohata, Y. Yokotani, Y. Oka, N. Juri, Y. Irie, Y. Harada, K. Nagao, S. Kohashi, N. Inomata, H. Uemura, S. Masuda, M. Gonda, S. Abe, A. Sakai, T. Kobayashi: Resources, data curation. K. Kamizaki: Investigation, methodology. S. Horie: Formal analysis, investigation. Y. Kogure: Formal analysis, investigation. K. Kataoka: Formal analysis, supervision. K. Fujita: Investigation, methodology. A. Machinaga: Resources, methodology. T. Imai: Resources, supervision, methodology. Y. Yanagimoto: Resources, data curation. Y. Nanno: Resources, data curation. K. Sofue: Data curation, validation. T. Kodama: Data curation, validation, investigation. N. Uza: Resources, supervision. T. Itoh: Data curation, validation, investigation, methodology. Y. Minami: Supervision, methodology. Y. Kodama: Conceptualization, resources, supervision, funding acquisition, writing–review and editing.

Acknowledgements

The authors thank S. Wakai and K. Sakou for their excellent technical assistance with the spatial transcriptomics and organoid experiments. We are grateful to the members of the Division of Gastroenterology, Department of Internal Medicine, Kobe University Graduate School of Medicine, for valuable discussions and clinical support and to the administrative staff for their assistance with regulatory and logistical coordination. We sincerely thank all patients who consented to participate in this study; their generous contribution to tissue specimens made this research possible. We thank Editage (www.editage.jp) for English language editing.

References

1. Siegel RL, Miller KD, Wagle NS, Jemal A (2023) Cancer statistics, 2023. CA Cancer J Clin 73:17–48
2. Rahib L et al (2014) Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States. Cancer Res 74:2913–2921
3. Kanno A et al (2018) Multicenter study of early pancreatic cancer in Japan. Pancreatology 18:61–67
4. Iiboshi T et al (2012) Value of cytodiagnosis using endoscopic nasopancreatic drainage for early diagnosis of pancreatic cancer: establishing a new method for the early detection of pancreatic carcinoma in situ. Pancreas 41:523–529
5. Hanada K et al (2019) Roles of ERCP in the early diagnosis of pancreatic cancer. Diagnostics 9:30
6. Hruban RH, Maitra A, Goggins M (2008) Update on pancreatic intraepithelial neoplasia. Int J Clin Exp Pathol 1:306–316
7. Wood LD, Canto MI, Jaffee EM, Simeone DM (2022) Pancreatic cancer: pathogenesis, screening, diagnosis, and treatment. Gastroenterology 163:386–402e1

8. Makohon-Moore AP et al (2018) Precancerous neoplastic cells can move through the pancreatic ductal system. Nature 561:201–205
9. Makohon-Moore A, Iacobuzio-Donahue CA (2016) Pancreatic cancer biology and genetics from an evolutionary perspective. Nat Rev Cancer 16:553–565
10. Gonda M et al (2024) Temporal progression of pancreatic cancer computed tomography findings until diagnosis: a large-scale multicenter study. United Eur Gastroenterol J 12:761–771
11. Neesse A, Algül H, Tuveson DA, Gress TM (2015) Stromal biology and therapy in pancreatic cancer: a changing paradigm. Gut 64:1476–1484
12. Feig C et al (2012) The pancreas cancer microenvironment. Clin Cancer Res 18:4266–4276
13. Kalluri R (2016) The biology and function of fibroblasts in cancer. Nat Rev Cancer 16:582–598
14. Öhlund D et al (2017) Distinct populations of inflammatory fibroblasts and myofibroblasts in pancreatic cancer. J Exp Med 214:579–596
15. Elyada E et al (2019) Cross-species single-cell analysis of pancreatic ductal adenocarcinoma reveals antigen-presenting cancer-associated fibroblasts. Cancer Discov 9:1102–1123
16. Biffi G et al (2019) IL1-induced JAK/STAT signaling is antagonized by TGFβ to shape CAF heterogeneity in pancreatic ductal adenocarcinoma. Cancer Discov 9:282–301
17. Helms EJ et al (2022) Mesenchymal lineage heterogeneity underlies nonredundant functions of pancreatic cancer-associated fibroblasts. Cancer Discov 12:484–501
18. Dominguez CX et al (2020) Single-cell RNA sequencing reveals stromal evolution into LRRC15 + myofibroblasts as a determinant of patient response to cancer immunotherapy. Cancer Discov 10:232–253
19. Collisson EA et al (2011) Subtypes of pancreatic ductal adenocarcinoma and their differing responses to therapy. Nat Med 17:500–503
20. Moffitt RA et al (2015) Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma. Nat Genet 47:1168–1178
21. Bailey P et al (2016) Genomic analyses identify molecular subtypes of pancreatic cancer. Nature 531:47–52
22. Cancer Genome Atlas Research Network (2017) Integrated genomic characterization of pancreatic ductal adenocarcinoma. Cancer Cell 32:185–203e13
23. Chan-Seng-Yue M et al (2020) Transcription phenotypes of pancreatic cancer are driven by genomic events during tumor evolution. Nat Genet 52:231–240
24. Ståhl PL et al (2016) Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. Science 353:78–82
25. Williams CG, Lee HJ, Asatsuma T, Vento-Tormo R, Haque A (2022) An introduction to spatial transcriptomics for biomedical research. Genome Med 14:68
26. Hwang WL et al (2022) Single-nucleus and spatial transcriptome profiling of pancreatic cancer identifies multicellular dynamics associated with neoadjuvant treatment. Nat Genet 54:1178–1191

