A.

Up-front resected non-neoadjuvant treated pancreatic cancers n = 305

→ Excluding 6 cases with metastatic disease at resection

Non-metastatic disease n= 299

→ Excluding 10 cases with 30-day or in-hospital mortality

Study cohort and overall survival (OS) analysis n=289

→ 25 matched WSS

- Cell subtype composition
- Tumor subtype
- Spatial analysis of co-expressor neighbors at gland and within gland level

Comparison with TMA:
- Cell subtype composition
- Tumor subtype

→ Excluding 4 cases with R2 resection

Study cohort for disease-free survival (DFS) analysis n=285

B.

Metastatic cohort

37 metastatic FFPE PDAC biopsies

24 fresh frozen metastatic biopsies

8-plex subtyping protein assay

14 bulk RNA sequencing, 10 single-cell RNA sequencing

Comparison with primary PDAC:
- Cell subtype composition
- Tumor subtype distribution
- Spatial analysis of co-expressor neighbors at gland and within gland level

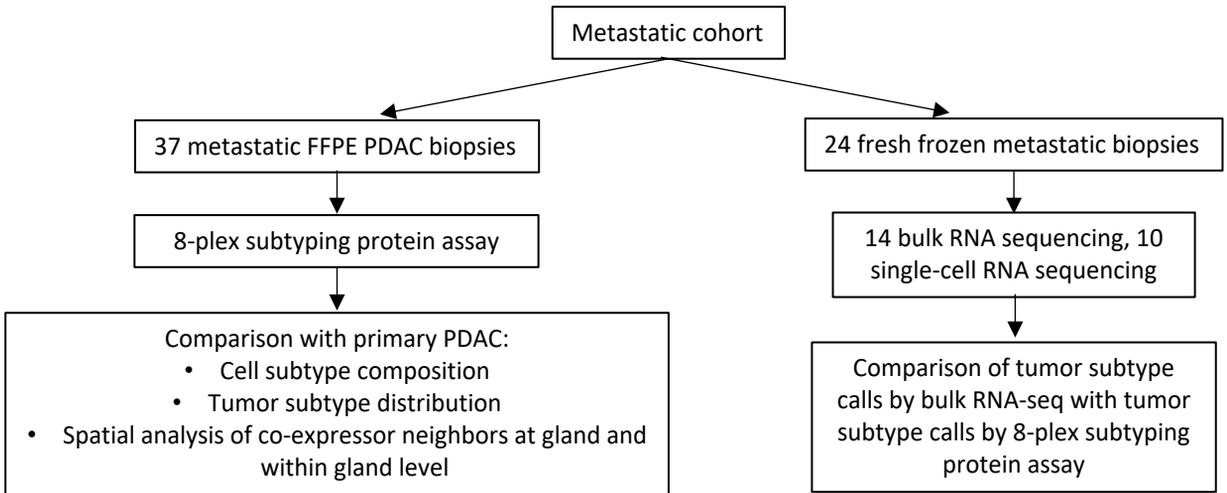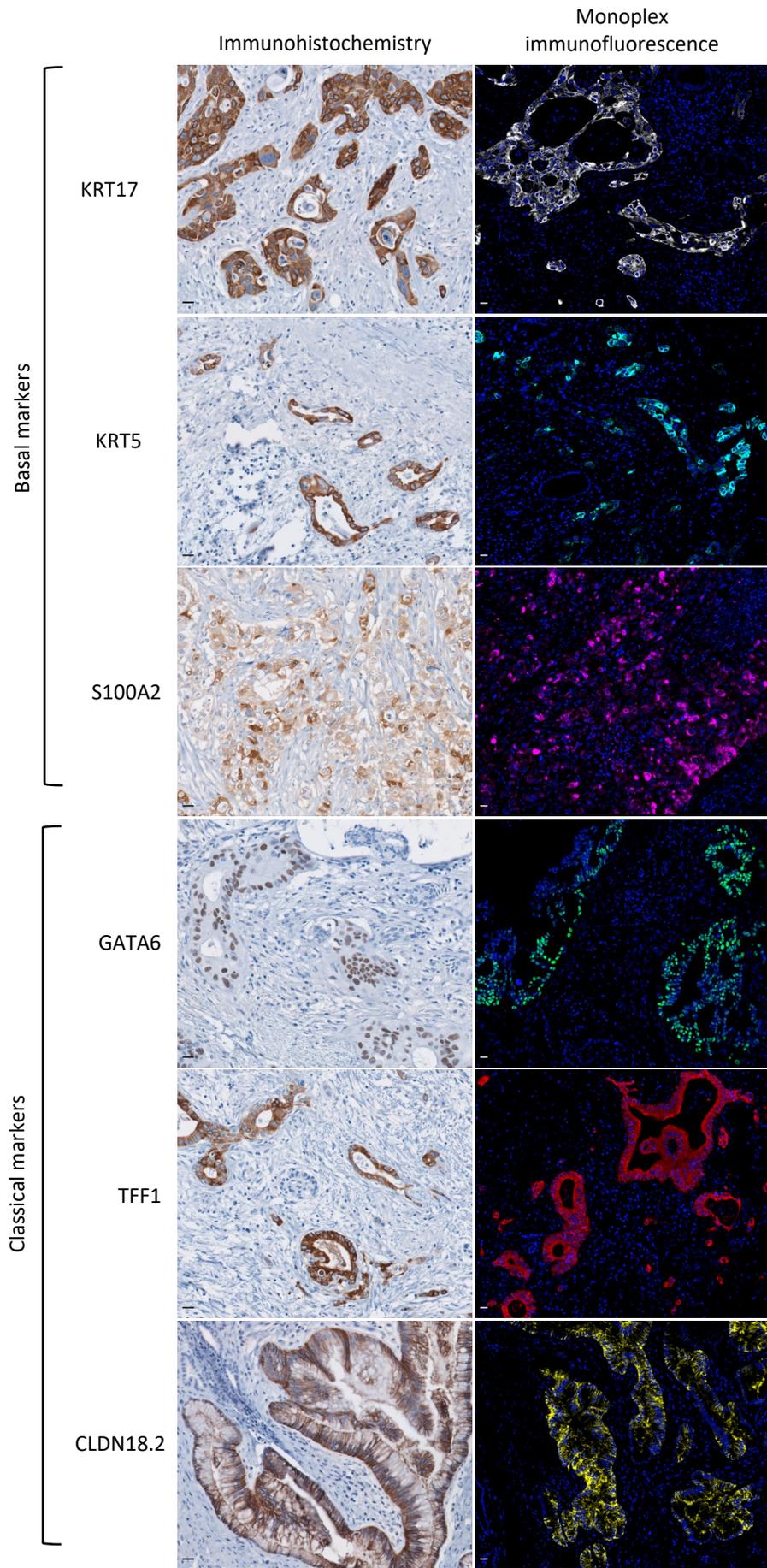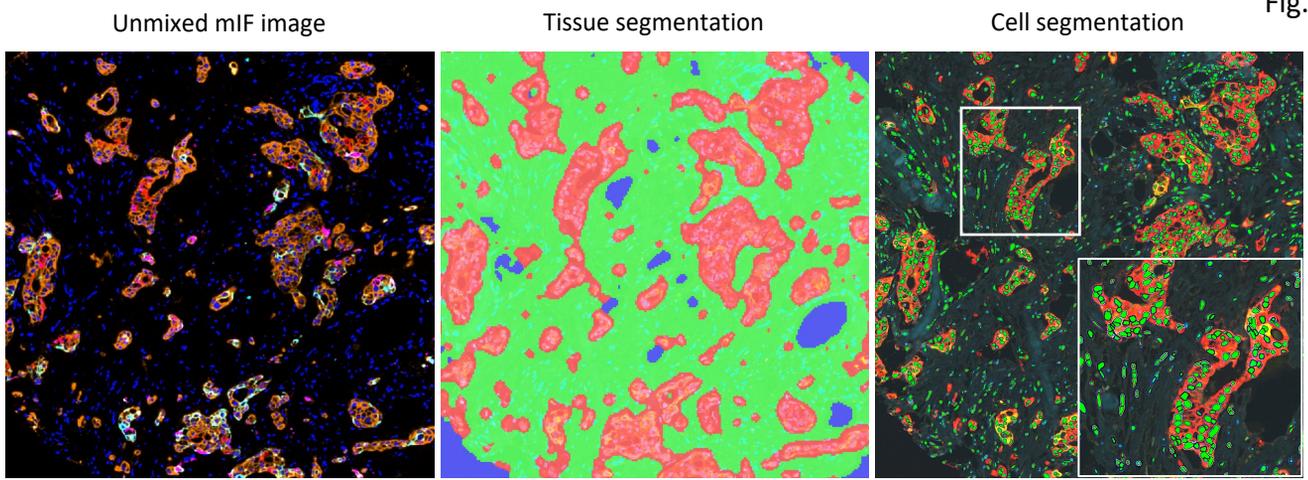Comparison of tumor subtype calls by bulk RNA-seq with tumor subtype calls by 8-plex subtyping protein assay
