Oncogenic GNAS drives a gastric pylorus program in intraductal papillary mucinous neoplasms of the pancreas

Vincent Quoc-Huy Trinh[1,2], Katherine E. Ankenbauer[3], Jiayue Liu[3], Maelle Batardiere[2], H. Carlo Maurer[4], Celina Copeland[3], Jahg Wong[2], Olivia Ben-Levy[3], Sabrina M. Torbit[3], Brenda Jarvis[3], Frank Revetta[5], Sergey Ivanov[3], Nidhi Jyotsana[3], Yuki Makino[6], Amanda M. Ruelas[3], Anna L. Means[1], Anirban Maitra[6,7], Marcus C.B. Tan[1,3,8,9,#], and Kathleen E. DelGiorno[1,3,8,9,#]

[1]Department of Surgery, Vanderbilt University Medical Center, Nashville, TN, USA
[2]Institute for Research in Immunology and Cancer and the Centre de recherche du Centre hospitalier de l'Université de Montréal, University of Montreal, Montreal, Quebec, Canada
[3]Department of Cell and Developmental Biology, Vanderbilt University, Nashville, TN, USA
[4]Department of Internal Medicine II, Klinikum Rechts der Isar, School of Medicine, Technical University of Munich, Germany
[5]Department of Pathology, Microbiology and Immunology, Vanderbilt University Medical Center, Nashville, TN, USA.
[6]Department of Translational Molecular Pathology, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA
[7]Sheikh Ahmed Pancreatic Cancer Research Center, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA
[8]Vanderbilt Ingram Cancer Center, Vanderbilt University Medical Center, Nashville, TN, USA
[9]Vanderbilt Digestive Disease Research Center, Vanderbilt University Medical Center, Nashville, TN, USA
Co-last, Co-Corresponding

* Correspondence to: KE DelGiorno, Department of Cell and Developmental Biology, Vanderbilt University, 465 21st Ave S, Nashville, TN 37232. E-mail: kathleen.delgiorno@vanderbilt.edu. MCB Tan, Division of Surgical Oncology, Section of Surgical Sciences, Vanderbilt University Medical Center, 2220 Pierce Ave, 597 Preston Research Building, Nashville, TN 37232-6860, USA. E-mail: marcus.c.tan@vumc.org

ABSTRACT

OBJECTIVE: Intraductal Papillary Mucinous Neoplasms (IPMNs) are cystic lesions and bona fide precursors for pancreatic ductal adenocarcinoma (PDAC). Recently, we showed that acinar to ductal metaplasia, an injury repair program, is characterized by a transcriptomic program similar to gastric spasmolytic polypeptide expressing metaplasia (SPEM), suggesting common mechanisms of reprogramming between the stomach and pancreas. The aims of this study were to assay IPMN for pyloric markers and to identify molecular drivers of this program.

DESIGN: We analyzed RNA-seq studies of IPMN for pyloric markers, which were validated by immunostaining in patient samples. Cell lines expressing *Kras*^{G12D} +/- *GNAS*^{R201C} were manipulated to identify distinct and overlapping transcriptomic programs driven by each oncogene. A PyScenic-based regulon analysis was performed to identify molecular drivers in the pancreas. Expression of candidate drivers was evaluated by RNA-seq and immunostaining.

RESULTS: Pyloric markers were identified in human IPMN. *GNAS*^{R201C} drove expression of these markers in cell lines and siRNA targeting of *GNAS*^{R201C} or *Kras*^{G12D} demonstrates that *GNAS*^{R201C} amplifies a mucinous, pyloric phenotype. Regulon analysis identified a role for transcription factors SPDEF, CREB3L1, and CREB3L4, which are expressed in patient samples. siRNA-targeting of *Spdef* inhibited mucin production.

CONCLUSION: *De novo* expression of a SPEM phenotype has been identified in pancreatitis and a pyloric phenotype in *Kras*^{G12D}-driven PanIN and *Kras*^{G12D}*;GNAS*^{R201C}-driven IPMN, suggesting common mechanisms of reprogramming between these lesions and the stomach. A transition from a SPEM to pyloric phenotype may reflect disease progression and/or oncogenic mutation. IPMN-specific *GNAS*^{R201C} amplifies a mucinous phenotype, in part, through SPDEF.

INTRODUCTION

Pancreatic ductal adenocarcinoma (PDAC) is the third leading cause of cancer-related deaths in the United States and is slated to become second by the year 2030[1]. This is largely due to late detection – pancreatic tumors are clinically silent, approximately 85% of patients present with incurable locally advanced or metastatic disease, and distant metastatic spread occurs when the tumors are small (<5mm). On the other hand, the time required for PDAC to emerge from a normal cell is long, some 15-20 years[2]. Thus, improving our understanding of the early events in neoplastic transformation is necessary and crucial to allow earlier diagnosis of PDAC, and also important therapeutically, to gain insights into how progression to cancer can be blocked or even reversed.

A major obstacle to the study of early events in pancreatic carcinogenesis is that the main precursor to PDAC, pancreatic intraepithelial neoplasia (PanIN), is not clinically manifested, but typically a histological footnote on surgical pathology reports. PanIN, like PDAC, is clinically silent and radiographically occult. Intriguingly, the opposite is true of intraductal papillary mucinous neoplasms (IPMN), which contribute to approximately 25% of cases of PDAC[3]. As IPMN are cystic rather than microscopic like PanIN, they are easily identified on abdominal imaging scans and as a result, 90% are diagnosed before cancer is present[4, 5]. Patients with IPMN typically have a long window of opportunity for surveillance and intervention before invasive disease develops.

Oncogenic mutations in *KRAS* are common in human PanIN lesions and have been shown to drive tumorigenesis through increasingly dysplastic grades of PanIN to PDAC in genetically engineered mouse models (GEMMs)[6]. Like PanIN, IPMN harbor *KRAS* mutations (~80%) but may additionally express oncogenic GNAS (~66%) driver mutations[7-9]. In GEMMs, combined expression of both *Kras* and *GNAS* has been shown to accelerate PDAC formation through a mixed phenotype of PanIN and IPMN[10]. IPMN may be further subclassified by a combination of morphology and molecular markers including gastric foveolar (MUC6+MUC5AC+), intestinal (MUC2+MUC5AC+), and pancreatobiliary, which is widely considered to be an advanced form of gastric-type IPMN[11, 12].

Metaplasia is thought to be an initiating event in both PanIN and IPMN formation. Metaplasia is a pathological term for the transdifferentiation of one cell type to another and is a form of plasticity common to injury and oncogene-induced disease progression in the gastrointestinal tract[13-15]. While it is largely thought to mitigate injury, it is also considered to be the first step in tumorigenesis in several organs. Metaplasia falling under the general rubric of pyloric-like has been reported in stomach injury and tumorigenesis, colonic adenoma formation, and pancreatic injury and tumorigenesis occurring through PanIN progression[16-19]. Gastric metaplasia, or spasmolytic polypeptide expressing metaplasia (SPEM), is the best characterized and reflects the transition of gastric chief cells (digestive enzyme producing cells) to a phenotype described by the expression of specific markers (e.g. MUC6, TFF2, AQP5, CD44v9, GKN3, etc.)[20-22]. When SPEM is accompanied by the presence of a foveolar pit cell lineage (TFF1, GKN1, GKN2, MUC5AC), the phenotype is a pyloric metaplasia due to a recapitulation of the pylorus region of the stomach[23, 24]. Recently, we combined

single cell RNA-sequencing (scRNA-seq), electron microscopy, and histopathology and showed that injury-induced acinar to ductal metaplasia (ADM) in the pancreas results in the formation of chemosensory tuft cells, hormone-producing enteroendocrine cells, and a population bearing canonical markers of SPEM[16]. ADM and PanIN resulting from oncogenic *Kras*^{G12D} expression are also characterized by SPEM with the additional formation of a separate, distinct foveolar pit cell-like population, reflecting pyloric metaplasia[16, 17]. The functional role of individual pyloric markers has previously been studied in GEMMs of gastric disease[25, 26], however, the program itself is hypothesized to represent a shared mechanism by which gastrointestinal organs respond to injury and tumorigenesis[24]. Here, we assayed IPMN for the expression of pyloric markers to determine if this program is enriched in or differentiates IPMN from PanIN lesions. We show that oncogenic GNAS is sufficient to drive a pyloric phenotype, including gastric markers, mucus production, and epithelial plasticity, and identify critical transcriptional regulators of this program.

MATERIALS AND METHODS:

Human Samples. A cohort of 41 patients with intraductal papillary mucinous neoplasms (IPMNs) ranging from low-grade (LG) dysplasia, high-grade (HG) dysplasia, to invasive were selected from Vanderbilt University Medical Center's institutional cohort with institutional review board approval (#110061). Normal stomach body, small intestine, and colon tissues were used as controls to threshold signal intensity.

Mice. Mice were housed in accordance with National Institutes of Health guidelines in an American Association for Accreditation of Laboratory Animal Care-accredited facility at M.D. Anderson Cancer Center. The M.D. Anderson Institutional Animal Care and Use Committees (IACUC) approved all animal studies. *LSL-Kras*$^{G12D/+}$, *Ptf1a*$^{Cre/+}$, *LSL-rtTA-TetO-GNAS*^{R201C} (*Kras;GNAS*) mice have previously been described[6, 10, 27]. Mice were given a diet containing doxycycline (Tusculum Feed Center, #9205-0827) beginning at 8 weeks of age for a period of either 10 or 20 weeks.

Multiplex Immunohistochemistry (MxIHC). Formalin-fixed paraffin-embedded tissues from the patient cohort were sectioned at 4 μm, heated to 60 °C for 30 min, deparaffinized in xylene, and rehydrated in an ethanol gradient. Slides were stained with Mayer hematoxylin (Sigma-Aldrich, MHS32), coverslipped with 30% glycerol (Sigma-Aldrich, G5516) in phosphate-buffered saline (PBS) and scanned with an Olympus VS200 slide scanner (Olympus, Tokyo, Japan). Slides were de-coverslipped by immersion in PBS and underwent antigen retrieval in 10 mM sodium citrate pH 6.0 with a microwave set at maximum power until boiling bubbles appeared, reduced to minimum power for 20 minutes, and left at RT for 30 minutes. Slides were treated with 3% hydrogen peroxide for 10 minutes, blocked with Protein Block, Serum-Free (Agilent Dako, X090930-2) for 10 minutes, and incubated with the primary antibody overnight. Secondary antibodies were incubated for 1 hour, and the signal was revealed with AEC+ High Sensitivity Substrate Chromogen (Dako Agilent, K3469). The slides were then coverslipped with 30% glycerol (Sigma-Aldrich, G5516) in phosphate-buffered saline (PBS). Slides were scanned with the VS200. Slides were de-coverslipped by immersion in PBS and underwent a 2-minute double-distilled water, 2 minutes 70% ethanol, 2 minutes 95% ethanol, 2 minutes ethanol, 2-minute double-distilled water sequence to eliminate the 3-amino-9-

ethylcarbazole (AEC) chromogen. Slides then restarted the previous sequence at the antigen retrieval step until all primary antibodies (**Table S1**) were performed. Scans were loaded in QuPath v0.4.0[28] and registered with the image-combiner v0.3.0 package.

Standard histological staining. Murine tissues were cut in 5 μm sections, mounted, and stained as previously described[16]. Sections were deparaffinized in xylene, rehydrated in a series of graded ethanols, and then washed in PBST and PBS. Endogenous peroxidase activity was blocked with a 1:50 solution of 30% H_2O_2:PBS followed by microwave antigen retrieval in 100 mM sodium citrate, pH 6.0. Sections were blocked with 1% bovine serum albumin and 5% normal goat serum in 10 mM Tris (pH 7.4), 100 mM $MgCl_2$, and 0.5% Tween-20 for 1hr at room temperature. Primary antibodies (**Table S1**) were diluted in blocking solution and were incubated on tissue sections overnight. Slides were then washed, incubated in streptavidin-conjugated secondaries (Abcam) and developed with DAB substrate (Vector). Periodic Acid Schiff/Alcian Blue (PAS/AB) staining was performed per the manufacturer's instructions (Abcam, ab245876). For standard IHC on human tissue sections, antigen retrieval was performed with pH 6.0 citrate buffer in a pressure cooker at 105°C for 15 minutes, with a 10-minute cool down. Blocking was performed in 0.03% H_2O_2 containing sodium azide for 5 minutes and primary antibodies (**Table S1**) were incubated for 60 minutes before detection (Dako EnVision+ System-HRP labeled Polymer) for 30 minutes and development for 5 minutes. All slides were scanned on the VS200.

IHC quantification. MxIHC Quantification. Areas measuring 250 x 250 μm^2 were acquired blindly on the hematoxylin layer, querying normal ducts, acinar-to-ductal metaplasia (ADM), low-grade dysplasia (LG), high-grade dysplasia (HG), and invasive disease under the supervision of a board-certified pathologist (VQT). Up to 3 areas of normal, ADM and INV, and up to 5 areas of LG and HG, were selected per patient (when present). An automated watershed threshold pixel detection was adjusted based on the staining patterns on the control stomach, small intestine, and colon tissues. Automated AEC deconvolution was performed in QuPath v0.4.0 by loading them as a H-DAB slide. An automatic calculation of the surface area with the AEC color deconvolution intensity over the threshold was then generated. Pseudocolors were given to each stain for the figures. **Monoplex Quantification of SPDEF, CREB3L1 and CREB3L4.** Scans of SPDEF, CREB3L1, and CREB3L4 were obtained with the VS200 and loaded in QuPath v0.4.3[28]. Color deconvolution of the hematoxylin stain was performed to select up to 3 areas of normal ducts, ADM, LG, HG, and INV. By consensus, one student and a board-certified pathologist (VQT) scored the nucleus and cytoplasm of cells on a scale of 0-1-2-3. **Cell type quantification in murine IPMN.** Scans of H&E staining from *Kras;GNAS* mice +/- doxycycline chow for 10 or 20 weeks (n = 3/condition, n = 12 total) were scored by a pathologist (VQT) and areas of IPMN and/or PanIN lesions were identified (n = 3, up to 6 regions/slide). Serial sections were stained for DCLK1 (tuft cells), synaptophysin (enteroendocrine cells), or PAS/AB (all mucins) and 1-6 lesions were identified per area of IPMN and/or PanIN (when present). All steps of analysis were performed blinded. Positive DCLK1 or synaptophysin cells were manually counted and divided by the number of nuclei per lesion to identify percent of that lesion constituted by a given cell type. To quantify PAS/AB mucin staining, lesions were annotated as regions of interest and manually drawn in FIJI. Total

signal and PAS/AB signal were thresholded to include only stained areas and percent positive area was calculated.

