

# Coordinated single-cell tumor micro-environment dynamics reinforce pancreatic cancer subtype

Received: 16 April 2022

Accepted: 14 August 2023

Published online: 26 August 2023

 Check for updates

Ki Oh<sup>1</sup>, Yun Jae Yoo<sup>1</sup>, Luke A. Torre-Healy<sup>1</sup>, Manisha Rao<sup>2,3</sup>, Danielle Fassler<sup>1</sup>, Pei Wang<sup>4</sup>, Michael Caponegro<sup>5</sup>, Mei Gao<sup>6</sup>, Joseph Kim<sup>6</sup>, Aaron Sasson<sup>7,8</sup>, Georgios Georgakis<sup>7,8</sup>, Scott Powers<sup>3,8</sup> & Richard A. Moffitt<sup>1,9,10</sup> ✉

Bulk analyses of pancreatic ductal adenocarcinoma (PDAC) samples are complicated by the tumor microenvironment (TME), i.e. signals from fibroblasts, endocrine, exocrine, and immune cells. Despite this, we and others have established tumor and stroma subtypes with prognostic significance. However, understanding of underlying signals driving distinct immune and stromal landscapes is still incomplete. Here we integrate 92 single cell RNA-seq samples from seven independent studies to build a reproducible PDAC atlas with a focus on tumor-TME interdependence. Patients with activated stroma are synonymous with higher myofibroblastic and immunogenic fibroblasts, and furthermore show increased M2-like macrophages and regulatory T-cells. Contrastingly, patients with ‘normal’ stroma show M1-like recruitment, elevated effector and exhausted T-cells. To aid interoperability of future studies, we provide a pretrained cell type classifier and an atlas of subtype-based signaling factors that we also validate in mouse data. Ultimately, this work leverages the heterogeneity among single-cell studies to create a comprehensive view of the orchestra of signaling interactions governing PDAC.

Pancreatic ductal adenocarcinoma (PDAC) is one of the leading causes of cancer death with a 5-year survival rate of only 10%<sup>1</sup>. Only 1 in 4 PDAC patients are diagnosed early enough for complete surgical resection and for most patients, chemotherapy eventually becomes the only option. Molecular analysis of PDAC tumors is often hampered by limited tumor cellularity and the presence of abundant stroma intermixed with endocrine, exocrine, and immune cells<sup>2</sup>. To better study PDAC neoplastic cells, alternative methods such as laser capture microdissection, organoids, and xenografts have been used<sup>3–6</sup>. We previously performed virtual microdissection on bulk RNA-seq samples establishing prognostic gene signatures for “basal-like” and

“classical” tumor subtypes, highlighting the importance of cancer-intrinsic heterogeneity across PDAC patients<sup>7</sup>. Additionally, we described “normal” and “activated” stromal subtypes, the latter having worse outcomes. Gene signatures such as these have been important in preliminary trials aimed at therapeutic decision support<sup>8–10</sup>. Exploration of the interplay of these signatures was recently explored and made accessible online<sup>11</sup>.

In recent years, the tumor microenvironment (TME) has been of great interest in the search for alternative therapeutic modalities<sup>12</sup>. Advances in microfluidics and single-cell RNA sequencing (scRNA-seq) have enabled high-throughput and high-resolution analysis of the

<sup>1</sup>Department of Biomedical Informatics, Stony Brook University, Stony Brook, NY, USA. <sup>2</sup>Department of Biomedical Engineering, Stony Brook University, Stony Brook, NY, USA. <sup>3</sup>Department of Pathology, Stony Brook University, Stony Brook, NY, USA. <sup>4</sup>Department of Cell Systems & Anatomy, University of Texas Health Science Center, San Antonio, TX, USA. <sup>5</sup>Department of Pharmacology, Stony Brook University, Stony Brook, NY, USA. <sup>6</sup>Department of Surgery, University of Kentucky and Markey Cancer Center, Lexington, KY, USA. <sup>7</sup>Department of Surgery, Stony Brook University, Stony Brook, NY, USA. <sup>8</sup>Stony Brook Cancer Center, Stony Brook University, Stony Brook, NY, USA. <sup>9</sup>Department of Hematology and Medical Oncology, Emory University, Atlanta, GA, USA. <sup>10</sup>Department of Biomedical Informatics, Emory University, Atlanta, GA, USA. ✉e-mail: [richard.austin.moffitt@emory.edu](mailto:richard.austin.moffitt@emory.edu)

cancer TME at an individual cell level<sup>13</sup>. This allows us to study tumor heterogeneity, while overcoming the challenges of blind-source separation<sup>14–17</sup>. Several single-cell studies of pancreatic tissues have helped to reveal previously unresolved biology and disease processes<sup>18–20</sup>. For example, novel progenitor-like duct cells were discovered that differentiate into mature ductal, acinar, or islet cells<sup>21</sup>. Cancer-associated fibroblasts (CAFs), including myofibroblastic CAFs (myCAFs), immunogenic CAFs (iCAFs), and antigen-presenting CAFs (apCAFs) were shown to play a role in extracellular matrix production, immunosuppression, vasculature remodeling, tumor proliferation and metastasis<sup>22,23</sup>. Specifically, these TME phenotypes are facilitated in part by signaling ligands that activate or suppress transcriptional programs. Tumor-derived Interleukin-1 (IL-1), for instance, induces a cytokine cascade consisting of LIF, IL6, and G-CSF in pancreatic stellate cells (PSCs) to form iCAF populations through JAK/STAT signaling<sup>24</sup>. Furthermore, subgroups of macrophages and lymphocytes have been described across multiple tumor types, adding to the phenotypic diversity and transcriptional heterogeneity that ultimately allows for TME plasticity.

With many new annotated cell types, the focus has shifted towards unraveling how cell-signaling mechanisms and networks of features collectively form heterogeneous tumors across PDAC patients. However, stark differences exist between cell type proportions obtained from different labs. Thus, the reproducibility of these novel cell types and unique signaling mechanisms that may exist among the tumor ecosystem and overall patient survival has not been sufficiently explored.

In this work, we leverage publicly available data sets to overcome collection and technical biases of individual analyses and create a single-cell atlas of the PDAC TME to provide a cohesive perspective of recent high-resolution findings. We demonstrate a close relationship between stromal and tumor subtypes which correlated with distinct signaling patterns across multiple cell compartments. We further highlight subtype-dependent cell-signaling interactions that may drive these distinct tissue phenotypes in hopes that targeting these interactions may lead to new therapeutic approaches.

## Results

### Integrated meta-analysis of single-cell PDAC provides a rich and detailed TME atlas

An integrated pancreas scRNA-seq dataset of normal and PDAC-derived patients was curated using local and public data sets processed through a Seurat pipeline for both discovery and validation sets (Fig. 1a, S1a, b). Patients across data sets contributed a unique composition of cell type subpopulations reflected by the initial study objectives, patient clinical conditions, and variable tissue processing, allowing us to observe a more comprehensive set of cell types not necessarily captured by any one study (Supplementary Data 1). UMAP representation of the atlas shows the successful integration of cells across patients and independent data sets (Fig. 1b, S1c) and provides rich compositional and phenotypic information from the PDAC tissue biopsies. Visualizing canonical markers identifies the major cell types (Fig. 1c) which were further refined into distinct subpopulations for closer analysis (Fig. 1d, S1b). These cell types were additionally verified to be consistent with previously established labels from the original studies.

**Fibroblastic Stroma:** The desmoplastic PDAC TME is commonly represented by an increase in both activated CAFs and stellate cell populations. Analysis of differential expression (Figure S1d) between stromal subclusters identified previously described myofibroblast CAFs (myCAFs), complement-secreting CAFs (csCAFs), and immunogenic CAFs (iCAFs)<sup>22,25</sup>. Two other previously described pancreatic stellate cell (PSC) related subpopulations<sup>26</sup> named here by smooth muscle (smPSC) and quiescent (qPSCs) were identified. Small populations of IL11+ CAFs, Schwann cells, and myocytes were observed

though not present in all patients and excluded from differential gene expression analysis. **Myeloid / Macrophage:** We detect 2 major dendritic cell (DCs) populations and several myeloid activation states. (Figure S1e). Two dendritic populations included a Regulatory T-Cell (T-Reg) recruiting LAMP3 + DC<sup>27</sup> and a monocyte-derived conventional CD1c+ DCs (mo-DCs). Resident macrophages showed complement-related expression, consistent with previous findings. Classically activated monocytes were identified by S100A8 expression<sup>28</sup>. We observed 3 tumor-associated macrophage (TAM) populations including a previously identified SPPI + TAM that expressed MIF, which has been associated with alternative activation. SPPI+ TAMs expressed the highest level of CXCL8, which has been reported at the invasive tumor front<sup>29</sup>. An M2-like TAM subpopulation appeared to be monocyte-derived (APOC1+)<sup>30</sup> and expressed granulin (GRN), which contributes to CD8+ T-cell exclusion in PDAC<sup>31</sup>. **Endothelium:** Endothelial cells were abundant within both PDAC and normal-derived pancreatic tissue. We identified 3 major cell subgroups based on previous single-cell studies, e.g., capillary 'Tip-like' EC<sup>32</sup>, venous EC<sup>33</sup>, and arterial EC<sup>34</sup> (Figure S1f). We identified a 'regulatory' endothelial subpopulation expressing anti-angiogenic markers (JAK/STAT inhibiting - SOCS3<sup>35</sup>, and SPRY1)<sup>36</sup>. **Lymphocytes:** We observe a dynamic landscape of both tumor-killing and tumor-permissive lymphocyte similarly observed in other solid malignancies. Collectively we identified B-lymphocytes, CD4+/CD8+ T-lymphocytes, and transitory populations such as progenitor lymphocytes and proliferating lymphocytes (Figure S1g). CD4+ T-cells included the LEF1+ Naive-CD4 and CTLA4 + T-Regs that expressed TIGIT. CD8 positivity was shared by early pre-dysfunctional TSC22D3+<sup>37</sup>, dysfunctional LAG3+ tumor-infiltrating lymphocytes (TILs)<sup>38</sup> but decreased in the NKG7+ cytotoxic effector CD8<sup>39</sup> population<sup>40</sup>. B-lymphocyte populations were comprised of immature (TCL1A+)<sup>41</sup>, mature memory B-cells, and plasma cells (MZB1+)<sup>42</sup>. A full list of compartment-specific gene expressions is provided in Supplementary Data 1.