27. Moncada R et al (2020) Integrating microarray-based spatial transcriptomics and single-cell RNA-seq reveals tissue architecture in pancreatic ductal adenocarcinomas. Nat Biotechnol 38:333–342
28. Carpenter ES et al (2023) Analysis of donor pancreata defines the transcriptomic signature and microenvironment of early neoplastic lesions. Cancer Discov 13:1324–1345
29. Braxton AM et al (2024) 3D genomic mapping reveals multifocality of human pancreatic precancers. Nature 629:679–687
30. Bell ATF et al (2024) PanIN and CAF transitions in pancreatic carcinogenesis revealed with spatial data integration. Cell Syst 15:753–769
31. Boj SF et al (2015) Organoid models of human and mouse ductal pancreatic cancer. Cell 160:324–338
32. Tiriac H et al (2018) Organoid profiling identifies common responders to chemotherapy in pancreatic cancer. Cancer Discov 8:1112–1129
33. Seino T et al (2018) Human pancreatic tumor organoids reveal loss of stem cell niche factor dependence during disease progression. Cell Stem Cell 22:454–467e6
34. Sato T et al (2009) Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. Nature 459:262–265
35. Bachem MG et al (1998) Identification, culture, and characterization of pancreatic stellate cells in rats and humans. Gastroenterology 115:421–432
36. Xu Z et al (2010) Role of pancreatic stellate cells in pancreatic cancer metastasis. Am J Pathol 177:2585–2596
37. Hayashi A et al (2020) A unifying paradigm for transcriptional heterogeneity and squamous features in pancreatic ductal adenocarcinoma. Nat Cancer 1:59–74
38. Landon-Brace N et al (2025) Analysis of an engineered organoid model of pancreatic cancer identifies hypoxia as a contributing factor in determining transcriptional subtypes. Sci Rep 15:23610
39. Thayer SP et al (2003) Hedgehog is an early and late mediator of pancreatic cancer tumorigenesis. Nature 425:851–856
40. Catenacci DV et al (2015) Randomized phase Ib/II study of gemcitabine plus placebo or vismodegib, a hedgehog pathway inhibitor, in patients with metastatic pancreatic cancer. J Clin Oncol 33:4284–4292
41. Ko AH et al (2016) A phase I study of FOLFIRINOX plus IPI-926, a hedgehog pathway inhibitor, for advanced pancreatic adenocarcinoma. Pancreas 45:370–375
42. Canto MI et al (2013) International Cancer of the Pancreas Screening (CAPS) Consortium summit on the management of patients with increased risk for familial pancreatic cancer. Gut 62:339–347
43. Canto MI et al (2018) Risk of neoplastic progression in individuals at high risk for pancreatic cancer undergoing long-term surveillance. Gastroenterology 155:740–751e2
44. Bankhead P et al (2017) QuPath: open source software for digital pathology image analysis. Sci Rep 7:16878

45. Hao Y et al (2021) Integrated analysis of multimodal single-cell data. Cell 184:3573–3587e29
46. Hao Y et al (2024) Dictionary learning for integrative, multimodal and scalable single-cell analysis. Nat Biotechnol 42:293–304
47. Hafemeister C, Satija R (2019) Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. Genome Biol 20:296
48. Korsunsky I et al (2019) Fast, sensitive and accurate integration of single-cell data with Harmony. Nat Methods 16:1289–1296
49. Efremova M, Vento-Tormo M, Teichmann SA, Vento-Tormo R (2020) CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. Nat Protoc 15:1484–1506
50. Jin S et al (2021) Inference and analysis of cell-cell communication using CellChat. Nat Commun 12:1088
51. Kogure Y et al (2022) Whole-genome landscape of adult T-cell leukemia/lymphoma. Blood 139:967–982
52. Shiraishi Y et al (2013) An empirical Bayesian framework for somatic mutation detection from cancer genome sequencing data. Nucleic Acids Res 41:e89
53. Bailey MH et al (2018) Comprehensive characterization of cancer driver genes and mutations. Cell 173:371–385e18
54. Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:550