Cell culture and *Gnas^{R201C}* induction. Murine cell lines generated from *Kras;GNAS* mice[10] (4838 and C241) were cultured in RPMI with L-glutamine (Corning 10-040-CV), 10% Tet system approved FBS (Gibco A4736401) and 1 x antibiotic/antimitotic (Corning 30-004-CI). To induce human *GNAS^{R201C}* expression, cells were stimulated with 1 μg/ml doxycycline (RPI Research Products International D43020) for 48-72 hours. Western blotting was conducted using Bio-Rad Mini-Protean 4-20% TGX gels, PVDF membranes and the Trans Blot Turbo system. Blocking was done in 5% non-fat milk. Primary antibodies are listed in **Table S1**. Blots were imaged using Immobilon HRP Substrate (WBKLS0500) and processed using an Amersham AI600 imaging system. For quantitative real-time polymerase chain reaction (qRT-PCR), RNA was isolated using Quick-RNA MiniPrep kits (Zymo Research R1055), and concentration was measured using NanoDrop 2000. qRT-PCR was performed using a Luna Universal One-Step RT-qPCR Kit (New England Laboratories, E3005) and analyzed using Bio-Rad CFX96 system and CFX Manager 3.1. Primers are listed in **Table S2**. For immunofluorescence, cells were seeded on a Matrigel-coated coverslip (1:50 dilution) (Corning, 356231) and fixed with 4% PFA (Electron Microscopy Services, 15712) for 15 minutes at room temperature. Following fixation, cells were washed 5 times in PBS and blocked and permeabilized using buffer containing 1% BSA (w/v), 5% donkey serum (v/v), 5% goat serum (v/v) and 0.3% Triton-X for 30 minutes at room temperature. Coverslips were washed in PBS 5 times. For primary incubation, primary antibodies (**Table S1**) were diluted in blocking buffer (5% donkey serum, 1% BSA and 0.05% Triton-X) and incubated at 37°C for 2 hours. Coverslips were washed 5 times in PBS and secondary antibody diluted in blocking buffer was added. Secondary antibodies were incubated at 37°C for 1 hour. Coverslips were washed 5 times in PBS and mounted with ProLong Gold Antifade Mountant with DAPI (Invitrogen, P36931) and imaged on the VS200 or an Echo Revolve microscope.

Immunofluorescence analysis. Cell line immunofluorescence images were analyzed for corrected total cell fluorescence (CTCF) using Fiji[29]. Briefly, representative ROIs in each image were selected using the freehand selection tool. The area, mean, integrated density, and raw integrated density were calculated for each ROI. The same values were also taken from the background. To calculate the CTCF, the following formula was used: raw integrated density – (area*average background mean). Images were analyzed for the area per cell in Fiji using custom macros. To measure this, the number of cells was calculated by adjusting the threshold (15, 255), using the Watershed feature to segment nuclei from each other and then using the analyze particles feature (size = 30 to infinity) to count the number of nuclei. To calculate the area, the image was thresholded once again (25, 255) and the analyze particles feature (size = 30 to infinity) was used to find the area of each stain. The area per cell was calculated by taking the total area and dividing it by the number of cells in each image.

siRNA knockdown of target genes. *Spdef* knockdown was accomplished in 4838 cells by siRNA mediated knockdown. siRNA for *Spdef* (#184152, Thermo-Fisher) and a nonspecific control (*Silencer* Negative Control

No. 1 siRNA, Thermo-Fisher) were transfected using lipofectamine RNAiMAX (#13778075, Thermo-Fischer). 4838 cells were seeded in 6cm dishes with 1μg/ml of doxycycline for 48 hours. At 40% confluency, cells were transfected for a final concentration of 20 pmol. After 72 hours, RNA was extracted, or cells were prepared for immunofluorescence. Lipid nanoparticle (LNP)-encapsulated siRNA to target *Kras* or *GNAS* was generated by microfluidic mixing of LNPs (Precision Nanosystems) and siRNA (Axolabs). 4838 cells were seeded in a 6 well plate. At 50% confluency, media was replaced with that containing 1 μg/ml doxycycline for 8 hours. At that time, 20 μl of LNP-siRNA was added for a final concentration of 2 μg/ml. Treated cells were collected 4 days later for analysis by qRT-PCR or western blotting. siRNA used is as follows (all LGC Axolabs): KRAS, sense strand 5′-ggGaacAACuaUaaagAasa-3′ and antisense strand 5′-dTUUCUUUAUAGUuGUUCCCusu-3′; GNAS, sense strand 5′-cuUcgcUGCugUguccUgsa-3′ and antisense strand 5′-dTCAGGACACAGCaGCGAAGusu-3′; and control Fluc, sense strand 5′-cuuAcGcuGAGuAcuucGAdTsdT-3′ and antisense strand 'UCGAAGuACUcAGCGuAAGdTsdT-3′.

Analysis of published single-cell RNA sequencing datasets. Processed count matrices from Bernard et al., were downloaded from the Gene Expression Omnibus (GEO) database[30]. Data analyses were executed using a dual-language approach, encompassing R (version 4.3.1, 2023-06-16) and Python (version 3.9.13, 2022-08-25), both tailored for a 64-bit macOS Ventura 13.4 platform. R integrated packages utilized in the study comprised `ggplot2 (3.4.3)`, `dplyr (1.1.3)`, `patchwork (1.1.3)`, `SeuratObject (4.1.3)`, `Seurat (4.3.0.1)`, `umap (0.2.10.0)`, `stxBrain.SeuratData (0.1.1)`, and `SeuratData (0.2.2)`. Python libraries of interest included `anndata (0.9.2)`, `scanpy (1.9.5)`, and `pandas (2.1.1)`. Seurat objects were established and normalized; individual Seurat objects representative of the six patient datasets from Bernard et al., were synthesized and post-merged into a unified object. Consolidated data were filtered to cells with a minimum of 300 genes, manifesting less than 15% mitochondrial genes (percent.mt<0.15). Emphasis was subsequently placed on identifying and leveraging highly variable features, with the selection of top 2,000 features (nFeature_RNA > 2000) via the `FindVariableFeatures()` function. Post scaling, data was subjected to Principal Component Analysis (PCA) to demarcate inherent patterns. Exploiting the dimensionality reduced space (spanning the first 20 principal components), cellular neighborhoods were mapped with the `FindNeighbors()` function (dims = 1:20,k.param = 20) and clusters were discerned using `FindClusters()` with a resolution parameter set at 0.5. For downstream analysis, the MuDataSeurat tool, accessible on [GitHub] (https://github.com/PMBio/MuDataSeurat), was employed to convert Seurat objects into the `.h5ad` format compatible with the `Scanpy` library. UMAPs labeled by sample or for individual gene markers ('CD44','MUC2','TFF2','PTPRC','EPCAM','AQP5','MUC5AC',etc.) were plotted for visualization.

Datasets from Bockerstett et al[21]., were downloaded from NCBI (NCBI accession numbers SRX7069341 and SRX7069340)[31] and underwent quality control (QC), normalization, and clustering utilizing Seurat. For both WT samples, the inclusion criteria were set for cells exhibiting more than 750 features (nFeature_RNA > 750), with mitochondrial gene content (percent.mt) exceeding 12.5% yet remaining below 60%. Clustering parameters specified a resolution of 0.6, employing dimensions 1 to 35, with 10 nearest neighbors (n.neighbors = 10) and a k parameter of 10 (k.param = 10). Dimensionality for subsequent analyses was

determined based on the elbow plot. WT2 followed a similar QC protocol, with an adjusted clustering resolution of 0.7 to accommodate the distinct cellular composition and dynamics inherent to the sample. Post QC, WT1 and WT2 samples were merged using Seurat's `merge()` function, and batch effects were subsequently addressed through the application of the FastMNN algorithm, available within the 'SeuratWrappers' package. The merged WT dataset, thereafter, referred to as wt.fmnn, was analyzed using FastMNN (`RunFastMNN(object.list = SplitObject(wt, split.by = "orig.ident"))`). Dimensionality reduction and cellular neighborhood mapping were achieved through `RunUMAP` and `FindNeighbors`, respectively, utilizing 50 dimensions (dims = 1:50) as dictated by the prior selection, with clustering achieved via `FindClusters` at a resolution of 0.5. An analogous methodology was applied to the two acute samples (Acute1 and Acute2), encompassing individual QC, processing, and integration, followed by batch effect mitigation using FastMNN. Cell types within both WT and acute integrated datasets were annotated based on a comprehensive set of gene signatures provided by the original article[21]. After the integration and annotation phases, the data were exported into the `.h5ad` format for compatibility with the Scanpy Python toolkit as stated before. Within Scanpy, specific gene signatures associated with pancreatitis, and *GNAS* and *Kras* knockdown, were overlaid onto the datasets using the `sc.tl.score_genes` function.

Bulk human RNA sequencing analysis. Compartment-specific gene expression profiles of human IPMN (n = 19), PanIN (n = 26) and PDAC (n = 197) were generated using laser capture microdissection with subsequent RNA sequencing as previously described[32-34]. Differentially expressed genes between human IPMN and PanIN were identified leveraging a generalized linear model as implemented in DESeq2 R package[35] and genes exhibiting a false discovery rate ≤ 0.05 were considered significantly differentially expressed.

Bulk murine RNA collection and sequencing analysis. RNA was extracted from either 4838 or C241 cells (+/- doxycycline) using Quick-RNA MiniPrep (Zymo Research R1055) and quality was measured using an Agilent Bioanalyzer; RNA Qubit assay was performed to measure RNA quantity. Poly(A) RNA enrichment was conducted using NEBNext Poly(A) mRNA Magnetic Isolation Module (NEB E7490), and the sequencing library was constructed using the NEBNext Ultra II RNA Library Prep Kit (E7765L) following the manufacturer's instructions. End repair, A-tailing, and adapter ligation was performed to generate the final cDNA library. Library quality was assessed using a Bioanalyzer and quantified using a qPCR-based method with the KAPA Library Quantification Kit (KK4873) and the QuantStudio 12K instrument. 150 bp paired-end sequencing was performed on the NovaSeq 6000 platform targeting 50M reads per sample. Raw sequencing data (FASTQ files) obtained from the NovaSeq 6000 was subjected to quality control analysis, including read quality assessment. Real Time Analysis Software (RTA) and NovaSeq Control Software (NCS) (1.8.0; Illumina) were used for base calling. MultiQC (v1.7; Illumina) was used for data quality assessments. Paired-end RNA sequencing reads (150bp long) were trimmed and filtered for quality using Trimgalore v0.6.7 (https://doi.org/10.5281/zenodo.7598955). Trimmed reads were aligned and counted using Spliced Transcripts Alignment to a Reference (STAR) v2.7.9a with the -quantMode GeneCounts parameter against the mm39 mouse genome and GENCODE comprehensive gene annotations (Relsease M31)[36]. ~50-100

million uniquely mapped reads were acquired per sample. Sample read counts were normalized and differential expression was performed using DESeq2 v1.34.0[35]. Genomic features counted fewer than five times across at least three samples were removed. False discovery rate adjusted for multiple hypothesis testing with Benjamini-Hochberg (BH) procedure p value < 0.05 and log2 fold change >1 was used to define differentially expressed genes. Gene set enrichment analysis (GSEA) was performed using the R package Clusterprofiler[37] with gene sets from the MSigDB database[38].

pySCENIC and gene regulatory network (GRN) inference. To infer the GRN in pyloric metaplasia of the pancreas, we performed SCENIC using pySCENIC functions on a scRNA-seq dataset of murine pancreatitis[16, 39]. This protocol allows the reconstruction of regulons (TF and known target genes) from gene co-expression data, assesses regulon activity in single cells, and can be used to find regulon-enriched cellular clusters. Specifically, we ran v0.11.0 of pySCENIC in a Singularity container built from the Docker Hub image, on ACCRE, Vanderbilt's High Performance Computing cluster. Following the quality control and feature selection performed in Seurat, we exported its raw counts to a matrix that was then converted to a LOOM file. Alongside a list of 1170 mouse TFs, this gene expression matrix served as input for calculating gene co-expression modules via GRNBoost2. To account for the stochastic nature of GRNBoost2, we calculated the co-expression modules 100 times and then retained only TF/target gene associations that exist in at least 80% of the runs. We then merged the results of these 100 runs as a left outer join operation and averaged the IM values reported for each association. This consensus GRN was then used as input for module pruning, where we filtered out indirect gene targets lacking the cis-regulatory motif associated with the TF. This step used SCENIC's RcisTarget and ranking databases for motifs (mm9-tss-centered-5kb-7species.mc9nr.feather) in the promoter of the genes [up to 500 base pairs (bp) around the transcriptional start site (TSS). The resulting coexpressed TF-target genes are then grouped into regulons. Lastly, the activity of the regulons was computed using SCENIC's AUCell function, which uses the "area under the curve" (AUC) to calculate whether a subset of the input gene set is enriched within the expressed genes for each cell. These activity data were further binarized (assigned an ON or OFF value, per regulon, per cell) by threshold on the AUC values of the given regulon. Both the AUCell and binarized regulon activity matrices were integrated into Seurat object via the "CreateAssayObject" function, for downstream analysis and visualization[40]. UMAP visuals of the binary and AUC metrics were created from Seurat's DimPlot() function. Heatmap visuals of binary regulon matrix were performed by ComplexHeatmap R package[41]. Network plots were created from Cytoscape from top 10% based on IM reported for TF/target association from the coexpression modules[42].

Statistical Analysis. Statistical analyses and data processing were performed in Image J or Prism (GraphPad, San Diego, CA). Statistical significance was calculated by either 2-tailed unpaired t-tests assuming equal variance or 1-way analysis of variance. Data are expressed as mean ± standard deviation.