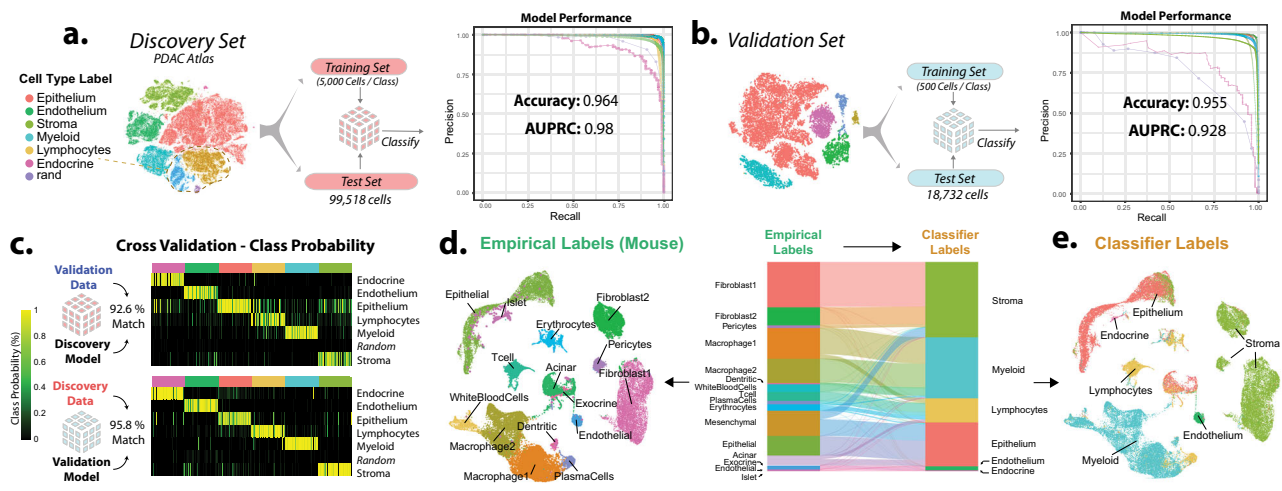
### Cross-TME analysis of secretome highlights key signaling axes and surface markers

We then surveyed all differentially expressed transcripts known to be secreted ligands and receptors across the cell types described above (Fig. 1e) using statistical criteria of (Minimum cell expression > 20% of cells, expression difference between groups >10%). Notable findings include T-cell recruiting CCL21<sup>43</sup> expression by qPSCs, IL6 expression by iCAFs, and INHBA by myCAFs. CXCL12, another chemotactic factor, was expressed by immunogenic CAFs and venous-ECs. The corresponding receptor, CXCR4, was highest in early (pre-activation) CD8+ T-cells and memory B-cells. The arterial, capillary, and venous endothelial subpopulations showed upregulation of angiogenic and mitogenic factors (i.e., VEGF, INSR), and immunomodulatory factors such as LIFR, CX3CL1, and CCL23. Monocyte-derived DCs expressed CSF2R, a receptor to CSF2 (GM-CSF) secreted by csCAFs and myCAFs. SPPI+ macrophages expressed MRC1 (M2 marker), and TREM2 which correlates with tumor infiltrating CD8 exhaustion<sup>44</sup>. LAMP3+ DCs had the highest expression of LGALS9 (Galectin-9), which promotes M2 polarization in macrophages expressing TIM-3 (HAVCR2)<sup>45</sup>. Lastly, we established a pan-TME panel of surface contact markers as a resource for stably expressed biomarkers that can identify cell types using orthogonal techniques (Supplementary Data 2, Figure S1i).

### Automated single-cell classifier for human and mouse data

To automate basic cell type annotations for pancreatic cancer single-cell data, we trained a multi-class random forest classifier using *singlecellnet*<sup>46</sup> with the discovery atlas (Fig. 2a). Testing against our internal held-out validation set of 99,518 cells, we achieved model accuracy of 96.4% and AUPRC of 98%. Conversely, we trained a second random forest model using a validation dataset from three independent studies (Fig. 2b). Finally, cross-validation of both models across





**Fig. 2 | Automated single-cell classifier produces stable results in human and mouse scRNA data.** **a** Training schema for the primary classifier trained on 5000 cells per cell type in the discovery set. Model performance for each cell type is shown on PR curve. **b** Secondary classifier built on an independent validation set accurately predicts the same cell type annotations. **c** Robust cross-validation of the two classifiers with unseen data are shown in the probability heatmap. **d** Sankey

diagram of classifier results with mouse pancreatic cancer samples shows concordance between empirical labels and classifier labels in both human and mouse single-cell data. Width of flow is proportional to total cells in each cell category. **e** UMAP plot of mouse scRNA-seq data colored by the original independent annotations and the automated classifier results exhibit robust results. Source data are provided as a Source Data file.

studies demonstrated high label correspondence (Validation: 92.6% Training: 95.8%) and probability of annotation calls within both models (Fig. 2c). To allow for compatibility with both human and mouse PDAC experimental models, we used only common features (homologous transcripts) for classification, allowing for recognition of broad cell types in both species (Fig. 2d, e).

### Normal and activated stromal signatures reflect a patient-level gradient between two distinct CAF populations

We sought to better align previously established subtypes with the more recent single-cell annotations of PDAC. We noticed a broad overlap of ‘Activated’ fibroblast signature expression involving myCAF, ‘reactive’ stroma, and classical CAFs (cCAFs)<sup>9,22,25</sup> (Fig. 3a, S2a). Close examination of immunogenic subpopulations showed iCAFs and the recently described csCAFs are indeed distinct groups with overlapping signatures (Fig. 3b). Localization of the apCAF-associated CD74 expression appeared to be highest in the Schwann cells (Fig. S2b) but an isolated group of cells representing the apCAFs was not found. The remaining cells of the stroma represent 2 PSC groups with CAF1/CAF2 signatures<sup>26</sup> (Fig. 3c). Furthermore, pseudobulk (aggregated) expression heatmaps of both stroma subtype signatures showed a gradient from normal stroma to gradually higher activated states. Consensus clustering divided patients into normal or activated stroma, and a mixed intermediate group (Fig. 3d, S2c, d). Stromal subtype signature scores highlighted the enrichment of the activated phenotype within myCAFs whereas the normal signature was found in PSCs (Fig. 3e).