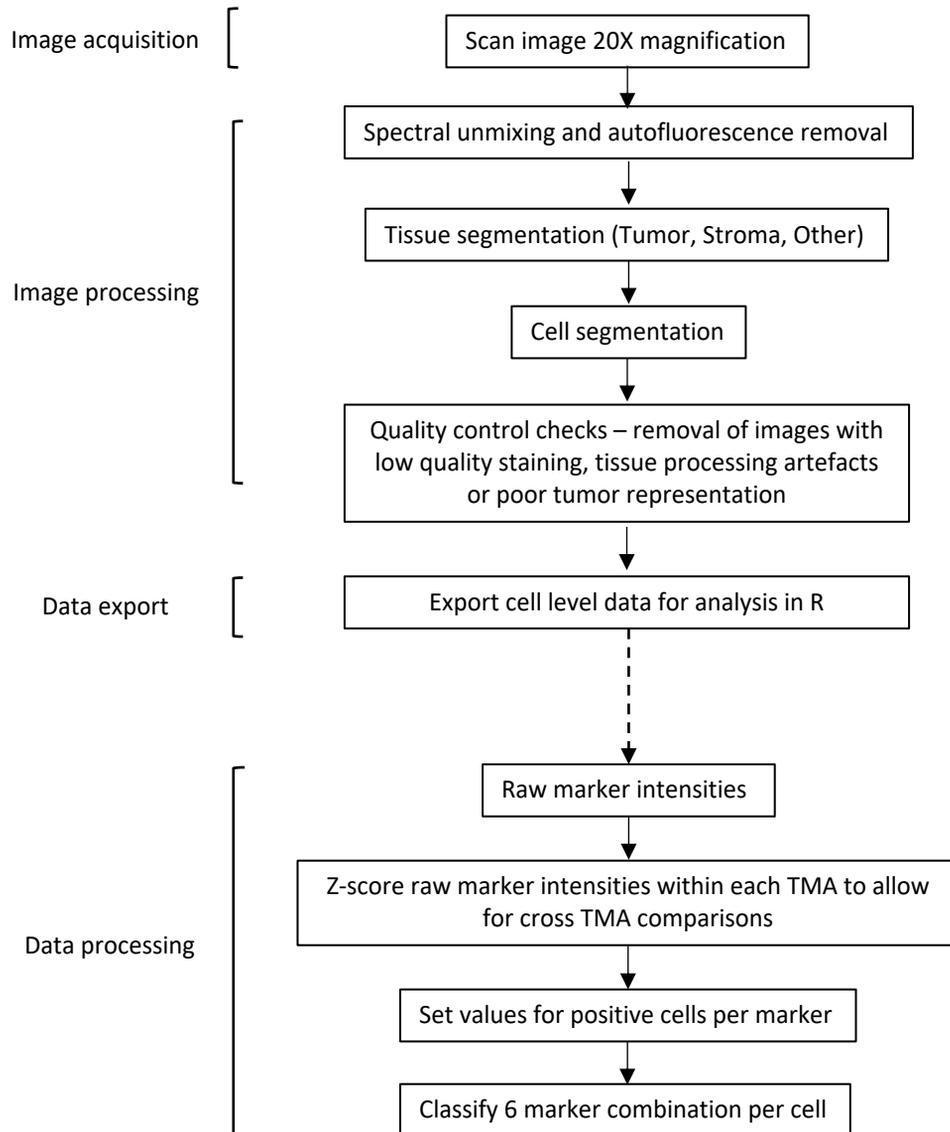
Supplementary figure S1. Study cohort summary. A: Primary resection cohort. 289 cases were included in the study cohort and overall survival analyses. DF/BWCC: Dana-Farber/Brigham and Women's Cancer Centre. URM: University of Rochester Medical Centre. SCI: Stanford Cancer Institute. B: Metastatic biopsy cohort. 37 cases were included in the study cohort and 24 matched fresh or frozen metastatic biopsies had RNA-sequencing data available (14 bulk-RNA sequencing, 10 single cell RNA-sequencing).