RESULTS

Human IPMN recapitulate pyloric metaplasia. Metaplastic tissue bearing markers of the gastric pylorus has been reported in injury and tumorigenesis in several gastrointestinal organs[16-19]. To evaluate expression of pyloric metaplasia markers in IPMN, we evaluated expression of *MUC5AC* (foveolar pit lineage marker), *TFF2*, *AQP5*, and *CD44* (SPEM markers), which produces splicing variant CD44v9, in published RNA-seq datasets derived from patient samples. Previously, Bernard et al., generated a single cell RNA-sequencing (scRNA-seq) dataset composed of epithelium and stroma from either low-grade or high-grade IPMN (n = 2 each) or IPMN associated PDAC (n = 2)(**Figure 1A-C**)[30]. Analysis of this dataset for molecular markers of IPMN subtype revealed widespread expression of *MUC5AC* throughout the epithelium of all samples, with *MUC6* and *MUC2* differentiating gastric- from intestinal-type IPMN, respectively (**Figure 1D**). Expression of *TFF2* and *AQP5* was enriched in the epithelium whereas *CD44* was identified in both the epithelium and stroma (**Figure 1E**). To evaluate a second dataset, we interrogated bulk RNA-seq generated by Maurer et al., of laser capture dissected epithelium or stroma from patient IPMN (n = 19), PanIN (n = 26), or PDAC (n = 197 epithelium, 124 stroma)[33, 34]. Like our analysis of IPMN scRNA-seq data, we identified expression of *TFF2* and *AQP5*, enriched in the epithelium of all disease states, and *CD44,* in both epithelial and stromal populations (**Figure 1F**). Finally, we examined expression of a SPEM gene signature generated from our previously published murine pancreatitis scRNA-seq dataset, in the Maurer dataset. We identified enrichment of several additional markers in pre-invasive PanIN and IPMN, as compared to PDAC (**Figure S1**)[16].

To confirm expression of pyloric metaplasia markers in patient IPMN, we conducted multiplex immunohistochemistry (MxIHC) on a cohort of 41 patients with IPMNs ranging from low-grade to high-grade dysplasia to locally invasive disease. Serial staining was conducted for MUC5AC, AQP5, and CD44v9 and expression scored in multiple tissue compartments within each sample. Signal for each stain was automatically detected using QuPath and overlaid to generate merged images (**Figure 2A**)[28]. Altogether, expression was evaluated in normal ducts (n = 95-109 regions of interest, ROIs), acinar to ductal metaplasia (ADM, n = 109 ROIs), low-grade IPMN (n = 110 ROIs), high-grade IPMN (n = 70 ROIs), and foci of invasive PDAC (n = 22 ROIs). All markers were detected; interestingly, protein localization changed between low- and high-grade IPMN (**Figure 2B**). In low-grade IPMNs, CD44v9 was consistently expressed within crypts the base of the IPMN epithelial component mimicking gastric SPEM, in a generally mutually exclusive pattern with MUC5AC (**Figure 2B**). AQP5 mostly correlated with CD44v9, as described in the SPEM signature at the protein level. In high-grade IPMN, both SPEM and gastric foveolar marker MUC5AC co-localized, recapitulating a pyloric phenotype (**Figure 2B**). Consistent with previous reports; we identified a significant increase in MUC5AC expression between normal ducts and IPMN and a significant decrease in expression with progression from low to high-grade IPMN and then to invasive[30]. Expression of both AQP5 and CD44v9 increased with disease progression, reaching significance in low and high-grade IPMN (**Figure 2C**). Collectively, these data demonstrate that a significant increase in pyloric metaplasia marker expression accompanies IPMN formation in patients.

Oncogenic GNAS drives a pyloric-type phenotype. While PanIN and IPMN have both been shown to express oncogenic *KRAS*, a large fraction of IPMN is additionally characterized by the expression of oncogenic *GNAS*[8]. To determine if *GNAS*^{R201C} expression contributes to the pyloric phenotype described above, we examined marker expression in two cell lines (4838, C241) derived from *LSL-Kras*$^{G12D/+}$, *Ptf1a*$^{Cre/+}$, *LSL-rtTA-TetO-GNAS*^{R201C} (*Kras;GNAS*) mice. Mice expressing *Kras*^{G12D} have been shown to largely form PanIN; additional expression of *GNAS*^{R201C} results in a mixed phenotype dominated by IPMN, cyst-like lesions and accelerated progression to PDAC[10]. Both cell lines harbor a *Kras*^{G12D} mutation but conditionally express human *GNAS*^{R201C} upon doxycycline (DOX) treatment[10]. In addition to increased GNAS^{R201C} expression with DOX, we also identified elevated mRNA levels of SPEM marker *Aqp5* (**Figure 3A-B**). To then test whether *GNAS*^{R201C} drives a pyloric-like transcriptomic program, the 4838 and C241 cell lines were treated +/- DOX and underwent RNA sequencing (n = 6, 3 biological replicates/condition) (**Figure 3C, Figure S2, File S1-2**). As expected, there was a significant increase in human (hu*GNAS*), but not murine *Gnas* (mu*Gnas*) expression (**Figure 3D**). Significant differences were identified between the two cell lines, likely due to biological heterogeneity or differences in tumor suppressor mutations (**Figure S2A**), however, upon DOX treatment, both cell lines upregulated expression of pyloric markers (**Figure 3E**). This included markers associated with both gastric foveolar pit cells (*Gkn1*, *Gkn2*, *Tff1, Mucl3*) and SPEM (*Tff2*, *Aqp5*, *Gkn3*) (**Figure 3E-F**). Interestingly, we also saw an increase in the expression of markers of gastric chief cells (*Pga5*, *Pgc*) (**Figure 3F**). To confirm the increase in expression of select markers with *GNAS*^{R201C}, we conducted immunofluorescence (IF) on 4838 and C241 cells +/- DOX. As shown in **Figure 3G-H**, we saw a significant increase in both AQP5 and CD44v9 in overall expression and intensity. Collectively, these data demonstrate that IPMN-driver *GNAS*^{R201C} induces a pyloric phenotype characterized by the upregulation of SPEM and foveolar pit cell markers.

Oncogenic GNAS drives a mucinous phenotype *in vivo*. In addition to pyloric markers, we noted what appeared to be an increase in mucin production by DOX-treated cells. Mucin production is characteristic of the gastric pylorus and increases under conditions of gastrointestinal injury or oncogenesis. Consistent with this, we found that *GNAS*^{R201C} mutation is associated with high expression of mucin genes in a small number of cancer cell lines (cBioPortal, **Figure S2F**). To determine if *GNAS*^{R201C} is sufficient to drive mucin production, we interrogated our RNA-seq dataset and found a significant increase in *Muc1*, *Muc5b*, and *Muc3a* expression in both cell lines, as well as an increase in *Muc5ac* specifically in C241. Interestingly, we also noted an increase in mucin N-acetylgalactosaminyltransferases *B4galnt3* and *Galnt2* (**Figure 4A**). Gene expression changes in additional mucins (e.g., *Muc20*, *Muc3*) and mucin modification genes (e.g., *B4galt1*, *St3gal4*, *Galnt6*) were identified as well (**File S1-2**). To confirm increased mucin production, both cell lines were stained with lectin GSII, which broadly labels mucins. We identified an increase in both expression and intensity in both cell lines with DOX treatment (**Figure 4B-C**). Consistent with RNA-seq data, MUC5AC expression increased with DOX treatment only in C241 cells (**Figure 4D-E**). Collectively, these data demonstrate that *GNAS*^{R201C} expression is sufficient to drive a pyloric phenotype as well as enhanced mucin production in IPMN-like pancreatic cancer cells.

Like IPMN, PanIN are also considered to be mucinous precursors to PDAC, however they are not known to express oncogenic *GNAS* mutations and are not defined by large pools of mucin. To identify potential differences between PanIN and IPMN, we performed differential gene expression analysis on the bulk PanIN (n = 26) and IPMN (n = 19) RNA-seq from the Maurer dataset[33, 34]. At the molecular level, both precursor lesions were very similar, however, we did identify gene expression differences suggesting divergence in cellular composition. PanIN had higher expression of tuft (*SUCNR1*, *AVIL*, *PTGS1*), and enteroendocrine (*HHEX*, *NEUROD1*) markers, while IPMN expressed higher levels of mucus cell markers (*MUC13*, *MUC5AC*, *MUC2*), consistent with our cell line data (**Figure 5A, File S3**).

To determine whether $GNAS^{R201C}$ preferentially drives mucus cell production and if PanIN and IPMN differ in cellular composition, we next evaluated cell type abundance in *KrasGNAS* mice. Pancreas tissues were collected from *KrasGNAS* mice treated with or without DOX for either 10 or 20 weeks (n = 3 mice/condition)[10]. H&E were then evaluated by a pathologist (VQT) for disease progression and cyst formation. When combined, the two groups were found to be equivalent in terms of disease progression, however, DOX treatment was found to enhance cyst formation (**Figure S3A**), as previously described[10]. Regions of cyst/IPMN and PanIN lesions were then identified by H&E (~3 regions of each/slide), with most mice harboring both lesion types (up to 6 areas identified/mouse). Serial sections were stained for tuft cell marker DCLK1[43], enteroendocrine cell (EEC) marker synaptophysin[44, 45], and for mucins (PAS/AB) and quantified. As shown in **Figure 5B**, $GNAS^{R201C}$ expression was sufficient to significantly decrease tuft (9.9%, 45 lesions vs. 5.2%, 52 lesions) and enteroendocrine cell numbers (3.1%, 42 lesions vs. 1.2%, 52 lesions) while significantly increasing mucin production (19.6%, 40 lesions vs. 26.6%, 52 lesions) in low grade lesions, consistent with our analysis of human data (**Figure 5A, File S3**). This phenotype, however, was lost over time with disease progression in mice treated +/- DOX for 20 weeks, harboring more advanced disease (**Figure 5B-C**). Interestingly, and in contrast to human cysts/IPMN, many of the larger cysts that formed were relatively devoid of all three cell types as compared to PanIN (**Figure S3B**); further differences between $GNAS^{R201C}$-induced disease in humans and mice remain to be assessed.

Oncogenic *Kras* and *GNAS* drive distinct and overlapping transcriptomic programs. While $GNAS^{R201C}$ predominantly drives the proliferation of mucin-releasing cells, both $Kras^{G12D}$ and $GNAS^{R201C}$ can drive IPMN formation. Further, both PanIN and IPMN are characterized by pyloric metaplasia. To identify distinct pathways driven by each oncogene, we treated 4838 cells +/- DOX with siRNA and performed RNA-seq (**Figures 6A, S3, File S4**). As shown in **Figure 6B**, siRNA treatment significantly reduced expression of *GNAS* (both human and murine) or *Kras* as compared to control. Interestingly, *Kras* knockdown resulted in a significant increase in murine *Gnas* expression, suggesting compensatory signaling between the two oncogenes. Gene signatures were created for knockdown of *Kras* or *GNAS* as compared to control and compared (**Figure 6C**). As shown in **Figure 6D**, targeting each oncogene decreased expression of both distinct and overlapping genes. *GNAS* knockdown reduced expression of many, but not all, genes upregulated by DOX treatment, confirming specificity (**Figure 6E**, **S4A, File S5**). Complete knockdown of all DOX-induced genes was not achieved, likely due to the partial knockdown of $GNAS^{R201C}$ or compensation

by murine *Gnas* (**Figure 6B**). Knockdown of *Kras* resulted in decreased expression of several pro-tumorigenic signaling pathways (*Myc*, *Hif1a*, MAPK signaling genes) and targeting either gene reduced expression of specific mucins or N-acetylgalactosaminyltransferases (**Figure 6E**, **File S6-7**). Consistent with *in vivo* phenotypes, targeting either *Kras* or *GNAS* was found to reduce expression of pyloric markers (*Aqp5*, *Gkn1*), as well as other genes identified in the DOX treatment signature (*Cftr*, *Prom1*, *Il33*) (**Figure 6B, 6E**). Co-targeting both *Kras* and *GNAS* in combination showed similar results (**Figure S4B-C**). Altogether, these data suggest that, while both oncogenic *Kras* and *GNAS* can drive a pyloric gene expression signature, additional expression of $GNAS^{R201C}$ amplifies this phenotype.

Finally, to better understand how closely the $GNAS^{R201C}$, $Kras^{G12D}$, or the pancreatitis-induced (SPEM-like, no oncogenic mutations) transcriptomic signatures resemble pyloric metaplasia, we overlaid these gene signatures on scRNA-seq of normal or injured (SPEM) murine stomach generated by Bockerstett et al[31]. In the normal stomach, the pancreatitis SPEM signature was enriched in gastric chief cells, while the oncogene-driven signatures were more similar to foveolar pit cells. In gastric injury, all signatures enriched in SPEM, with the oncogene-driven gene signatures shifting towards foveolar pit cell enrichment, reflecting pyloric metaplasia (**Figure 6F**). The implications of this shift are currently unknown but may represent changes associated with tumor progression.

Identification of master regulators of pyloric metaplasia. Previously, we showed through scRNA-seq, immunostaining, and electron microscopy that pancreatitis is characterized by a cell population resembling SPEM[16]. To identify master regulator transcription factors driving this phenotype, we performed a PyScenic-based Regulon analysis on our scRNA-seq dataset (~13,000 cells), which identifies candidate factors by expression of known downstream target genes (**Figure 7A**). Among the candidate regulators, we identified the transcription factor *Spdef*, which has been shown to be the master regulator of mucin-producing goblet cell formation in the stomach, intestines, and lung (**Figure 7B**)[46-48]. Consistent with a role for driving mucin production, we found that *GNAS* mutations in human cell lines are associated with both elevated mucin and *SPDEF* expression (**Figure S2F**, **Figure 7C**) and that $GNAS^{R201C}$ expression is sufficient to drive *Spdef* expression in *Kras;GNAS* murine cell lines (4838 and C241) (**Figure 7E**). Among the top 10% of predicted *Spdef* target genes, we identified several SPEM markers (*Gkn3*, *Muc6*) as well as additional transcription factors predicted to drive this phenotype (*Creb3l1*, *Creb3l4*) (**Figure 7D**, **File S8**). Like *Spdef*, $GNAS^{R201C}$ was sufficient to increase expression of both *Creb3l1* (4838 cells) and *Creb3l4* (4838 and C241 cells) (**Figure 7E**). Regulon analysis predicted that both *Creb3l1* and *Creb3l4* regulate expression of canonical SPEM markers (*Aqp5*, *Tff2*, *Muc6*, among others) as well the expression of each other (*Spdef*, *Creb3l1* and *Creb3l4*) (**Figure S5-6, FileS9, File S10**). To compare these data to $GNAS^{R201C}$-induced changes in gene expression, we first overlaid the DOX-on gene signatures from both the 4838 and C241 cell lines onto the pancreatitis scRNA-seq dataset and identified enrichment in the SPEM cluster (**Figure 7F-G**); expression patterns closely resembled that for *Spdef* (**Figure 7B**). We next examined changes in expression for *Spdef*, *Creb3l1*, and *Creb3l4* target genes with DOX treatment and $GNAS^{R201C}$ expression. We identified an increase in target genes with DOX treatment in both cell lines for all three transcription factors, including SPEM markers

(**Figure 7H**, **S7**, **S8**), suggesting that *GNAS*^{R201C} drives expression of a pyloric metaplasia program through the activity of one or more of these master regulator transcription factors. Consistent with this, treatment of 4838 cells with *Spdef* siRNA significantly decreased mucin marker GSII as compared to control (**Figure 7I-K**).

