HHS Public Access

Author manuscript

Cancer Discov. Author manuscript; available in PMC 2026 June 23.

AI-Powered Deep Visual Proteomics Reveals Critical Molecular Transitions in Pancreatic Cancer Precursors

Jimin Min[1], **Lisa Schweizer**[2], **Gijs Zonderland**[2], **Benson Chellakkan Selvanesan**[3], **Julie H. Thomsen**[2], **Lukas Oldenburg**[2], **Seong-Woo Bae**[4], **Bongjun Kim**[1], **Sharía D. Hernández**[2], **Gabriela Jez**[2], **Vincent Bernard**[3], **Benjamin J. Swanson**[5], **Kelsey A. Klute**[6], **Huamin Wang**[7], **Thomas C. Caffrey**[8], **Paul M. Grandgenett**[8], **Michael A. Hollingsworth**[8], **Ishani Ummat**[2], **Maximilian T. Strauss**[2], **Andreas Mund**[2], **Anirban Maitra**[1,9]

[1]Laura and Isaac Perlmutter Cancer Center, Department of Medicine, New York University Grossman School of Medicine, NYU Langone Health, New York, New York.

[2]Resolute Bio, Copenhagen, Denmark.

[3]Department of Radiation Oncology, The University of Texas MD Anderson Cancer Center, Houston, Texas.

[4]Division of Precision Medicine, Department of Medicine, New York University Grossman School of Medicine, New York, New York.

[5]Department of Pathology, Microbiology, and Immunology, University of Nebraska Medical Center, Omaha, Nebraska.

[6]Department of Internal Medicine, University of Nebraska Medical Center, Omaha, Nebraska.

[7]Department of Pathology, The University of Texas MD Anderson Cancer Center, Houston, Texas.

[8]The Eppley Institute for Research in Cancer and Allied Diseases, Fred & Pamela Buffet Cancer Center, University of Nebraska Medical Center, Omaha, Nebraska.

[9]Department of Pathology, New York University Grossman School of Medicine, NYU Langone Health, New York, New York.

Abstract

Corresponding Authors: Andreas Mund, Resolute Bio, Ryesgade 113, København Ø 2100, Denmark. amund@resolute.bio; and Jimin Min, New York University Grossman School of Medicine, Smilow 906, 522 1st Ave, New York, NY 10016. jimin.min@nyulangone.org.
J. Min and L. Schweizer contributed equally to this article.

Authors' Contributions

J. Min: Conceptualization, data curation, investigation, visualization, writing–original draft, writing–review and editing. **L. Schweizer:** Conceptualization, data curation, formal analysis, investigation, visualization, writing–review and editing. **G. Zonderland:** Data curation. **B.C. Selvanesan:** Data curation. **J.H. Thomsen:** Figure generation, writing–review and editing of the manuscript. **L. Oldenburg:** Formal analysis, methodology. **S.-W. Bae:** Data curation. **B. Kim:** Data curation, methodology. **S.D. Hernández:** Pathology, writing–review and editing of the manuscript. **G. Jez:** MS data acquisition. **V. Bernard:** Methodology. **B.J. Swanson:** Resources. **K. A. Klute:** Resources. **H. Wang:** Resources. **T.C. Caffrey:** Resources. **P.M. Grandgenett:** Conceptualization, resources, writing–review and editing. **M.A. Hollingsworth:** Conceptualization, resources. **I. Ummat:** Funding acquisition, writing–review and editing. **M.T. Strauss:** Formal analysis, methodology. **A. Mund:** Conceptualization, supervision, project administration, writing–original draft, writing–review and editing. **A. Maitra:** Conceptualization, supervision, project administration, writing–review and editing.

Pancreatic ductal adenocarcinoma (PDAC) evolves through precursors, yet the protein programs governing early progression remain poorly defined. We applied Deep Visual Proteomics (DVP)—integrating computational pathology, laser microdissection, and mass spectrometry (MS)—to profile normal ducts, acinar-to-ductal metaplasia (ADM), low-grade (LG) and high-grade (HG) pancreatic intraepithelial neoplasia (PanIN), and invasive carcinoma from organ donors and patients with PDAC. Quantifying 9,181 proteins from ~100 cells per region, we uncovered a molecular field effect in histologically normal ducts and proteomic divergence of LG-PanINs by cancer context. We identified four stage-associated molecular programs. Stress adaptation and immune engagement emerged early in cancer-associated normal ducts. Metabolic reprogramming initiated in normal ducts and intensified across PanIN progression. Mitochondrial remodeling became prominent in HG-PanINs before invasion. MS detected KRAS hotspot mutant peptides within incidental precursor lesions from cancer-free individuals. These findings demonstrate that molecular reprogramming precedes histologic transformation, creating opportunities for earlier detection of lethal cancer.

SIGNIFICANCE: Artificial intelligence (AI)–guided DVP represents the first in-depth assessment of the proteomic landscapes observed during the multistep progression of pancreatic adenocarcinoma, including histologically normal ducts, ADM, and LG- and HG-PanIN lesions. These data represent a unique resource of candidate biomarkers and interception targets against this lethal disease.

INTRODUCTION

Pancreatic ductal adenocarcinoma (PDAC) remains one of oncology's most formidable challenges, with a dismal 5-year survival rate below 10% owing to late detection, therapeutic resistance, and rapid metastatic spread (1). The canonical progression model depicts a stepwise transition from histologically unremarkable normal ductal epithelium to noninvasive precursor lesions known as pancreatic intraepithelial neoplasia (PanIN) to invasive carcinoma, driven by the sequential accumulation of oncogenic *kRAS* activation and inactivation of *CDkN2A*, *TP53*, and *SMAD4* (2, 3). Acinar-to-ductal metaplasia (ADM) reflects the ductal reprogramming of acinar cells and is widely viewed as an early state that can seed PanIN formation in *kRAS*-driven tumorigenesis (4). In contrast to PanIN lesions, the biology of ADMs in humans remains poorly described, with most knowledge stemming from murine models (5).

PanINs are graded histologically as low-grade (LG) or high-grade (HG) based on architectural and cytologic atypia. Although LG-PanINs retain relatively preserved ductal architecture, HG-PanIN lesions exhibit marked nuclear pleomorphism, loss of polarity, and cribriform growth patterns that approach carcinoma *in situ* (6, 7). Understanding the molecular transitions between these grades is critical for identifying lesions with high malignant potential. Although HG-PanIN is widely considered the immediate noninvasive precursor to PDAC, its clinical interpretation is challenging because of the diagnostic difficulties in distinguishing HG-PanIN from the “cancerization” of the duct due to adjacent PDAC (7–9).

Recent studies in cancer-free organ donor samples have demonstrated a remarkably high prevalence of incidental PanIN (iPanIN) lesions (~60% of the general population), including in younger adults (10). Given the relatively uncommon incidence of PDAC, most iPanINs (LG-iPanIN) will not progress to invasive neoplasia during an individual's lifetime. It remains unknown whether LG-iPanINs harbor a distinct molecular signature compared with precursor lesions adjacent to PDAC [cancer-associated PanINs (cPanIN) or LG-cPanINs]. A central unresolved question is whether HG-cPanINs represent a quantitative intensification of LG programs or a qualitatively distinct molecular state, a debate that has not been systematically resolved at the protein level in human tissue. Bridging this knowledge gap is crucial: meaningful early detection and cancer interception strategies must identify which precursor lesions are most likely to progress while avoiding surveillance and overdiagnosis pitfalls (11, 12).

Technical challenges compound these knowledge gaps. Low-volume liquid biopsies often fail to detect circulating tumor DNA in early-stage disease; bulk omics approaches sacrifice critical spatial context; antibody-based methods offer limited proteomic coverage, constraining discovery potential; and mRNA levels do not necessarily correspond to protein abundance. Moreover, given the low accessibility of organ donor pancreas samples, nearly all prior studies have predominantly analyzed precursor lesions from patients already diagnosed with PDAC (13, 14), potentially missing *de novo* molecular events in cancer-free individuals. Growing evidence also indicates heterogeneous evolutionary trajectories profoundly influenced by the precancer microenvironment (10, 13, 15). The concept of "field cancerization"—molecular alterations extending beyond visible lesions into the surrounding normal-appearing tissues—plays a critical role in PDAC development, potentially driving the high recurrence rates following surgical resection (16).

We overcome these constraints through Deep Visual Proteomics (DVP), integrating high-resolution microscopy, artificial intelligence (AI)–powered computational pathology, targeted laser microdissection, and ultrasensitive mass spectrometry (MS) to enable comprehensive spatial protein analysis within precisely defined cell populations (17–19). In this study, we create the first spatially resolved proteomic cartography of multistep PDAC development by analyzing precisely defined tissue compartments across the spectrum from normal ducts to noninvasive precursor lesions to invasive carcinoma. We hypothesized that molecular reprogramming might precede histologic changes and that the cancer context could profoundly influence the cellular state of otherwise histologically normal tissue. By comparing normal ducts and precursor lesions from both cancer-free individuals and patients with PDAC, we identify critical molecular transitions during pancreatic carcinogenesis and establish a framework for risk stratification. We also profile protein signatures of human ADM. This work addresses fundamental questions about the earliest events in PDAC development, with direct implications for transforming clinical practice from late-stage intervention to early detection and targeted prevention (20).

RESULTS

DVP Workflow and Cohort Overview

To study the spatial and molecular progression of PDAC at its earliest stages, we implemented DVP (18), an integrated spatial proteomic approach combining complementary cutting-edge technologies (Fig. 1A). We analyzed 115 discrete histologically annotated and microdissected regions from 20 individuals, including 10 cancer-free organ donors (incidental or iCohort) and 10 patients with PDAC (cancer-associated or cCohort), capturing the full spectrum from normal ducts to invasive carcinoma (Supplementary Table S1). Our DVP workflow uses 3-μm-thick sections of formalin-fixed paraffin-embedded (FFPE) tissues stained with hematoxylin and eosin (H&E). After acquiring high-resolution images, we used H-optimus-0 (RRID: SCR_028225), a pathology foundation model trained on more than 500,000 histopathology slides, to identify phenotypic clusters across the progression spectrum. This AI-powered approach enhanced our ability to precisely distinguish morphologically similar but biologically distinct tissue compartments at cellular resolution.

Guided by the pathology foundation model and pathologist annotation, we selected seven biological groups for proteomic profiling, with the prefixes “i” and “c” corresponding to incidental organ donor–associated and cancer-associated lesions, respectively: normal ducts (iNormal and cNormal), precursor lesions (iADM, LG-iPanIN, LG-cPanIN, and HG-cPanIN), and invasive carcinoma (PDAC and “Tumor”; Fig. 1B–D; Supplementary Fig. S1A–S1C). As previously described, HG-PanINs are essentially never present in the absence of cancer (21), and thus, no HG-iPanINs were identified in our series. Using laser microdissection with high spatial precision, we isolated small regions containing approximately 100 phenotypically matched cells for each sample. This targeted approach enabled us to overcome the limitations of bulk tissue analysis, which can mask critical signals from rare cell populations and dilute compartment-specific molecular signatures. We performed subsequent proteomic profiling on the Orbitrap Astral mass spectrometer, quantifying 9,181 unique proteins (Fig. 1E) and 117,178 unique peptide precursors (Fig. 1F) across all tissue states. The proteins spanned six orders of magnitude in abundance, demonstrating the exceptional sensitivity and dynamic range of our approach (Fig. 1G; Supplementary Fig. S2A; Supplementary Table S1).

Our quality control analyses confirmed the presence of peptides corresponding to pivotal PDAC-associated genomic drivers across the proteomic abundance range, including KRAS and TP53, and established cancer cell markers like HMGA2 (22) and stem cell markers like DCLK1 (Fig. 1G; Supplementary Fig. S2A; ref. 23). We reliably identified canonical PanIN markers, including ANXA10, S100P, CLDN18, and MUC5AC, at consistent levels across PanIN lesions regardless of cancer context (Supplementary Fig. S2B). Notably, a comparison with recent work on mice (24) revealed that our dataset captured both shared and human-specific signatures of PanIN initiation (Supplementary Fig. S2C). Quality control analyses of epithelial (KRT7 and EPCAM), connective tissue (DCN), and acute inflammation markers (C3, C7, and CD14) confirmed the expected enrichment of epithelial populations in the microdissected samples, while also showing the presence of

limited stromal and immune components in epithelial-proximal regions (Supplementary Fig. S2D). Batch effect analyses demonstrated consistent data quality following normalization (Supplementary Fig. S3A–S3C), and hierarchical clustering confirmed clear stratification across tissue states (Supplementary Fig. S3D). This integrated methodology allowed us to analyze thousands of proteins within spatially and phenotypically defined cell populations, directly linking molecular profiles to cellular identity and tissue context across multistep human pancreatic neoplasia.