Immunohistochemistry

Monoplex immunofluorescence

Basal markers

KRT17

KRT5

S100A2

Classical markers

GATA6

TFF1

CLDN18.2

Supplementary figure S2. Single marker immunohistochemistry and monoplex immunofluorescence in primary PDAC specimens. Left panel, immunohistochemistry (top to bottom): KRT17, KRT5, S100A2, GATA6, TFF1, CLDN18.2. Right manel, monoplex immunofluorescence (top to bottom): KRT17 (Opal 520), KRT5 (Opal 480), S100A2 (Opal 650), GATA6 (Opal 540), TFF1 (Opal 690), CLDN18.2 (Opal 570).
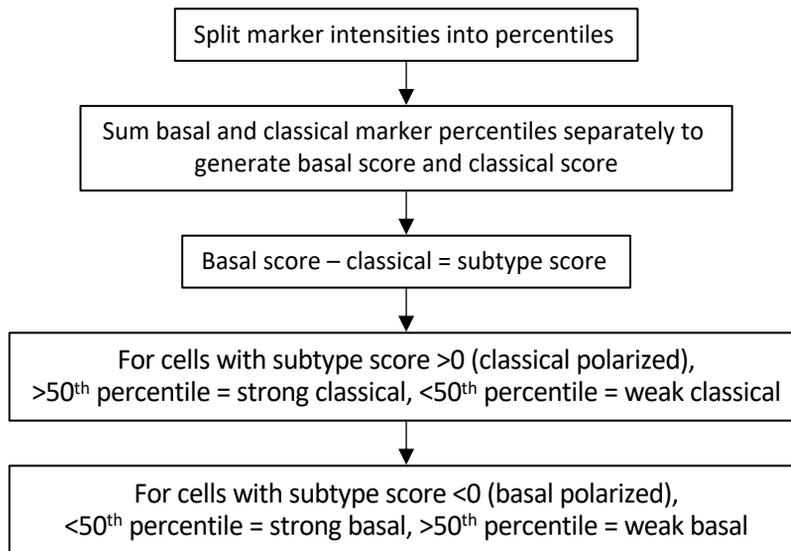
A.

| Unmixed mIF image | Tissue segmentation | Cell segmentation |

B.

Image acquisition
Scan image 20X magnification

Image processing
Spectral unmixing and autofluorescence removal

Tissue segmentation (Tumor, Stroma, Other)

Cell segmentation

Quality control checks – removal of images with low quality staining, tissue processing artefacts or poor tumor representation

Data export
Export cell level data for analysis in R

Data processing
Raw marker intensities

Z-score raw marker intensities within each TMA to allow for cross TMA comparisons

Set values for positive cells per marker

Classify 6 marker combination per cell

Supplementary figure S3. Image and data processing. A: (Left to right) Representative images of unmixed multiplex immunofluorescence image, tissue segmentation (red: tumor, green: stroma: blue: other/non-tissue) and cell segmentation (green: nuclei). B: Workflow for image processing and data analysis.
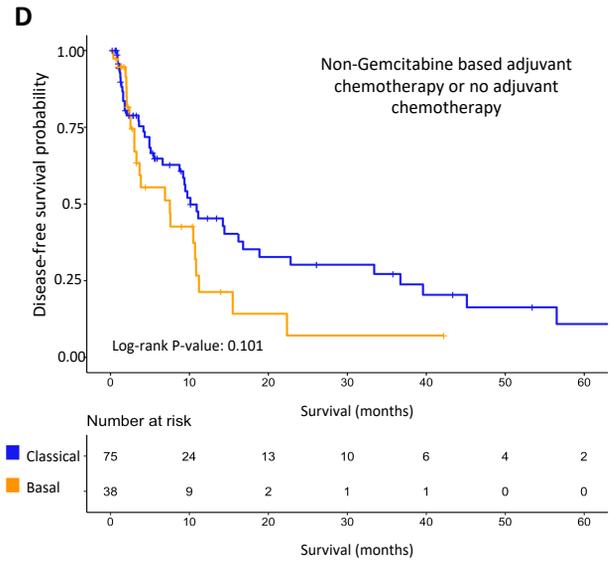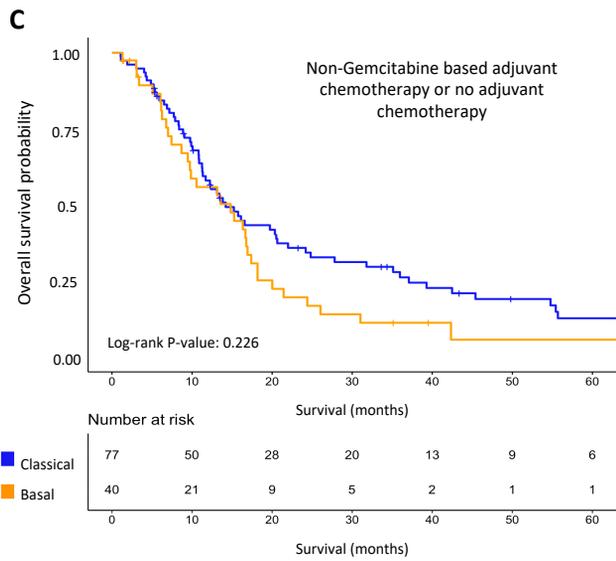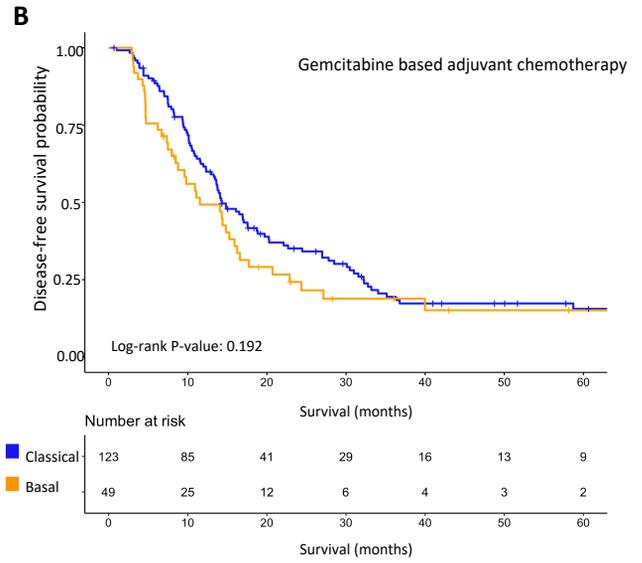
A.