To investigate the relevance of SPDEF, CREB3L1, and CREB3L4 to human IPMN and PDAC, we next examined expression in patient samples. First, we interrogated the bulk RNA-seq dataset generated by Maurer et al., of IPMN (n = 19), PanIN (n = 26), and PDAC (n = 197 epithelium, 124 stroma)[33, 34]. Expression of all three transcription factors was elevated in the epithelium of all disease states, as well as elevated *CREB3L1* in the stroma (**Figure 8A**). To examine protein expression, we next conducted IHC on 23 tissue samples collected from patients with IPMN +/- associated PDAC and scored expression. Regions of interest encompassed normal ducts (n = 43-45), acinar to ductal metaplasia (ADM, n = 39-41), LG IPMN (n = 51-54), HG IPMN (n = 27-33), and invasive PDAC (n = 9-10). Low expression of all three transcription factors was detected in a portion of normal ducts with expression increasing in ADM and reaching significance for SPDEF and CREB3L1 (**Figure 8B-C**). While CREB3L4 expression remained low throughout IPMN to PDAC progression, SPDEF remained high, though expression did not significantly increase between LG IPMN and PDAC. CREB3L1 expression, though, did significantly increase with disease progression from normal ducts through IPMN to PDAC (**Figure 8B-C**). Collectively, these data suggest a role for predicted master regulators of pyloric metaplasia in IPMN progression in patients.

DISCUSSION

PanIN and IPMN, the two most prevalent PDAC precursor lesions, are defined by mucin-producing cells harboring defining genetic alterations. The involvement of pyloric metaplasia has been described previously in murine PanIN and we confirm a similar process here in human and murine IPMNs[16, 17]. Pyloric metaplasia-defining MUC5AC, CD44v9, and AQP5 are strongly expressed in acinar-to-ductal metaplasia and their co-expression is identified in dysplastic stages of IPMN progression (**Figure 2**). Furthermore, PanIN and IPMN pyloric metaplasia cells harbor similar whole-transcriptomic signatures, reciprocating gastric SPEM and foveolar pit cell lineages and raising the possibility of a conserved program between these organs[16, 17]. The identification of similar processes of injury and repair between gastrointestinal organs could lead to the discovery of targetable pathways for multiple inflammatory or pre-malignant conditions.

IPMNs differ clinically from PanIN through their capacity to be detected by routine imaging due to their larger size and cystic nature, filled with mucin. To the best of our knowledge, there is no insight as to why these precursors behave distinctly yet ultimately complete their sequence by transforming into adenocarcinomas with ductal features. We discovered that both oncogenic KRAS and GNAS drive pyloric metaplasia signatures, yet oncogenic GNAS^{R201C} amplifies a mucinous phenotype, consistent with a role described for GNAS^{R201H}[49]. *In vivo*, oncogenic GNAS also increased the relative proportion of mucin-expressing cells in early lesions likely leading to the formation of cysts. Concomitantly, other cell types associated with KRAS activation, such as tuft and enteroendocrine cells, show a relative reduction following

GNAS activation. Previously, we showed that tuft cell ablation accelerated pancreatic tumorigenesis through PanIN progression[50]; the lower abundance of tuft cells in *KrasGNAS* mice may partially explain the accelerated phenotype in these mice[10].

As oncogenic GNAS mutations are largely restricted to IPMNs in humans, our finding suggests that GNAS could be driving the formation of these disease-defining cysts. These findings concur with those by Liffers et al., which show that human PanIN and gastric-type IPMN are remarkably similar at the transcriptomic level but highlight MUCL3 as a distinctive marker expressed only in gastric IPMNs[11]. Our findings show that oncogenic GNAS drives MUCL3 expression and is the likely root cause for this difference (**Figure 3**). A core limitation of this interpretation is the existence of IPMNs lacking GNAS mutations. We are hard pressed to understand what drives cyst formation in these cases, though our data suggest that KRAS is sufficient and that these IPMN may be MUCL3-negative. Studies in mice have suggested the formation of IPMN-like, AQP5+ lesions driven be oncogenic KRAS and nuclear expression of GSK3β[51]. Further, a recent study shows that loss of RNF43 expression in the context of oncogenic KRAS expression is sufficient to drive IPMN/cyst formation and accelerated tumorigenesis in mice[52]. A follow-up study highlighting the features, size, and mucin content of patient IPMN could shed further light into the discrepancies between IPMNs and PanIN.

Through siRNA testing and regulon analyses, we found that oncogenic GNAS likely drives this mucin-producing phenotype through a master-regulator SPDEF-CREB3L1-CREB3L4 axis. Indeed, these markers increase in expression as early as the acinar-to-ductal metaplasia phase, and CREB3L1 expression continues to increase with disease progression. SPDEF activity is a conserved process identified in the gastric, intestinal, and pulmonary epithelial response to injury and dysplasia, and was recently shown to drive a mucinous phenotype in PanIN[46-48, 53]. GNAS mutations have been suggested to drive a mucinous phenotype in other conditions, such as mucinous carcinomatosis of the appendix, describing a common role for this oncogene in tumorigenesis[54]. This further corroborates that not only is pyloric metaplasia conserved between organs, but also that distinct master regulators underly these processes and the transition between cell states. While some studies have been conducted looking at the functional role for individual pyloric markers in disease progression[26], it remains to be determined if SPEM and the subsequent addition of a foveolar pit lineage, or pyloric metaplasia, is protective or reflects a detectable sign of disease progression.

FIGURE LEGENDS

Figure 1. Analysis of IPMN RNA-seq datasets reveals markers of pyloric metaplasia. SCANPY plots of (**A**) scRNA-seq data from Bernard et al.[30], of tissue from patients with low grade (LG) or high grade (HG) IPMN (n = 2 each), or pancreatic cancer associated with IPMN (PDAC, n = 2). Expression of (**B**) Epithelial marker *EPCAM*, (**C**) immune cell marker *PTPRC*/CD45, (**D**) molecular markers of IPMN subtype or (**E**) SPEM markers trefoil factor 2 (*TFF2*), aquaporin 5 (*AQP5*), and *CD44*. (**F**) Barplots comparing expression (log2TPM) of microdissected epithelium and matched stroma from Maurer et al.[33], from patients with IPMN (n = 19), PanIN (n = 26), and PDAC (n = 197 epithelium, 124 stroma) for *TFF2*, *AQP5*, and *CD44*. **, $p < 0.01$; ****, $p < 0.001$.

Figure 2. Pyloric metaplasia markers are upregulated in human IPMN. (**A**) Pseudo-colored immunohistochemical staining for pyloric metaplasia markers MUC5AC (red), AQP5 (yellow) or CD44v9 (green), top row, and automated detection of signal (bottom row) by QuPath to merge MxIHC data[28]. (**B**) Examples of hematoxylin staining or merged MxIHC staining of low grade (left) or high grade (right) IPMN. Scale bars, 20 μm. (**C**) Quantification of staining in (A-B) for 41 IPMN patients including normal ducts (n = 95-109), acinar to ductal metaplasia (n = 109), low grade IPMN (n = 110), high grade IPMN (n = 70), and invasive IPMN (n = 22). *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.005$; ****, $p < 0.001$.

Figure 3. Oncogenic GNAS drives a pyloric metaplasia program in IPMN. (**A**) qPCR for human *GNAS* (hu*GNAS*) and *Aqp5* in murine cell lines treated +/- doxycycline (DOX)(n = 3 technical replicates). (**B**) Western blot for human $GNAS^{R201C}$. (**C**) Principal component analysis (PCA) of RNA-seq conducted on both cell lines +/- DOX (n = 6, 3 biological replicates). (**D**) Barplots of murine (ms) or human (hu) *GNAS* expression, determined by RNA-seq, in both cell lines. (**E**) Heatmap of top upregulated genes in both cell lines with DOX. Upregulated genes include gastric pit cell markers *Mucl3*, *Gkn1* and *Gkn2*. (**F**) Barplots showing upregulation of gastric pit cell marker *Tff1*, pyloric markers *Tff2*, *Aqp5*, and *Gkn3*, and gastric chief cell markers *Pgc* and *Pga5* with DOX treatment. (**G**) Immunofluorescence and quantification of signal area and intensity for AQP5 and (**H**) CD44v9 in both cell lines +/- DOX (n = 3 biological replicates). Scale bars, 25 μm. Black bars, control; white bars, DOX treatment. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.005$; ****, $p < 0.001$.

Figure 4. Oncogenic GNAS drives mucin production. (**A**) Barplots showing upregulation of mucin genes *Muc1*, *Muc5b*, and *Muc3a* and mucin modification genes *B4galnt3* and *Galnt2* in both 4838 and C241 with DOX treatment and $GNAS^{R201C}$ expression, and *Muc5ac* in C241 alone. (**B**) Representative IF and (**C**) quantification of signal area and intensity for broad mucin marker GSII (n = 3 biological replicates). (**D**) Representative IF and (**E**) quantification of signal area and intensity for mucin MUC5AC (n = 3 biological replicates). Scale bars, 25 μm. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.005$; ****, $p < 0.001$.

Figure 5. $GNAS^{R201C}$ expression preferentially drives mucus cell formation *in vivo*. (**A**) Heat map showing expression of select tuft cell, enteroendocrine cell (EEC), or mucus cell gene markers from Maurer et al.[33], from the epithelium of human IPMN (n = 19), PanIN (n = 26), or PDAC (n = 197)[16, 33]. (**B**) Quantification of tuft cell number (DCLK1+), enteroendocrine cell number (Synaptophysin+), or mucus cell area (PAS/AB+) per epithelial component of lesions in *KrasGNAS* mice +/- DOX as determined by t test. (**C**) Representative images of H&E, DCLK1 or synaptophysin IHC, or PAS/AB (mucin) staining in PanIN or cyst/IPMN from *KrasGNAS* mice +/- DOX. Scale bars, 100 μm for the top three rows, 200 μm for the bottom row. *, $p < 0.05$; **, $p < 0.01$; ****, $p < 0.001$.

Figure 6. Oncogenic Kras and GNAS drive distinct and overlapping transcriptomic programs. (**A**) PCA of RNA-seq conducted on cell line 4838 treated +/- DOX followed by CTRL siRNA or siRNAs targeting *Kras*, *GNAS*, or both *Kras*+*GNAS* (n = 3 per condition). (**B**) Barplots showing expression of hu*GNAS*, ms*Gnas*, *Kras*, or *Aqp5* under all conditions. (**C**) Heatmap showing select genes upregulated by DOX or down regulated by *GNAS*, *Kras*, or *Kras*+*GNAS* siRNA combined. (**D**) Venn diagram of genes down regulated in DOX treated cells with either *Kras* or *GNAS* siRNA. (**E**) Select genes down regulated by only *Kras* or *GNAS* siRNA or common to both. (**F**) Uniform manifold approximation and projection (UMAP) of scRNA-seq data collected from normal murine stomach (top) or a model of acute gastric injury (bottom), overlaid with either murine pancreatitis, *GNAS*-off or *Kras*-off gene signatures. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.005$; ****, $p < 0.001$.

Figure 7. Master regulator transcription factor SPDEF gene targets are upregulated by oncogenic GNAS. (**A**) UMAP of scRNA-seq data collected from a murine model of pancreatitis showing the formation of a gastric SPEM[16]. (**B**) RNA expression or binarized activity of *Spdef* predicted by regulon analysis overlayed on the UMAP from (A). (**C**) Expression of *SPDEF* is enriched in human cancer cell lines bearing *GNAS* mutations. (**D**) Plot of *Spdef* targets predicted by regulon analysis identifying gastric SPEM markers (*Gkn3*, *Muc6*) and transcription factors *Creb3l1* and *Creb3l4*. (**E**) Barplots showing increased expression of *Spdef*, *Creb3l1*, and *Creb3l4* in IPMN cell lines with DOX treatment. (**F**) DOX/GNAS-on gene signatures from IPMN cell line 4838 or (**G**) C241 overlayed on the UMAP in (A) showing overlap with *Spdef* expression. (**H**) Heatmap of *Spdef* target gene expression in control or DOX-treated 4838 cells. (**I**) qPCR for *Spdef*, (**J**)

immunofluorescence for GSII and (**K**) quantification of GSII area or intensity in 4838 cells treated with DOX and either control (CTRL) or *Spdef* siRNA. Scale bars, 5 μm. **, $p < 0.01$; ***, $p < 0.005$; ****, $p < 0.001$.

Figure 8. Predicted pyloric metaplasia master regulator transcription factors are expressed in human IPMN. (**A**) Barplots comparing expression (Log2TPM) of microdissected epithelium and stroma from Maurer et al.[33] from IPMN (n = 19), PanIN (n = 26), and PDAC (n = 197 epithelium, 124 stroma) for *SPDEF*, *CREB3L1*, and *CREB3L4*. (**B-C**) Representative immunohistochemistry and quantification for SPDEF, CREB3L1, and CREB3L4 conducted on 23 IPMN patient specimens including normal ducts (n = 95-109), acinar to ductal metaplasia (n = 109), low grade IPMN (n = 110), high grade IPMN (n = 70), and pancreatic cancer (Invasive, n = 22). *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.005$; ****, $p < 0.001$.