Figures

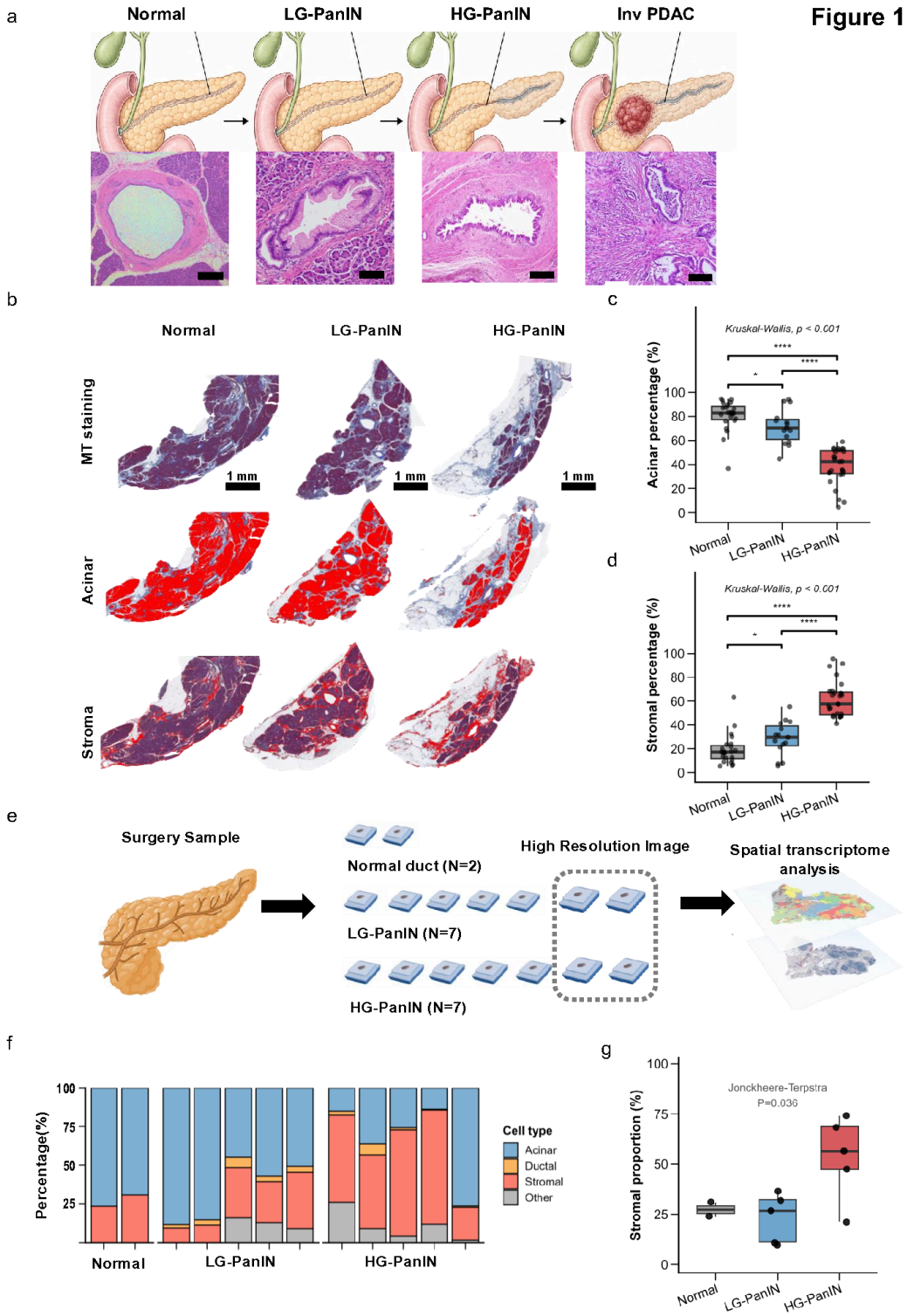


Figure 1

Spatial transcriptomic analysis reveals marked stromal remodeling at the HG-PanIN stage.

(a) Schematic of the stepwise PanIN progression model (top) with representative H&E images (bottom). Black arrows indicate stepwise progression from normal pancreas through LG-PanIN and HG-PanIN to invasive PDAC. (b) Masson's trichrome staining of normal, LG-PanIN, and HG-PanIN tissues (top) with automated image segmentation identifying acinar/parenchymal (Biebrich scarlet-positive, red) and stromal/collagen (aniline blue-positive, blue) compartments (middle and bottom). Unstained (white) areas, including adipose tissue and tissue voids, were excluded from quantification. (c) Quantification of acinar tissue proportions across PanIN stages. Box plots show median and interquartile range with individual data points (Normal, n = 23; LG-PanIN, n = 15; HG-PanIN, n = 27). Kruskal–Wallis test, $P < 0.001$. Pairwise comparisons by Dunn's post hoc test with Bonferroni correction: $P < 0.05$, $^{**}P < 0.0001$. (d) Quantification of stromal tissue proportions across PanIN stages. Box plots show median and interquartile range with individual data points (normal, n = 23; LG-PanIN, n = 15; HG-PanIN, n = 27). Kruskal–Wallis test, $P < 0.001$. Pairwise comparisons by Dunn's post hoc test with Bonferroni correction: $P < 0.05$, $^{****}P < 0.0001$. (e) Study design for spatial transcriptomic profiling. Surgically resected specimens were analyzed using 10x Genomics Visium (Normal, n = 2; LG-PanIN, n = 5; HG-PanIN, n = 5) and Visium HD (LG-PanIN, n = 2; HG-PanIN, n = 2) platforms. (f) Stacked bar chart showing the cell type composition (Acinar, Ductal, Stromal, and Other) for each sample, grouped by diagnosis category. Cell types were assigned based on cluster-level marker gene scoring. (g) Box plot of stromal proportion across diagnosis categories. Individual data points represent individual samples. Stromal remodeling was markedly expanded at the HG-PanIN stage, whereas LG-PanIN showed an intermediate but not statistically significant increase over normal tissue. Jonckheere–Terpstra trend test, $P = 0.036$. Scale bars, 200 μm (a) and 1 mm (b).