Split marker intensities into percentiles

Sum basal and classical marker percentiles separately to generate basal score and classical score

Basal score – classical = subtype score

For cells with subtype score >0 (classical polarized),
>50th percentile = strong classical, <50th percentile = weak classical

For cells with subtype score <0 (basal polarized),
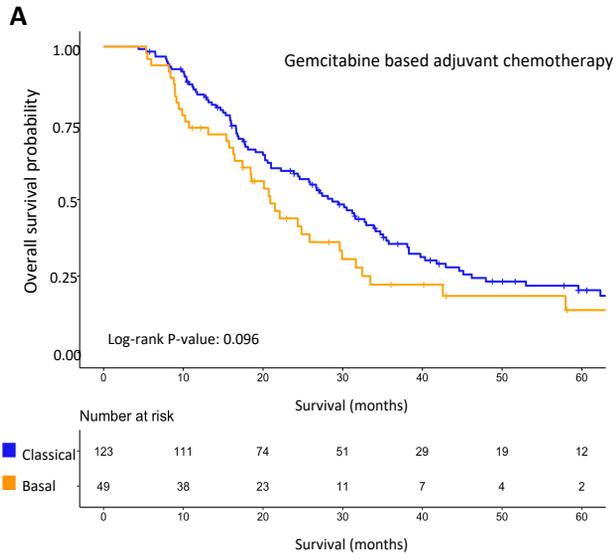<50th percentile = strong basal, >50th percentile = weak basal

Supplementary figure S4. Co-expressor cell polarization methodology. A: Data analysis workflow for co-expressor cell polarization.
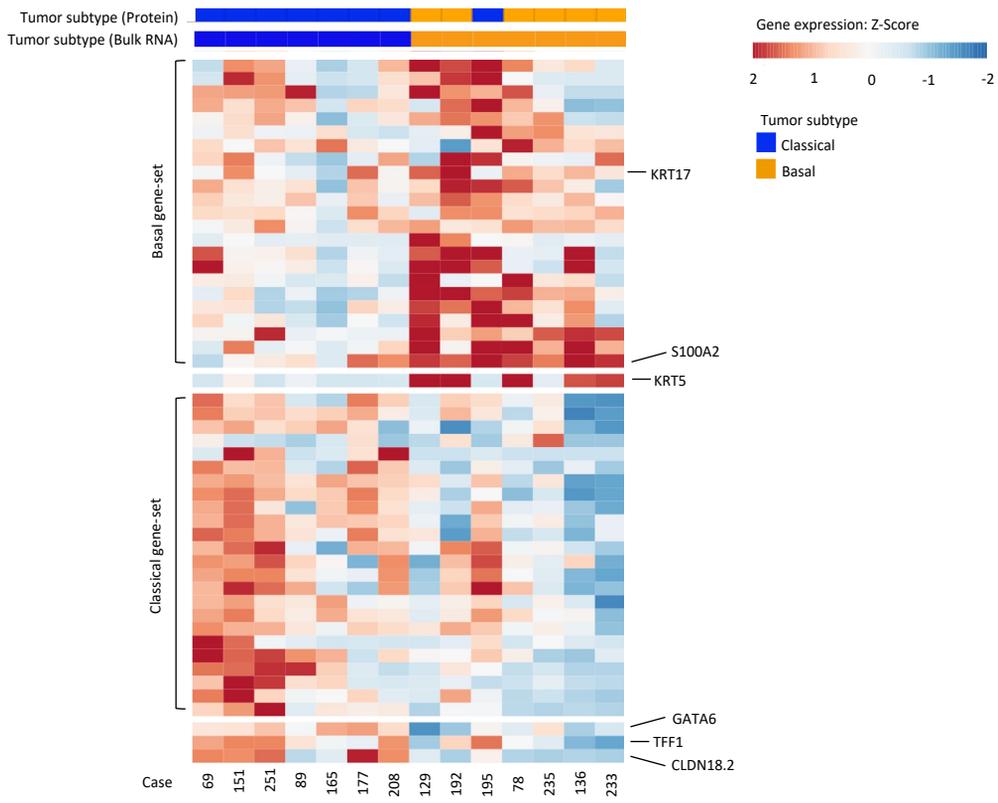
Supplementary figure S5. Orthogonal validation of multiplex immunofluorescence with RNA-sequencing. A: Comparison of multiplex immunofluorescence (mIF) and bulk RNA-sequencing of 14 metastatic biopsies. Top: Tumor subtype by multiplex immunofluorescence. Middle: Tumor subtype by bulk RNA-seq. Bottom: gene expression using bulk RNA-seq.