Figure S1. Expression of pyloric metaplasia markers in human pre-malignant lesions and PDAC. Heat map showing expression of SPEM markers identified in a murine model of pancreatitis in a previously reported dataset of laser capture dissected epithelium from patient IPMN (n = 19), PanIN (n = 26), and PDAC (n = 197)[16, 33, 34].

Figure S2. *GNAS^{R201C}* expression drives transcriptomic changes in PDAC cell lines. (**A**) Heat map of differentially expressed genes between cell lines 4838 and C241. (**B**) GSEA analysis of gene expression signature changes in 4838 or (**C**) C241 cells with DOX treatment. (**D**) GO analysis of gene expression signature changes in 4838 or (**E**) C241 cells with DOX treatment. (**F**) *MUC5AC*, *MUC5B*, and *MUC13* expression in PDAC tumors from cBioPortal with wild type (n = 96) or mutant *GNAS$^{R201C/H}$* (n = 4)[55]. ***, $p < 0.005$; ****, $p < 0.001$.

Figure S3. *GNAS^{R201C}* expression drives a cystic phenotype *in vivo*. (**A**) Quantification of highest-grade lesion or cystic score in *KrasGNAS* mice +/- DOX (combined 10- and 20-weeks treatment). Grade, 1 = 1/3 of tissue is low grade; 2 = 2/3 of tissue is low grade; 3 = high grade; 4 = invasive. Cystic score, 0 = no cysts; 1 = mild; 2 = extensive. (**B**) Quantification of tuft cell number (DCLK1), enteroendocrine number (SYP, synaptophysin), or mucinous area (PAS/AB) in lesions from *KrasGNAS* mice +/- DOX (for 10 or 20 weeks) plotted by lesion type (PanIN or IPMN/cyst). *, $p < 0.05$; ***, $p < 0.005$; ****, $p < 0.001$.

Figure S4. Validation of *KRAS* and *GNAS*-specific siRNA. (**A**) Screening of *Kras* siRNA in human pancreatic cancer cell lines PANC1 or MiaPaCa2 by qPCR. (**B**) Selected *Kras* siRNA reduces pERK and *KRAS^{G12D}* expression in 4838 and C241 cells by Western blot. (**C**) Screening of *GNAS* siRNA in 4838 or C241 cells by qPCR. (**D**) Selected GNAS siRNA reduces *GNAS^{R201C}* expression in 4838 and C241 cells by Western blot. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.005$; ****, $p < 0.001$.

Figure S5. Oncogenic *Kras* and *GNAS* drive distinct and overlapping transcriptomic programs. Heatmap, Venn diagram or top 40 gene target list for differentially expressed genes between (**A**) genes upregulated with DOX and downregulated with *GNAS* siRNA, (**B**) DOX treated cells treated with either *Kras* siRNA or *Kras* and *GNAS* siRNA in combination, or (**C**) DOX treated cells with either *GNAS* siRNA or *Kras* and *GNAS* siRNA in combination.

Figure S6. Creb3l1 and Creb3l4 are predicted regulators of pancreatitis induced SPEM. RNA expression or binarized activity of (**A**) *Creb3l1* or (**B**) *Creb3l4* predicted by Regulon analysis overlayed on the UMAP from Figure 7A.

Figure S7. Creb3l1 and Creb3l4 are predicted regulators of pyloric-type metaplasia. Plots of either (**A**) *Creb3l1* or (**B**) *Creb3l4* predicted target genes generated from PyScenic Regulon analysis of scRNA-seq data generated from murine pancreatitis[16]. Select markers are circled in red.

Figure S8. Spdef and Creb3l4 target genes increase with *GNAS^{R201C}* expression. Heat maps of (**A**) SPDEF target genes overlaid on RNA-seq data of C241 cells +/- DOX and *GNAS^{R201C}* expression or (**B**) CREB3L4 target genes overlaid on RNA-seq data of either 4838 (left) or C241 (right) cells +/- DOX. Top 10% of gene targets are shown. Select markers are bolded.

Figure S9. Creb3l1 target genes increase with *GNAS*^{R201C} expression. Heat maps of CREB3L1 target genes overlaid on RNA-seq data of either 4838 (left) or C241 (right) cells +/- DOX and *GNAS*^{R201C} expression. Top 10% of gene targets are shown. Pyloric-type metaplasia markers are highlighted.

Antibody	Company	Catalog #	Dilution, WB	Dilution, IHC or IF
AQP5	Sigma-Aldrich	HPA065008	N/A	1:500
CD44v9 (human)	Cosmo Bio	LKG-M003	N/A	1:250
CD44v10e16 (mouse)	Cosmo Bio	LKG-M002	N/A	1:250
CREB3L1	Sigma	HPA024069	N/A	1:300
CREB3L4	Sigma	HPA038122	N/A	1:300
DCLK1	Abcam	Ab37994	N/A	1:500
ERK (phosphorylated)	Cell Signaling	4370	1:1000	N/A
GAPDH	Cell Signaling	5174	1:1000	N/A
GNAS (R201C)	GeneTex	GTX135412	1:1000	N/A
GSII	Invitrogen	L21415	N/A	1:250
KRAS (G12D)	Cell Signaling	14429	1:1000	N/A
MUC5AC	Invitrogen	MA-12178	N/A	1:500
SPDEF	Lifespan Bioscience	LS-C-499857	N/A	1:300
Synaptophysin (SYP)	Cell Marque	336R-94	N/A	1:500

Table S1. Primary antibodies used for western blot (WB), immunohistochemistry (IHC), or immunofluorescence (IF) studies.

Gene	Forward Primer	Reverse Primer
huGNAS	CCT GAG TGT GAT GAA CGT GCC TG	CGA AGC AGG TCC TGA TCG CTC
RPLPO	GCC AAT AAG GTG CCA GCT G	CTC CCA CCT TGT CTC CAG TC.
Aqp5	GCC ATC TTG TGG GGA TCT AC	CCC AGA AGA CCC AGT GAG AG
Kras	GAG AGG CCT GCT GAA AAT GAC TG	GTC CCT CAT TGC ACT GTA CTC CTC
Spdef	TCC TCT CTG CTC ACT CTG AA	AGA GCT CAT GTG TAT CCC TAG A

Table S2. Primers for qRTPCR.

DATA AVAILABILITY STATEMENT

Sequencing data generated in this study is available in the gene expression omnibus (GSE244280).

ACKNOWLEDGEMENTS

The authors would like to thank Sydney Batts and JoAnna Dennis for technical assistance and James Goldenring for critical reading of the manuscript. The Vanderbilt Creative Data Solutions Shared Resource (RRID:SCR_022366) performed and/or assisted with RNA-Seq data processing, analysis, and deposition to GEO. The Vanderbilt Institute of Nanoscale Science and Engineering core facility assisted with the fabrication of LNP-siRNAs. The Vanderbilt University Medical Center VANTAGE Core provided technical assistance for sample preparation and sequencing, and VANTAGE is supported in part by Clinical and Translational Science Award Grant 5UL1 RR024975-03, Vanderbilt Ingram Cancer Center Grant P30 CA68485, Vanderbilt Vision Center Grant P30 EY08126, and National Institutes of Health/National Center for Research Resources Grant G20 RR030956

CONFLICTS OF INTEREST

AM is listed as an inventor on a patent that has been licensed by Johns Hopkins University to ThriveEarlier Detection. AM serves as a consultant for Tezcat Biotechnology.

FUNDING

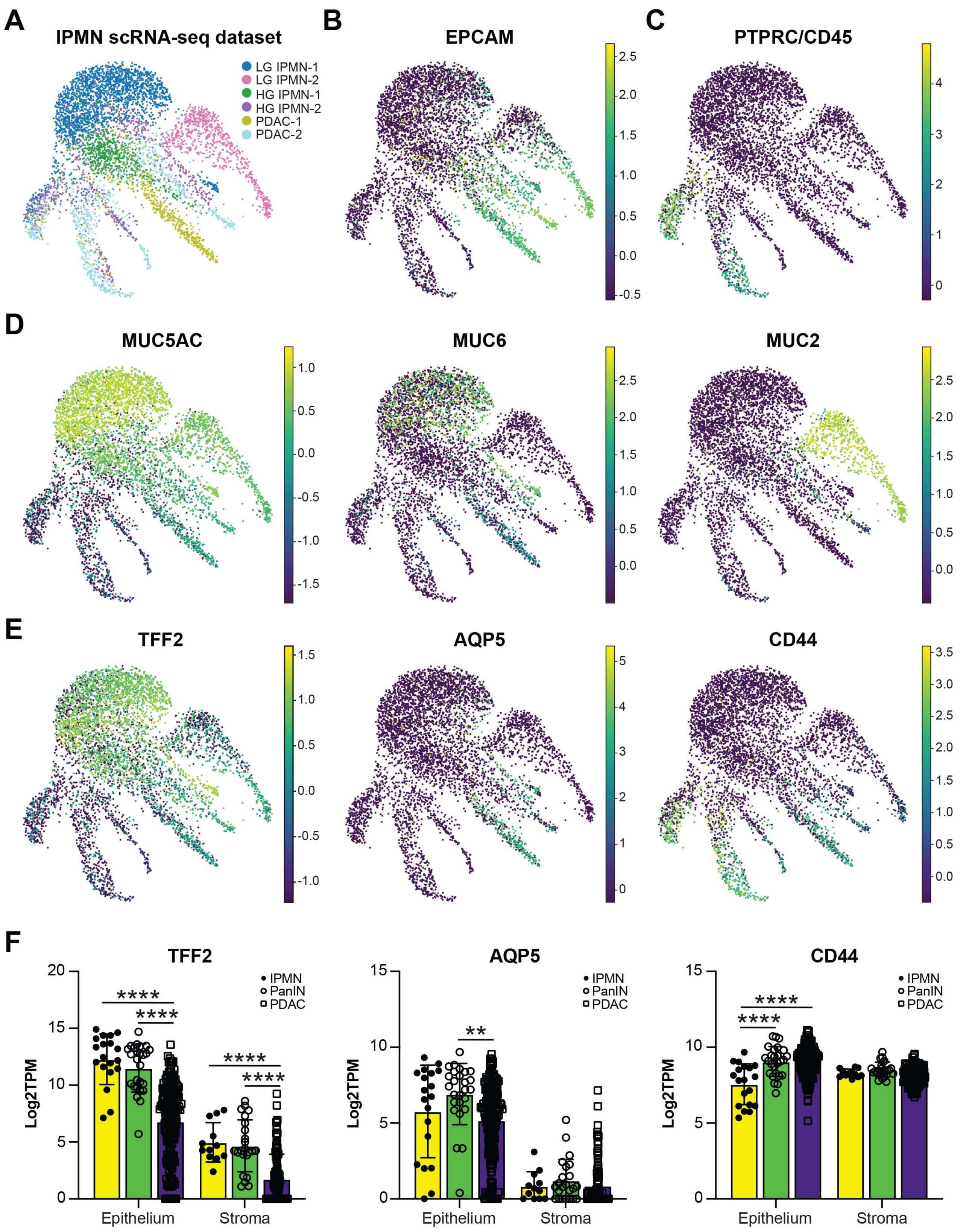
VQT is supported by the Canadian Cancer Society Breakthrough Grant (BTG-23), the Institute for Research in Immunology and Cancer start-up funds, the Faculty of Medicine of the University of Montreal Salary Support Grant for Clinical Scholars, the Fonds de Recherche Québec Santé Clinical Scholar Establishment Fund, the Fonds de Recherche Québec Santé Clinical Scholar J1 Salary Support Grant, and the McLaughlin Foundation Fellowship Scholarship. AMR was supported by NIH T32GM139400. AM is supported by the MD Anderson Pancreatic Cancer Moon Shot Program, the Sheikh Khalifa Bin Zayed Al-Nahyan Foundation and NIH (U01CA200468, U54CA274371, R01CA220236). MCBT is supported by a Vanderbilt Digestive Disease Research Center Pilot and Feasibility Grant (P30 DK058404), Vanderbilt Supporting Careers in Research for Interventional Physicians and Surgeons (SCRIPS) Faculty Research Award [VUMC66796 (1018894)], and a Nikki Mitchell Foundation Pancreas Club Seed Grant. MCBT and NJ are supported by The Department of Defense (DOD W81XWH2211121-1). HCM received support the KKF Clinician-Scientist program, TUM. The DelGiorno laboratory is supported by the Vanderbilt Ingram Cancer Center Support Grant (NIH/NCI P30 CA068485), the Vanderbilt-Ingram Cancer Center SPORE in Gastrointestinal Cancer (NIH/NCI 5P50 CA236733), the Vanderbilt Digestive Disease Research Center (NIH/NIDDK P30 DK058404), an American Gastroenterological Association Research Scholar Award (AGA2021-13-02), NIH/NIGMS R35 GM142709, The Department of Defense (DOD W81XWH2211121), The Sky Foundation, Inc (AWD00000079), and Linda's Hope (Nashville, TN).