### Cellular composition of PDAC TME is correlated with stromal subtype

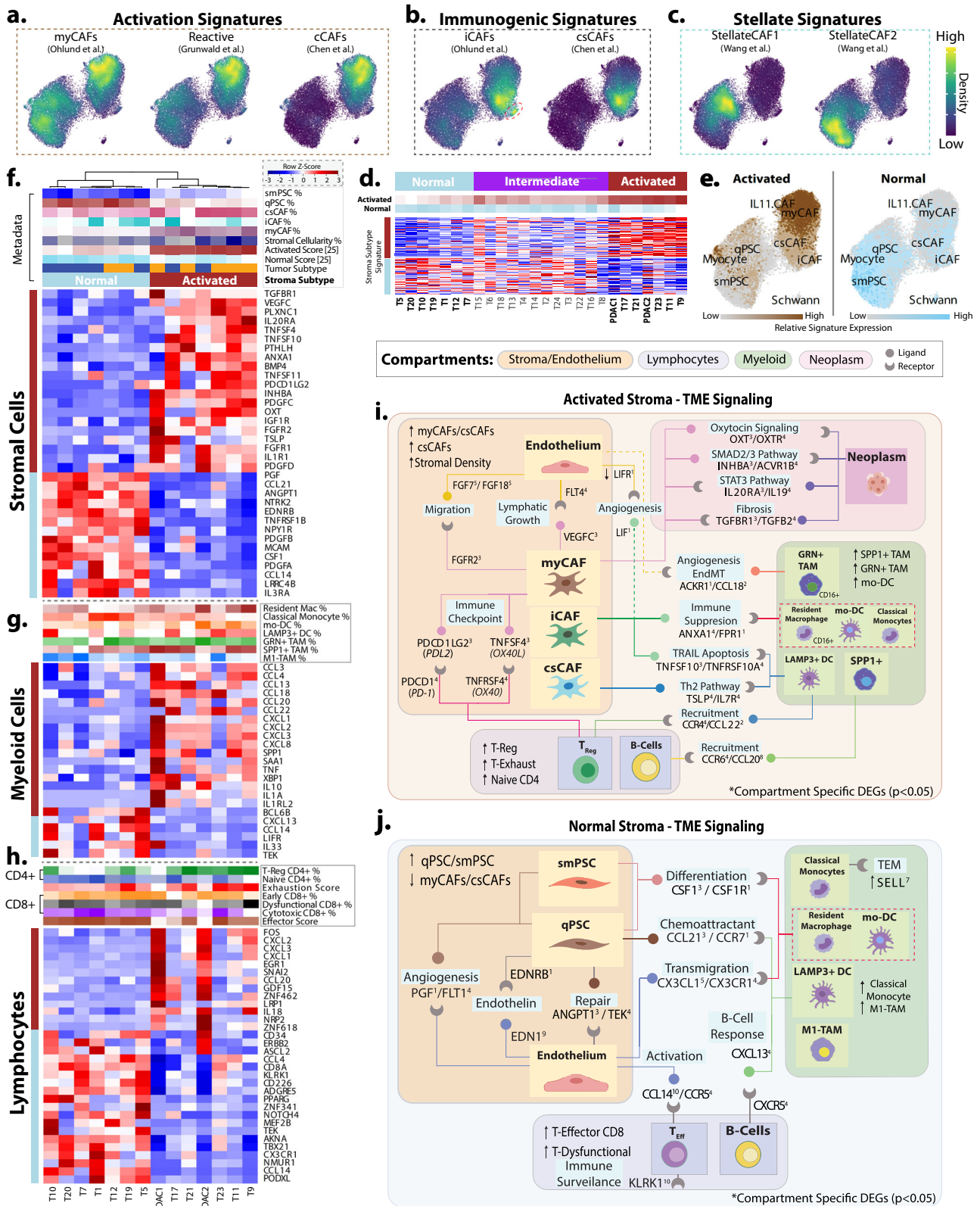
Based on the distinct enrichment of subtype signatures across the CAF and PSC subpopulations, we hypothesized that differential signaling based on stromal subtypes eventually leads to distinct TME compositions and tumorigenic phenotypes (Fig. 3f–h). We observed patients of the predominantly activated stroma phenotype had with a higher percentage of myCAFs and csCAFs within the stroma while normal subtype patients had higher PSCs. Similar comparisons in the myeloid compartment showed higher SPP1+ and GRN+ TAMs, resident macrophages, and mo-DCs with activated stroma while classically activated monocytes and tumoricidal macrophages (M1-like, CIBERSORT)

were higher in normal subtypes. Lymphocytes of the activated cohort showed higher proportions of CD4+ T-Regs, and higher exhausted T-cell signature scores whereas normal patients had higher overall T-cells, and an increase in CD8+ T (cytotoxic and dysfunctional), naive-CD4+ T-cells % and effector-related signature scores<sup>47</sup>.

### Stromal subtypes distinguish anti-tumor and pro-tumor phenotypes governed by TME signaling

In the activated subtype, myCAFs show expression consistent with secreted INHBA which is involved in the SMAD2/SMAD3 pathway in the epithelium. Additionally, the cancer progression and chemoresistance-related OXT/OXTR (oxytocin) genes were highly expressed between CAFs and the neoplastic cells<sup>48</sup>. Myofibroblast CAFs (myCAFs) and endothelial TGFBR1 were identified as a potential targets for TGFB2 from basal-like tumor cells<sup>49</sup>. IL19-induced profibrotic STAT3 pathway is shown by IL20RA in activated myCAFs<sup>50</sup>. Immune checkpoint molecules PDCD1LG2 (PDL2)<sup>51</sup> and TNFSF4 (OX40L)<sup>52</sup> were expressed by myCAFs while CD8+ T-Cells and T-Regs expressed the matching receptor pair (PD1 and OX40) respectively. Myeloid cells of the activated TME portrayed M2 polarization marked by the expression of CD209 in GRN+ TAMs (Fig. 3g). The myCAFs and endothelium, exhibited upregulation of FGF7/FGFR2 and VEGFC/FLT4 signaling that facilitates cell migration and lymphatic growth respectively<sup>53</sup>(Fig. 3f). The endothelium appeared to facilitate chemotaxis via IL1B signaling in the activated TME (Figure S2c) and additionally expressed macrophage regulating genes including MIF and MARCO.

Overall signaling across cells of the normal stroma cohort showed a contrasting anti-tumorigenic immune phenotype. In particular, the recruitment of cytotoxic T-Cells by the endothelium can be inferred by the upregulated CXCL1/CXCLR1 axis<sup>54</sup>. PSCs exhibited recruitment and differentiation through CCL21/CCR7 with LAMP3+DC, and CSF1/CSF1R with both resident macrophages and mo-DCs. Further, we observed higher expression of NOTCH signaling and adhesion genes which may regulate TME density thereby modulating macrophage phenotypes<sup>55</sup>. Notably, PSCs expressed angiotensin (ANGPT1) that targets endothelial TIE2 (TEK) and facilitates vessel repair which favors the TME for drug delivery and immune infiltration<sup>56</sup>. The normal endothelium also showed upregulation of myeloid recruiting Fractalkine (CX3CL1)<sup>54</sup>. Endothelial LIFR expression was high in this cohort but down-regulated by LIF by iCAFs in the activated TME<sup>57</sup> (Figure S2g).



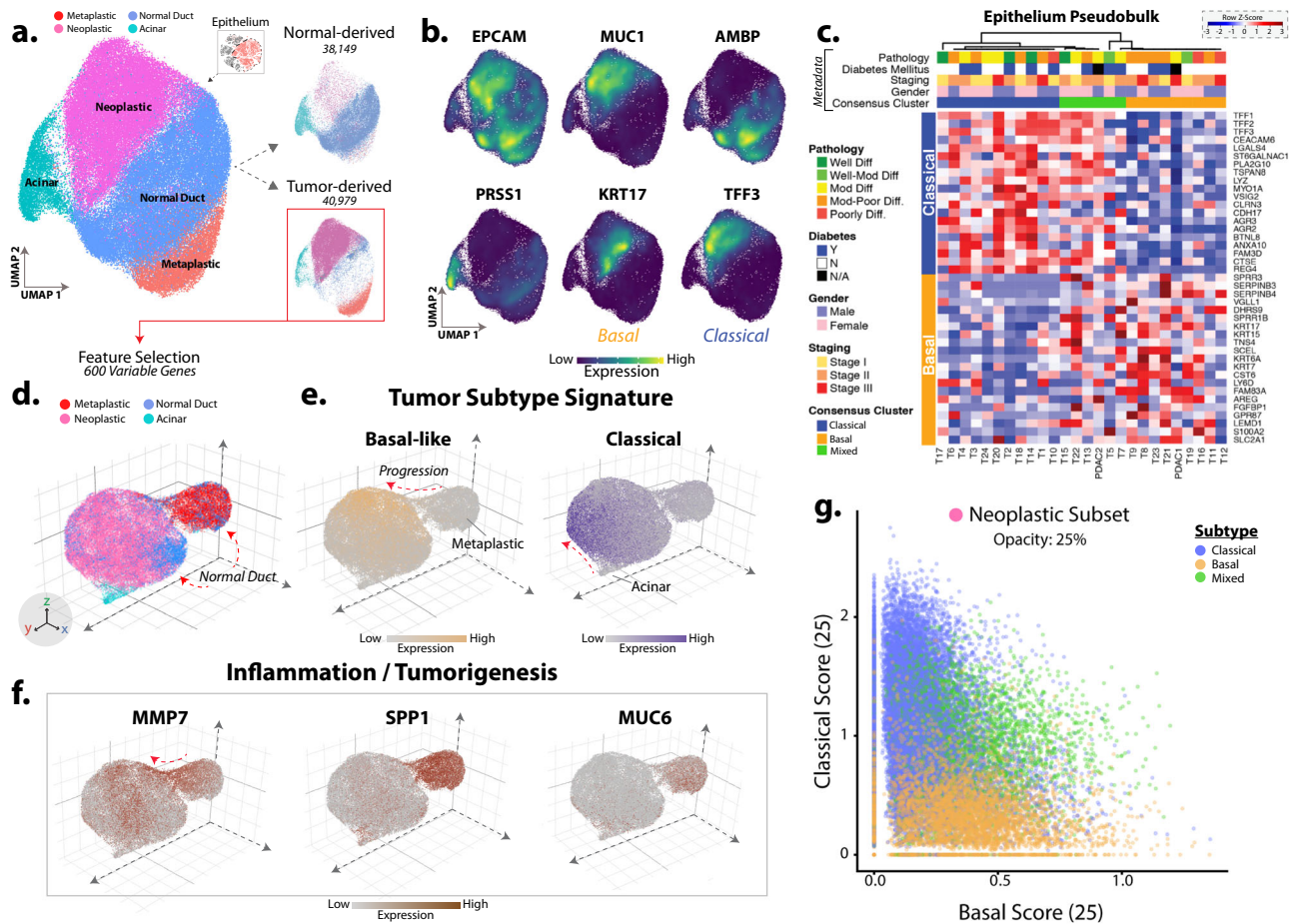
We compared the average expression of our subtype-associated panel of signaling genes across all patients (including intermediate), showing a coarse gradient of expression across subtypes (Figure S2h). Furthermore, the expression of this gene set was concordantly expressed by the stroma and myeloid compartments of the validation cohort (Figure S2i). An overall model for these interactions, based on empirical expression data, is shown in Fig. 3i, j.