Spatial Proteomic Dissection Reveals Functional Niches and Intratumoral Heterogeneity in PDAC

To resolve intratumoral heterogeneity, we performed multiregion proteomic sampling across three PDAC tumors with matched LG-cPanIN lesions. Using local, slide-specific clustering of the whole-tissue morphology with H-optimus-0, we selected two phenotypically distinct regions within the epithelial tumor compartment per case (epithelioid vs. fusiform tumor morphologies) for targeted microdissection (Fig. 2A–C). All slides were subjected to human pathology review. This design enriches tumor–cell state differences while retaining spatial context and complements heterogeneity observations previously made primarily at the transcriptional level (25). Principal component analyses (PCA) demonstrated pronounced molecular divergence between individual regions in all three cases (Fig. 2D–F). Importantly, the differences were not restricted to desmoplastic extracellular matrix (ECM) proteins but extended to intracellular stress, RNA regulatory, and mitochondrial programs.

In patient 5, the desmoplastic region (region 1) was enriched for matrix and structural proteins (e.g., LAMA4, COL18A1, FLNA, and ACTA1), along with cytoskeletal and adhesion-associated regulators (e.g., TNS1 and PDLIM family members). In contrast, a spatially distinct region (region 2) showed a higher abundance of RNA-processing factors (e.g., SF3B1, SF3B3, and HNRNPU) and mitochondrial or stress-associated proteins (e.g., NDUFS8, HTRA2, and PPOX; Fig. 2D). In patient 2, regions separated along an axis from nuclear RNA-binding and splicing proteins (e.g., NONO, HNRNPK, and SF3A2) in region 7 to cytoskeletal and contractile components (e.g., MYL12A, MYL12B, and PDLIM proteins) in region 3 (Fig. 2E). In patient 3, region 0 was enriched for lysosomal and vacuolar proteins (e.g., ATP6V1H and TECPR2) and immune/secretory-associated factors (e.g., CD5L, CFH, STX5, and TMED5), whereas region 6 showed higher levels of mitochondrial and metabolic proteins (e.g., TFAM, ECHS1, and HSPD1; Fig. 2F).

In these molecularly distinct tumor regions, cancer-related markers such as TP53 and DCLK1 exhibited divergent expression patterns (Supplementary Fig. S4A and S4B), consistent with region-specific tumor cell states. Given the limited number of tumors profiled, this analysis serves as a proof-of-concept demonstration of AI-guided spatial sampling for intratumoral heterogeneity.

These spatially resolved molecular signatures uncover functional niches that conventional bulk analysis would obscure and demonstrate that PDAC heterogeneity spans both extracellular remodeling and tumor-cell intrinsic stress and metabolic programs.

Cancer-Specific Molecular Signatures Differentiate PDAC from Precancers

Our analysis of the total data variance across all regions, including both tumor and nontumor compartments, revealed a clear separation of the tumor proteome from normal ducts, iADM, and PanINs (Fig. 3A). Notably, HG-cPanINs occupied a distinct position intermediate between LG-cPanINs and tumors, suggesting that HG lesions represent a distinct molecular state from that of the tumor proteome and, by extension, are unlikely to simply represent “cancerization” of the duct. Hierarchical clustering confirmed the stratification of HG-cPanINs and tumors (Supplementary Fig. S3D). The key contributors to the tumor-enriched proteome included LGALS1 (Galectin-1), a well-established regulator of tumor–stroma interactions (26, 27), and SERPINH1 (HSP47), an ECM-associated chaperone involved in collagen processing (Fig. 3B; refs. 28, 29). Intriguingly, COLGALT1—a collagen glycosyltransferase implicated in collagen posttranslational modification—was upregulated specifically in tumor compartments, consistent with intensified collagen remodeling during invasion (Fig. 3B). Differential expression and pathway enrichment analyses highlighted tumor-enriched programs, including ECM remodeling, MAPK/MET signaling, and enhanced cell motility (Supplementary Tables S2 and S3). The tumor proteome was also enriched for DNA replication/repair factors implicated in chemotherapy response (e.g., TOP2A and FEN1) and mediators of tumor–stroma crosstalk (e.g., POSTN and S100A4), underscoring the biological relevance of our dataset (Supplementary Tables S2 and S3).

The tumor samples displayed molecular hallmarks consistent with late-stage progression. Elevated TP53 levels appeared specifically in tumors and HG-cPanINs, likely representing the stabilization of the mutant protein (Supplementary Fig. S5A). DCLK1, a putative cancer stem cell marker, showed an increase from LG-cPanIN to tumor (Supplementary Fig. S5B). We identified a distinct set of proteins shared uniquely by HG-cPanINs and tumors (Supplementary Fig. S5C; Supplementary Table S4). Of note, although these proteins may be uniquely expressed, they may also have abundances below the MS detection limit; nonetheless, such pronounced differences in detectable abundance suggest biological relevance. These proteins, uniquely shared between HG-cPanINs and invasive tumors, included transcription factors and signaling regulators (RUNX2 and FOSL1), tumor suppressors and growth regulators (RASSF2 and UNC5B), cell-cycle proteins (CDC20), and IFN-related proteins (IFITM3; Supplementary Fig. S5D, representative examples). A smaller set of proteins was uniquely detected only in tumors, including ANGPTL4 and PRR11, supporting their potential utility as invasion-associated markers (Supplementary Fig. S5E). Although HG-cPanIN lesions share some aspects of this tumor-associated proteome, the exclusive appearance of tumor-unique factors at the PDAC stage likely reflects the acquisition of cell-autonomous proliferative and invasive capabilities. This final transformation to invasive carcinoma likely represents the culmination of the extended evolutionary timeline estimated in genetic studies, in which a median time of several years from the most recent common ancestor to clinically apparent PDAC has been proposed (30).

DVP Reveals Divergent Molecular Programs within Morphologically Equivalent Lesions

Focusing on nontumor lesions, PCA revealed that lesion morphology and disease association were the dominant factors shaping the molecular landscape of pancreatic tissue (Fig.

3C). The primary axis of variation (dimension 1, accounting for 14.4% of the total variance) separated histologic classifications of cells (Normal, iADM, and PanIN), with HG-cPanINs showing the most distinct separation. The secondary axis (dimension 2, accounting for 7.1% of variance) stratified cancer-associated samples ("cCohort") from cancer-free samples ("iCohort"). This analysis revealed that despite their morphologic similarity, lesions from cancer-free and cancer-associated contexts exhibited profound molecular differences, underscoring how conventional histopathology can miss critical disease-relevant changes. Together, these axes capture major components of variance linked to lesion class and cancer context. iADM samples occupy an intermediate region of the embedding, overlapping most strongly with LG-iPanINs and extending toward LG-cPanINs, but showing limited proximity to iNormal/cNormal, consistent with metaplasia spanning multiple transitional states.

To define stage-structured protein modules while explicitly capturing similarities between HG-cPanIN and PDAC, we performed a targeted ANOVA across tissue states and extracted proteins with defined progression patterns by *post hoc* Tukey testing (Fig. 3D; Supplementary Table S5). Clusters 1 and 2 comprised proteins that increased from histologically normal ducts through PanIN to HG-cPanIN and PDAC, including stress-response factors (e.g., HSP90B2P, Fig. 3E) and cell-interaction/progenitor features (e.g., CD44, Fig. 3F), together with increased ribosomal/translation machinery. Cluster 3 peaked in early precursors (iADM and PanIN lesions) and was dominated by mucin/secretory and glycosylation proteins (e.g., MGAT3, GALNT4, FUT3, and GOLM1, Fig. 3G), consistent with an early secretory and glycosylation shift in precursor lesions. Cluster 4 captured the coordinated loss of differentiated pancreatic function (e.g., NFIC, Fig. 3H).

Global differential abundance analysis revealed proteomic differences consistent with a cancer-associated field effect in histologically normal ducts. Despite matched morphology, cNormal differed from iNormal by 192 proteins, with 68% (131 proteins) increased in cNormal (Fig. 3I; Supplementary Table S2). Upregulated proteins included stress/secretory-pathway and epithelial remodeling factors (e.g., AGR2), as well as immune surveillance and antigen-presentation components (e.g., B2M). We also observed a higher abundance of several exocrine digestive enzymes (CTRB1, CTRC, and CELA2A) in cNormal, consistent with altered periductal composition and/or epithelial remodeling in the cancer-bearing pancreas. Importantly, cNormal regions were microdissected from tissue sections spatially separated from the tumor (different blocks), arguing against tumor-cell contamination as the source of this signature.

Comparing histologically matched LG-iPanIN and LG-cPanIN lesions, we identified 301 differentially expressed proteins; 88% (264 proteins) were higher in LG-cPanIN, underscoring the impact of cancer context on LG precursor states (Fig. 3J; Supplementary Table S2). LG-cPanINs were enriched for ECM remodeling and adhesion proteins (e.g., CD44, LUM, BGN, COL6A2, and FBLN5), immune and IFN-response factors (e.g., B2M, STAT1, and immunoglobulins), vesicle trafficking proteins (e.g., RAB1B and TMED7), oxidative-stress regulators (e.g., MGST1), and glycolytic mediators (e.g., PGAM1). Together, these patterns indicate that field cancerization imprints a distinct

immunometabolic and microenvironment-responsive proteome on otherwise histologically indistinguishable LG lesions.

Proteomic Signatures of Cancer-Associated HG PanIN Lesions

The HG-cPanINs formed a molecularly distinct, tumor-proximal precursor state. Compared with cNormal ducts, HG-cPanINs showed 1,075 differentially abundant proteins, and 951 proteins differed between LG-cPanIN and HG-cPanIN lesions (Fig. 4A and B; Supplementary Table S2). HG-cPanINs exhibited increased abundance of nutrient acquisition and mitochondrial membrane biology proteins [e.g., transferrin receptor protein (TFRC), TOMM20/TOMM40, and LETM1] together with lipogenesis and glycolysis regulators [e.g., fatty acid synthase (FASN) and phosphofructokinase (PFKP)], whereas LG-cPanINs retained higher levels of detoxification and exocrine enzymes (e.g., GSTA2, CYP2C8, PRSS1/PRSS2, and CPA1; Fig. 4B; Supplementary Table S2). Metabolic rewiring dominated the HG-cPanIN proteome. FASN, a key enzyme in *de novo* lipogenesis, was markedly upregulated in HG-cPanINs, as was PFKP, consistent with enhanced glycolytic flux (Fig. 4C). Together with increased TFRC, these changes support a high bioenergetic and biosynthetic demand in advanced dysplasia. Pathway enrichment analysis highlighted exceptionally strong activation of mitochondrial gene expression and translation modules in HG-cPanINs, including mitochondrial translation initiation, elongation, termination, and respiratory electron transport (Fig. 4D and E; Supplementary Table S3). These data suggest that HG-cPanINs display a mitochondria-enriched proteomic signature before invasion, consistent with increased mitochondrial demand. Components of the PINK1/Parkin mitophagy pathway emerged prominently at the HG-cPanIN stage, including elevated SQSTM1/p62, VDAC1, and PGAM5 (Fig. 4F). Coupled with mitochondrial translation and oxidative phosphorylation programs, this signature suggests dependence on mitochondrial quality control to sustain dysplastic growth, nominating a potential vulnerability window for interception.

Molecular Hallmarks of the Cancer Field Effect and Early PDAC Evolution

Based on our comprehensive proteomic profiling, we delineated four recurring molecular programs that characterize the cancer field effect and progression to PDAC: (i) stress adaptation, (ii) immune engagement, (iii) metabolic reprogramming, and (iv) mitochondrial remodeling (Fig. 5A). Representative proteins within these programs included stress/secretory adaptation factors (e.g., CYB5A, AGR2, TAGLN, PLP2, and WTIP), immune and inflammatory mediators [e.g., MHC class I (MHC-I) components, STAT1, STAT3, α−2-macroglobulin (A2M), IRF9, C3, C7, and CRP], metabolic nodes (e.g., PGAM1, PDK4, RPN1, PFKP, and FASN), and mitochondrial fitness/quality control proteins (e.g., SQSTM1, VDAC1, PGAM5, CKMT1A, GSR, TST, and ACADS). Rather than proving temporal order, these cross-sectional data show that these programs are already detectable in morphologically normal ducts in cancer-bearing pancreata and become progressively more prominent across precursor lesions and PDAC, supporting their candidacy as early biomarkers and potential intervention points.

Stress Adaptation

Pathway enrichment analysis (Fig. 5B and C) indicated activation of stress-adaptation programs in cNormal ducts compared with iNormal, including endosomal/vacuolar trafficking, cargo concentration in the endoplasmic reticulum, and amyloid fiber formation signatures (Supplementary Table S3). Consistent with these pathways, cNormal displayed increased abundance of stress-response proteins (e.g., CYB5A, WTIP, and TAGLN) and ER/secretory trafficking factors (e.g., AGR2 and PLP2; Fig. 5D and E; Supplementary Table S2). Many of these markers remained elevated in LG-cPanIN relative to LG-iPanIN (Fig. 5D and E; Supplementary Table S2), suggesting that adaptive responses to microenvironmental stress persist through preinvasive progression.