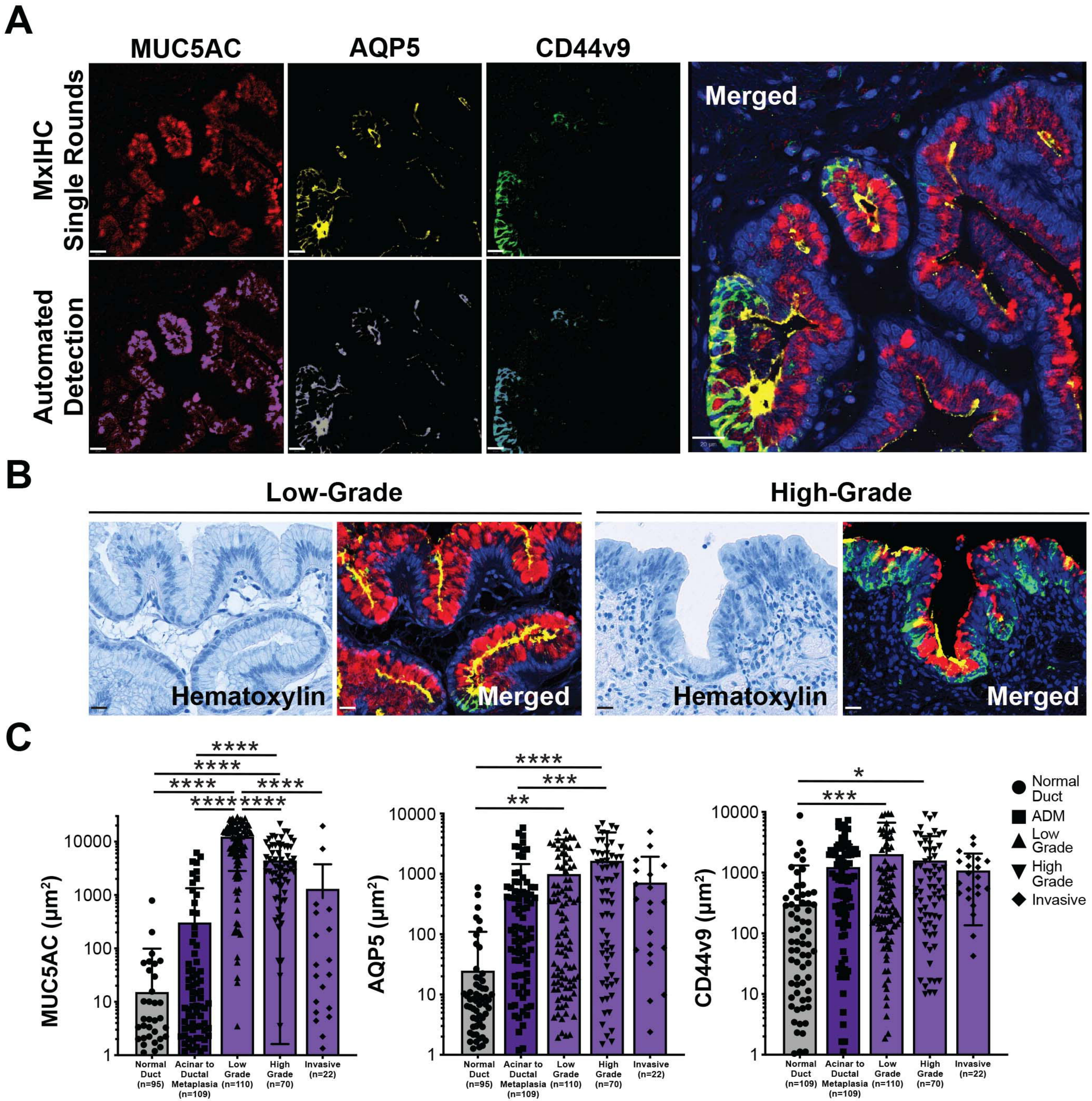
REFERENCES

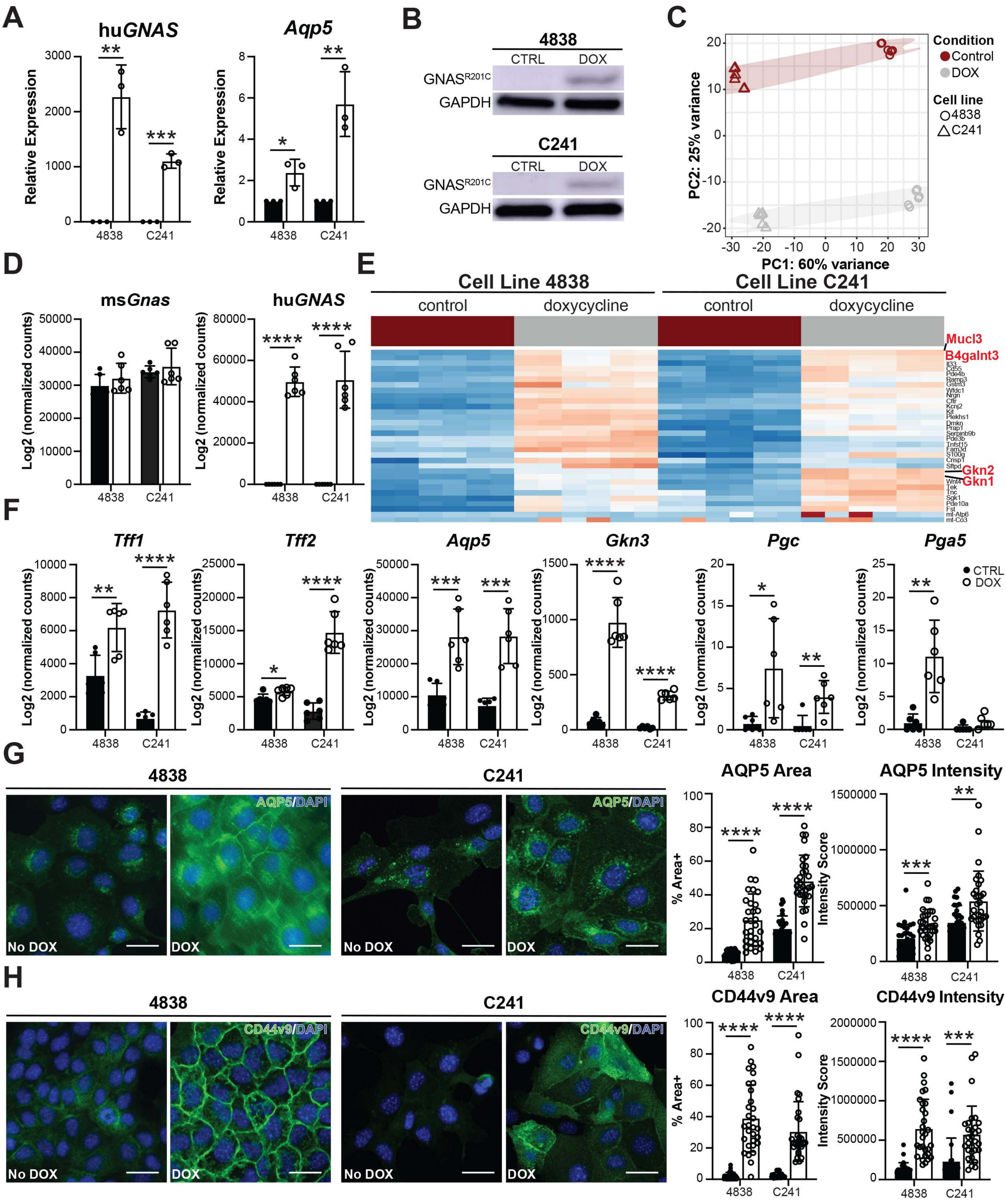
1. Rahib L, Wehner MR, Matrisian LM, et al. Estimated Projection of US Cancer Incidence and Death to 2040. JAMA Netw Open 2021;4:e214708.
2. Yachida S, Jones S, Bozic I, et al. Distant metastasis occurs late during the genetic evolution of pancreatic cancer. Nature 2010;467:1114-7.
3. Maitra A, Fukushima N, Takaori K, et al. Precursors to invasive pancreatic cancer. Adv Anat Pathol 2005;12:81-91.
4. Del Chiaro M, Ateeb Z, Hansson MR, et al. Survival Analysis and Risk for Progression of Intraductal Papillary Mucinous Neoplasia of the Pancreas (IPMN) Under Surveillance: A Single-Institution Experience. Ann Surg Oncol 2017;24:1120-1126.
5. Laffan TA, Horton KM, Klein AP, et al. Prevalence of unsuspected pancreatic cysts on MDCT. AJR Am J Roentgenol 2008;191:802-7.
6. Hingorani SR, Petricoin EF, Maitra A, et al. Preinvasive and invasive ductal pancreatic cancer and its early detection in the mouse. Cancer Cell 2003;4:437-50.
7. Springer S, Masica DL, Dal Molin M, et al. A multimodality test to guide the management of patients with a pancreatic cyst. Sci Transl Med 2019;11.
8. Tan MC, Basturk O, Brannon AR, et al. GNAS and KRAS Mutations Define Separate Progression Pathways in Intraductal Papillary Mucinous Neoplasm-Associated Carcinoma. J Am Coll Surg 2015;220:845-854 e1.
9. Springer S, Wang Y, Dal Molin M, et al. A combination of molecular markers and clinical features improve the classification of pancreatic cysts. Gastroenterology 2015;149:1501-10.

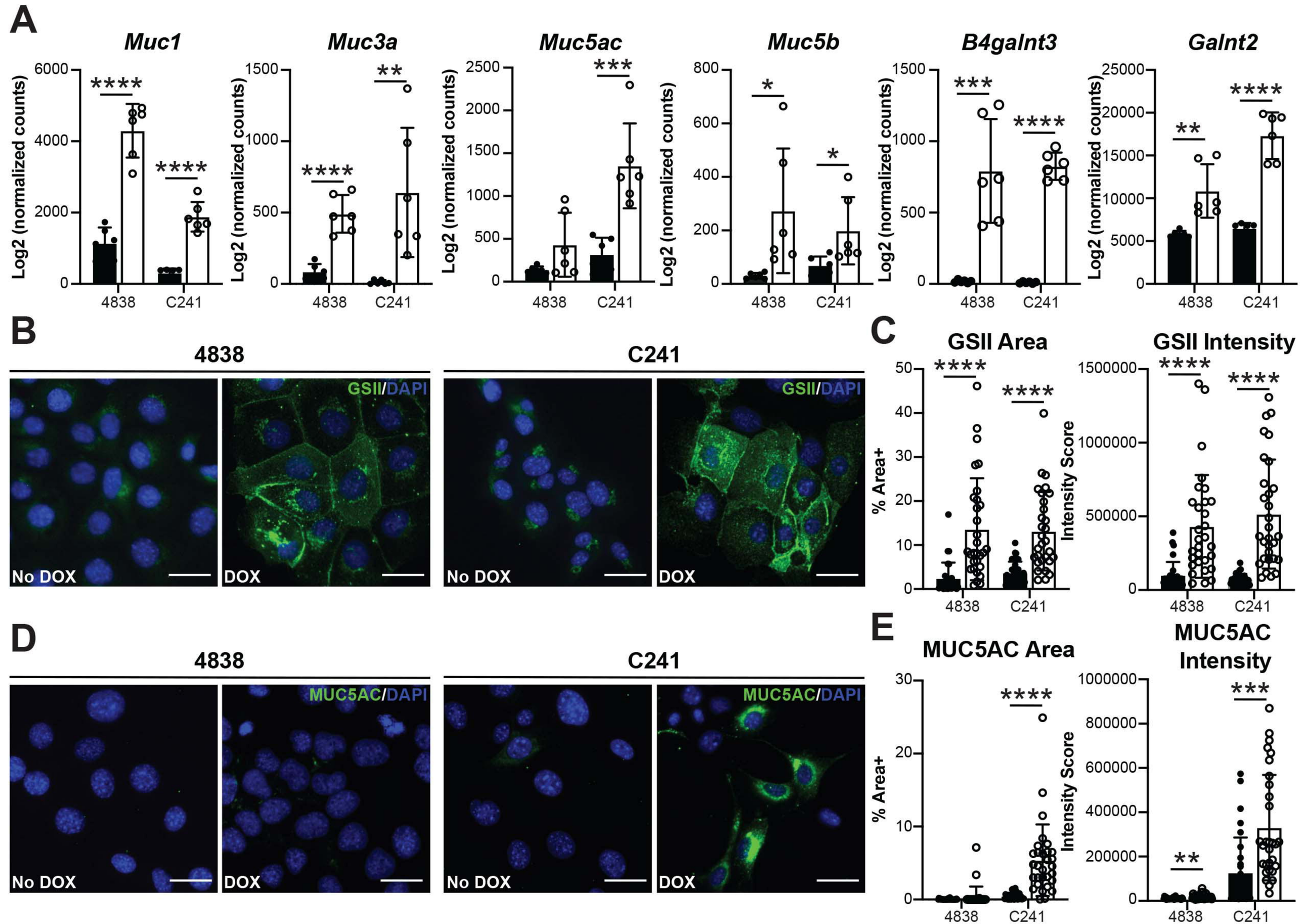
10. Ideno N, Yamaguchi H, Ghosh B, et al. GNAS(R201C) Induces Pancreatic Cystic Neoplasms in Mice That Express Activated KRAS by Inhibiting YAP1 Signaling. Gastroenterology 2018;155:1593-1607 e12.
11. Liffers ST, Godfrey L, Frohn L, et al. Molecular heterogeneity and commonalities in pancreatic cancer precursors with gastric and intestinal phenotype. Gut 2023;72:522-534.
12. Iyer MK, Shi C, Eckhoff AM, et al. Digital spatial profiling of intraductal papillary mucinous neoplasms: Toward a molecular framework for risk stratification. Sci Adv 2023;9:eade4582.
13. Alison MR. The cellular origins of cancer with particular reference to the gastrointestinal tract. Int J Exp Pathol 2020;101:132-151.
14. Murtaugh LC, Keefe MD. Regeneration and repair of the exocrine pancreas. Annu Rev Physiol 2015;77:229-49.
15. Storz P. Acinar cell plasticity and development of pancreatic ductal adenocarcinoma. Nat Rev Gastroenterol Hepatol 2017;14:296-304.
16. Ma Z, Lytle NK, Chen B, et al. Single-Cell Transcriptomics Reveals a Conserved Metaplasia Program in Pancreatic Injury. Gastroenterology 2022;162:604-620 e20.
17. Schlesinger Y, Yosefov-Levi O, Kolodkin-Gal D, et al. Single-cell transcriptomes of pancreatic preinvasive lesions and cancer reveal acinar metaplastic cells' heterogeneity. Nat Commun 2020;11:4516.
18. Chen B, Scurrah CR, McKinley ET, et al. Differential pre-malignant programs and microenvironment chart distinct paths to malignancy in human colorectal polyps. Cell 2021;184:6262-6280 e26.
19. Schmidt PH, Lee JR, Joshi V, et al. Identification of a metaplastic cell lineage associated with human gastric adenocarcinoma. Lab Invest 1999;79:639-46.
20. Nam KT, Lee HJ, Sousa JF, et al. Mature chief cells are cryptic progenitors for metaplasia in the stomach. Gastroenterology 2010;139:2028-2037 e9.
21. Bockerstett KA, Lewis SA, Noto CN, et al. Single-Cell Transcriptional Analyses Identify Lineage-Specific Epithelial Responses to Inflammation and Metaplastic Development in the Gastric Corpus. Gastroenterology 2020;159:2116-2129 e4.
22. Lee SH, Jang B, Min J, et al. Up-regulation of Aquaporin 5 Defines Spasmolytic Polypeptide-Expressing Metaplasia and Progression to Incomplete Intestinal Metaplasia. Cell Mol Gastroenterol Hepatol 2022;13:199-217.
23. Goldenring JR, Mills JC. Cellular Plasticity, Reprogramming, and Regeneration: Metaplasia in the Stomach and Beyond. Gastroenterology 2022;162:415-430.
24. Goldenring JR. Pyloric metaplasia, pseudopyloric metaplasia, ulcer-associated cell lineage and spasmolytic polypeptide-expressing metaplasia: reparative lineages in the gastrointestinal mucosa. J Pathol 2018;245:132-137.
25. Meyer AR, Engevik AC, Willet SG, et al. Cystine/Glutamate Antiporter (xCT) Is Required for Chief Cell Plasticity After Gastric Injury. Cell Mol Gastroenterol Hepatol 2019;8:379-405.
26. Wada T, Ishimoto T, Seishima R, et al. Functional role of CD44v-xCT system in the development of spasmolytic polypeptide-expressing metaplasia. Cancer Sci 2013;104:1323-9.
27. Iglesias-Bartolome R, Torres D, Marone R, et al. Inactivation of a Galpha(s)-PKA tumour suppressor pathway in skin stem cells initiates basal-cell carcinogenesis. Nat Cell Biol 2015;17:793-803.
28. Bankhead P, Loughrey MB, Fernandez JA, et al. QuPath: Open source software for digital pathology image analysis. Sci Rep 2017;7:16878.
29. Shihan MH, Novo SG, Le Marchand SJ, et al. A simple method for quantitating confocal fluorescent images. Biochem Biophys Rep 2021;25:100916.
30. Bernard V, Semaan A, Huang J, et al. Single-Cell Transcriptomics of Pancreatic Cancer Precursors Demonstrates Epithelial and Microenvironmental Heterogeneity as an Early Event in Neoplastic Progression. Clin Cancer Res 2019;25:2194-2205.
31. Bockerstett KA, Lewis SA, Wolf KJ, et al. Single-cell transcriptional analyses of spasmolytic polypeptide-expressing metaplasia arising from acute drug injury and chronic inflammation in the stomach. Gut 2020;69:1027-1038.
32. Maurer HC, Olive KP. Laser Capture Microdissection on Frozen Sections for Extraction of High-Quality Nucleic Acids. Methods Mol Biol 2019;1882:253-259.
33. Maurer C, Holmstrom SR, He J, et al. Experimental microdissection enables functional harmonisation of pancreatic cancer subtypes. Gut 2019;68:1034-1043.
34. Maurer HC, Curiel-Garcia A, Holmstrom S, et al. Ras-dependent activation of BMAL2 regulates hypoxic metabolism in pancreatic cancer. bioRxiv 2023.
35. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15:550.