**Tumor subtype signatures segregate PDAC single cells and correlate with distinct TME activity**

Epithelial cells aggregated from both normal-derived and PDAC-derived tissue were subclustered for finer cell type annotation, revealing acinar, normal duct, inflammatory metaplastic, and neoplastic groups. Splitting the data by sample origin (normal or tumor) showed enrichment of neoplasia and metaplasia in tumor samples (Fig. 4a). UMAP density

**Fig. 3 | Stroma-specific variations across patients highlight distinct phenotypic groups.** **a** Density enrichment of myCAFs, csCAFs, and reactive stroma signatures shows overlapping expression. **b** Enrichment of iCAF and csCAF signatures. Red circle indicates specific clustering of iCAF cells. **c** Enrichment of PSC signatures representing two separate subgroups. **d** Gene expression heatmap using Normal and Activated signatures (250 genes) clusters patients into three interpretable groups (normal, intermediate, activated) Heatmap cell colors are shown as z-score between -3 to +3. **e** Overlaying of the Moffitt Normal (blue) /Activated (brown) signature scores identified two distinct localizations of the prognostic phenotypes. **f** Pseudobulk heatmap showing differential secretome factors in the stroma between stromal subtypes. Sample tracks at the top show percentage of subpopulations in each patient stroma. **g** Heatmap showing differential secretome

expression in the myeloid (monocyte/macrophages) between stromal subtypes. **h** Differential secretome/cell-contact related gene expression in the lymphocytes (T-lymphocytes) between stromal subtypes. **i** Graphical summary of the interaction network of TME signaling activities upregulated in the activated stroma. Circles and crescents represent ligands and paired receptors. Up/Down arrows indicate the compartmental subpopulation balances observed. Ligands are directional and activate the compartment and subpopulation where the corresponding receptor is distinctly expressed. **j** Graphical interaction network of upregulated signaling within the normal subtype TME. Reference to DEG: 1Fig. 1e, 2Fig S1e, 3Fig 3f, 5Fig S2g, 6Fig S1g, 7Fig S3f, 8Fig 3g, 9Fig S1f, 10Fig 3h. Heatmap colors represented by z score expression -3 to +3. Source data are provided as a Source Data file.

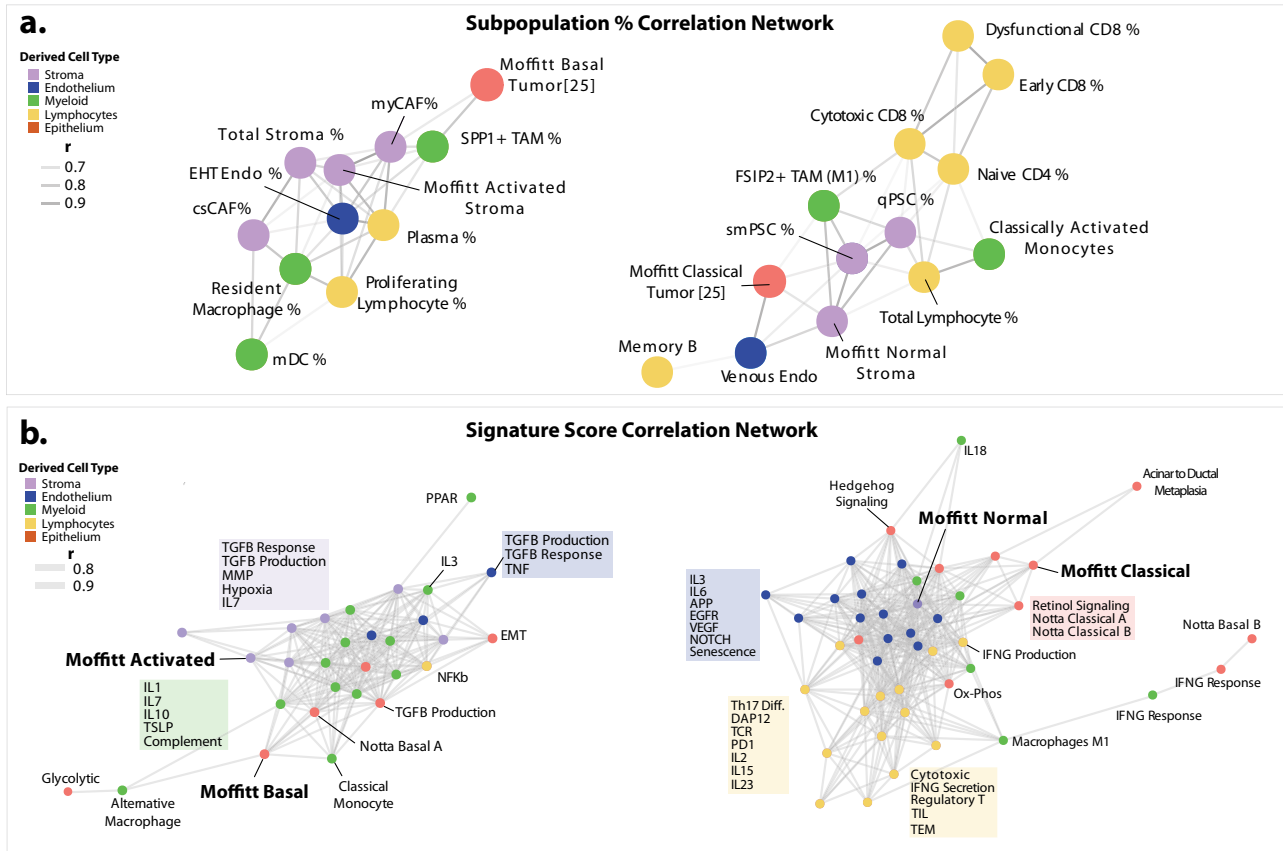


**Fig. 4 | Primary PDAC subtypes show distinct localization within the epithelia.** **a** UMAP projection of the integrated epithelium of both normal and PDAC-derived samples (left). Splitting this subset by source shows distinct subpopulation grouping (right). **b** Density plots show general localization of common PDAC-related markers (EPCAM: Epithelium, MUC1: Malignancy, AMBP: Normal-like, PRSS1: Acinar, KRT17: Basal-like, TFF3: Classical). **c** Expression heatmap shows the average epithelial expression (pseudobulk) for each patient (column) using the top 25 genes for both basal and classical signatures (rows). Column tracks show the associated clinical metadata and consensus clustering subtype results (blue: classical, orange: basal, green: mixed). **d** UMAP of PDAC-derived epithelia projected

onto a 3D space models the separation and gradation of the transforming cells. **e** Localization of the basal-like and classical signatures scores highlight distinct regions and juxtaposition with normal duct and acinar cells. **f** Differential genes (MMP7, SPP1, MUC6) present as markers for inflammation and early precursor cells in the metaplastic subpopulation. **g** Scatterplot of the subtype scores for each single-cell in the neoplastic subset. Cells are colored by the originating patient's subtype designation which shows distinct basal and classical groups with an additional density in between (opacity 25%). Source data are provided as a Source Data file.

plots show clear localization of canonical genes that correspond with our cell annotations (Fig. 4b). Notably, the expression of *KRT17* (basal) and *TFF3* (classical) markers were strictly isolated to the neoplastic cluster. Epithelial analysis using pseudobulk identified patients of basal or classical phenotypes and in some cases, both were detected within the same patient's tissue confirming recent PDAC organoids work studies of subtype dichotomy<sup>58</sup> (Fig. 4c). We further analyzed the tumor

epithelia using the top 1,000 variable genes and projected the data onto a 3D UMAP space (Fig. 4d). This perspective shows bridging between metaplasia and neoplasia, possibly highlighting late tumorigenic differentiation. Normal duct cells were also observed at the edge of the metaplastic cluster hinting at fundamental transitions from the major normal duct cluster (Figure S3a). The metaplastic cluster presented with early malignancy markers such as *MMP7*<sup>59,60</sup> and *MUC6*<sup>61</sup> (Figure S3c)



**Fig. 5 | Cross-compartmental feature analysis shows two polar TME phenotype profiles.** **a** Positively correlated ( $r > 0.6$ ) feature nodes visualized on a correlation-based network layout. Node colors represent the derived Cell Type I (legend). Opacity of connections (edges) represent Pearson correlation  $r$  value. **b** Positively correlated ( $r > 0.7$ ) feature nodes visualized on a correlation-based network layout.

Node colors represent the derived Cell Type I (legend). Opacity of connections (edges) represents Pearson correlation  $r$  value. Distant nodes were adjusted to fit. The comprehensive heatmap representation of this data is presented in Supplementary Fig. 4 with significant  $p$  values denoted by asterisk. Source data are provided as a Source Data file.

and uniquely expressed tumor suppressor genes that point to a homeostatic mechanism against cellular damage or mutation (Fig. 4f). Next, we subset the neoplastic subpopulation and scored each single-cell for basal/classical subtype expression to show that cancer cells match the patient phenotype accordingly while some patients have an admixture of both basal and classical as others have observed (Fig. 4g). Cells of the ‘mixed’ subtype patients represented an intermediate state rather than merely an admixture of both neoplastic subtypes. This observation of both ‘hybrid’ tumor patients and ‘intermediate’ subtypes is consistent with a recent histological study<sup>62</sup>.