Immune Engagement

The analysis of the cancer field effect revealed activation of inflammatory signaling networks and increased antigen presentation via MHC-I in cNormal ducts (Fig. 5B and F; Supplementary Table S3). These immune alterations were accompanied by elevated acute-phase proteins (e.g., CRP and complement components) and increased STAT signaling proteins relative to the incidental lesions from the cancer-free cohort (iCohort; Supplementary Fig. S6A). Detailed protein-level analysis revealed a progressive inflammatory gradient across precursor states (Fig. 5G). Multiple acute-phase reactants and complement components increased from iNormal to cNormal and were most pronounced in LG-cPanIN lesions, including CRP, SERPINA1, SERPING1, complement proteins (C3, C4A, C7, and C9), haptoglobin, and A2M (Fig. 5G; Supplementary Table S2). Together, these data indicate that LG-cPanINs occupy an inflamed epithelial ecosystem before overt malignant transformation. IFNγ signaling emerged as a dominant feature of the cancer-associated precursor proteome, peaking in LG-cPanIN lesions (Fig. 5H; Supplementary Table S3). This signature included transcriptional regulators (e.g., STAT1 and IRF9), antigen-presentation machinery (e.g., B2M and HLA class I/II components), and IFN-stimulated effectors (e.g., GBP4, OAS1/2/3, and TRIM14/25). These patterns are consistent with genomic reports of a large burden of independent LG-PanIN clones in cancer-adjacent pancreas (on the order of ~10^3 lesions; ref. 14), suggesting that only a subset of lesions ultimately progresses to invasive disease.

Metabolic Reprogramming

Relative to cancer-free tissue (iNormal and LG-iPanIN), cNormal and LG-cPanIN exhibited an immunometabolic shift marked by increased PGAM1 and decreased PDK4 (Fig. 3I and J; Supplementary Fig. S6B; Supplementary Table S2), consistent with enhanced glycolytic flux and altered regulation of pyruvate entry into the tricarboxylic acid cycle. The transition from cNormal to LG-cPanIN was accompanied by coordinated metabolic and biosynthetic remodeling, including upregulation of fatty-acid transport and β-oxidation regulators (e.g., SLC27A2 and OXCT1) and mitochondrial energy-handling and redox enzymes (e.g., CKMT1A, HCCS, and SCO2; Fig. 5I), reinforcing the central role of metabolic remodeling in early preneoplastic transformation. Concomitantly, the transition toward invasive PDAC was accompanied by coordinated changes in mitochondrial and invasion-associated proteins (Supplementary Fig. S6C and S6D), as detailed in the following

section. Finally, glycosylation remained a consistent feature of PanINs (Supplementary Fig. S7A), accompanied by increased abundance of glycosylation and ER-associated proteins, including RPN1 and B4GALT1 (Supplementary Fig. S7B and S7C), consistent with aberrant glycan remodeling in PDAC.

Mitochondrial Remodeling

The transition from LG-cPanIN to invasive PDAC was marked by coordinated changes in mitochondrial proteins alongside the induction of invasion-associated factors (Fig. 5J; Supplementary Table S2). Multiple mitochondrial proteins (e.g., TST, GSR, ACADS, and CKMT1A) decreased, whereas invasion-promoting proteins such as FSCN1, STMN1, and S100A4 increased (exemplarily shown in Supplementary Fig. S6C and S6D). In parallel, tumors showed strong enrichment of desmoplastic ECM and stromal markers, including FN1, POSTN, and VIM, consistent with the acquisition of an invasive remodeling program that is distinct from PanIN-intrinsic epithelial changes.

We next asked whether these metabolic changes coincided with the shifts between classical and basal-like PDAC lineage states (25, 31). Both LG-iPanIN and LG-cPanIN lesions retained a predominantly classical signature, including strong PanIN-associated CLDN18 expression and relatively stable GATA6 and HNF4A levels (Supplementary Fig. S8A), whereas basal-like markers (e.g., HMGA2, KRT17, and S100A2) were most evident at the invasive tumor stage (Supplementary Fig. S8B). Consistent with this lineage shift, tumors showed lower expression of mitochondrial pyruvate carrier 1 relative to LG-cPanINs (Supplementary Table S2), aligning with reports linking basal-like PDAC to diminished mitochondrial pyruvate import (32).

The Proteomic Landscape of Human ADM

Although murine ADM has been extensively profiled by single-cell approaches (33, 34), human ADMs—and their relationship to LG-PanIN lesions—remain poorly defined. Comparing cancer-free iNormal ducts with iADM and LG-iPanIN regions, differential expression identified 129 upregulated proteins in iADM and 90 proteins in LG-iPanIN, with a shared core of 72 proteins induced in both lesions (Fig. 6A–C; Supplementary Table S2). CKMT1A was among the most strongly induced proteins across both lesions (Fig. 6A–C), suggesting that altered mitochondrial energy handling is an early feature of ductal reprogramming. Secretory-axis proteins, such as the Golgi-associated protein GOLM1 and the glycosyltransferase B4GALT5, were also consistently upregulated in both lesions, further highlighting early activation of secretory and glycosylation pathway flux during preneoplastic initiation (Fig. 6D and E). More broadly, both precursor states exhibited coordinated induction of oxidative stress and metabolic reprogramming proteins (e.g., HK1, NQO1, and GPX2), secretory pathway components (e.g., AGR2), and ductal/PanIN-associated epithelial barrier and adhesion factors (e.g., CLDN18/CLDN2, MUC1/MUC6, and S100P; Supplementary Table S2). A direct comparison between iADM and LG-iPanIN lesions uncovered divergent programs (37 proteins; Fig. 6F; Supplementary Table S2). iADM preferentially expressed metabolic reprogramming and ductal/progenitor-associated markers, exemplified by PHGDH (Fig. 6F and G). In contrast, LG-iPanINs selectively upregulated chromatin regulation and epithelial architecture proteins, including

CENPV and LIN7A (Fig. 6F and H), consistent with the establishment of the structured epithelial program characteristic of stable precursor lesions.

Orthogonal Validation of Field Effect and Early Precursor Signatures

To orthogonally validate our proteomic observations, we integrated our spatial DVP data with the published spatial transcriptomics from cancer-free organ donors and PDAC-adjacent pancreas (Supplementary Fig. S9A; ref. 10). Across matched biological groups, the majority of biological targets discussed in our study correlated positively (Supplementary Fig. S9B), including several proteins showcasing high levels of confirmatory signals (e.g., PTPN1 and SPINT2, Supplementary Fig. S9C). Notably, for stress/secretory pathway factors (CYB5A and AGR2), metabolic regulators such as PDK4, and immune-response components (STAT1), proteomics provided particularly robust signals while following similar trends compared with the transcriptome (Supplementary Fig. S9D).

We next performed multiplex immunofluorescence (mIF) to validate epithelial-intrinsic differences during tumorigenesis and to compare LG-iPanINs and LG-cPanINs, using an independent organ donor cohort and additional patient samples. When quantifying the percentage of marker-positive epithelial cells (Supplementary Fig. S10A), we observed a significant increase in $CKMT1A^{+}$ cells across early precursor groups (iADM, LG-iPanIN, and LG-cPanIN) in both iCohorts and cCohorts (Supplementary Fig. S10B and S10C), consistent with our proteomic findings (Supplementary Fig. S6C). The $FSCN1^{+}$ cells were selectively increased in tumors compared with cNormal and LG-cPanIN (Supplementary Fig. S10D and S10E), also in accordance with the proteome (Supplementary Fig. S6D). Field effect–associated proteins, including CYB5A (stress adaptation, Fig. 5D) and STAT1 (immune engagement, Supplementary Fig. S6A), confirmed an enrichment in LG-cPanINs relative to LG-iPanINs (Supplementary Fig. S10F–S10H). Metabolic reprogramming markers, PGAM1 and PDK4 (Supplementary Fig. S6B), displayed a consistent inverse trend between the two LG-PanIN groups, with PGAM1 increased and PDK4 decreased in LG-cPanINs (Supplementary Fig. S10I–S10K). Together, these orthogonal validations support the four core programs that emerge early in pancreatic tumorigenesis—stress adaptation, immune engagement, metabolic reprogramming, and mitochondrial remodeling—and anchor them to measurable protein markers in archival tissue.

Mutation-Specific Peptide Profiling of KRAS in Incidental Precursor Lesions

It has been reported that PanINs in patients with neoplastic pathologies in the pancreas, even when histologically LG, frequently harbor oncogenic *kRAS* hotspot mutations (14). To determine whether *kRAS* mutations are also present in precursor lesions from otherwise cancer-free individuals, we performed a targeted mutation-specific search, enabled by the consistent detection of *kRAS* peptides in our dataset, together with orthogonal Droplet Digital PCR (ddPCR). Exploratory ddPCR analysis using a *kRAS* G12/G13 probe set (including G12A, G12C, G12D, G12R, G12S, G12V, and G13D) on consecutive sections from a subset of individuals revealed *kRAS* G12/G13 mutation signals in incidental precursor lesions (Supplementary Fig. S11A and S11B), indicating that oncogenic *kRAS* mutations are present within the lesional epithelium of pancreata in cancer-free organ donors. Next, we assessed whether DVP could directly detect *kRAS* hotspot variants at

the peptide level. Mutant *kRAS* tryptic peptides (predominantly G12D and G12V, with occasional G13D) were identified across iADM and LG-iPanIN, whereas these variants were largely absent from iNormal ducts (Fig. 7A; Supplementary Fig. S11C; Supplementary Table S6). Among iADM lesions, 44% (4/9) of microdissected regions from two donors carried *kRAS* G12D. In LG-iPanINs, multiple *kRAS* variants were detected in spatially separated lesions, consistent with the previously reported polyclonal nature of precursor lesions (Fig. 7A; Supplementary Fig. S11C; ref. 14).

Overall, 46% of LG-iPanINs (12/26 lesions across four/eight donors) harbored *kRAS* mutations, including G12D (41.7%, 5/12 lesions, three donors), G12V (16.7%, 2/12 lesions, two donors), and G13D (50%, 6/12 lesions, two donors), with one lesion exhibiting polyclonality (coexpression of G12D and G13D, donor 1; Fig. 7B; Supplementary Table S6). Notably, several individuals with *kRAS* mutations were younger than 40 years of age (e.g., G13D in donor 6, age 38 years; G12D in donor 8, age 32 years), whereas multiple distinct mutations were more frequent in individuals in their 60s, suggesting a potential age-associated increase in mutational burden (Fig. 7B). Although cross-sectional sampling and FFPE-based validation limit temporal inference, ddPCR orthogonally confirmed the presence of G12D in iADM and LG-iPanIN, but not in iNormal ducts (Supplementary Fig. S11D and S11E). Collectively, these data indicate that similar to cPanINs, incidental precursor lesions from organ donors also harbor *kRAS* mutations and may, therefore, possess latent malignant potential. Our preliminary data also suggest that the frequency and spectrum of allelic variants in iPanINs might vary from those observed in cPanINs, although this requires validation in larger sample sets. Collectively, we demonstrate that DVP can recover actionable mutant peptides from minute archival lesions, enabling integrated proteogenotyping of early pancreatic neoplasia.

DISCUSSION

Our AI-powered spatial proteomic map provides a highresolution, protein-level characterization of both incidental and cancer-associated precursors in the development of PDAC. Our work defines molecular changes across five critical transitions: normal-to-ADM, ADM-to-PanIN, incidental-to-cancer-associated LG-PanIN, LG-to-HG PanIN, and PanIN-to-invasive carcinoma. Our central discovery—that histologically normal ducts from patients with cancer display significantly different molecular profiles than those from cancer-free individuals despite identical histopathologic presentation—represents an important advance with immediate implications for risk stratification and early intervention strategies.

Histologically normal ducts in cancer-bearing pancreas exhibited extensive molecular alterations despite their normal appearance under conventional microscopy, consistent with a field cancerization signature encompassing stress adaptation, immune engagement, and metabolic reprogramming. AGR2 was strongly induced in morphologically normal ducts (iNormal vs. cNormal, 5.4-fold higher in cNormal), indicating that cancer-associated proteomic reprogramming extends beyond histologically apparent lesions. CYB5A was similarly increased (8.1-fold), potentially reflecting a compensatory response to inflammatory and oxidative stress. These observations align with field cancerization

models in which paracrine signaling and stromal/inflammatory cues reshape the epithelial ecosystem and may contribute to postresection recurrence risk.

Our findings align with recent spatial transcriptomic analyses by Bell and colleagues (13) and Carpenter and colleagues (10), who documented transcriptional reprogramming in the PDAC microenvironment. However, our proteomic approach extends these observations by revealing actionable protein-level alterations with higher resolution between morphologically similar ductal lesions, particularly demonstrating more pronounced molecular differences between LG-iPanIN and LG-cPanIN lesions than reported at the transcriptome level. Although Carpenter and colleagues concluded that similarities far outweighed any differences between sporadic and tumor-associated PanINs at the RNA level (10), our proteomic data reveal substantial molecular divergence between these histologically similar lesions.