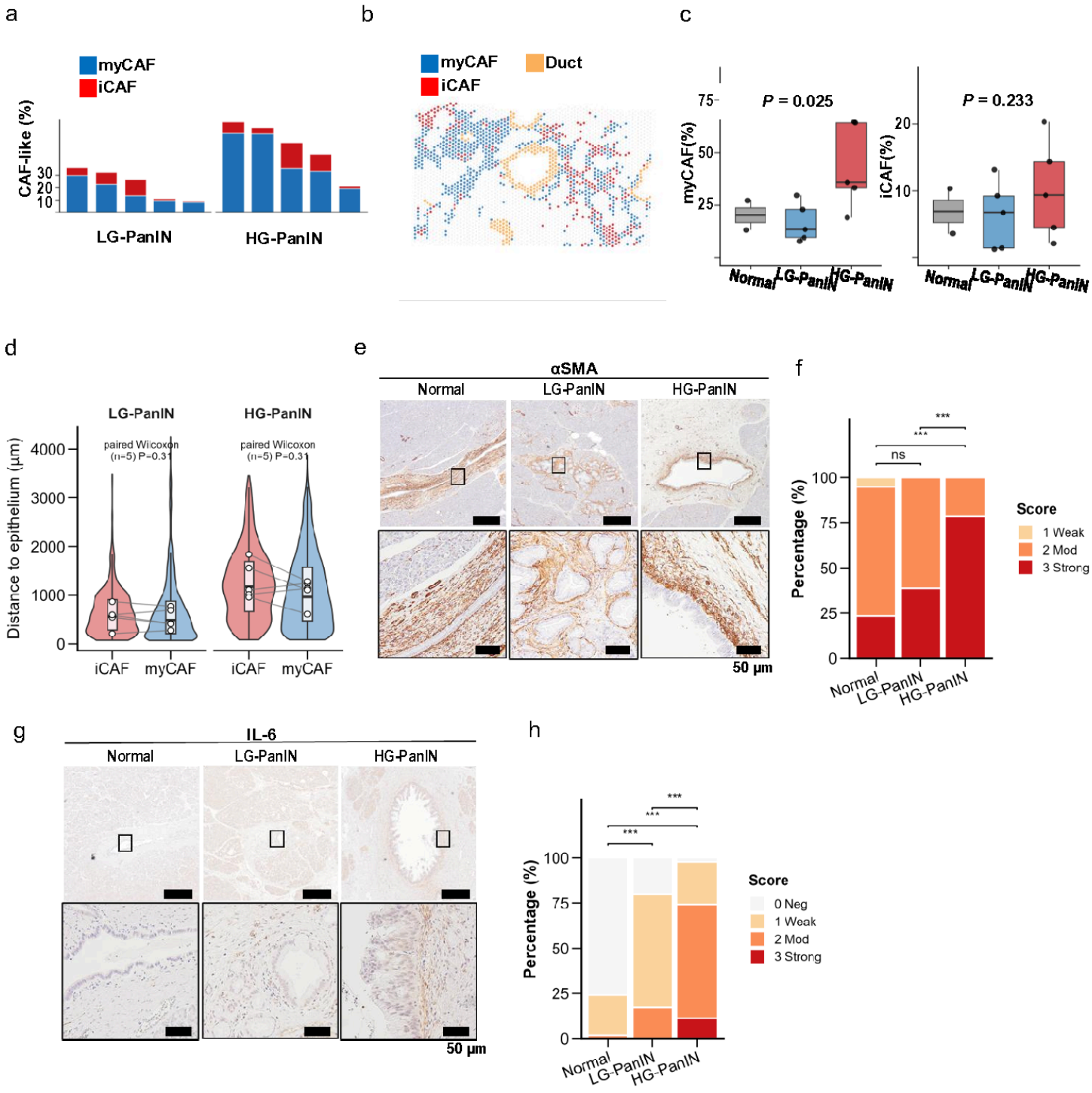


Figure 2

Spatial organization of CAF-like fibroblast programs during PanIN progression.

(a–d) Standard Visium analysis; (e–h) immunohistochemistry. Visium spots (55 μm) are multicellular: iCAF-like and myCAF-like labels indicate program enrichment, not individual-cell identities; no spot deconvolution was applied. (a) Proportions of inflammatory (iCAF-like, red) and myofibroblastic (myCAF-like, blue) cancer-associated fibroblast program-positive spots in low-grade (LG) and high-grade (HG) pancreatic intraepithelial neoplasia (PanIN), with five sections per stage. Bars represent sections ordered by descending histologically annotated stromal proportion; case-to-sample mapping and shared patient origins are in Supplementary Table 8. (b) Representative LG-PanIN section (7528-22H03674_11): ductal, yellow (189 spots); myCAF-like, blue (642); iCAF-like, red (257). (c) Program-positive spot proportions across normal, LG-PanIN and HG-PanIN samples (n = 2, 5 and 5 sections). Points represent sections. Jonckheere–Terpstra trend tests: myCAF-like, $P = 0.025$; iCAF-like, $P = 0.233$. These trends are exploratory because sections are partially nested within patients. (d) Distances to the nearest epithelial spot. Violin distributions show spots; connected points show per-section medians (n = 5 sections per stage). Paired Wilcoxon comparisons of iCAF-like and myCAF-like distances: $P = 0.31$ in both stages. Spot distributions are for visualization only. (e,g) Representative α-smooth muscle actin (αSMA; e) and interleukin-6 (IL-6; g) staining with magnified insets. (f,h) Stromal expression-score distributions: 0, negative; 1, weak; 2, moderate; 3, strong. Duct counts for normal/LG-PanIN/HG-PanIN: αSMA, 77/31/74; IL-6, 58/35/89. Pairwise comparisons use Dunn's test with Bonferroni correction: ns, not significant; ***$P \leq 0.001$; full statistics are provided in Supplementary Tables 2 and 3. Scale bars, 200 μm in low-magnification images and 50 μm in insets (e,g).