**Differential autocrine signal analysis highlights drivers of subtype-dependent TME phenotypes**

Recent analysis of metastatic PDAC revealed how subtype-dependent autocrine factors can influence the TME phenotype<sup>49</sup>. Applying this concept to primary PDAC tissue, we present a list of differentially expressed secreted factors between basal and classical neoplastic cells (Figure S3d). We detected known basal-specific factors such as SPPI<sup>63</sup> and other secreted factors that may orchestrate broader subtype-dependent changes across the ecosystem. To decrease the contribution of intermediate patient signals, the 6 intermediate patients were not included in this comparison. Particularly, factors related to epithelial-mesenchymal transition, TNFA, and IL6/JAK/STAT signaling were elevated in basal cancer cells along with signatures for inflammatory response and IFNG response.

Given the overlap of tumor and stromal phenotypes in our analysis, we wanted to outline the potential mechanisms by which tumor subtypes influence stromal phenotypes (Figure S3e). Classical cancer

patients showed elevation of pancreatic intraepithelial neoplasm (PanIN) related hedgehog signaling<sup>61</sup> genes (SHH, IHH) in their stroma. Correspondingly, PTCH1 was uniquely expressed in csCAFs and myCAFs pointing to an interaction between classical tumor-related desmoplasia<sup>64,65</sup>. Interestingly IL7, which increases immune effector cell function is unique to these patients and may lead classical tumors to have a similar cytotoxic activity seen in the normal subtype<sup>66</sup>. Fibroblast growth factor receptor (FGFR2), which promotes cancer stemness<sup>67</sup> was highly expressed in classical epithelium<sup>67</sup>.

Basal-like patients had increased myCAF and csCAFs proportions and relatedly, higher activated stroma expression while classical patients showed higher PSCs. Basal-like patient stroma showed higher expression of CCL18, a chemokine known to increase T-Regs<sup>68</sup> and M2-like macrophages. This immunosuppressive pattern was observed in the activated TME and CCL18 may serve as one of many mechanisms bridging PDAC tumor subtypes with stromal subtypes. BTNL8, known to stimulate proliferation and cytokine production of naive T-cells was elevated in the stroma of classical patients and may be connected to the overall increase in CD3D + T-Cells and decreased naive T-cells in these patients. Lastly, we provide differentially expressed surface-related genes for both basal and classical cancer cells to aid future investigation (Figure S3f).

**Cross-compartment-based features correlations across patients**

To explore the interconnectedness of the TME, we asked how the presence of one cell type may associate with phenotypes in other compartments. Correlation analysis of subpopulation proportions revealed two major modules of correlated features (Fig. 5a, S4b). In the

first module, both the normal stroma and classical cell type abundances were positively correlated. PSC stromal proportions were also significantly correlated with Classical epithelial proportions, consistent with our normal subtype observations. Notably, this group also had positive associations with cytotoxic CD8<sup>+</sup> T-cell and the FSIP2+ TAMs. The second module grouped basal and activated signatures together along with an increase in SPPI+ TAMs, myCAFs, T-Regs, and immunogenic CAFs.

We furthermore extended this analysis with a list of gene signatures curated from recent publications (Supplementary Data 3) representing canonical PDAC subtypes, pathways, immunomodulation and metabolic signatures (Fig. 5b, S4c). When looking at compartment-specific expression signatures instead of cell type proportions, we saw ‘basal-like/activated’ signatures were associated with evidence of increased epithelial/endothelial TGF $\beta$  production while the TGF $\beta$  response signature was elevated in the stroma. Additionally, matrix metalloproteinase signatures within the stroma, and IL17 signaling by myeloid cells was observed (Fig. 5b). The other ‘classical/normal’ TME profile included the classical and normal subtype signatures as well as IFNG response signature elevated in the myeloid and epithelial compartment. Signatures for acinar to ductal metaplasia, hedgehog and retinol signaling were included in this module. Metabolic comparisons showed the glycolysis signature group with the basal module while oxidative phosphorylation associated with the classical module as previously described<sup>69</sup>. Lymphocyte signatures included transendothelial migration, tumor-infiltrating lymphocytes<sup>70</sup>, and PD1 signaling. Lastly, organoid-derived signatures related to positive chemotherapy response<sup>71</sup> were associated with normal/classical TME. Taken as a whole, this analysis provides a uniting model of correlated cell type-specific signaling describing two drastically different TMEs in PDAC.

## Discussion

This work represents a large effort in leveraging scRNA-seq data to globally correlate published gene signatures with cell type-specific tumor composition and TME signaling. Our work indicates that previous tumor characterizations through molecular, metabolic, or histological approaches should not be viewed individually as disconnected observations. Prior work has suggested that cells of origin<sup>72</sup> or lineage specifiers may control tumor subtype<sup>73</sup>. Additionally, recent multiplex immunofluorescence revealed striking subtype intra-tumoral heterogeneity hinting at intermediate TME landscapes<sup>74</sup>. Here we show co-varying modules (basal and activated vs. classical and normal) appear to orchestrate two distinct cross-compartmental paradigms, although the causal nature of this phenomenon remains unclear. While fibroblast growth factors, T-Reg dominance, and TAMs (SPPI+, GRN+) defined the activated TME, the normal TME had marked T-effector dysfunction, and endothelial transmigration and repair.

Our analysis not only corroborates established PDAC signaling phenotypes, but also highlights less studied yet interesting mechanisms that may open doors to new combination therapies. Identified as a key immunological receptor in our study, CXCR4 antagonism has been shown to allow for immune infiltration and synergistically effects with immune checkpoint inhibitors<sup>75-77</sup>. Inflammatory M1-like macrophages, previously shown to be dependent on NAD<sup>+</sup> supply, was associated with the NMNAT3+ TAMs<sup>78</sup>. CTLA4/TIGIT+ regulatory T-lymphocytes may point to decreased effector function and increased immunosuppressive activity of lymphocytes in patients with greater activated stroma<sup>79</sup>. We identified IL2RG as a target enriched in classical tumor epithelium and has previously shown to attenuate PanIN growth through JAK3 suppression in orthogonally implanted pancreatic cancer<sup>80</sup>. Similarly, FGFR2 which promotes tumor stemness, could be targeted to block the FGF7 expressed from CAFs<sup>81</sup>. Moreover, the differential secretome signatures established in this work can be utilized to assess subtype-dependent secretome activity in

single-cell or bulk RNA studies. By providing a pretrained cell type classifier, our work provides the building blocks of a translational platform to classify single cells, and construct TME profiles that may better inform clinicians about potential therapeutic opportunities.

Our approach to cell type labeling was conservative, aiming to avoid clusters only detectable in single studies. While we took care to avoid this pitfall, a major limitation was our reliance on batch correction of the harmony R package. With any batch correction procedures, there is a risk of introducing artifactual signals. Of note, the minor IL11+ CAFs were not identified across all studies and may be a product of the cluster-based approach in the setting of read count sparsity. A minor ‘EHT-EC’ endothelial subpopulation appeared to be low-quality cells without clear distinguishing markers. In addition, the difficulty in identifying apCAFs in our stroma meta-analysis may be related to differences in tissue pre-processing such as fibroblast disaggregation. Lastly, the validation cohort in this study was comparatively small but future studies incorporating more patients will establish the consistency of the TME patterns observed in this study.

While this manuscript derives many conclusions from expression data alone, targeted validations continue to appear across the literature. In a recent publication, Zeng et al. showed an *in vitro* study and immunofluorescent staining highlighting CCL18 signaling and NF $\kappa$ B-signaling by tumor-associated macrophages (TAMs) as an activator of stromal cells associated with poor prognosis in breast cancer patients. Appropriately, in the aggressive activated stroma concept described here, we describe a GRN+ TAM-specific CCL18 signaling via ACKR1 to induce angiogenesis and endothelial-mesenchymal transition (Fig. 3j). Furthermore, our signature correlation network for the activated & basal tumor profile highlighted NF $\kappa$ B-signaling (specifically by lymphocytes) as a strong correlated TME feature (Fig. 5b). Additionally, Davidson et al. using CRISPR, shRNA, and flow cytometry in mouse models, showed that the immune checkpoint PD-L2 (PDCD1LG2) is predominantly found in the stroma of pancreatic cancer. Although the literature shows a heavy focus on Programmed Death Ligand expression in the context of tumor cells and lymphocytes, we found preferentially PD-L2 expression by myCAFs of the ‘activated’ TME and proposed a potential signaling axis with Regulatory T-Cells via PD1 (PDCD1) (Fig. 3i)<sup>82</sup>. These recent independent findings support our work, which we expect to help provide important context to future molecular studies<sup>83-85</sup>.

Importantly, combination therapies that target TME mediators are being investigated to expand treatment options beyond cytotoxic agents<sup>86-88</sup>. This is in part due to the increasing knowledge of the cancer ecosystem and improved methods for high-resolution molecular data acquisition of patient tissues and model systems. However, our ability to characterize and match patient tumors with optimal treatments hinders the development of novel therapies<sup>89</sup>. This study outlines a basic framework to process a single patient tumor into compartment-specific analysis that leverages pertinent gene signature related to molecular subtypes, and pathologic signaling in PDAC and other solid tumors. For example, the recent failure of a phase II clinical trial using CSF1R inhibitor, Cabiralizumab<sup>90</sup>, may have been related to the ‘normal stroma’ association of the CSF1/CSF1R signaling axis and non-specific targeting based on the broadly shared expression of CSF1R seen across myeloid subpopulations<sup>91</sup>. Future work incorporating cell-specific expression into therapeutic design may provide important context that is missing currently.