The coordinated inflammatory signatures in cancer-associated tissues raise a crucial question: Does inflammation contribute to neoplastic transformation or primarily reflect a response to emerging clones? We observed coordinated activation of STAT signaling pathways—including increased STAT3 and STAT1 in LG-cPanIN lesions—together with enrichment of IFN signaling and acute inflammatory response programs in cancer-associated normal ducts. These data indicate that inflammatory and IFN programs are already engaged at early, preinvasive stages and may create a permissive or selective microenvironment that shapes which precursor clones persist and progress. This interpretation is consistent with the experimental evidence that inflammatory cues can accelerate KRAS-driven pancreatic tumorigenesis (15).

A major advance of this study is the first systematic proteome-level characterization of HG-PanINs. We demonstrate that HG-cPanINs represent a molecularly distinct state characterized by a tumor-like proteomic program rather than a simple intensification of LG programs. The emergence of mitochondrial translation/respiration, *de novo* lipogenesis (FASN), and mitophagy components (SQSTM1, VDAC1, and PGAM5) suggests that HG-cPanINs operate under heightened bioenergetic and quality control demands, potentially creating metabolic vulnerabilities before invasion. Proteins uniquely shared by HG-cPanINs and tumors (e.g., UBE2C, CDC20, RUNX2, and IFITM3) provide a focused candidate panel for identifying HG dysplasia in preinvasive tissue.

Our mutation-specific peptide search provides a complementary, protein-level view of *kRAS* variation in early pancreatic neoplasia. Across microdissected regions of incidental lesions from cancer-free organ donors, we detected *kRAS* hotspot mutant peptides in iADM and LG-iPanIN lesions, consistent with the heterogeneous precursor evolution described by genomic studies (14). A subset of these findings aligns with orthogonal ddPCR data, in which we observed *kRAS* mutation status at the DNA level in these incidental precursor lesions. Because our study is cross-sectional and DNA validation from region-matched FFPE tissue can be technically challenging, we interpret these detections as evidence of feasibility rather than definitive lineage tracing. Nevertheless, the ability to recover mutant peptides from minute archival lesions highlights the potential of integrated proteogenotyping

to refine the molecular staging of precursors—particularly in settings in which DNA quality or quantity is limiting.

These discoveries enable new strategies for early detection. We propose a field effect biomarker panel comprising CYB5A, WTIP, B2M, and AGR2 for identifying high-risk field cancerization before visible lesions appear, potentially through the analysis of pancreatic fluid or liquid biopsies. Genetic studies estimate a median latency period of several years from the common ancestor to clinically apparent cancer (35), presenting a window to identify at-risk patients far earlier than current imaging-based approaches permit. The distinct molecular profiles of LG-cPanIN versus LG-iPanIN lesions provide criteria for risk stratification beyond histologic grading. Markers such as STAT1, CD44, and IFN signaling components could help distinguish lesions with more cancer-associated molecular features from those that seem incidental, informing surveillance strategies and intervention timing. An immunometabolic axis involving CYB5A/STAT1/PGAM1/PDK4 provides a candidate biomarker panel with mechanistic interpretability. These features nominate a mechanistically interpretable biomarker panel for identifying progression-permissive LG lesions in clinical specimens, and they highlight that LG histology does not equate to uniform molecular risk.

Several limitations warrant consideration. First, our study is cross-sectional and samples different lesion stages across individuals; future longitudinal and multilesion sampling within patients will be needed to directly map evolutionary trajectories. Second, although we enriched for ductal epithelium by laser microdissection, microscopic admixture with periductal stroma or acinar material remains possible and may contribute to some stage-associated proteins. However, this composite profile is relevant for signatures such as immune engagement and ECM remodeling. Nonetheless, we prioritized epithelial-intrinsic and biologically interpretable candidates and deprioritized obvious blood-contamination markers (e.g., hemoglobins) in biomarker-focused analyses. Integration with single cell–level approaches will be essential to further deconvolve epithelial versus microenvironmental contributions. Third, intralesional mutational heterogeneity within the ~100-cell regions profiled could not be resolved at single-cell resolution. Fourth, invasive PDAC exhibits substantial regional heterogeneity, which constrains one-to-one mapping of precursor states to tumor endpoints and motivates future studies integrating proteomics with spatial multiomics at higher spatial granularity. Future work should also validate these findings in larger cohorts, establish causal relationships between molecular alterations and disease progression, and develop clinical assays for promising biomarkers.

Although H-optimus-0 facilitated the identification of small PDAC precursor lesions and normal ducts, pathologist annotations remain the reference standard. Quantitative evaluations showed substantial agreement between computational clustering and expert annotations, while also indicating that expert review remains necessary for complete and biologically appropriate region selection (Supplementary Fig. S12A–S12C; “Methods”). In this setting, model-based clustering serves as a scalable preselection step for pathologist confirmation rather than a replacement for expert interpretation. For clinical application, prospective validation on independent cohorts with standardized annotation protocols would be required.

The integration of AI-driven image analysis with DVP, as demonstrated here for pancreatic cancer precursors, offers a generalizable framework applicable to other malignancies. Since its initial development (18), DVP has contributed to the growing field of spatial proteomics—the 2024 *Nature* Method of the Year (17)—and has been successfully applied to study tumorigenesis in other heterogeneous cancers, including borderline and LG serous ovarian tumors (36). Foundation model–based approaches have the potential to scale tumor heterogeneity analyses to thousands of samples while substantially reducing analysis time; our case-based analyses serve as proof of concept for such future large-scale applications. The compatibility of our pipeline with FFPE archival tissue, high sensitivity for minute protein quantities, and high-throughput resolution of tumor heterogeneity position this approach as a powerful tool for investigating precursor biology across diverse cancer types.

In conclusion, our spatial proteomics platform unveils a novel molecular landscape of PDAC precursors and significantly enhances our understanding of neoplastic transformation trajectories. Molecular reprogramming at the protein level precedes histologic changes, establishing a new paradigm for risk assessment and early intervention. By providing the proteomic counterpart to established genetic models of clonal evolution and spread, our work completes a more comprehensive picture of PDAC development. The molecular transitions we identified—from cancer-associated field effects to mitochondria-dependent programs in HG-PanINs—provide a framework for developing stage-specific biomarkers and therapeutic targeting strategies. This approach offers the potential to transform PDAC clinical management from late-stage treatment to early interception, when curative intervention remains possible.

METHODS

Study Design and Cohort

The study cohort included tissue specimens from 22 cancer-free organ donors and 13 patients with PDAC, including the validation cohort (Supplementary Table S7). Tissues used for MS-based proteomics were procured within 4 hours post mortem through the University of Nebraska Medical Center (UNMC) Normal Organ Recovery Program, in cooperation with Live On Nebraska, and through the Rapid Autopsy Program, or were obtained from surgical resections at UNMC and MD Anderson Cancer Center (MDACC). Organ donor pancreas tissues used for orthogonal validation were obtained from Donor Network West. Group assignment was based on clinical and histopathologic criteria. To minimize bias, all samples were double-pseudonymized prior to data generation and downstream analyses. The FFPE tissue blocks were sectioned at 3-μm thickness and mounted on precoated 2-mm PEN membrane slides (MicroDissect GmbH). A total of 24 tissue specimens were selected based on the availability and quality of FFPE blocks suitable for laser microdissection.

All lesions were reviewed by A. Maitra, an expert in pancreatic pathology with more than two decades of expertise in PDAC and PanINs (37–39), as “ground truth” for histologic classification. Donor pancreata were selected based on the absence of clinical pancreatic disease and were histologically normal apart from incidental precursor lesions. In organ donor pancreas, normal duct epithelium (iNormal), acinar-to-ductal metaplasia (iADM), and iPanINs (LG-iPanINs) were identified. In the DVP workflow, two organ donor samples

were excluded after quality control due to the absence of target lesions in consecutive sections, resulting in a total of 10 organ donor cases included in the downstream ansalysis. Additionally, nontumor tissues from 10 patients with PDAC were included in the proteomic analysis. Five of these samples contained LG-cPanINs and the corresponding normal ducts (cNormal), whereas the remaining five patient samples contributed HG-cPanIN lesions, with multiple subregions sampled for proteomic analysis. Tumor regions were analyzed from three PDAC cases with matched LG-cPanINs. We note that LG-cPanINs and the corresponding tumor samples were collected from different tissue sections on separate blocks. This spatial separation ensures that the molecular signatures identified in LG-cPanINs represent true field effects rather than contamination from adjacent tumor cells. A clear patient-to-sample relationship is provided in Supplementary Table S8. For orthogonal validation of protein expression through mIF, an independent cohort of 10 organ donors was included. For cancer-associated samples, eight samples from patients with PDAC were used in the validation, which consisted of five cases from the proteomic analysis and three additional new cases (Supplementary Table S7). The organ donor cohort spanned 23 to 69 years of age, and the patient cohort spanned 49 to 84 years of age. Both female and male subjects were included across all cohorts, with sex considered a biological variable.

All procedures were conducted in accordance with the Declaration of Helsinki and approved by the UNMC (IRB 091–01) and MDACC (IRB LAB05–0854) Institutional Review Board, with written informed consent obtained prior to enrollment. All patient and donor data were deidentified and anonymized. Cohort metadata, including age and gender, are provided in Supplementary Table S7.

AI Image Analysis

We prepared and stained with H&E and imaged as described in the Method Supplement. Each acquired whole-slide image (WSI) was stored in ".mrxs" format and processed using a custom tile extraction pipeline. Initially, each WSI was downsampled to a lower resolution to enable tissue-background segmentation. The resulting tissue regions were partitioned into nonoverlapping tiles of 224 × 224 pixels at a full resolution of 0.24 microns per pixel (mpp), corresponding to a physical tile size of 54 × 54 μm. This tile size matches the native input format of vision transformer–based foundation models, ensuring compatibility with the pretrained architecture of H-optimus-0. A tile resolution of 0.24 mpp was selected, as preliminary analyses indicated that this scale sufficiently preserves morphology relevant to downstream tasks. In the iCohort and cCohort, a total of 1,618,248 tiles were extracted, resulting in an average of 77,059 tiles per slide. In the HG-cPanIN cohort, a total of five slides were processed with 156,948 tiles. The extracted tiles were passed to H-optimus-0 (RRID: SCR_028225), a state-of-the-art foundation model for computational pathology that demonstrates strong performance on clinical benchmarking datasets (40, 41). The model consists of 1.1 billion parameters and was pretrained on more than 500,000 WSIs with no further fine-tuning on custom data. Model inference was performed on a workstation with two NVIDIA RTX 4090 GPUs. For further details on H&E tiling, tile embedding, and alignment to proteomic measurements, see Mathian and colleagues (42). These technical specifications, combined with the public availability of H-optimus-0, enable independent replication of our analytic framework.

We performed two clustering strategies on tiles depending on the cohort: (i) global clustering for cCohort and iCohort, including the HG-cPanIN subgroup, and (ii) local clustering for the PDAC cohort. In the global setting, Faiss k-means clustering was applied to tile-level feature vectors aggregated across all slides within a cohort and $k \in \{2,\ldots,50\}$. To identify optimal k clusters that reflect areas of interest, k was adjusted interactively based on expert input on a subset of slides from the respective cohort. Based on this input, we set k = 23 for the iCohort, k = 25 for the HG-cPanIN cohort, and k = 10 for the cCohort and assigned tiles to clusters based on the cosine similarity of tile features and cluster centroids. This global strategy ensures the detection of consistent morphologic patterns shared across the entire cohort. We applied local clustering to the PDAC cohort on a per-slide basis. In this study, the number of clusters varied within $k \in \{2,\ldots,10\}$ per slide until the tumor region split into two distinct phenotype clusters. This approach allows replicate sampling of intraslide tumor heterogeneity, enabling the investigation of morphologically distinct tumor cell states within individual patients rather than cross-patient comparisons. All slides were subjected to human pathology review. Throughout the article, we refer to each cluster as a "region." Guided by H&E-based clustering, we exported image masks for the collection of 100 phenotypically matched cells from the respective biological lesions and normal ducts. To equalize the protein input amount in the mass spectrometer, an additional size filter was used to select 100 phenotypically matched cells matching a total area threshold of 18,590 μm^2.

Evaluation of k for Lesion-Specific Clustering

To determine the optimal number of clusters (k) for the k-means clustering of tile-level features, we developed an annotation-guided evaluation framework. This analysis aimed to determine the cohort-specific values of k that produce the highest agreement between a single cluster and the corresponding pathologist-annotated region, as measured by the F1 score. We demonstrated this approach for three lesion types combined: PanIN, epithelial linings (normal ducts), and ADM. An expert pathologist annotated the regions of interest (ROI) across all slides of each cohort using polygon annotations stored in GeoJSON format. Each tile was then spatially matched to the annotation polygons. A tile was classified as lying within an annotated region if more than half of its area overlapped with an annotation polygon; all remaining tiles were classified as outside the annotation. This overlap threshold ensures that only tiles predominantly covered by annotated tissue are treated as ground-truth positives, avoiding bias from tiles that only marginally intersect an annotation boundary. For every candidate k (k = 2–70), global k-means clustering was performed on the pooled tile features of each cohort. For each resulting clustering, we identified the single best-matching cluster as the one containing the largest number of tiles that colocalized with the annotation (i.e., the cluster maximizing the true positive count). The tiles assigned to this best-matching cluster were then treated as the predicted positive set, and precision, recall, and F1 were computed. These metrics were plotted as a function of k for each cohort (Supplementary Fig. S12). The optimal k was selected as the value that maximized F1 within a plateau; when multiple values of k achieved comparable F1, the smallest k was preferred to minimize model complexity. This led to optimal values of k = 47 for iPanIN, k = 62 for cPanIN, and k = 50 for HG-cPanIN. Although the present analysis was performed on the full set of annotated slides per cohort, the framework is designed to generalize such that pathologist

annotations on a representative subset of slides are sufficient to optimize k, which can then be applied cohort-wide. This makes the approach scalable to larger cohorts in which exhaustive annotation would be impractical.