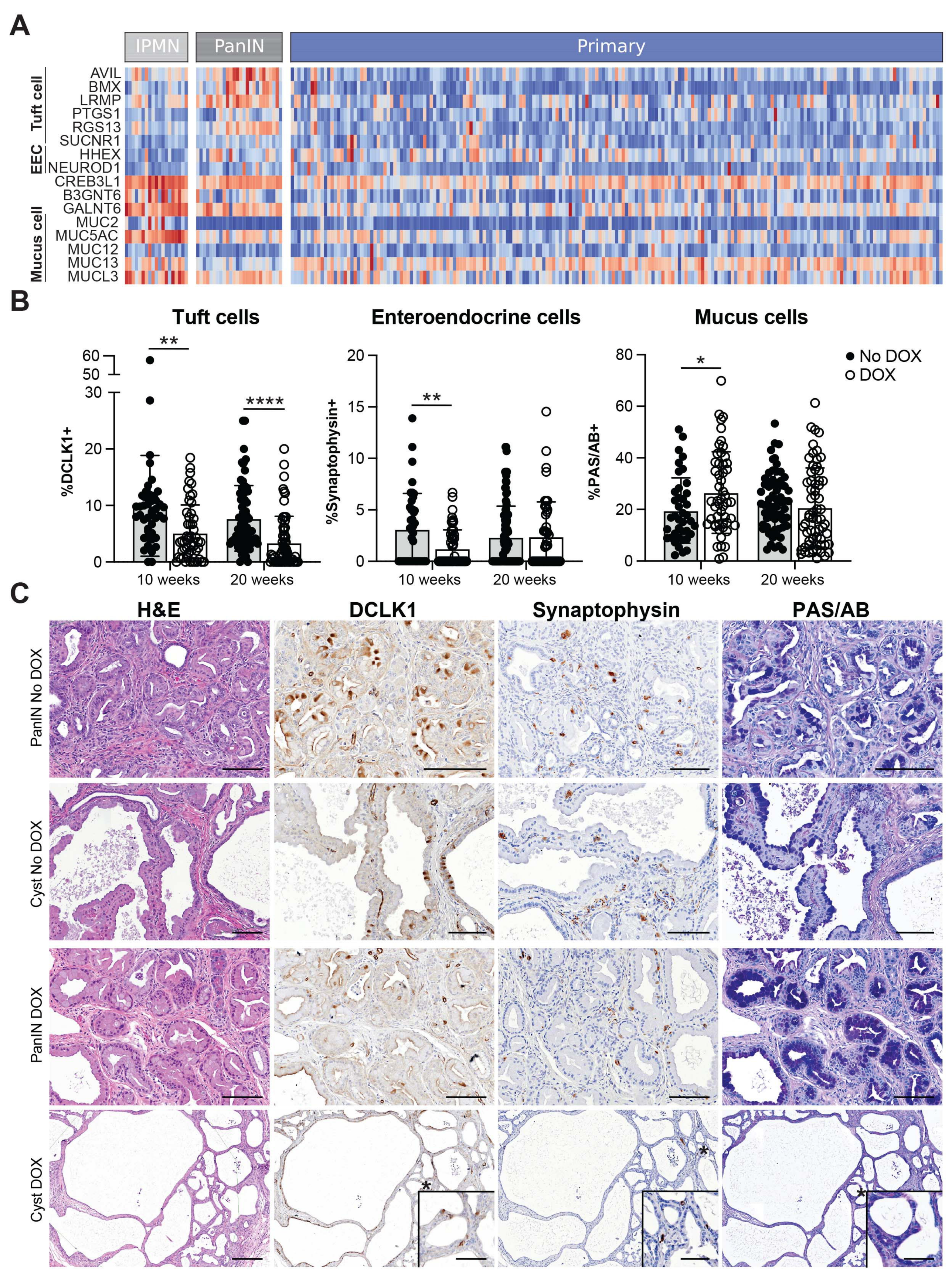
36. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013;29:15-21.
37. Yu G, Wang LG, Han Y, et al. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 2012;16:284-7.
38. Liberzon A, Birger C, Thorvaldsdottir H, et al. The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst 2015;1:417-425.
39. Aibar S, Gonzalez-Blas CB, Moerman T, et al. SCENIC: single-cell regulatory network inference and clustering. Nat Methods 2017;14:1083-1086.
40. Stuart T, Butler A, Hoffman P, et al. Comprehensive Integration of Single-Cell Data. Cell 2019;177:1888-1902 e21.
41. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics 2016;32:2847-9.
42. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003;13:2498-504.
43. Delgiorno KE, Hall JC, Takeuchi KK, et al. Identification and manipulation of biliary metaplasia in pancreatic tumors. Gastroenterology 2014;146:233-44 e5.
44. Farrell AS, Joly MM, Allen-Petersen BL, et al. MYC regulates ductal-neuroendocrine lineage plasticity in pancreatic ductal adenocarcinoma associated with poor outcome and chemoresistance. Nat Commun 2017;8:1728.
45. Caplan LR, Vavinskaya V, Gelikman DG, et al. Enteroendocrine Cell Formation Is an Early Event in Pancreatic Tumorigenesis. Front Physiol 2022;13:865452.
46. Horst D, Gu X, Bhasin M, et al. Requirement of the epithelium-specific Ets transcription factor Spdef for mucous gland cell function in the gastric antrum. J Biol Chem 2010;285:35047-55.
47. Gregorieff A, Stange DE, Kujala P, et al. The ets-domain transcription factor Spdef promotes maturation of goblet and paneth cells in the intestinal epithelium. Gastroenterology 2009;137:1333-45 e1-3.
48. Chen G, Korfhagen TR, Xu Y, et al. SPDEF is required for mouse pulmonary goblet cell differentiation and regulates a network of genes associated with mucus production. J Clin Invest 2009;119:2914-24.
49. Komatsu H, Tanji E, Sakata N, et al. A GNAS mutation found in pancreatic intraductal papillary mucinous neoplasms induces drastic alterations of gene expression profiles with upregulation of mucin genes. PLoS One 2014;9:e87875.
50. DelGiorno KE, Chung CY, Vavinskaya V, et al. Tuft Cells Inhibit Pancreatic Tumorigenesis in Mice by Producing Prostaglandin D2. Gastroenterology 2020.
51. Ding L, Roeck K, Zhang C, et al. Nuclear GSK-3beta and Oncogenic KRas Lead to the Retention of Pancreatic Ductal Progenitor Cells Phenotypically Similar to Those Seen in IPMN. Front Cell Dev Biol 2022;10:853003.
52. Hosein AN, Dangol G, Okumura T, et al. Loss of Rnf43 Accelerates Kras-Mediated Neoplasia and Remodels the Tumor Immune Microenvironment in Pancreatic Adenocarcinoma. Gastroenterology 2022;162:1303-1318 e18.
53. Tonelli C, Yordanov GN, Hao Y, et al. A mucus production programme promotes classical pancreatic ductal adenocarcinoma. Gut 2024.
54. Alakus H, Babicky ML, Ghosh P, et al. Genome-wide mutational landscape of mucinous carcinomatosis peritonei of appendiceal origin. Genome Med 2014;6:43.
55. Bailey P, Chang DK, Nones K, et al. Genomic analyses identify molecular subtypes of pancreatic cancer. Nature 2016;531:47-52.

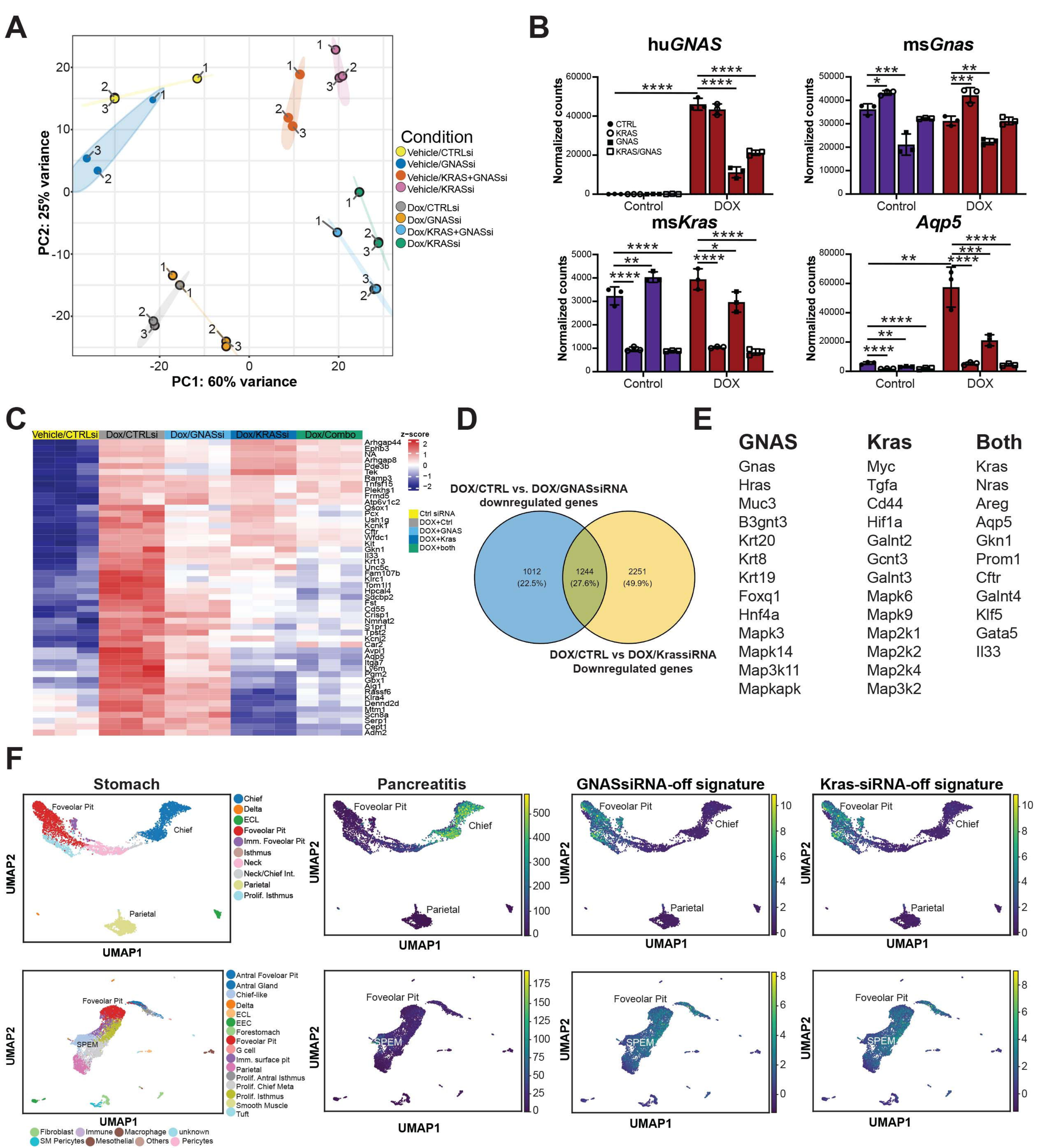

A
IPMN scRNA-seq dataset
LG IPMN-1
LG IPMN-2
HG IPMN-1
HG IPMN-2
PDAC-1
PDAC-2
B
EPCAM
C
PTPRC/CD45
D
MUC5AC
MUC6
MUC2
E
TFF2
AQP5
CD44
F
TFF2
AQP5
CD44
IPMN
PanIN
PDAC
Log2TPM
Epithelium
Stroma
****
**