## Methods

### Ethics statement

This study was approved by Stony Brook University Human Subjects Committee (IRB), Board Ref# 2017-4223-F and 2017-4223-RI. The patients provided informed written consent to perform the comparative transcriptomic analysis.



**PDAC tissue single-cell processing.** Two PDAC patients (PDAC1 - female, PDAC2 - male) were included through approved Institutional Review Board (IRB) study #11006941 at Stony Brook Hospital. Freshly resected human PDAC tissue was acquired in the grossing room immediately following resection via Whipple or distal pancreatectomy. Bulk tissue specimen was rapidly dissociated into single-cell suspension to reduce transcriptional artifacts of handling and time. Subsequent lysis and single-cell prep were performed as detailed previously<sup>17</sup>. The primary tumor specimen and the metastatic lesions were delivered from the surgical pathology core, once the margins were declared negative, in advanced DMEM/F12 on ice. The specimens were washed in ice-cold PBS and minced into 5-mm<sup>3</sup> sections. For each tissue section, a small section was fixed in 10% formalin for histological studies. Both the tissue specimens were processed simultaneously to avoid the batch effect. The remainder tissue specimens were minced into smaller pieces of 0.5–1 mm<sup>3</sup> and digested for 45 min at 37 °C, in a solution containing 5 mg/mL Collagenase Type II (Invitrogen) and 1 mg/mL dispase (Invitrogen), in the presence of 2.5% fetal bovine serum (FBS) (Gemini Bioproducts). The tissues were additionally digested with 1 mg/mL DNase (Stem Cell Technologies) for 15 min. The digested cells were pelleted and washed two times, and any visible blood cells were removed using the ACK lysis buffer (Sigma-Aldrich). These suspensions were filtered separately through a 70- $\mu$ m cell strainer to get rid of undigested tissue chunks. If the suspension needed further declumping, it was digested for a minute in TryPLE express (Invitrogen), neutralized with 10% FBS, and filtered through a 40- $\mu$ m cell strainer. The final cell suspensions were pelleted by centrifuging at 300 g for 5 min and resuspended in PBS with 0.1% BSA (Life Technologies) to a final concentration of 10,000,000 cells/mL. The cell viability was examined using trypan blue exclusion (Invitrogen). The Cell Ranger Single-Cell Suite 3.01 was used for demultiplexing, barcode assignment, and raw sequencing processing. The processed and unprocessed data for these patient samples is now provided as well under GEO repository (Accession #: GSE231535). Sex or gender was not used as criteria during data collection.

**Curation of available human scRNA-seq data sets.** Publicly available scRNA-seq data on pancreas was obtained from multiple sources. Raw sequencing reads from Peng et al. (GSA: CRA001160) were aligned to genes using Alevin<sup>92</sup> and GRCh38 while all other data sets were integrated beginning with *gene x cell* count matrices. Additional data sets were included: Qadir et al. “GSE131886”, Moncada et al. “GSE111672”, Muraro et al. “GSE85241”, Segerstolpe et al. “EMTAB-5061”, and Lin et al. “GSE154778”.

**Processing and integration of human scRNA-seq Data.** Seurat (v3.0.2 & v4.0) R package and custom scripts were used for single-cell analysis. 4 independent single-cell RNA-seq data sets were processed for quality control and preliminary analysis. Default parameters were used unless specified (min.cell = 10, min.genes = 300, percent.mt <10%). Cells with abnormally high or low feature counts, percent mitochondrial gene expression >10% were filtered. Ribosomal and mitochondrial genes were removed from features. Protein coding mRNA genes were selected as features for isolating the top 8000 highly variable genes, and the top 30 PCA dimensions were used for downstream integration. Using Harmony R package<sup>93</sup>, technical factors between data sets were corrected to construct a well-integrated embedding across cell types and patients. Batch-corrected harmony PCA components were then used as input for UMAP, tSNE, and FindNeighbor function. The sensitivity for FindClusters was set to 0.2 using the Louvain method. FindAllMarkers was used to annotate the coarse cell type [Level 1]. Subsequent subpopulation analysis of major cells groups (stroma, epithelium, lymphocytes, myeloid) was further dissected by applying this variable feature selection, PCA, Harmony, and dimensional reduction

procedure over each cell type. Compartment-specific parameters are as follows: Myeloid (4000 variable genes, 10 dims), Epithelium (3000 genes, 8 dims), Tumor-Only subset (600 genes, 5 dims), Stroma (6000 genes, 20 dims), Endothelial (4000 genes, 12 dims), Lymphocytes (4000 genes, 18 dims). Visualization of gene expression and signature scores throughout the study used Seurat’s FeaturePlot function and nebulosa R package<sup>94</sup>.

**Curation and processing of mouse scRNA-seq data sets.** Curation of the mouse dataset included publicly available data sets from three different studies<sup>23,95,96</sup>. Mouse data sets were transformed and scaled separately before merging (i.e. Hosein, Elyada, and Gabitova-Cornell) for normalization using 3000 variable features and scaled the data sets. We selected 6,000 variable features across all data sets for integration features. 50 PCA components were used for harmony integration. UMAP parameters were as follows: assay = “SCT”, reduction = “harmony”, dims = 1:10, n.neighbors = 80, n.component = 2, min.dist = 0.3, and negative.sample.rate = 10. Top 10 dimensions were used to find nearest neighbors and lastly 0.2 for cluster resolution. Finally, we used the FindAllMarkers function to help annotate mouse cell clusters.

**Streamlined single-cell PDAC classifier.** Using the SingleCellNet R package (v0.4.1), we trained a random forest model on the discovery atlas dataset<sup>46</sup>. Default parameters were used with the following exceptions (nTopGenes=20, nTopGenePairs=40, nTrees=20). For each cell type group, 5,000 *training* cells were used to establish top gene pairs. The remaining 99,518 *validation* cells were used for internal performance metrics. For cross-species utilization of the classifier, we first converted mouse gene names into human gene names through “mouse2human” function from the R package “homologene”; if there were multiple mouse gene names aligned with one human gene name, we aggregated all into one human gene name. 14,084 common protein-coding genes within the human atlas genes and the mouse genes were set as the feature search space (original gene counts: 19,554).

**Differential secretome analysis of TME.** The atlas was down-sampled to 3000 cells for each of the subpopulations with the exception of minor cell populations that were not completely resolved (IL11+ CAFs, EHT-EC). The FindAllMarkers function was run with parameters (min.pct >0.2, min.diff.pct >0.1, method = “MAST”). Top 100 differential genes based on log-fold change were filtered against a cell-signaling database of known secretome (685 genes), cell-contact (240 genes), ECM (85 genes) added from CellChat v0.0.2<sup>97,98</sup>.

**Subtype-dependent pseudobulk analysis of TME cell signaling.** Individual “pseudobulk signatures” of patients’ stroma, myeloid, endothelium, lymphocytes, and epithelial compartments were calculated by aggregating all cell expression per compartment. DESeq2 was used to define differential genes between patients with different stromal subtypes and tumor subtypes. Cell-signaling genes for subsequent translational analysis were derived once again from the CellChat v0.0.2 table of ligand/receptor interactions. Generally, candidate signaling factors activity from subpopulations was mechanistically associated with the cell type uniquely expressing the corresponding pair. Intermediate patients were excluded from both stroma and tumor comparisons. For tumor secretome analysis, intermediate patients were removed from analysis to focus on the strongest basal or classical samples. Additional code and analysis are available at <https://github.com/rmoffitt/scOh>.

**Compartment phenotype correlations and networks.** Cell type-specific features were calculated for all patients using an extensively curated list of gene signatures with a focus on PDAC and immune-related pathways (Supplementary Data 3). For signatures that did not

have a clear cell type association, the cell type that had the highest expression across patients was selected. In a second experiment, cell type percentages were calculated for each compartment for each patient. This percentage is the fraction of a patient sample's cell type (e.g. Stroma) that is made up of a particular cell subpopulation (e.g. myCAF). For example, the myCAF % feature represents the # of myCAFs a patient sample contains, divided by the patient's total # of stromal cells. Further, Cytotoxic CD8+ T-Cell % is derived by a patient's Cytotoxic CD8+ T-Cell count divided by the patient's total # of lymphocytes. Two correlation matrices, for both signature features and cell type percentages, were produced using corr R package. Features were sorted using 'First Principal Component' - sorting. Significance (designated by \*) was set to  $p < 0.05$ . Correlation networks were produced using ggraph R package with an  $r > 0.6$  and  $r > 0.7$  filter for subpopulation % and signature networks, respectively. "kk" was used as the graph layout.

**Statistics and reproducibility.** The design of the study was exploratory in nature, aimed at investigating the association of tumor subtypes with individual patients' microenvironments. No statistical method was used to predetermine sample size and data sets were integrated as they became available. To promote reproducibility and transparency in our research, we have made all data and code used for analyses publicly available.

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

Transcriptomic data from newly sequenced patients are available at GEO Accession #: [GSE231535](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE231535). Additional publicly available data sets include Qadir et al. ([GSE131886](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE131886)), Moncada et al. ([GSE111672](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE111672)), Muraro et al. ([GSE85241](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE85241)), Segerstolpe et al. ([EMTAB-5061](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=EMTAB-5061)), and Lin et al. ([GSE154778](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE154778)). Mouse data utilized in the classifier experiments include Hosen et al. ([GSE12558](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE12558)), Elyada et al. ([GSE129455](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE129455)), and Gabitov-Cornell et al. ([GSE156210](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE156210)) Source data are provided in this paper.