Data Analysis and Statistics

All analyses were performed in the R statistical environment (version 4.3.1) using custom scripts and established packages for proteomics and pathway analysis. MS intensities were $\log_2$-transformed and normalized across samples; proteins were filtered for missingness and quantified at the protein-group level (Supplementary Table S1). For differential abundance testing between biological groups, we used moderated linear modeling with empirical Bayes shrinkage, and we controlled the false discovery rate using the Benjamini–Hochberg adjustment (Supplementary Table S2). For pairwise group comparisons displayed in boxplots, unpaired Student *t* tests were performed assuming equal variances; significance thresholds were: ns, *P* values >0.05; *, *P* values ≤0.05; **, *P* values ≤0.01; ***, *P* values ≤0.001; and ****, *P* values ≤0.0001. For multigroup testing across the PanIN continuum, we applied targeted ANOVA with multiple testing correction and clustered significant proteins to define stage-structured modules (Supplementary Table S5). Gene set enrichment analysis was performed using curated pathway collections, and enriched pathways were reported with adjusted *P* values (Supplementary Table S3). Unless otherwise stated, statistical tests were two-sided. GraphPad Prism was used for the generation of *kRAS* variant plots.

Search for KRAS Variants

Mutational FASTA sequences were generated by *in silico* mutagenesis of the *kRAS* reference sequence by substituting the reference amino acid at the specified residue position with the corresponding mutant residue (e.g., G12D). Mutated sequences were annotated with headers reflecting the introduced mutation. The set of mutations was selected based on the most frequently observed *kRAS* variants reported in Fig. 3 of Huang and colleagues (43), encompassing colorectal, pancreatic, and lung cancers: G12A, G12C, G12D, G12R, G12V, G13C, G13D, Q61H, Q61R, and A146T. These variants represent the most frequently observed mutations in *kRAS*-driven cancers. Combined with the common FASTA file, the custom file was used as an additional reference during data processing in DIA-NN 1.8.1 for the main quantification of this article. To increase the identification confidence of altered peptides, *kRAS* variants were searched without the match-between-runs function. The DIA-NN output file “report.tsv” was used for data evaluation after additional filtering for peptides quantified with equal to or less than 1% on the “Q.Value,” “Global.Q.Value,” and the “Global.PG.Q.Value.” Detailed sample counts included in the analysis, together with the corresponding Q-values for each *kRAS* variant in samples from both the iCohort and cCohort, are provided in Supplementary Table S6. Because the presence of *kRAS* mutations in PanINs from patients with PDAC has already been well documented (14), *kRAS* mutation analyses were conducted exclusively on incidental precursor lesions in the cancer-free donor cohort. For visualization of the altered KRAS peptide with the amino acid sequence “LVVVGAGGVGK,” hits were either counted per biological cohort or statistical q-values were plotted. The bar plots were generated using GraphPad Prism 10.

mIF

The mIF was performed on FFPE tissue sections from an independent organ donor cohort ($n = 10$) and additional patient samples ($n = 8$, including three new cases) to validate epithelial-intrinsic differences identified in our proteomic analyses. The sections under-went heat-induced epitope retrieval at 120°C using a citrate-based target retrieval solution (Agilent, S1699), followed by blocking with 5% goat serum in PBS for 1 hour at ambient temperature. Primary antibodies were applied overnight at 4°C and detected with secondary antibodies incubated for 1 hour at ambient temperature. The tissues were mounted with ProLong Diamond antifade reagent containing DAPI (Invitrogen). The images were acquired using a ZEISS fluorescence microscope with AxioVision software (RRID: SCR_002677) at 20x magnification. Ductal ROIs were annotated by histopathology, and per-ROI fractions of marker-positive ductal cells were quantified using QuPath following DAPI-based nuclear segmentation (Supplementary Fig. S10A). Group comparisons were performed using an unpaired Student *t* test with the Welch correction. Detailed antibody information is provided in Supplementary Table S9.

ddPCR

The genomic DNA was extracted from laser-microdissected regions from consecutive FFPE sections using the QIAamp DNA FFPE Advanced Kit (QIAGEN). The ddPCR was performed using the *kRAS* G12/G13 Screening Kit (Bio-Rad), which enables multiplex detection of *kRAS* hotspot variants (G12V, G12D, G12R, G12C, G12S, G12A, and G13D), or with a variant-specific *kRAS* G12D probe (Bio-Rad; identifier dHsaMDV2510596), on a QX200 ddPCR System (Bio-Rad), following established protocols (44, 45). Each assay included a *kRAS* wild-type control from a healthy individual and a mutant control from a pancreatic cancer cell line to estimate false-positive rates across samples, ensure the absence of contamination, and verify probe performance and ddPCR efficiency on each plate. Given the small and heterogeneous precursor regions analyzed, the samples were classified as mutation-positive if at least one mutant droplet was detected. The results are presented as bar plots showing the number of microdissected regions per biological group, together with the corresponding percentage of mutation-positive regions, generated using GraphPad Prism 10.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Tissues used in this study were obtained from the University of Nebraska Medical Center through the Rapid Autopsy Program and Normal Organ Recovery Program in cooperation with Live On Nebraska, as well as from the MD Anderson Cancer Center. We gratefully acknowledge and thank the donors, patients, and their families for their invaluable contributions to scientific research. This work was partially funded by Resolute Bio. A. Maitra was supported by NIH grants U54CA274371, U01CA200468, and U24CA274274. P.M. Grandgenett and M.A. Hollingsworth were supported by U01CA210240, P30CA36727, R50CA211462, U01CA284086, and U54CA274329. J. Min was supported by a Pilot/Feasibility Award from the Texas Medical Center Digestive Disease Center (P30DK056338) and a Seed Grant from the Hirshberg Foundation for Pancreatic Cancer Research. H. Wang was supported by P01CA117969 and P50CA221707. V. Bernard was supported by the 2025 Lustgarten Foundation-AACR Career Development Award for Pancreatic Cancer Research, in honor of John Robert Lewis (25-20-67-BERN). We thank Constantin Ammar for the constructive discussion about the *kRAS* variant search and

A. T. Reckard for technical assistance. The authors used Claude (Anthropic) to improve the quality of the writing for specific sections of the manuscript.

Authors' Disclosures

L. Schweizer reports employment with Resolute Bio during the conduct of the study and a patent for EP P2710EP00, 2025, pending. G. Zonderland reports employment with Resolute Bio during the conduct of the study. J.H. Thomsen reports employment with Resolute Bio during the conduct of the study. L. Oldenburg reports employment with Resolute Bio during the conduct of the study. S.D. Hernandez reports employment with Resolute Bio during the conduct of the study. G. Jez reports employment with Resolute Bio during the conduct of the study. B.J. Swanson reports other support from Cogent Biosciences outside the submitted work. I. Ummat reports employment with Resolute Bio during the conduct of the study and a patent for EP P2710EP00, 2025, pending. M.T. Strauss reports a patent for EP P2710EP00, 2025, pending and employment with Resolute Bio. A. Mund reports employment with Resolute Bio during the conduct of the study and a patent for EP P2710EP00, 2025, pending. A. Maitra reports royalties from a patent licensed to Exact Sciences from Johns Hopkins University. No disclosures were reported by the other authors.

Data Availability

Pseudonymized MS-based proteomics data of this study have been deposited in the ProteomeXchange Consortium via the PRIDE partner repository (RRID: SCR_003411) with the dataset identifier PXD072559. Imaging raw files and all custom code used for data analysis have been deposited on Zenodo (RRID: SCR_004129) under accession 19110683 (https://doi.org/10.5281/zenodo.19110683), enabling end-to-end reproduction of the analysis from raw images to final results. The repository includes a self-contained Jupyter Notebook implementing the morphology-guided clustering workflow (H-optimus-0 feature extraction, k-means clustering, and sample selection) with all study-specific parameters and thresholds, a Jupyter Notebook benchmarking clustering performance against pathologist annotations, as well as all R scripts for the proteomics data analysis, including a README mapping each script to its corresponding figure.

REFERENCES

1. Halbrook CJ, Lyssiotis CA, Pasca di Magliano M, Maitra A. Pancreatic cancer: advances and challenges. Cell 2023;186:1729–54. [PubMed: 37059070]
2. Fischer CG, Wood LD. From somatic mutation to early detection: insights from molecular characterization of pancreatic cancer precursor lesions. J Pathol 2018;246:395–404. [PubMed: 30105857]
3. Ying H, Kimmelman AC, Bardeesy N, Kalluri R, Maitra A, DePinho RA. Genetics and biology of pancreatic ductal adenocarcinoma. Genes Dev 2025;39:36–63. [PubMed: 39510840]
4. Giroux V, Rustgi AK. Metaplasia: tissue injury adaptation and a precursor to the dysplasia-cancer sequence. Nat Rev Cancer 2017;17: 594–604. [PubMed: 28860646]
5. Li AL, Sugiura K, Nishiwaki N, Suzuki K, Sadeghian D, Zhao J, et al. FRA1 controls acinar cell plasticity during murine $Kras^{G12D}$-induced pancreatic acinar to ductal metaplasia. Dev Cell 2024;59: 3025–42.e7. [PubMed: 39178842]
6. Hruban RH, Adsay NV, Albores-Saavedra J, Compton C, Garrett ES, Goodman SN, et al. Pancreatic intraepithelial neoplasia: a new nomenclature and classification system for pancreatic duct lesions. Am J Surg Pathol 2001;25:579–86. [PubMed: 11342768]
7. Ren B, Liu X, Suriawinata AA. Pancreatic ductal adenocarcinoma and its precursor lesions: histopathology, cytopathology, and molecular pathology. Am J Pathol 2019;189:9–21. [PubMed: 30558727]
8. Hosoda W, Chianchiano P, Griffin JF, Pittman ME, Brosens LA, Noë M, et al. Genetic analyses of isolated high-grade pancreatic intraepithelial neoplasia (HG-PanIN) reveal paucity of alterations in TP53 and SMAD4. J Pathol 2017;242:16–23. [PubMed: 28188630]