Figure 3

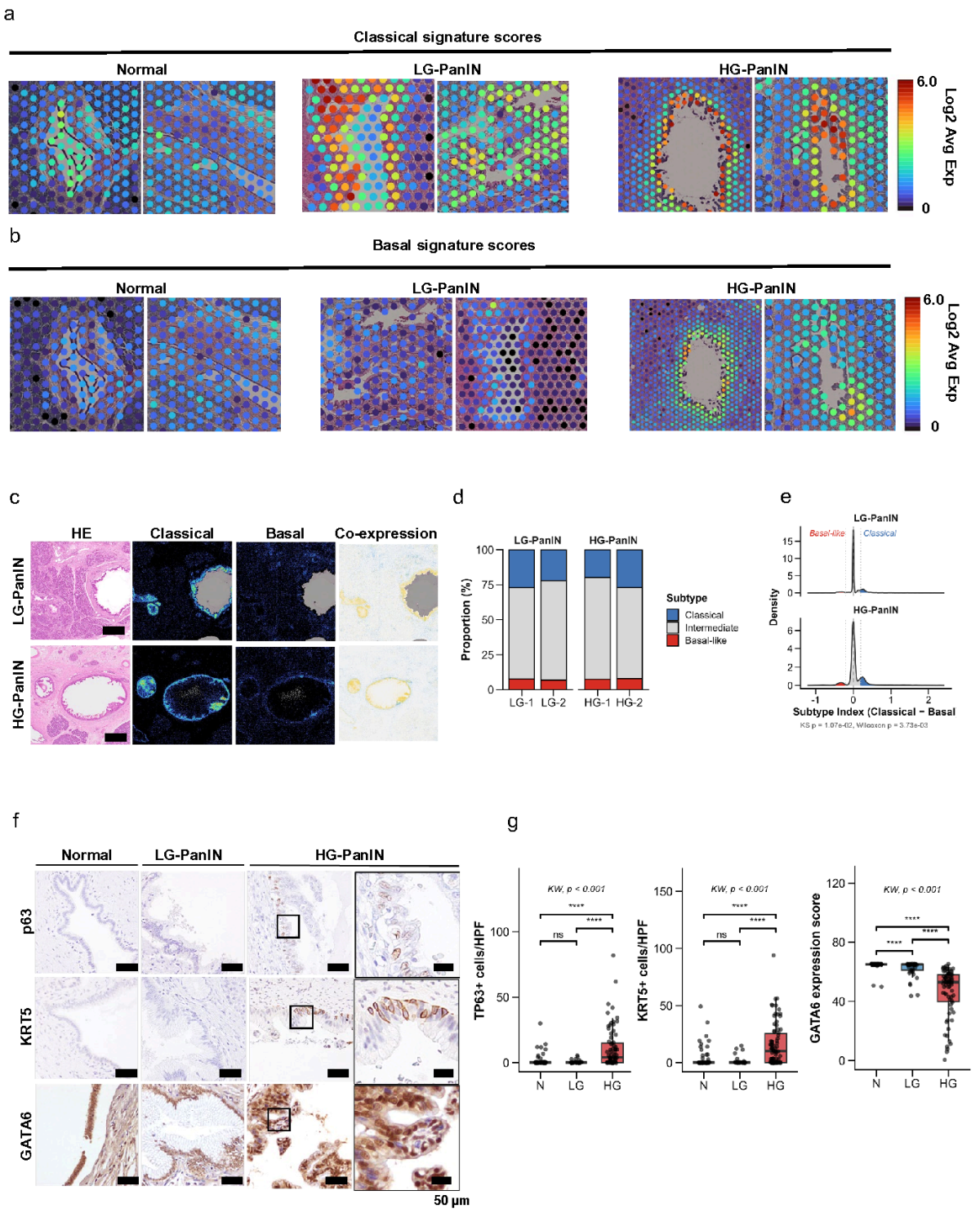


Figure 3

Focal basal-like epithelial drift during PanIN progression.

(a and b) Spatial mapping of classical (a) and basal-like (b) signature scores across Normal, LG-PanIN, and HG-PanIN tissues using Visium. Color scale represents log2 average expression. (c) Validation using Visium HD. H&E images with corresponding classical signature, basal-like signature, and co-expression

mapping in LG-PanIN and HG-PanIN tissues, demonstrating spatial heterogeneity at single-cell resolution. (d) Quantification of epithelial subtype composition from Visium HD analysis. Stacked bar plots showing the percentage of classical, intermediate, and basal-like populations in HG-PanIN versus LG-PanIN. Threshold-based categorical proportions appear visually similar between stages; statistical testing of the underlying continuous distribution is shown in panel (e). The dominant intermediate population (~67%–69%) reflects continuous phenotypic transition rather than discrete subtype switching. (e) Density plot of classical-to-basal spectrum (Subtype Index = Classical score − Basal score) demonstrating a leftward shift in HG-PanIN compared to LG-PanIN, indicating modest basal-like drift along a continuous spectrum. (f) Immunohistochemical validation for basal markers (p63 and KRT5) and classical marker (GATA6) in Normal, LG-PanIN, and HG-PanIN tissues. Insets show high-magnification views. (g) Quantification of IHC markers across epithelial classifications. (Left) TP63-positive cells per HPF (Normal, n = 76; LG-PanIN, n = 36; HG-PanIN, n = 74). (Middle) KRT5-positive cells per HPF (Normal, n = 68; LG-PanIN, n = 37; HG-PanIN, n = 78). (Right) GATA6 retained-expression score (Normal, n = 69; LG-PanIN, n = 40; HG-PanIN, n = 68). GATA6 Expression Score was calculated as max(GATA6-negative cells) minus GATA6-negative cells per HPF; higher values indicate retained GATA6 expression and a more classical-like phenotype. All panels: Kruskal–Wallis test, $P < 0.001$. Pairwise comparisons by Dunn's post hoc test: ns, not significant; ****$P < 0.0001$. Scale bars, 200 μm (c and f low-magnification panels) and 50 μm (f insets).