### Code availability

Output of gene expression results, Seurat objects, patient metadata markdowns in R code, and classifier with vignette documentation can be found on GitHub: <https://github.com/rmoffitt/scOh>.

### References

- Siegel, R. L., Miller, K. D. & Jemal, A. Cancer statistics, 2020. *CA Cancer J. Clin.* **70**, 7–30 (2020).
- Logsdon, C. D. et al. Molecular profiling of pancreatic adenocarcinoma and chronic pancreatitis identifies multiple genes differentially regulated in pancreatic cancer (Cancer Research (2003) (2649-2657)). *Cancer Res.* **63**, 3445 (2003).
- Maurer, H. C. et al. Experimental microdissection enables functional harmonisation of pancreatic cancer subtypes. *Gut* **68**, 1034–1043 (2019).
- Wolff, R. A. et al. Dynamic changes during the treatment of pancreatic cancer. *Oncotarget* **9**, 14764–14790 (2018).
- Boj, S. F. et al. Organoid models of human and mouse ductal pancreatic cancer. *Cell* **160**, 324–338 (2015).
- Laise, P. et al. Pancreatic ductal adenocarcinoma comprises coexisting regulatory states with both common and distinct dependencies. *bioRxiv* <https://www.biorxiv.org/content/10.1101/2020.10.27.357269v1> (2020).
- Moffitt, R. A. et al. Virtual microdissection identifies distinct tumor- and stroma- specific subtypes of pancreatic ductal adenocarcinoma. *Nat. Genet.* **47**, 1168–1178 (2015).
- Aung, K. L. et al. Genomics-driven precision medicine for advanced pancreatic cancer - early results from the COMPASS trial. *Clinical Cancer Res.* **24**, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5968824/> (2017).
- Grünwald, B. T. et al. Spatially confined sub-tumor microenvironments orchestrate pancreatic cancer pathobiology. *bioRxiv* <https://www.biorxiv.org/content/10.1101/2021.02.18.431890v1> (2021).
- Rashid, N.U. et al. Purity independent subtyping of tumors (PurlST), a clinically robust, single-sample classifier for tumor subtyping in pancreatic cancer. *Clin. Cancer Res.* (2019) <https://doi.org/10.1158/1078-0432.CCR-19-1467>.
- Torre-Healy, L. A. et al. Open-source curation of a pancreatic ductal adenocarcinoma gene expression analysis platform (pdacr) supports a two-subtype model. *Commun. Biol.* **6**, 163 (2023).
- Whiteside, T. L. The tumor microenvironment and its role in promoting tumor growth. *Oncogene* **27**, 5904–5912 (2008).
- Tang, F. et al. mRNA-Seq whole-transcriptome analysis of a single cell. *Nat. Methods* **6**, 377–382 (2009).
- Trapnell, C., Cacchiarelli, D. & Qiu, X. Monocle: cell counting, differential expression, and trajectory analysis for single-cell RNA-Seq experiments. *Bioconductor* 10 <http://bioconductor.org/packages/devel/bioc/vignettes/monocle/inst/doc/monocle-vignette.pdf> (2017).
- MD, C. et al. A distinct microglial subset at the tumor-stroma interface of glioma. *Glia* **69**, 1767–1781 (2021).
- Azizi, E. et al. Single-cell map of diverse immune phenotypes in the breast tumor microenvironment. *Cell* **174**, 1293–1308.e36 (2018).
- Rao, M. et al. Comparative single-cell RNA sequencing (scRNA-seq) reveals liver metastasis-specific targets in a patient with small intestinal neuroendocrine cancer. *Cold Spring Harb. Mol. Case Stud.* **6**, 1–9 (2020).
- Muraro, M. J. J. et al. A single-cell transcriptome atlas of the human pancreas. *Cell Syst.* **3**, 385–394.e3 (2016).
- Baron, M. et al. A single-cell transcriptomic map of the human and mouse pancreas reveals inter- and intra-cell population structure. *Cell Syst.* **3**, 346–360.e4 (2016).
- Segerstolpe, Å. et al. Single-cell transcriptome profiling of human pancreatic islets in health and type 2 diabetes. *Cell Metab.* **24**, 593–607 (2016).
- Qadir, M. M. F. et al. Single-cell resolution analysis of the human pancreatic ductal progenitor cell niche. *Proc. Natl Acad. Sci. USA* **117**, 10876–10887 (2020).
- Öhlund, D. et al. Distinct populations of inflammatory fibroblasts and myofibroblasts in pancreatic cancer. *J. Exp. Med.* **214**, jem.20162024 (2017).
- Elyada, E. et al. Cross-species single-cell analysis of pancreatic ductal adenocarcinoma reveals antigen-presenting cancer-associated fibroblasts. *Cancer Discov.* **9**, 1102–1123 (2019).
- Biffi, G. et al. IL1-induced JAK/STAT signaling is antagonized by TGFβ to shape CAF heterogeneity in pancreatic ductal adenocarcinoma. *Cancer Discov.* **9**, 282–301 (2019).
- Chen, K. et al. Single-cell RNA-seq reveals dynamic change in tumor microenvironment during pancreatic ductal adenocarcinoma malignant progression. *EBioMedicine* **66**, 103315 (2021).
- Wang, Y. et al. Single-cell analysis of pancreatic ductal adenocarcinoma identifies a novel fibroblast subtype associated with poor prognosis but better immunotherapy response. *Cell Discov.* **7**, 36 (2021).
- Zhang, Q. et al. Landscape and dynamics of single immune. *Cells* **179**, 829–845.e20 (2019).
- Yang, J. et al. Calcium-binding proteins S100A8 and S100A9: investigation of their immune regulatory effect in myeloid cells. *Int J. Mol. Sci.* **19**, 1833 (2018).