9. Hutchings D, Waters KM, Weiss MJ, Wolfgang CL, Makary MA, He J, et al. Cancerization of the pancreatic ducts: demonstration of a common and under-recognized process using immunolabeling of paired duct lesions and invasive pancreatic ductal adenocarcinoma for p53 and Smad4 expression. Am J Surg Pathol 2018;42:1556–61. [PubMed: 30212393]
10. Carpenter ES, Elhossiny AM, Kadiyala P, Li J, McGue J, Griffith BD, et al. Analysis of donor pancreata defines the transcriptomic signature and microenvironment of early neoplastic lesions. Cancer Discov 2023;13:1324–45. [PubMed: 37021392]
11. Pian LL, Song MH, Wang TF, Qi L, Peng TL, Xie KP. Identification and analysis of pancreatic intraepithelial neoplasia: opportunities and challenges. Front Endocrinol (Lausanne) 2024;15:1401829. [PubMed: 39839479]
12. Srivastava S, Koay EJ, Borowsky AD, De Marzo AM, Ghosh S, Wagner PD, et al. Cancer overdiagnosis: a biological challenge and clinical dilemma. Nat Rev Cancer 2019;19:349–58. [PubMed: 31024081]
13. Bell ATF, Mitchell JT, Kiemen AL, Lyman M, Fujikura K, Lee JW, et al. PanIN and CAF transitions in pancreatic carcinogenesis revealed with spatial data integration. Cell Syst 2024;15:753–69.e5. [PubMed: 39116880]
14. Braxton AM, Kiemen AL, Grahn MP, Forjaz A, Parksong J, Mahesh Babu J, et al. 3D genomic mapping reveals multifocality of human pancreatic precancers. Nature 2024;629:679–87. [PubMed: 38693266]
15. Liu Y, Deguchi Y, Wei D, Liu F, Moussalli MJ, Deguchi E, et al. Rapid acceleration of KRAS-mutant pancreatic carcinogenesis via remodeling of tumor immune microenvironment by PPARdelta. Nat Commun 2022;13:2665. [PubMed: 35562376]
16. Curtius K, Wright NA, Graham TA. An evolutionary perspective on field cancerization. Nat Rev Cancer 2018;18:19–32. [PubMed: 29217838]
17. Method of the Year 2024: spatial proteomics. Nat Methods 2024; 21:2195–6. [PubMed: 39643689]
18. Mund A, Coscia F, Kriston A, Hollandi R, Kovács F, Brunner AD, et al. Deep Visual Proteomics defines single-cell identity and heterogeneity. Nat Biotechnol 2022;40:1231–40. [PubMed: 35590073]
19. Nordmann TM, Mund A, Mann M. A new understanding of tissue biology from MS-based proteomics at single-cell resolution. Nat Methods 2024;21:2220–2. [PubMed: 39643675]
20. Haldar SD, Vilar E, Maitra A, Zaidi N. Worth a pound of cure? Emerging strategies and challenges in cancer immunoprevention. Cancer Prev Res (Phila) 2023;16:483–95. [PubMed: 37001882]
21. Basturk O, Adsay NV. Early cancerous lesions of the pancreas and ampulla: current concepts and challenges. Gastroenterol Clin North Am 2024;53:57–84. [PubMed: 38280751]
22. Yamamoto N, Dobersch S, Loveless I, Samraj AN, Jang GH, Haraguchi M, et al. HMGA2 expression predicts subtype, survival, and treatment outcome in pancreatic ductal adenocarcinoma. Clin Cancer Res 2025;31:733–45. [PubMed: 39680021]
23. Westphalen CB, Takemoto Y, Tanaka T, Macchini M, Jiang Z, Renz BW, et al. Dclk1 defines quiescent pancreatic progenitors that promote injury-induced regeneration and tumorigenesis. Cell Stem Cell 2016;18: 441–55. [PubMed: 27058937]
24. Xu Y, Wang X, Li Y, Mao Y, Su Y, Mao Y, et al. Multimodal single cell-resolved spatial proteomics reveal pancreatic tumor heterogeneity. Nat Commun 2024;15:10100. [PubMed: 39572534]
25. Pei G, Min J, Rajapakshe KI, Branchi V, Liu Y, Selvanesan BC, et al. Spatial mapping of transcriptomic plasticity in metastatic pancreatic cancer. Nature 2025;642:212–21. [PubMed: 40269162]
26. Orozco CA, Martinez-Bosch N, Guerrero PE, Vinaixa J, Dalotto-Moreno T, Iglesias M, et al. Targeting galectin-1 inhibits pancreatic cancer progression by modulating tumor-stroma crosstalk. Proc Natl Acad Sci U S A 2018;115:E3769–E3778. [PubMed: 29615514]
27. Vinaixa J, Martínez-Bosch N, Gibert J, Manero-Rupérez N, Santofimia-Castaño P, Baudou FG, et al. Nuclear Galectin-1 promotes KRAS-dependent activation of pancreatic cancer stellate cells. Proc Natl Acad Sci U S A 2025;122:e2424051122. [PubMed: 40172967]
28. Maitra A, Iacobuzio-Donahue C, Rahman A, Sohn TA, Argani P, Meyer R, et al. Immunohistochemical validation of a novel epithelial and a novel stromal marker of pancreatic

ductal adenocarcinoma identified by global expression microarrays: sea urchin fascin homolog and heat shock protein 47. Am J Clin Pathol 2002;118: 52–9. [PubMed: 12109856]

29. Yoneda A, Minomi K, Tamura Y. Heat shock protein 47 confers chemoresistance on pancreatic cancer cells by interacting with calreticulin and IRE1alpha. Cancer Sci 2021;112:2803–20. [PubMed: 34109710]
30. Yachida S, Jones S, Bozic I, Antal T, Leary R, Fu B, et al. Distant metastasis occurs late during the genetic evolution of pancreatic cancer. Nature 2010;467:1114–7. [PubMed: 20981102]
31. Williams HL, Dias Costa A, Zhang J, Raghavan S, Winter PS, Kapner KS, et al. Spatially resolved single-cell assessment of pancreatic cancer expression subtypes reveals Co-expressor phenotypes and extensive intratumoral heterogeneity. Cancer Res 2023;83:441–55. [PubMed: 36459568]
32. Ali HA, Karasinska JM, Topham JT, Johal D, Kalloger S, Metcalfe A, et al. Pancreatic cancer tumor organoids exhibit subtype-specific differences in metabolic profiles. Cancer Metab 2024;12:28. [PubMed: 39363341]
33. Ma Z, Lytle NK, Chen B, Jyotsana N, Novak SW, Cho CJ, et al. Single-cell transcriptomics reveals a conserved metaplasia program in pancreatic injury. Gastroenterology 2022;162:604–20.e20. [PubMed: 34695382]
34. Schlesinger Y, Yosefov-Levi O, Kolodkin-Gal D, Granit RZ, Peters L, Kalifa R, et al. Single-cell transcriptomes of pancreatic preinvasive lesions and cancer reveal acinar metaplastic cells' heterogeneity. Nat Commun 2020;11:4516. [PubMed: 32908137]
35. Makohon-Moore AP, Matsukuma K, Zhang M, Reiter JG, Gerold JM, Jiao Y, et al. Precancerous neoplastic cells can move through the pancreatic ductal system. Nature 2018;561:201–5. [PubMed: 30177826]
36. Schweizer L, Kenny HA, Krishnan R, Kelliher L, Bilecz AJ, Heide J, et al. Spatial proteo-transcriptomic profiling reveals the molecular landscape of borderline ovarian tumors and their invasive progression. Cancer Cell 2025;43:1495–511.e7. [PubMed: 40578359]
37. Hruban RH, Takaori K, Klimstra DS, Adsay NV, Albores-Saavedra J, Biankin AV, et al. An illustrated consensus on the classification of pancreatic intraepithelial neoplasia and intraductal papillary mucinous neoplasms. Am J Surg Pathol 2004;28:977–87. [PubMed: 15252303]
38. Maitra A, Adsay NV, Argani P, Iacobuzio-Donahue C, De Marzo A, Cameron JL, et al. Multicomponent analysis of the pancreatic adenocarcinoma progression model using a pancreatic intraepithelial neoplasia tissue microarray. Mod Pathol 2003;16:902–12. [PubMed: 13679454]
39. Maitra A, Fukushima N, Takaori K, Hruban RH. Precursors to invasive pancreatic cancer. Adv Anat Pathol 2005;12:81–91. [PubMed: 15731576]
40. Breen J, Allen K, Zucker K, Godson L, Orsi NM, Ravikumar N. A comprehensive evaluation of histopathology foundation models for ovarian cancer subtype classification. NPJ Precis Oncol 2025;9:33. [PubMed: 39885243]
41. Campanella G, Chen S, Singh M, Verma R, Muehlstedt S, Zeng J, et al. A clinical benchmark of public self-supervised pathology foundation models. Nat Commun 2025;16:3640. [PubMed: 40240324]
42. Mathian É, Oldenburg L, Chelebian E, Schweizer L, Zonderland G, Egebjerg K, et al. In: Erdt M, Chen Y, Wesarg S, Drechsler K, Freiman M, Thomas S, et al., editors. TileDVP: decoding the tissue proteome from H&E images. Cham (Switzerland): Springer Nature; 2026. p. 21–31.
43. Huang L, Guo Z, Wang F, Fu L. KRAS mutation: from undruggable to druggable in cancer. Signal Transduct Target Ther 2021;6:386. [PubMed: 34776511]
44. Bernard V, Kim DU, San Lucas FA, Castillo J, Allenson K, Mulu FC, et al. Circulating nucleic acids are associated with outcomes of patients with pancreatic cancer. Gastroenterology 2019;156:108–18.e4. [PubMed: 30240661]
45. Semaan A, Bernard V, Wong J, Makino Y, Swartzlander DB, Rajapakshe KI, et al. Integrated molecular characterization of intraductal papillary mucinous neoplasms: an NCI cancer moonshot precancer atlas Pilot Project. Cancer Res Commun 2023;3:2062–73. [PubMed: 37721516]

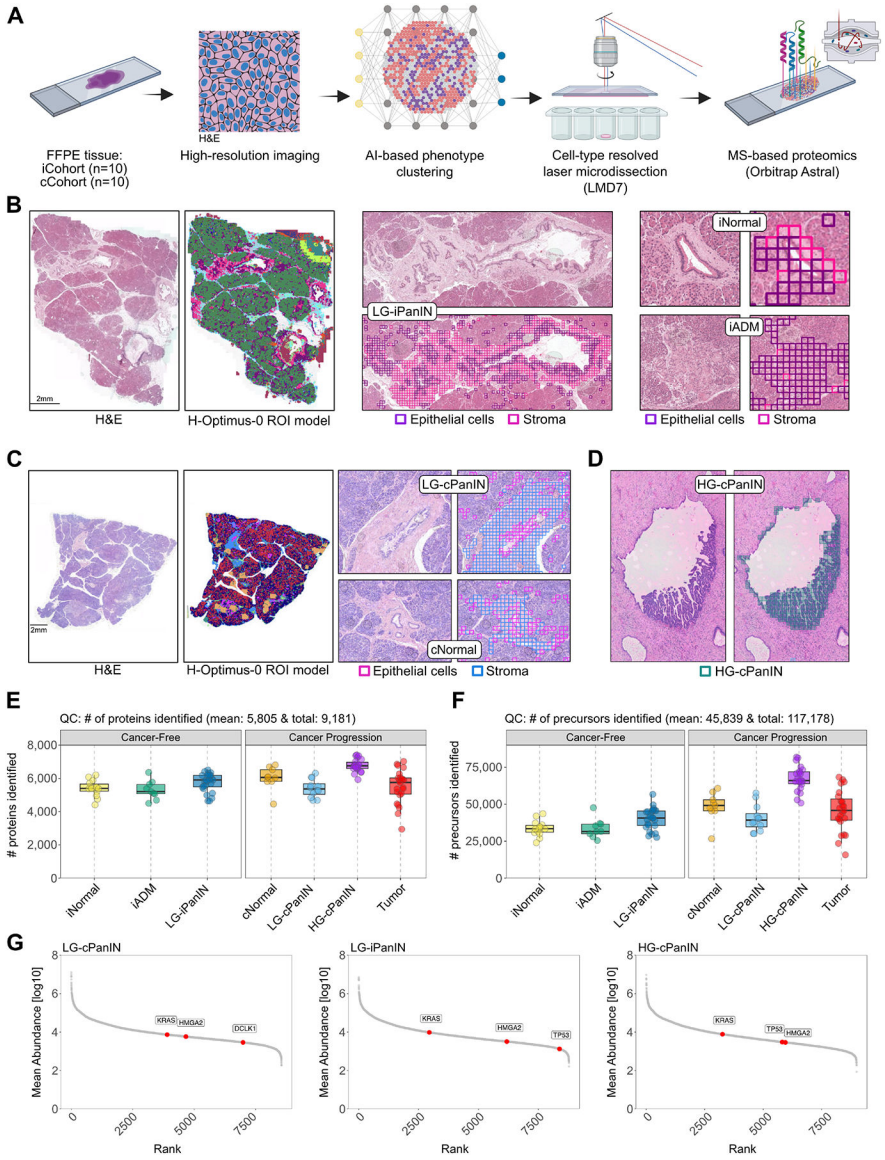

Figure 1.

DVP workflow and cohort overview. **A**, Schematic overview of the DVP workflow. FFPE tissue specimens were H&E stained. High-content images were analyzed using AI-based phenotype clustering to extract tissue morphologies and disease-associated pathologies. Cells were extracted using a Leica LMD7 laser-microdissection instrument, and cell type–specific proteomes were acquired on an Orbitrap Astral mass spectrometer. **B–D,** Pathology foundation models with representative images. The H-optimus-0 ROI model was applied to H&E images to identify LG-iPanIN, iNormal, and iADM lesions (**B**), LG-cPanIN and cNormal (**C**), and HG-cPanIN (**D**). **E** and **F,** Number of proteins (**E**) and peptide precursors (**F**) quantified across all cohort groups using MS-based proteomics. **G,** Mean ranked abundance of proteins found in LG-iPanIN, LG-cPanIN, and HG-cPanIN groups. The abundance of key markers is highlighted in red (KRAS, HMGA2, DCLK1, and TP53).

Normal, normal ductal epithelium; Tumor, tumor tissue. Sample overview: Details on sample counts and origins are provided in Supplementary Table S8.