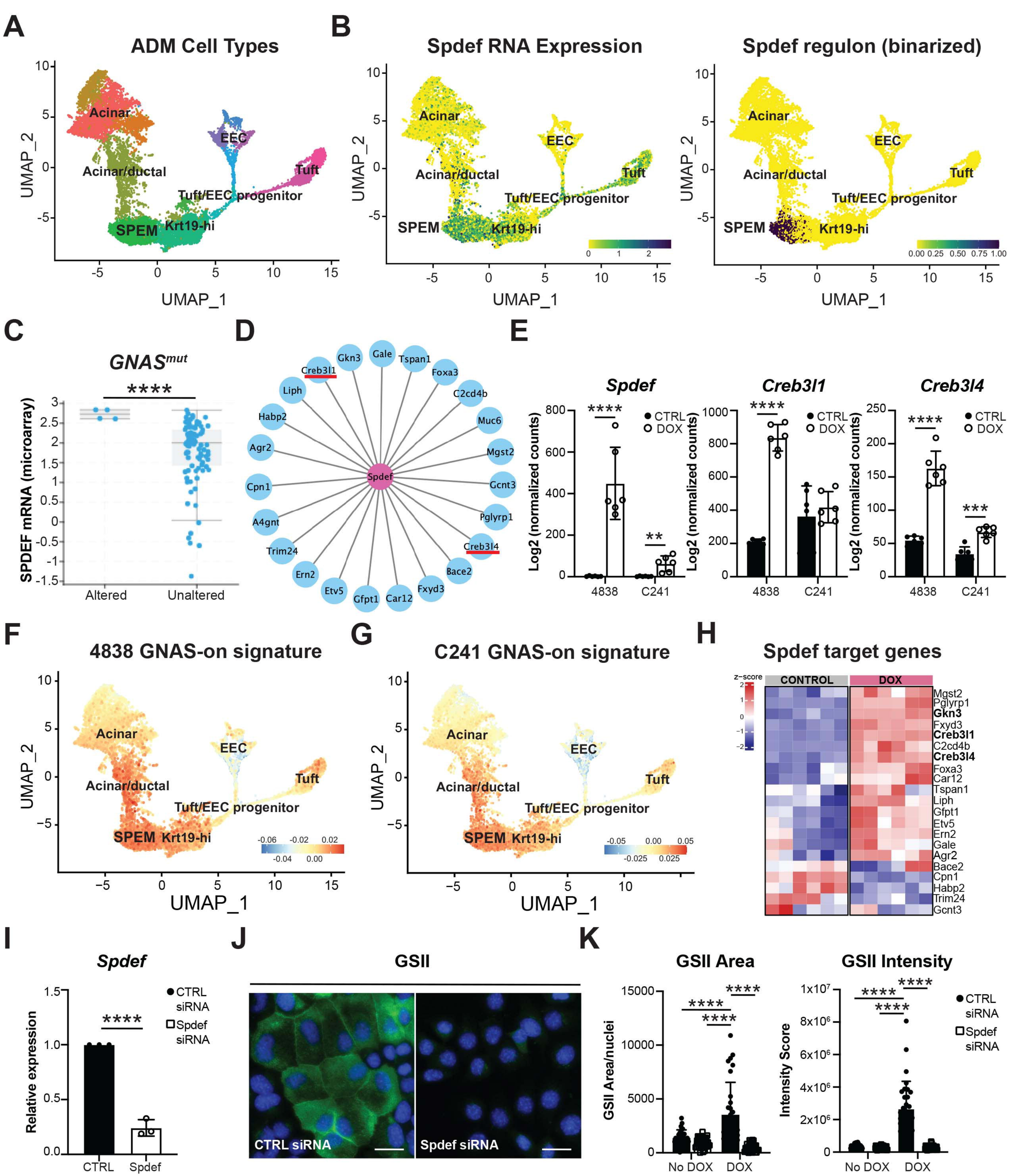

A
MUC5AC
AQP5
CD44v9
MxIHC Single Rounds
Automated Detection
Merged
20 µm
B
Low-Grade
High-Grade
Hematoxylin
Merged
Hematoxylin
Merged
C
MUC5AC (µm²)
AQP5 (µm²)
CD44v9 (µm²)
10000
1000
100
10
1
****
***
**
*
Normal Duct (n=95)
Acinar to Ductal Metaplasia (n=109)
Low Grade (n=110)
High Grade (n=70)
Invasive (n=22)
Normal Duct (n=109)
Normal Duct
ADM
Low Grade
High Grade
Invasive

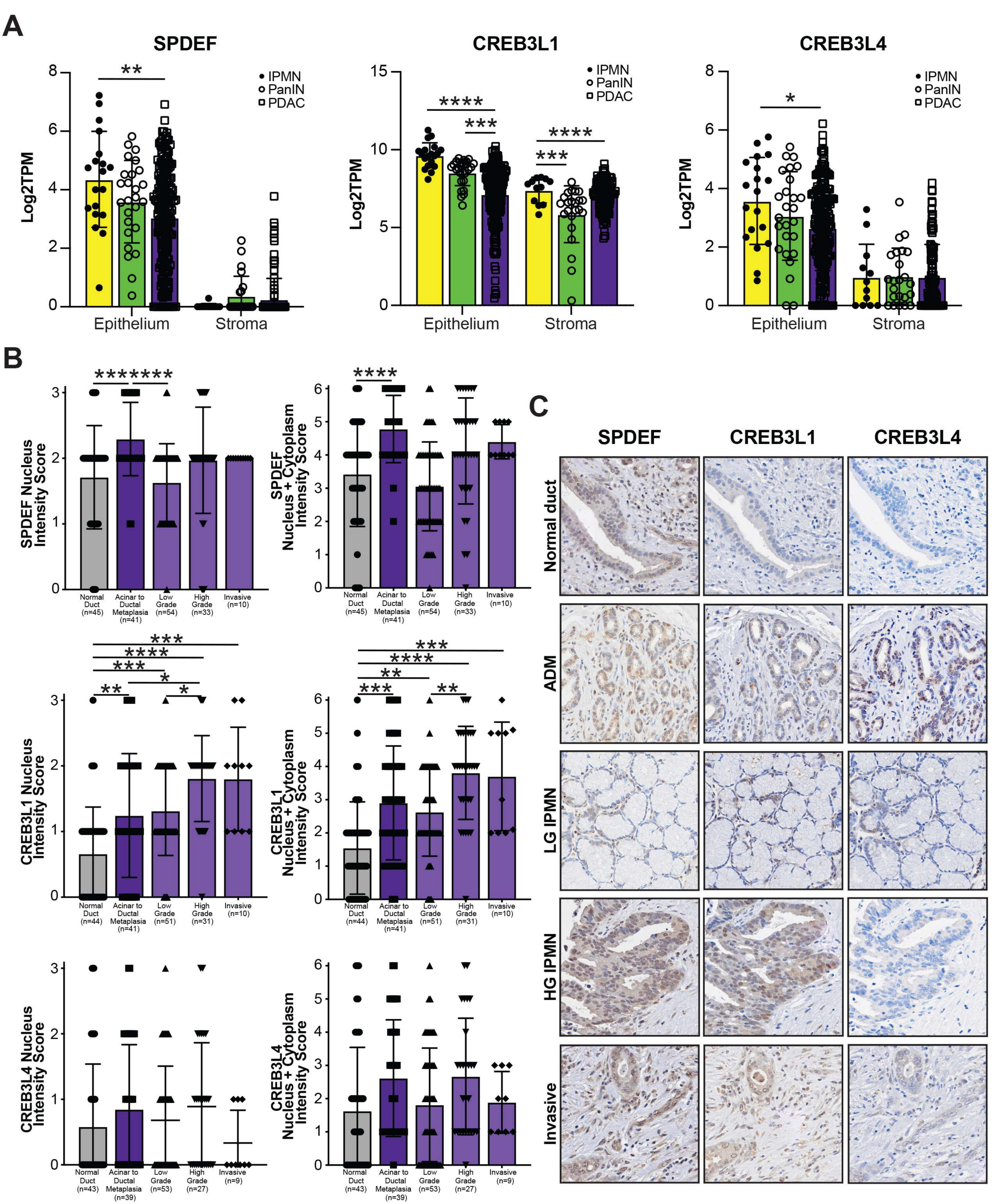

A
huGNAS
Aqp5
Relative Expression
4838
C241
B
4838
CTRL
DOX
GNASR201C
GAPDH
C241
C
Condition
Control
DOX
Cell line
4838
C241
PC2: 25% variance
PC1: 60% variance
D
msGnas
huGNAS
Log2 (normalized counts)
E
Cell Line 4838
Cell Line C241
control
doxycycline
Mucl3
B4galnt3
Gkn2
Gkn1
F
Tff1
Tff2
Aqp5
Gkn3
Pgc
Pga5
CTRL
DOX
G
4838
C241
AQP5 Area
AQP5 Intensity
AQP5/DAPI
No DOX
DOX
% Area+
Intensity Score
H
CD44v9 Area
CD44v9 Intensity
CD44v9/DAPI


A
Muc1
Muc3a
Muc5ac
Muc5b
B4galnt3
Galnt2
Log2 (normalized counts)
4838
C241
B
4838
C241
GSII/DAPI
No DOX
DOX
C
GSII Area
GSII Intensity
% Area+
Intensity Score
D
MUC5AC/DAPI
E
MUC5AC Area
MUC5AC Intensity


A
IPMN
PanIN
Primary
Tuft cell
EEC
Mucus cell
AVIL
BMX
LRMP
PTGS1
RGS13
SUCNR1
HHEX
NEUROD1
CREB3L1
B3GNT6
GALNT6
MUC2
MUC5AC
MUC12
MUC13
MUCL3
B
Tuft cells
Enteroendocrine cells
Mucus cells
%DCLK1+
%Synaptophysin+
%PAS/AB+
10 weeks
20 weeks
No DOX
DOX
C
H&E
DCLK1
Synaptophysin
PAS/AB
PanIN No DOX
Cyst No DOX
PanIN DOX
Cyst DOX


A
PC2: 25% variance
PC1: 60% variance
Condition
Vehicle/CTRLsi
Vehicle/GNASsi
Vehicle/KRAS+GNASsi
Vehicle/KRASsi
Dox/CTRLsi
Dox/GNASsi
Dox/KRAS+GNASsi
Dox/KRASsi
B
huGNAS
msGnas
msKras
Aqp5
Normalized counts
CTRL
KRAS
GNAS
KRAS/GNAS
Control
DOX
C
Vehicle/CTRLsi
Dox/CTRLsi
Dox/GNASsi
Dox/KRASsi
Dox/Combo
z-score
Ctrl siRNA
DOX+Ctrl
DOX+GNAS
DOX+Kras
DOX+both
D
DOX/CTRL vs. DOX/GNASsiRNA downregulated genes
1012 (22.5%)
1244 (27.6%)
2251 (49.9%)
DOX/CTRL vs DOX/KrassiRNA Downregulated genes
E
GNAS
Gnas
Hras
Muc3
B3gnt3
Krt20
Krt8
Krt19
Foxq1
Hnf4a
Mapk3
Mapk14
Map3k11
Mapkapk
Kras
Myc
Tgfa
Cd44
Hif1a
Galnt2
Gcnt3
Galnt3
Mapk6
Mapk9
Map2k1
Map2k2
Map2k4
Map3k2
Both
Kras
Nras
Areg
Aqp5
Gkn1
Prom1
Cftr
Galnt4
Klf5
Gata5
Il33
F
Stomach
Pancreatitis
GNASsiRNA-off signature
Kras-siRNA-off signature
UMAP1
UMAP2
Foveolar Pit
Chief
Parietal
SPEM
Chief
Delta
ECL
Foveolar Pit
Imm. Foveolar Pit
Isthmus
Neck
Neck/Chief Int.
Parietal
Prolif. Isthmus
Antral Foveloar Pit
Antral Gland
Chief-like
Delta
ECL
EEC
Forestomach
Foveolar Pit
G cell
Imm. surface pit
Parietal
Prolif. Antral Isthmus
Prolif. Chief Meta
Prolif. Isthmus
Smooth Muscle
Tuft
Fibroblast
Immune
Macrophage
unknown
SM Pericytes
Mesothelial
Others
Pericytes


A
ADM Cell Types
Acinar
Acinar/ductal
EEC
Tuft
Tuft/EEC progenitor
SPEM
Krt19-hi
UMAP_1
UMAP_2
B
Spdef RNA Expression
Spdef regulon (binarized)
C
GNASmut
****
SPDEF mRNA (microarray)
Altered
Unaltered
D
Spdef
Creb3l1
Creb3l4
Gkn3
Gale
Tspan1
Foxa3
C2cd4b
Muc6
Mgst2
Gcnt3
Pglyrp1
Bace2
Fxyd3
Car12
Gfpt1
Etv5
Ern2
Trim24
A4gnt
Cpn1
Agr2
Habp2
Liph
E
Spdef
Creb3l1
Creb3l4
Log2 (normalized counts)
CTRL
DOX
4838
C241
**
***
F
4838 GNAS-on signature
G
C241 GNAS-on signature
H
Spdef target genes
z-score
CONTROL
DOX
Mgst2
Pglyrp1
Gkn3
Fxyd3
Creb3l1
C2cd4b
Creb3l4
Foxa3
Car12
Tspan1
Liph
Gfpt1
Etv5
Ern2
Gale
Agr2
Bace2
Cpn1
Habp2
Trim24
Gcnt3
I
Spdef
Relative expression
CTRL siRNA
Spdef siRNA
CTRL
Spdef
J
GSII
CTRL siRNA
Spdef siRNA
K
GSII Area
GSII Area/nuclei
GSII Intensity
Intensity Score
No DOX
DOX


A
SPDEF
CREB3L1
CREB3L4
IPMN
PanIN
PDAC
Log2TPM
Epithelium
Stroma
B
SPDEF Nucleus Intensity Score
SPDEF Nucleus + Cytoplasm Intensity Score
CREB3L1 Nucleus Intensity Score
CREB3L1 Nucleus + Cytoplasm Intensity Score
CREB3L4 Nucleus Intensity Score
CREB3L4 Nucleus + Cytoplasm Intensity Score
Normal Duct (n=45)
Acinar to Ductal Metaplasia (n=41)
Low Grade (n=54)
High Grade (n=33)
Invasive (n=10)
Normal Duct (n=44)
Low Grade (n=51)
High Grade (n=31)
Normal Duct (n=43)
Low Grade (n=53)
High Grade (n=27)
Invasive (n=9)
C
Normal duct
ADM
LG IPMN
HG IPMN
Invasive