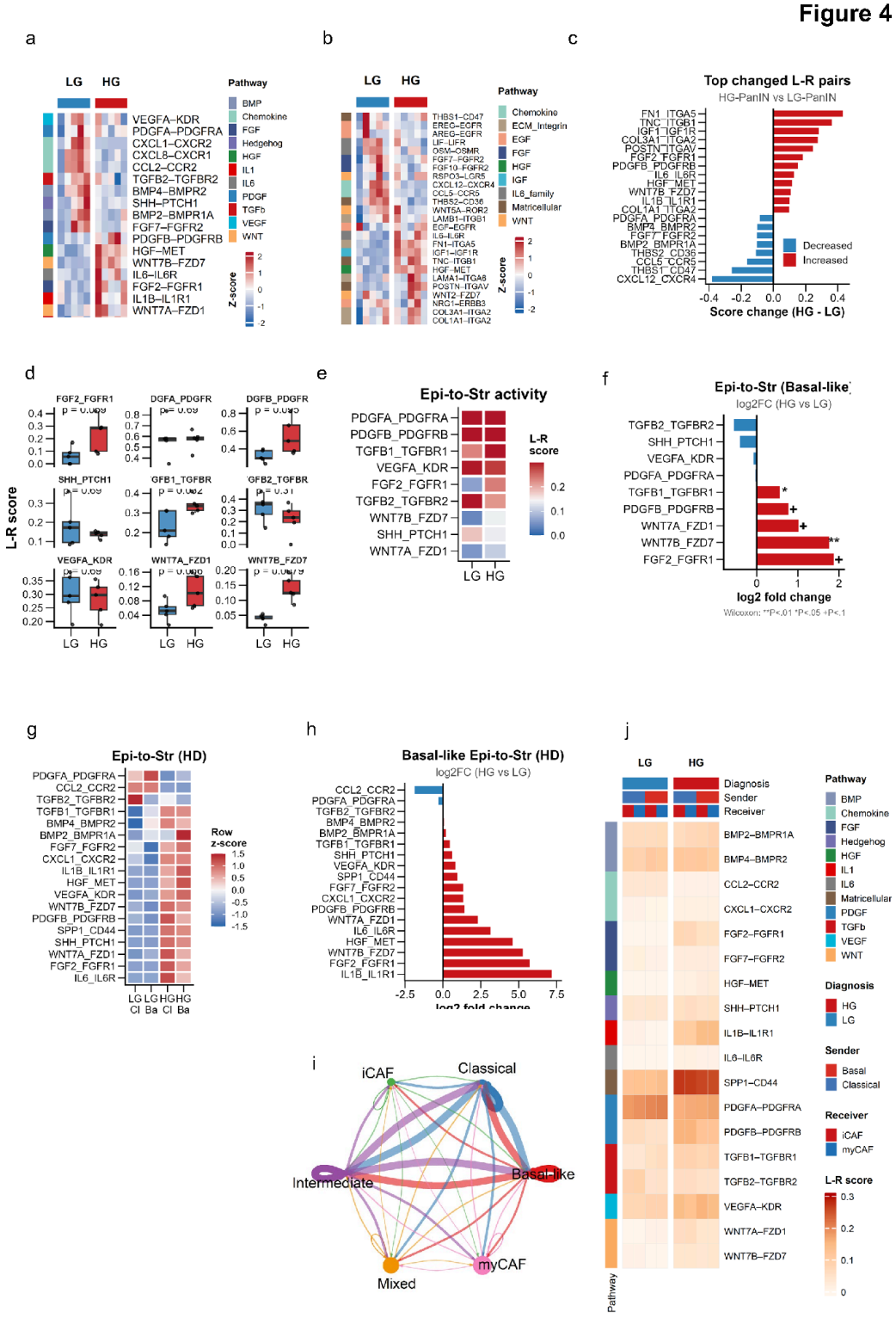


Figure 4

Candidate epithelial–stromal interactions during PanIN progression.

(a–f) Standard Visium analysis of 10 pancreatic intraepithelial neoplasia (PanIN) sections: five low-grade (LG) and five high-grade (HG). Forty-four curated ligand–receptor (L–R) pairs span 17 pathways. Scores are geometric means of pseudobulk ligand and receptor expression in pathologist-annotated source and target compartments. They estimate candidate interactions, not signaling flux or pathway dependency. (a,b) Row-standardized heatmaps of epithelial-to-stromal (a) and stromal-to-epithelial (b) scores; columns represent sections, grouped by stage, and rows represent L–R pairs. Blue/red denotes lower/higher standardized scores. (c) Twenty pairs with the largest absolute score changes; red/blue indicates increases/decreases in HG-PanIN relative to LG-PanIN. (d) Scores for nine epithelial-to-stromal pairs: FGF2–FGFR1, PDGFA–PDGFRA, PDGFB–PDGFRB, SHH–PTCH1, TGFB1–TGFBR1, TGFB2–TGFBR2, VEGFA–KDR, WNT7A–FZD1 and WNT7B–FZD7. Points represent sections; Wilcoxon rank-sum comparisons. (e) Mean scores for these nine pairs by stage. (f) Corresponding log2 fold changes. (g–j) Visium HD analysis of four specimens from independent patients (n = 2 per stage). (g) Row-standardized epithelial-to-stromal scores by stage and classical or basal-like epithelial subtype. (h) Log2 fold changes for basal-like epithelial-to-stromal scores in HG-PanIN versus LG-PanIN. Sample-level Wilcoxon comparisons are exploratory with only two specimens per group; effect estimates require independent validation. (i) Candidate communication between classical, intermediate and basal-like epithelium and inflammatory (iCAF-like), myofibroblastic (myCAF-like) and mixed cancer-associated fibroblast programs; arrow thickness denotes interaction score. (j) Scores across epithelial-subtype/fibroblast-program combinations, grouped by stage, sender and receiver, with pathway annotations. Standard Visium stage comparisons are also exploratory because the 10 sections originate from four patients, two contributing both stages (Supplementary Table 8). Statistical procedures and multiple-testing corrections are described in Online Methods.