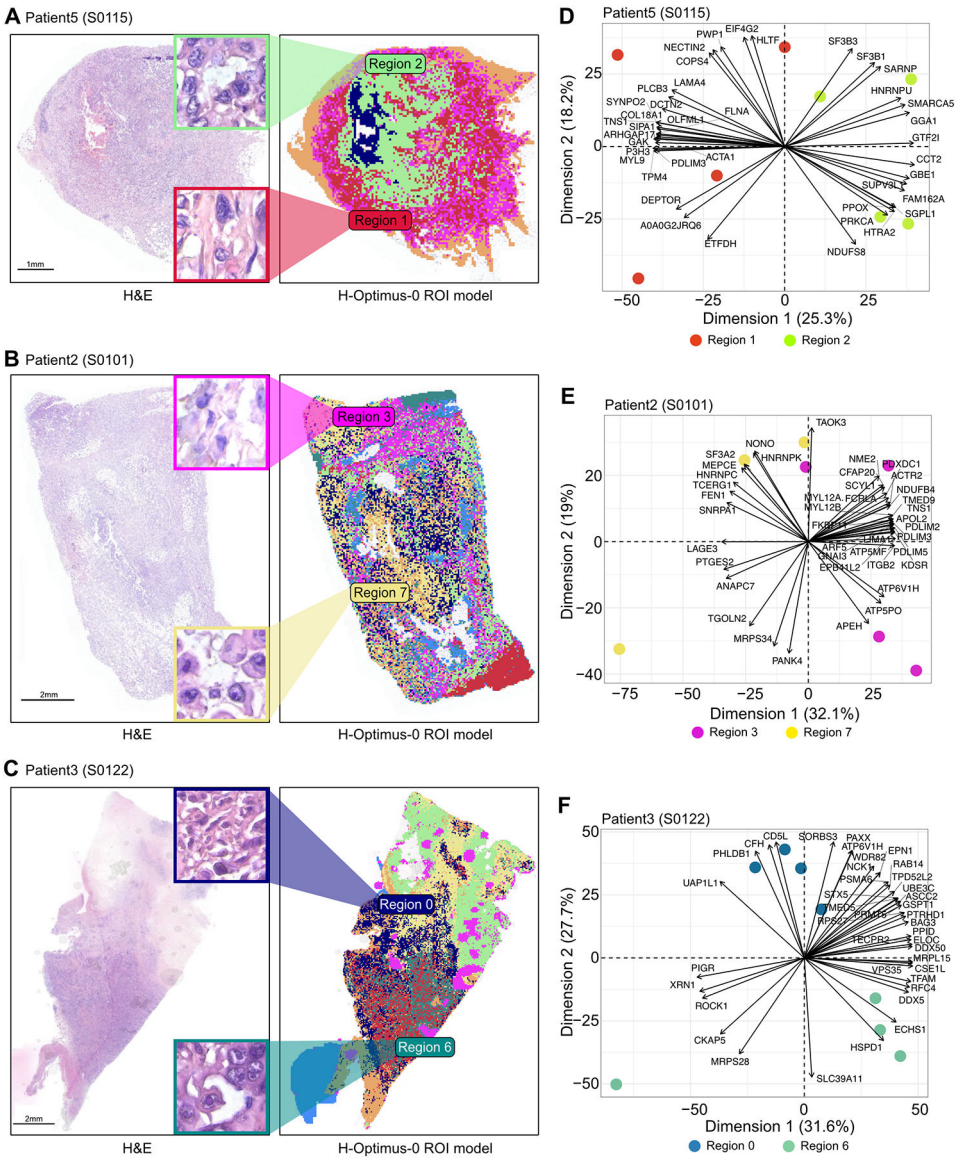

Figure 2.

Protein signatures associated with PDAC intratumoral heterogeneity. **A–C,** H&E images with a clustering analysis using an H-optimus-0 model, with representative enlarged areas showing distinct intratumor heterogeneity in patient 5 (regions 1 and 2, **A**), patient 2 (regions 3 and 7, **B**), and patient 3 (regions 0 and 6, **C**). **D,** PCAs showing differential protein expression between the regions in patient 5, with region 1 enriched for ECM-remodeling proteins, cytoskeletal, and adhesion proteins, whereas region 2 showed a higher abundance of RNA-processing factors and mitochondrial or stress-associated proteins. **E,** PCA showing regions separated along an axis ranging from nuclear RNA-binding and splicing proteins in region 7 to cytoskeletal and contractile components in region 3 of patient 2. **F,** PCA showing one region displayed a higher abundance of lysosomal/vacuolar and immune-associated proteins (region 0), whereas the other was enriched for mitochondrial and metabolic proteins (region 6) in patient 3.

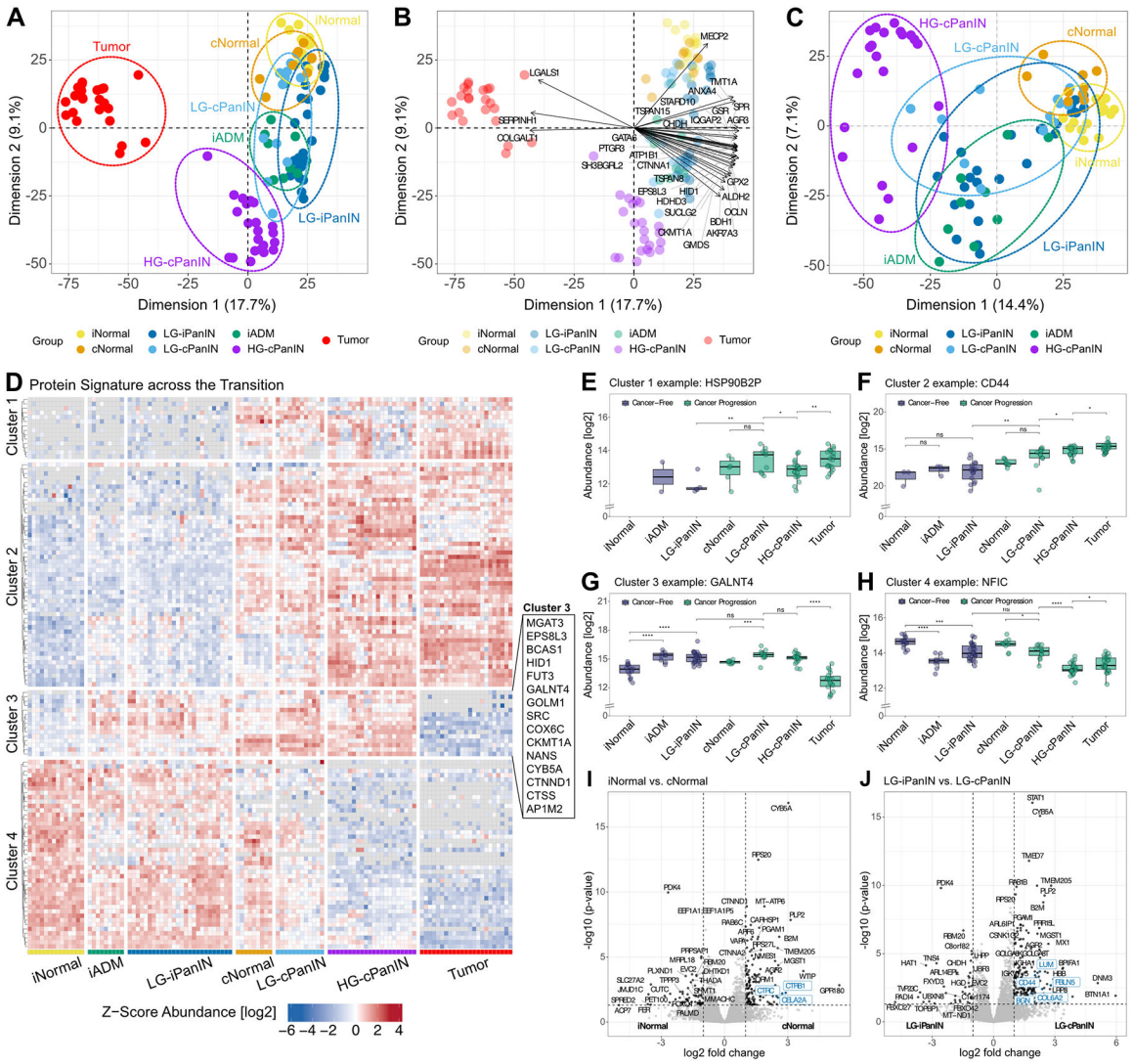

Figure 3.

Molecular signatures across the precancer-to-PDAC transition. **A–C,** PCAs of indicated cohort groups: PCA of all cohort groups including i/cNormal, iADM, LG-i/cPanIN, HG-cPanIN, and Tumor (**A**) with its major drivers of data variance highlighted in **B**, PCA of i/cNormal, iADM, LG-i/cPanIN, and HG-cPanIN (**C**). Dimensions of highest variance are shown (dimensions 1 and 2). **D,** Heatmap showing protein signatures associated with normal-to-PanIN-to-PDAC tumorigenesis. ANOVA significance analysis (*P* value < 1e–10) and post-Tukey extraction of proteins with significant differences identified proteins that significantly altered between cancer-free and cancer-associated samples while exhibiting similar abundance between cNormal and LG-cPanIN or between HG-cPanIN and PDAC. Proteins from cluster 3, which are associated with early precursor programs, are highlighted. Colors represent z-score–normalized protein abundances (log_2). **E–H,** Boxplots of representative proteins from each cluster: HSP90B2P (cluster 1; **E**), CD44 (cluster 2; **F**), GALNT4 (cluster 3; **G**), and NFIC (cluster 4; **H**). An unpaired Student *t* test was performed assuming equal variances for compared groups. ns, *P* values >0.05; *, *P* values ≤0.05; **,

P values ≤0.01; ***, *P* values ≤0.001; ****, *P* values ≤0.0001. **I** and **J,** Differential protein expression of morphologically equivalent lesions between the cancer-free (incidental) and cancer-associated cohorts: iNormal vs. cNormal (**I**) and LG-iPanIN vs. LG-cPanIN (**J**). Significantly regulated proteins [fold change (log_2) > 2; *P* value <0.05] are highlighted in black. Three digestive enzyme precursors are highlighted in blue in (**I**). Proteins associated with ECM remodeling and adhesion are highlighted in blue in **J**. Normal, normal ductal epithelium; Tumor, tumor tissue. Sample overview: Details on sample counts and origins are provided in Supplementary Table S8.

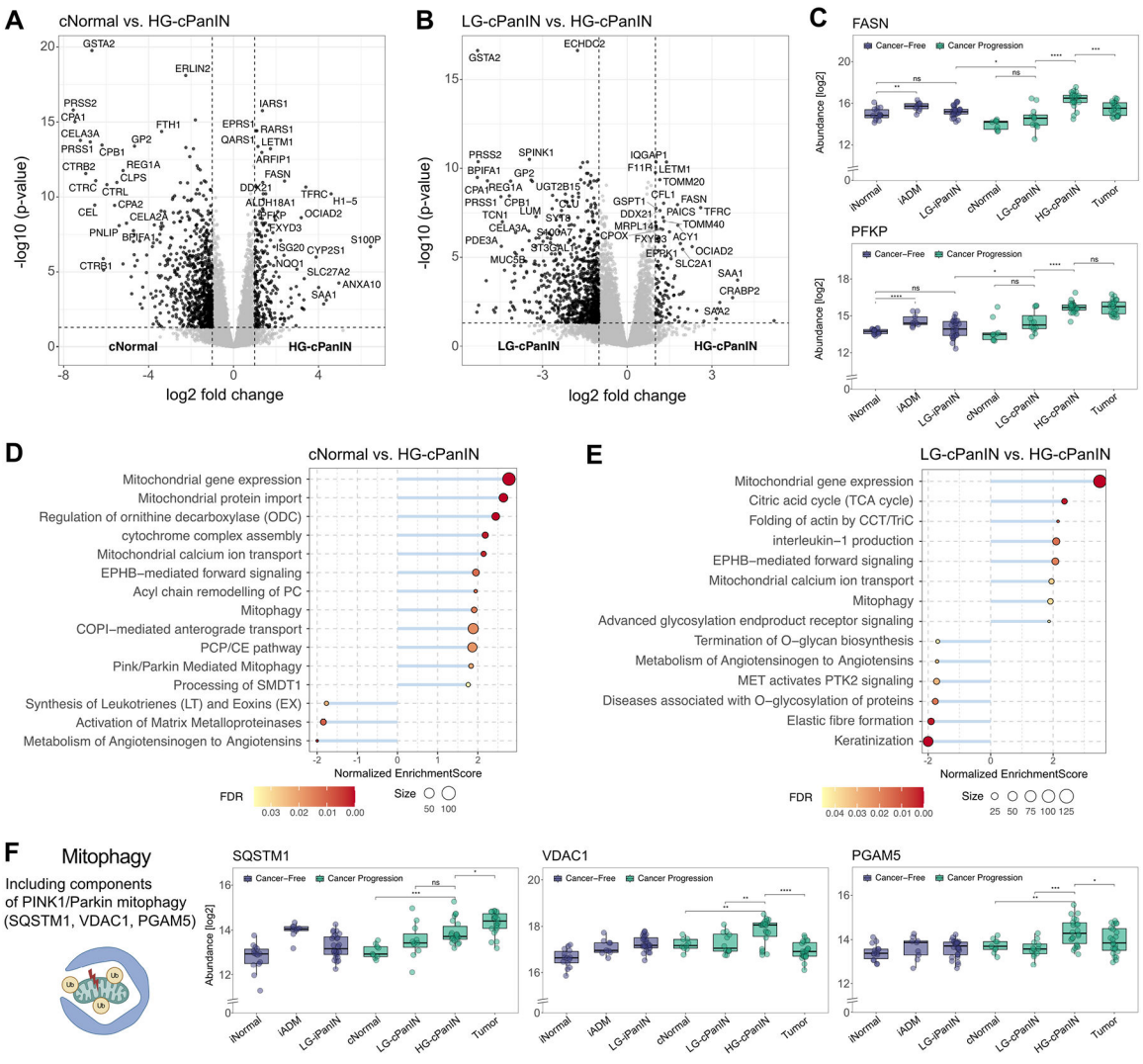

Figure 4.