A

Overall survival probability

Gemcitabine based adjuvant chemotherapy

Log-rank P-value: 0.096

Survival (months)

Number at risk

	0	10	20	30	40	50	60
Classical	123	111	74	51	29	19	12
Basal	49	38	23	11	7	4	2

Survival (months)

B

Disease-free survival probability

Gemcitabine based adjuvant chemotherapy

Log-rank P-value: 0.192

Survival (months)

Number at risk

	0	10	20	30	40	50	60
Classical	123	85	41	29	16	13	9
Basal	49	25	12	6	4	3	2

Survival (months)

C

Overall survival probability

Non-Gemcitabine based adjuvant chemotherapy or no adjuvant chemotherapy

Log-rank P-value: 0.226

Survival (months)

Number at risk

	0	10	20	30	40	50	60
Classical	77	50	28	20	13	9	6
Basal	40	21	9	5	2	1	1

Survival (months)

D

Disease-free survival probability

Non-Gemcitabine based adjuvant chemotherapy or no adjuvant chemotherapy

Log-rank P-value: 0.101

Survival (months)

Number at risk

|	0	10	20	30	40	50	60
Classical	75	24	13	10	6	4	2
Basal	38	9	2	1	1	0	0

Survival (months)

Supplementary figure S6. Kaplan-Meier curves for overall survival and disease-free survival by basal-classical expression subtype according to adjuvant treatment received. A: Association of tumor subtype with overall survival (OS) in patients treated with adjvuant gemcitabine or gemcitabine combination). B: Association of tumor subtype with disease-free survival (DFS) in patients treated with adjuvant gemcitabine or gemcitabine combination). C: Association of tumor subtype with OS in patients treated with non-gemcitabine based adjuvant chemotherapy or no adjuvant chemotherapy. D: Association of tumor subtype with DFS in patients treated with non-gemcitabine adjuvant chemotherapy or no chemotherapy. Log-rank p-values shown for comparison of survival among patients with basal and classical tumors.
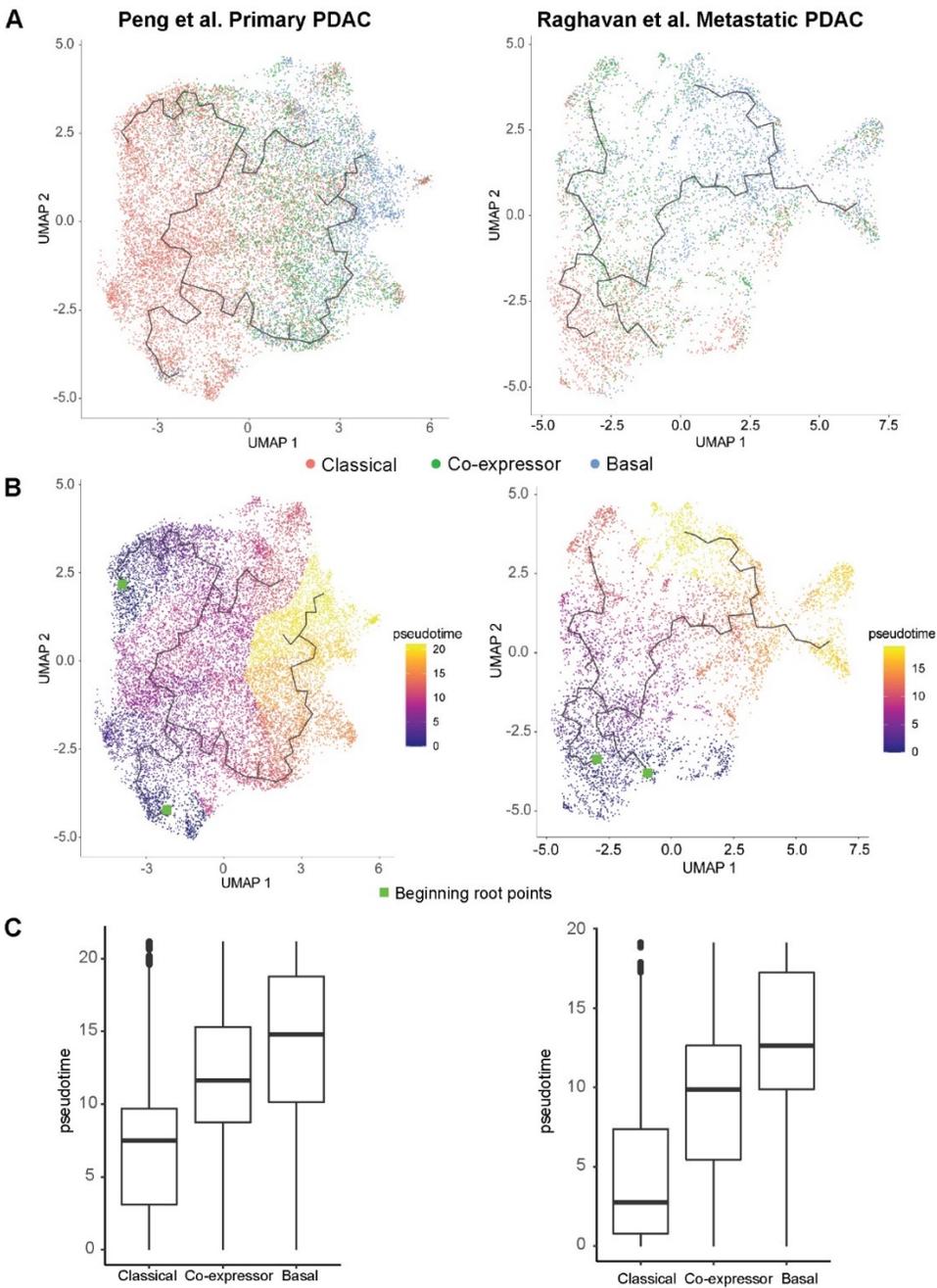
Fig.S7
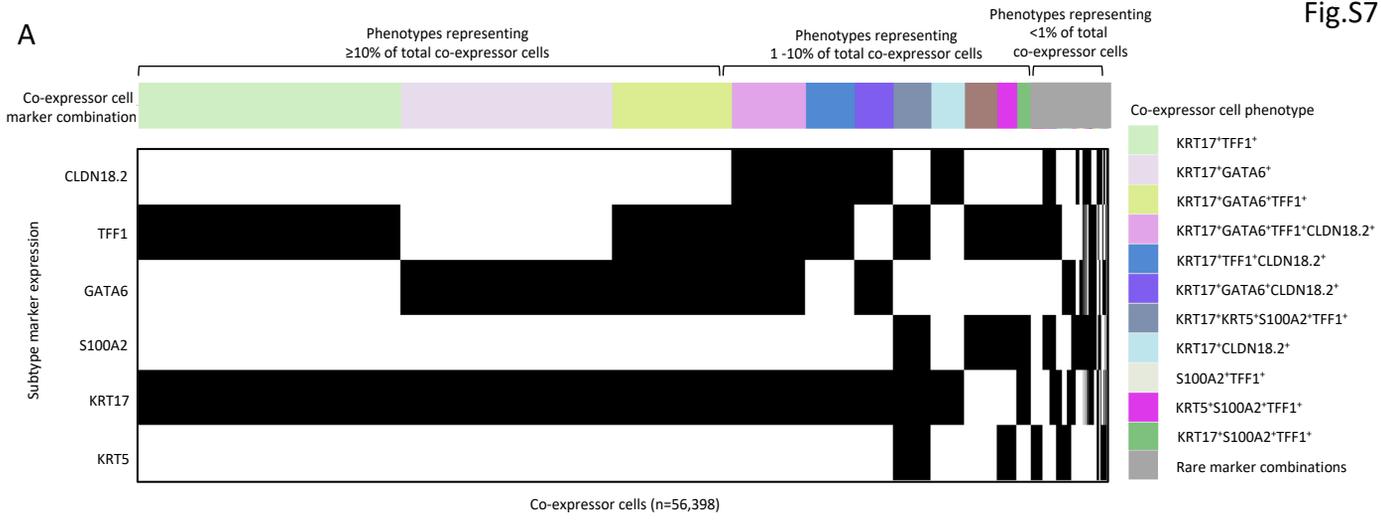
Supplementary figure S7. Co-expressor cell phenotype frequencies. The combinations KRT17$^+$TFF1$^+$, KRT17$^+$GATA6$^+$ and KRT17$^+$GATA6$^+$TFF1$^+$ represented the three most common phenotypes for co-expressor cells (n=56,398) in our primary resection cohort.