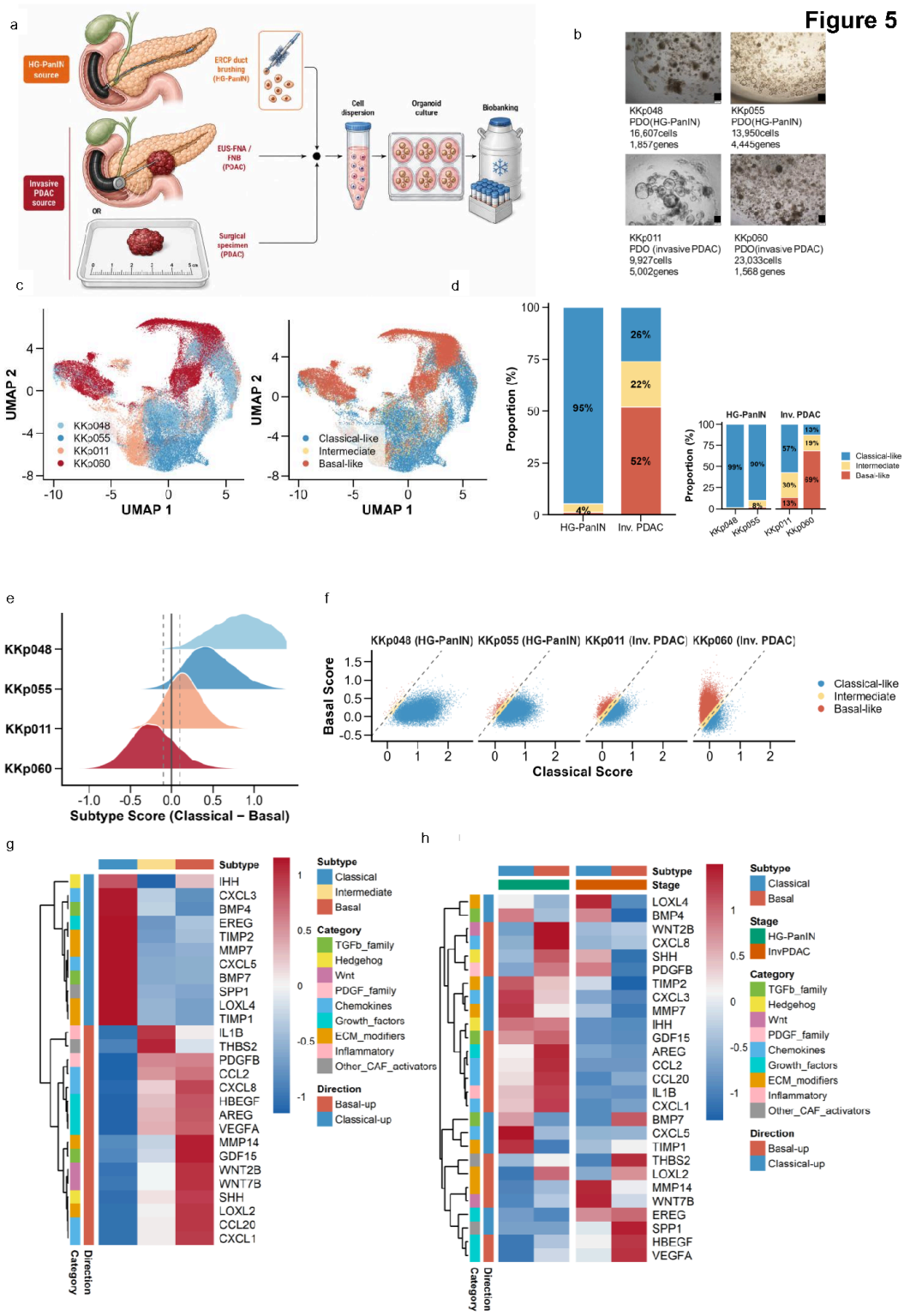


Figure 5

Patient-derived organoid single-cell analysis supports epithelial subtype heterogeneity and reveals subtype-specific fibroblast-engaging ligand expression.

(a) Schematic of patient-derived organoid (PDO) establishment. HG-PanIN cells were acquired via pancreatic duct brushing during ERCP, while invasive PDAC cells were obtained from EUS-guided tissue acquisition or surgical specimens, followed by cell dispersion, organoid culture, and biobanking (see also Extended Data Fig. 7). (b) Representative bright-field images of established patient-derived organoids. Two HG-PanIN-derived organoids (KKp048: 16,607 cells, 1,857 genes; KKp055: 13,950 cells, 4,445 genes) and two invasive PDAC-derived organoids (KKp011: 9,927 cells, 5,002 genes; KKp060: 23,033 cells, 1,568 genes) were subjected to single-cell RNA sequencing. (c) UMAP visualization of single-cell transcriptomes. Left panel colored by sample identity (KKp048, KKp055, KKp011, KKp060); right panel colored by epithelial subtype: Classical-like (blue), Intermediate (yellow), and Basal-like (red). (d) Quantification of epithelial subtype composition. Left: aggregated comparison between HG-PanIN and invasive PDAC. Right: individual sample breakdown. HG-PanIN organoids are predominantly Classical-like and contain a small but non-zero Basal-like fraction that is barely visible at this scale (KKp048: 98.6% Classical-like, 0.2% Basal-like; KKp055: 90.3% Classical-like, 1.9% Basal-like), while invasive PDAC organoids show substantial Basal-like enrichment (KKp011: 57.2% Classical-like, 13.2% Basal-like; KKp060: 12.7% Classical-like, 68.8% Basal-like). (e) Density plot of score difference (Classical minus Basal) across samples, showing leftward shift from HG-PanIN to invasive PDAC. (f) Scatter plots of Basal score versus Classical score for each sample (KKp048, KKp055, KKp011, KKp060), colored by epithelial subtype (Classical-like, Intermediate, Basal-like), illustrating the continuous distribution of epithelial subtypes. (g) Heatmap of fibroblast-engaging ligand expression in HG-PanIN organoids, stratified by all three epithelial subtypes (Classical-like, Intermediate, and Basal-like). Ligands are categorized by signaling pathway (TGFβ_family, Hedgehog, Wnt, PDGF_family, Chemokines, Growth_factors, ECM_modifiers, Inflammatory, Other_CAF_activators). Direction arrows indicate Basal-up or Classical-up enrichment. (h) Heatmap of fibroblast-engaging ligand expression across disease progression (HG-PanIN to invasive PDAC), comparing organoids stratified by epithelial subtype (Classical and Basal) and disease stage. Ligands are categorized by signaling pathway. Direction arrows indicate Basal-up or Classical-up enrichment at each stage. Scale bars (b), 200 μm.