Proteomic signatures of HG-cPanIN lesions. **A** and **B,** Differential protein expression within the cancer-associated cohort comparing cNormal vs. HG-cPanIN (**A**) and LG-cPanIN vs. HG-cPanIN (**B**). Significantly regulated proteins [fold change (log_2) > 2; *P* value <0.05] are highlighted in black. **C,** Representative metabolic proteins exhibiting upregulated expression patterns between cNormal and LG-cPanIN to HG-cPanIN: FASN (top) and PFKP (bottom). An unpaired Student *t* test was performed assuming equal variances for compared groups. ns, *P* values >0.05; *, *P* values ≤0.05; **, *P* values ≤0.01; ***, *P* values ≤0.001; ****, *P* values ≤0.0001. **D** and **E,** Representative biological pathways enriched using gene set enrichment analysis comparing cNormal vs. HG-cPanIN (**D**) and LG-cPanIN vs. HG-cPanIN (**E**). FDR values are visualized by a color gradient, whereas the number of proteins enriched for a pathway is indicated by point size. **F,** Representative proteins from PINK1/Parkin mitophagy exhibiting upregulated expression patterns between cNormal and LG-cPanIN to HG-cPanIN: SQSTM1, VDAC1, and PGAM5. Statistical testing was performed as in **C**. Normal, normal ductal epithelium; TCA, tricarboxylic acid; Tumor,

tumor tissue; CCT: Chaperonin containing T-Complex Protein 1. Sample overview: Details on sample counts and origins are provided in Supplementary Table S8.

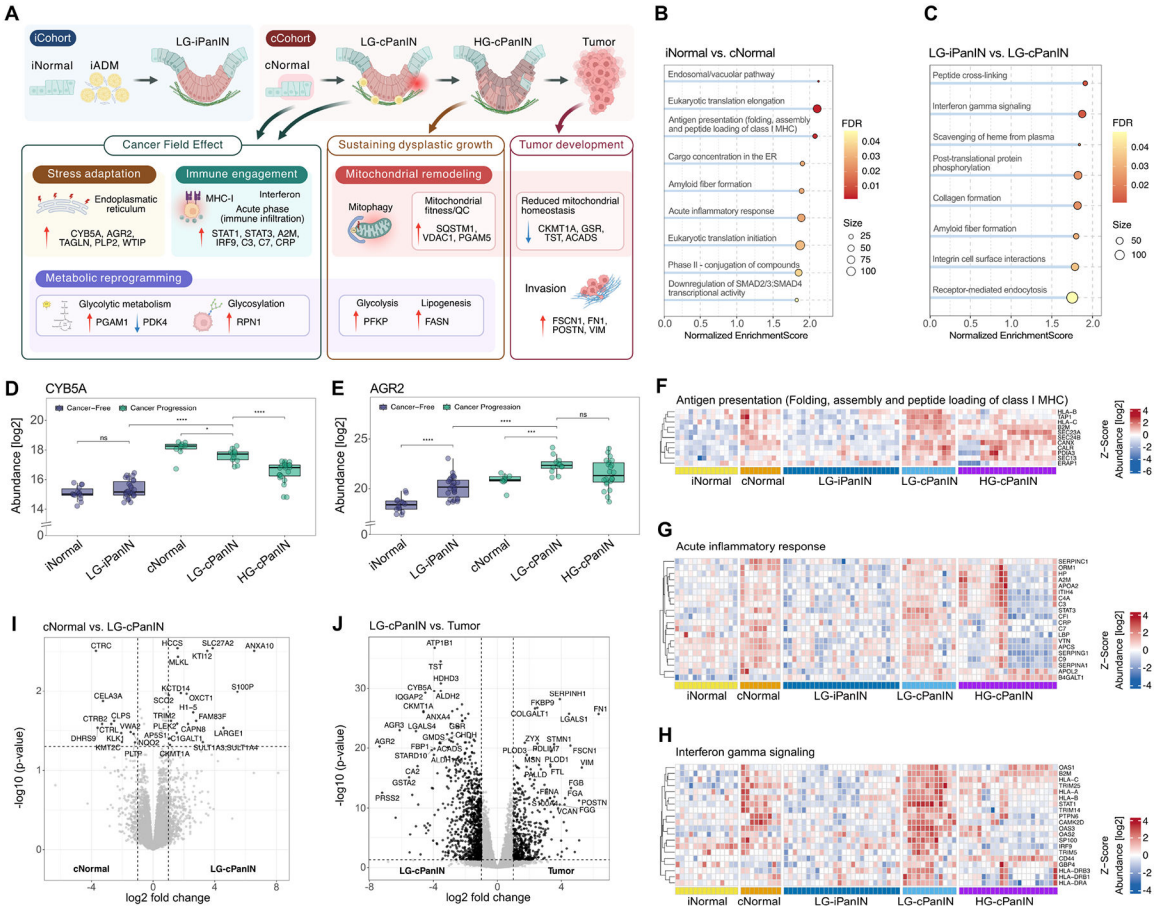

Figure 5.

Molecular hallmarks of the cancer field effect and early PDAC evolution. **A,** Schematic summary of molecular programs and key representative proteins involved in the multistep progression of human pancreatic neoplasia. **B** and **C,** Representative biological pathways enriched using gene set enrichment analysis comparing iNormal vs. cNormal (**B**) and LG-iPanIN vs. LG-cPanIN (**C**). FDR values are visualized by a color gradient, whereas the number of proteins enriched for a pathway is indicated by point size. **D** and **E,** Representative markers, CYB5A (**D**) and AGR2 (**E**), associated with the cancer field effect (stress/secretory adaptation) in cancer-associated normal ducts and PanIN lesions. An unpaired Student *t* test was performed assuming equal variances for the compared groups. ns, *P* values >0.05; *, *P* values ≤0.05; ***, *P* values ≤0.001; ****, *P* values ≤0.0001. **F–H,** Heatmap showing proteins enriched for the pathways “antigen presentation (folding, assembly, and peptide loading of class I MHC)” (**F**), “acute inflammatory response” (**G**), and “IFNγ signaling” (**H**). Colors represent z-score–normalized protein abundances (log_2). **I** and **J,** Differential protein expression within the cancer-associated cohort comparing cNormal vs. LG-cPanIN (**I**) and LG-cPanIN vs. Tumor (**J**). Significantly regulated proteins [fold change (log_2) > 2; *P* value <0.05] are highlighted in black. Normal, normal ductal epithelium; Tumor, tumor tissue. Sample overview: Details on sample counts and origins are provided in Supplementary Table S8. [**A,** Created in BioRender, RRID:SCR_018361].

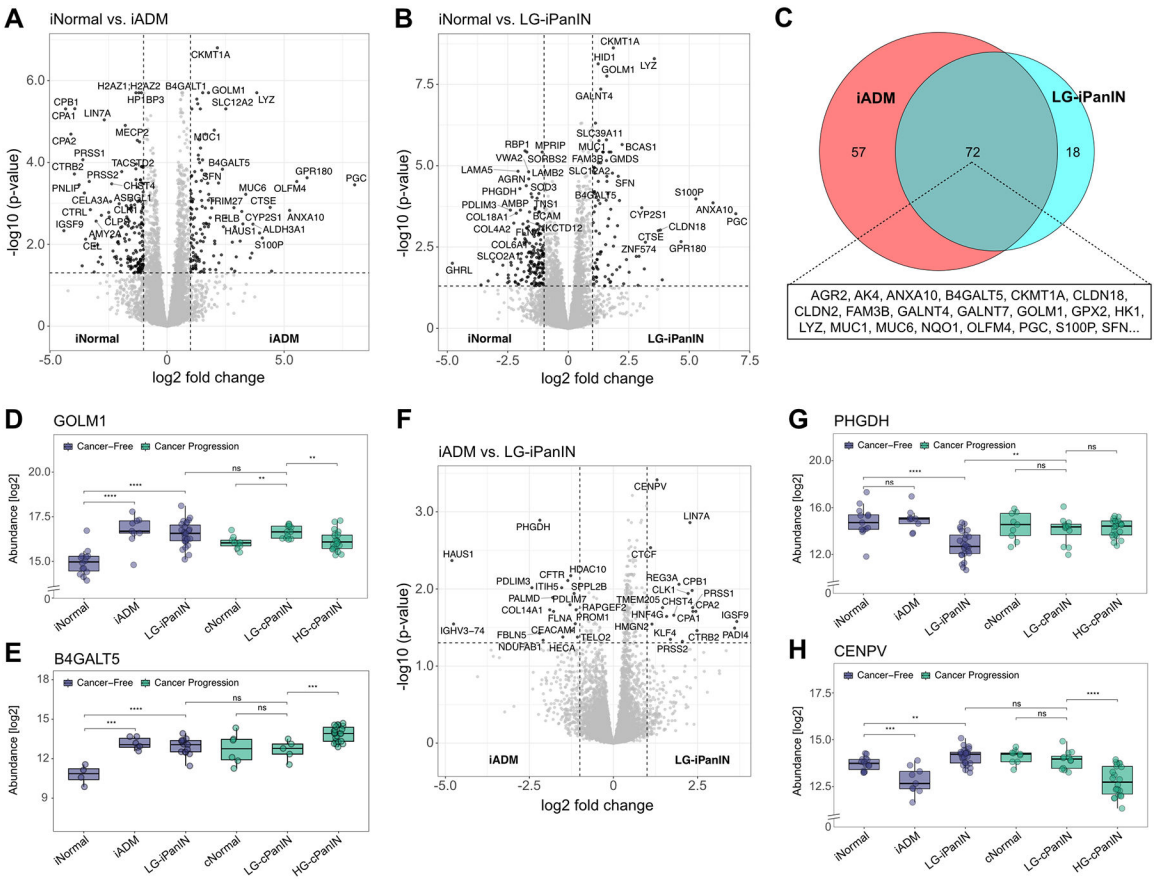

Figure 6.

The proteomic landscape of human ADM. **A** and **B,** Differential protein expression of iNormal vs. iADM (**A**) and iNormal vs. LG-iPanIN (**B**). Significantly regulated proteins [fold change (log_2) > 2; *P* value <0.05] are highlighted in black. **C,** Venn diagram showing unique and overlapping upregulated proteins in iADM and LG-iPanIN compared with iNormal. Representative overlapping proteins are highlighted. **D** and **E,** Representative protein markers enriched in iADM lesions compared with iNormal ducts: GOLM1 (**D**) and B4GALT5 (**E**). An unpaired Student *t* test was performed assuming equal variances for compared groups. ns, *P* values >0.05; **, *P* values ≤0.01; ***, *P* values ≤0.001; ****, *P* values ≤0.0001. **F,** Differential protein expression of iADM vs. LG-iPanIN. Significantly regulated proteins [fold change (log_2) > 2; *q* values <0.05] are highlighted in black. **G** and **H,** Representative proteins exhibiting divergent expression patterns between iADM and LG-iPanIN: PHGDH (**G**) and CENPV (**H**). Statistical testing was done as in **D** and **E**. Normal, normal ductal epithelium; Tumor, tumor tissue. Sample overview: Details on sample counts and origins are provided in Supplementary Table S8.

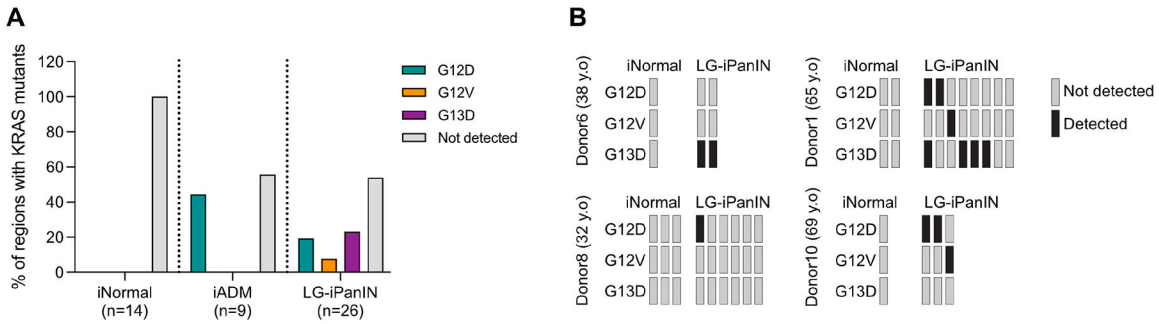

Figure 7.

Mutation-specific peptide profiling of *KRAS* in incidental precursor lesions. **A,** Detection of *KRAS* hotspot variants at the peptide level by DVP in iNormal, iADM, and LG-iPanIN, including G12D, G12V, and G13D. The bar plot shows the percentage of microdissected regions harboring *KRAS* mutants. **B,** Summary diagrams showing *KRAS* G12D, G12V, and G13D mutation profiles in four representative cancer-free individuals of different ages. Each column represents a single microdissected region. Black indicates regions harboring *KRAS* G12D, G12V, or G13D mutations, and gray indicates regions in which these *KRAS* mutations were not detected.