A.

Cell subtype ● Basal cell ● Co-expressor cell ● Classical cell
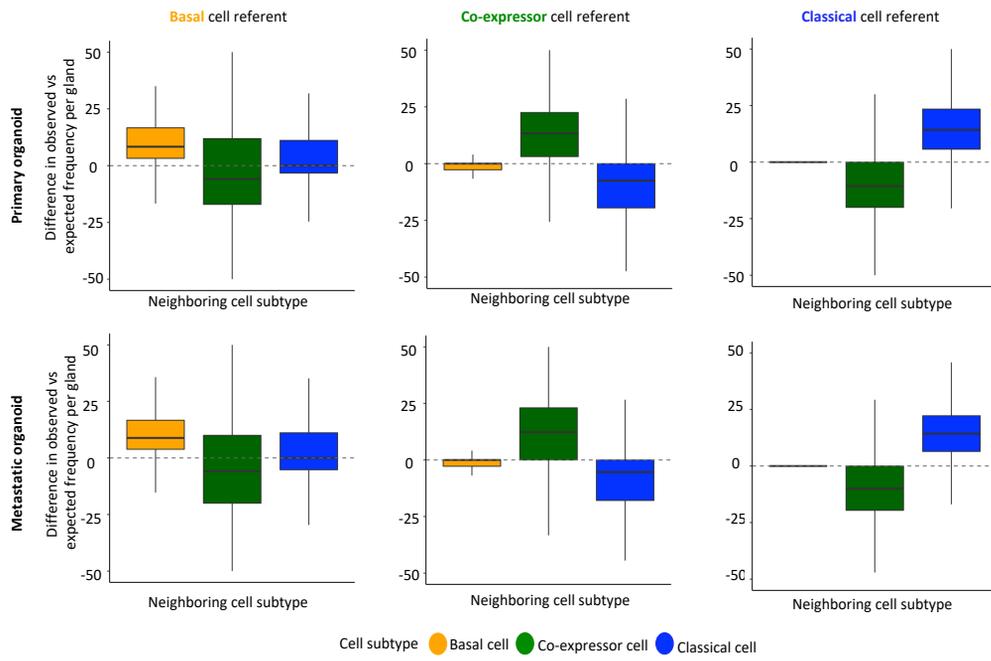
Supplementary figure S8. Direct neighbor analysis in patient-derived organoids. Direct neighbor analysis with comparison of expected and observed frequencies of subtypes by cell of reference and specimen type (primary and metastatic patient-derived organoids). Basal cells neighbor basal cells and classical cells neighbor classical cells more frequently than expected by chance. Difference between observed an expected frequencies per gland unit depicted (Mann-Whitney test between observed and expected frequencies).

A

Peng et al. Primary PDAC   Raghavan et al. Metastatic PDAC

● Classical   ● Co-expressor   ● Basal

B

■ Beginning root points

C

Supplementary figure S9. Trajectory and Pseudotime Analysis of the Peng et al. and Raghavan et al. PDAC datasets. A. UMAP embedding of tumor cells from Peng et al. (left) and Raghavan et al. (right), showing the location of the classical, basal, and co-expressor cell types, overlayed with the predicted cell trajectory from Monocle3, as indicated by the black line. B. The same UMAP embeddings as in A, now colored with the pseudotime value of each cell, showing an increase in pseudotime as the cell trajectory moves from classical to basal cell types through co-expressor cell intermediates. Starting root points for the pseudotime analysis were chosen in the classical cell populations, as indicated by green square. C. Quantification of pseduotime values for each cell type indicated in B, showing that the co-expressor cells contain intermediate pseudotime values between the classical and basal transcriptional cell types in this analysis (all pvalues <0.0001 for Kruskal-Wallis test and pairwise comparisons using Wilcoxon rank sum test).