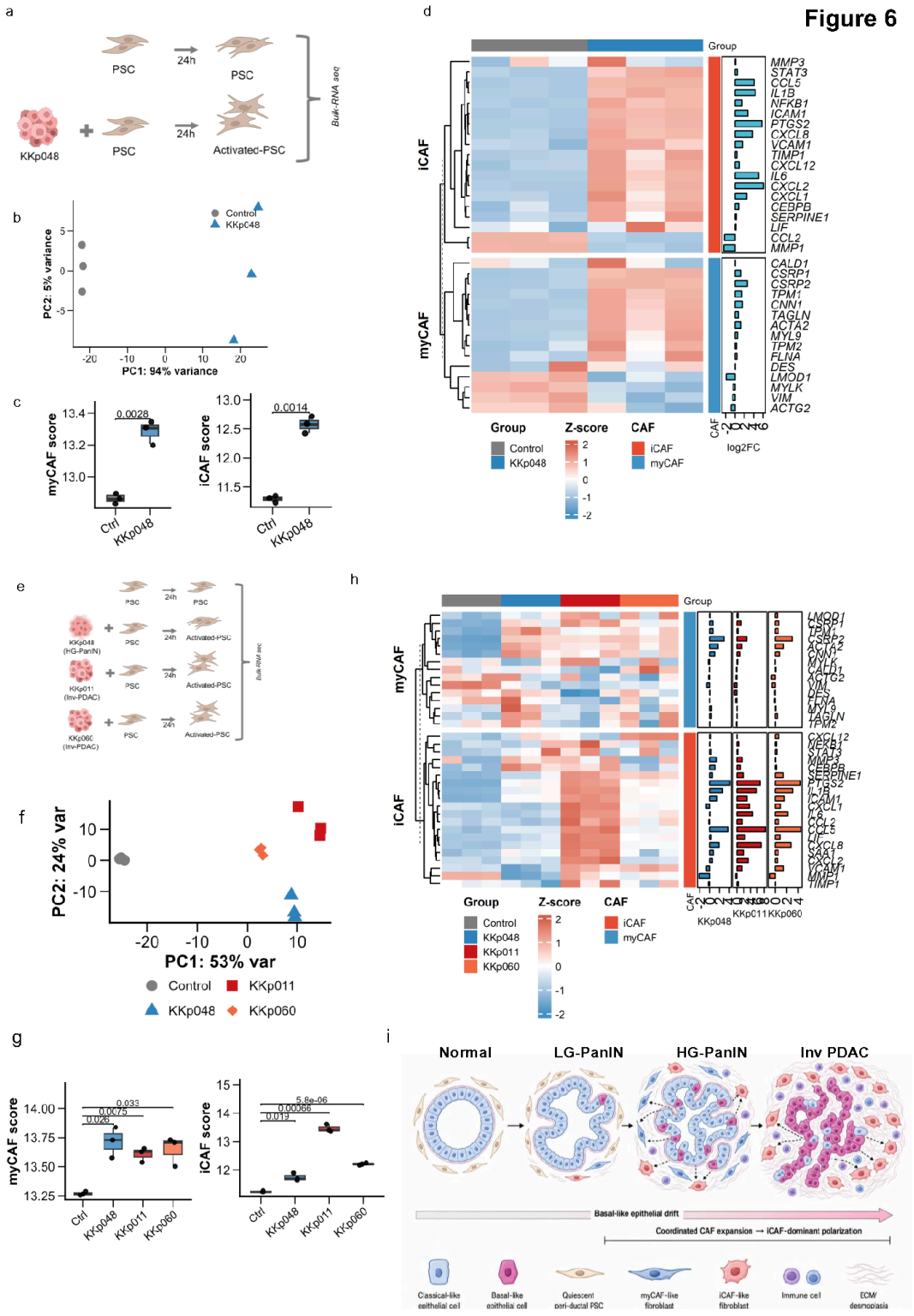


Figure 6

A high-grade precursor organoid engages fibroblast programs in pancreatic stellate cells.

(a–d) Initial conditioned-medium (CM) experiment using KKp048, a high-grade pancreatic intraepithelial neoplasia (HG-PanIN)-derived organoid line. (a) Primary human pancreatic stellate cells (PSCs) received KKp048 CM or control medium for 24 h before bulk RNA sequencing (n = 3 independent culture replicates per group). (b) Principal component analysis (PCA); PC1 explains 94% of variance. (c) Myofibroblastic (myCAF-like) and inflammatory (iCAF-like) cancer-associated fibroblast program scores: P = 0.0028 and P = 0.0014, respectively. (d) Row-standardized marker-expression heatmap with log2 fold-change annotations. (e–h) Independent extended experiment with control medium, KKp048 CM and invasive pancreatic ductal adenocarcinoma (PDAC) organoid CM from KKp011 or KKp060 (n = 3 independent culture replicates per condition; 24 h exposure). KKp048 is the sole HG-PanIN line used in both CM experiments. (e) Experimental design. (f) PCA; PC1 and PC2 explain 53% and 24% of variance. (g) Program scores. MyCAF-like comparisons against control: KKp048, P = 0.026; KKp011, P = 0.0075; KKp060, P = 0.033. Corresponding iCAF-like comparisons: P = 0.019, P = 0.00066 and P = 5.8×10^{-6}. (h) Row-standardized marker expression and fold changes relative to control; inflammatory programs are more pronounced with invasive PDAC CM. (i) Proposed model of epithelial–stromal remodeling from normal ducts through low-grade and high-grade PanIN to invasion. The schematic distinguishes classical-like and basal-like epithelium, contractile PSCs, inflammatory and myofibroblastic fibroblast programs, and fibrosis. Program scores in c and g are means of variance-stabilized marker-gene expression per culture replicate; points represent replicates and comparisons use two-sided, unpaired Student's t-tests. Matrigel-only CM served as the organoid control. With three culture replicates per condition and one HG-PanIN donor line, results provide proof of principle. Independent reproduction of the KKp048 effect does not establish generalizability across precursor donors. See Online Methods for experimental and statistical details.

Supplementary Files

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

- ExtendedDataFigure10.pdf
- ExtendedDataFigure8.pdf
- ExtendedDataFigure7.pdf
- ExtendedDataFigure5.pdf
- ExtendedDataFigure9.pdf
- ExtendedDataFigure4.pdf
- InventoryofSupportingInformationNatureCancer.docx
- ExtendedDataFigure6.pdf
- SupplementaryTablesNatureCancer.docx
- ExtendedDataFigure2.pdf
- ExtendedDataFigure1.pdf

- ExtendedDataFigure3.pdf