29. Wu, L. et al. Spatially-resolved transcriptomics analyses of invasive fronts in solid tumors. *bioRxiv* <https://www.biorxiv.org/content/10.1101/2021.10.21.465135v1> (2021).
30. Lauer, S. J. et al. Two copies of the human apolipoprotein C-I gene are linked closely to the apolipoprotein E gene. *J. Biol. Chem.* **263**, 7277–7286 (1988).
31. Quaranta, V. et al. Macrophage-derived granulin drives resistance to immune checkpoint inhibition in metastatic pancreatic cancer. *Cancer Res.* **78**, 4253–4269 (2018).
32. Fleming, R. E., Crouch, E. C., Ruzicka, C. A. & Sly, W. S. Pulmonary carbonic anhydrase IV: developmental regulation and cell-specific expression in the capillary endothelium. *Am. J. Physiol.* **265**, L627–L635 (1993).
33. Thiriot, A. et al. Differential DARC/ACKR1 expression distinguishes venular from non-venular endothelial cells in murine tissues. *BMC Biol.* **15**, 1–19 (2017).
34. Su, T. et al. Single-cell analysis of early progenitor cells that build coronary arteries. *Nature* **559**, 356–362 (2018).
35. Stahl, A. et al. SOCS3 is an endogenous inhibitor of pathologic angiogenesis. *Blood* **120**, 2925 (2012).
36. Lee, S. et al. Sprouty1 inhibits angiogenesis in association with up-regulation of p21 and p27. *Mol. Cell Biochem.* **338**, 255–p261 (2010).
37. Yan, M. et al. Dynamic regulatory networks of T cell trajectory dissect transcriptional control of T cell state transition. *Mol. Ther. Nucleic Acids* **26**, 1115–1129 (2021).
38. Fei, Q. et al. High-dimensional single-cell analysis delineates radiofrequency ablation induced immune microenvironmental remodeling in pancreatic cancer. *Cell Death Dis.* **11**, 1–13 (2020).
39. Ng, S. S. et al. The NK cell granule protein NKG7 regulates cytotoxic granule exocytosis and inflammation. *Nat. Immunol.* **21**, 1205–1218 (2020).
40. Xiao, Z., Mescher, M. F. & Jameson, S. C. Detuning CD8 T cells: down-regulation of CD8 expression, tetramer binding, and response during CTL activation. *J. Exp. Med.* **204**, 2667–2677 (2007).
41. Punt, S. et al. Whole-transcriptome analysis of flow-sorted cervical cancer samples reveals that B cell expressed TCL1A is correlated with improved survival. *Oncotarget* **6**, 38681–38694 (2015).
42. Reimold, A. M. et al. Plasma cell differentiation requires the transcription factor XBP-1. *Nature* **412**, 300–307 (2001).
43. Turnquist, H. R. et al. CCL21 induces extensive intratumoral immune cell infiltration and specific anti-tumor cellular immunity. *Int. J. Oncol.* **30**, 631–639 (2007).
44. Binnewies, M. et al. Targeting TREM2 on tumor-associated macrophages enhances immunotherapy. *Cell Rep.* **37**, 109844 (2021).
45. Zhang, W., Zhang, Y. & Fang, Q. [Effect of Galectin-9/Tim-3 pathway on the polarization of M1/M2 subtype in murine macrophages induced by lipopolysaccharide]. *Zhonghua Wei Zhong Bing Ji Jiu Yi Xue* **30**, 836–841 (2018).
46. Tan, Y. & Cahan, P. SingleCellNet: a computational tool to classify single cell RNA-seq data across platforms and across species. *Cell Syst.* **9**, 207–213.e2 (2019).
47. Yu, X., Zhang, L., Chaudhry, A., Rapaport, A. S. & Ouyang, W. Unravelling the heterogeneity and dynamic relationships of tumor-infiltrating T cells by single-cell RNA sequencing analysis. *J. Leukoc. Biol.* **107**, 917–932 (2020).
48. Harricharran, T. & Ogunwobi, O. O. Oxytocin and oxytocin receptor alterations, decreased survival, and increased chemoresistance in patients with pancreatic cancer. *Hepatobiliary Pancreat. Dis. Int.* **19**, 175–180 (2020).
49. Raghavan, S. et al. Microenvironment drives cell state, plasticity, and drug response in pancreatic cancer. *Cell* **184**, 6119–6137.e26 (2021).
50. Chakraborty, D. et al. Activation of STAT3 integrates common pro-fibrotic pathways to promote fibroblast activation and tissue fibrosis. *Nat. Commun.* **8**, 1–16 (2017).
51. Latchman, Y. et al. PD-L2 is a second ligand for PD-1 and inhibits T cell activation. *Nat. Immunol.* **2**, 261–268 (2001).
52. Li, Y. et al. Stress-induced upregulation of TNFSF4 in cancer-associated fibroblast facilitates chemoresistance of lung adenocarcinoma through inhibiting apoptosis of tumor cells. *Cancer Lett.* **497**, 212–220 (2021).
53. Niu, J. et al. Keratinocyte growth factor/fibroblast growth factor-7-regulated cell migration and invasion through activation of NF- $\kappa$ B transcription factors \*. *J. Biol. Chem.* **282**, 6001–6011 (2007).
54. Imaizumi, T., Yoshida, H. & Satoh, K. Regulation of CX3CL1/fractalkine expression in endothelial cells. *J. Atheroscler. Thromb.* **11**, 15–21 (2004).
55. Larsen, A. M. H. et al. Collagen density modulates the immunosuppressive functions of macrophages. *J. Immunol.* **205**, 1461–1472 (2020).
56. Jin-Sung Park, A. et al. Normalization of tumor vessels by Tie2 activation and Ang2 inhibition enhances drug delivery and produces a favorable tumor microenvironment. *Cancer Cell* **30**, 953–967 (2016).
57. Hilton, D. J. LIF: lots of interesting functions. *Trends Biochem Sci.* **17**, 72–76 (1992).
58. Juiz, N. et al. Basal-like and Classical cells coexistence in pancreatic cancer revealed by single cell analysis. *bioRxiv* <https://doi.org/10.1101/2020.01.07.897454>. (2020).
59. Sawey, E. T., Johnson, J. A. & Crawford, H. C. Matrix metalloproteinase 7 controls pancreatic acinar cell transdifferentiation by activating the Notch signaling pathway. *Proc. Natl Acad. Sci.* **104**, 19327–19332 (2007).
60. Fukuda, A. et al. Stat3 and MMP7 contribute to pancreatic ductal adenocarcinoma initiation and progression. *Cancer Cell* **19**, 441 (2011).
61. Prasad, N. B. et al. Gene expression profiles in pancreatic intraepithelial neoplasia reflect the effects of hedgehog signaling on pancreatic ductal epithelial cells. *Cancer Res.* **65**, 1619–1626 (2005).
62. Delecourt, F. PACpAlnt: a deep learning approach to identify molecular subtypes of pancreatic adenocarcinoma on histology slides. *BioRxiv* <https://doi.org/10.1101/2022.01.04.474951> (2022)
63. Zhou, D. C. et al. Spatial drivers and pre-cancer populations collaborate with the microenvironment in untreated and chemoresistant pancreatic cancer. *bioRxiv* <https://www.biorxiv.org/content/10.1101/2021.01.13.426413v1> (2021).
64. Steele, N. G. et al. Inhibition of hedgehog signaling alters fibroblast composition in pancreatic cancer. *Clin. Cancer Res.* **27**, 2023–2037 (2021).
65. Gu, D., Schlotman, K. E. & Xie, J. Deciphering the role of hedgehog signaling in pancreatic cancer. *J. Biomed. Res.* **30**, 353 (2016).
66. Colpitts, S. L., Dalton, N. M. & Scott, P. IL-7 receptor expression provides the potential for long-term survival of both CD62L high central memory T cells and Th1 effector cells during leishmania major infection. *J. Immunol.* **182**, 5702–5711 (2009).
67. Quan, M. Y. et al. An FGFR/AKT/SOX2 signaling axis controls pancreatic cancer stemness. *Front. Cell Dev. Biol.* **8**, 287 (2020).
68. Chenivresse, C. et al. Pulmonary CCL18 recruits human regulatory T cells. *J. Immunol.* **189**, 128–137 (2012).
69. Espiau-Romera, P., Courtois, S., Parejo-Alonso, B. & Sancho, P. Molecular and metabolic subtypes correspondence for pancreatic ductal adenocarcinoma classification. *J. Clin. Med.* **9**, 1–21 (2020).
70. Chen, H., Pan, Y., Jin, X. & Chen, G. An immune cell infiltration-related gene signature predicts prognosis for bladder cancer. *Sci. Rep.* **11**, 1–13 (2021).

71. Tiriach, H. et al. Organoid profiling identifies common responders to chemotherapy in pancreatic cancer. *Cancer Discov.* **8**, 1112–1129 (2018).
72. Sarai, K. et al. Cell of origin affects tumour development and phenotype in pancreatic ductal adenocarcinoma. *Gut* **68**, 487–498 (2018).
73. Hayashi, A. et al. A Unifying paradigm for transcriptional heterogeneity and squamous features in pancreatic ductal adenocarcinoma. *Nat. Cancer* **1**, 59–74 (2020).
74. Williams, H. L. et al. Spatially resolved single-cell assessment of pancreatic cancer expression subtypes reveals co-expressor phenotypes and extensive intratumoral heterogeneity. *Cancer Res.* **83**, 441–455 (2023).
75. Song, J. S. et al. A highly selective and potent CXCR4 antagonist for hepatocellular carcinoma treatment. *Proc. Natl Acad. Sci. USA* **118**, e2015433118 (2021).
76. Santagata, S. et al. CXCR4 and CXCR7 signaling pathways: a focus on the cross-talk between cancer cells and tumor microenvironment. *Front. Oncol.* **11**, 1225 (2021).
77. Luker, G. D. et al. At the bench: pre-clinical evidence for multiple functions of CXCR4 in cancer. *J. Leukoc. Biol.* **109**, 969–989 (2021).
78. Cameron, A. M. et al. Inflammatory macrophage dependence on NAD<sup>+</sup> salvage is a consequence of reactive oxygen species-mediated DNA damage. *Nat. Immunol.* **20**, 420–432 (2019).
79. Catakovic, K., Klieser, E., Neureiter, D. & Geisberger, R. T cell exhaustion: from pathophysiological basics to tumor immunotherapy. *Cell Commun. Signal.* **2017 15:1** **15**, 1–16 (2017).
80. Ayars, M. et al. IL2RG, identified as overexpressed by RNA-seq profiling of pancreatic intraepithelial neoplasia, mediates pancreatic cancer growth. *Oncotarget* **8**, 83370–83383 (2017).
81. Carter, E. P. et al. Dissecting FGF signalling to target cellular crosstalk in pancreatic cancer. *Cells* **10**, 847 (2021).
82. Gao, X. et al. ETV4 promotes pancreatic ductal adenocarcinoma metastasis through activation of the CXCL13/CXCR5 signaling axis. *Cancer Lett.* **524**, 42–56 (2022).
83. Klussmeier, A., Aurich, S., Niederstadt, L., Wiedenmann, B. & Grötzing, C. Secretin receptor as a target in gastrointestinal cancer: expression analysis and ligand development. *Biomedicines* **10**, 536 (2022).
84. Chen, K. et al. Single cell RNA-seq reveals the CCL5/SDC1 receptor-ligand interaction between T cells and tumor cells in pancreatic cancer. *Cancer Lett.* **545**, 215834 (2022).
85. Cheng, M. et al. Immunosuppressive role of SPP1-CD44 in the tumor microenvironment of intrahepatic cholangiocarcinoma assessed by single-cell RNA sequencing. *J. Cancer Res. Clin. Oncol.* 1–16 (2022) <https://doi.org/10.1007/S00432-022-04498-W/FIGURES/7>.
86. Waibl Polania, J., Lerner, E. C., Wilkinson, D. S., Hoyt-Miggelbrink, A. & Fecci, P. E. Pushing past the blockade: advancements in T cell-based cancer immunotherapies. *Front. Immunol.* **12**, 4862 (2021).
87. Xiang, X., Wang, J., Lu, D. & Xu, X. Targeting tumor-associated macrophages to synergize tumor immunotherapy. *Signal Transduct. Target. Ther.* **6**, 1–12 (2021).
88. Abdelfattah, N. et al. Single-cell analysis of human glioma and immune cells identifies S100A4 as an immunotherapy target. *Nat. Commun.* **13**, 1–18 (2022).
89. Matrisian, L. M. & Berlin, J. D. The past, present, and future of pancreatic cancer clinical trials. *Am. Soc. Clin. Oncol. Educ. Book* **35**, e205–e215 (2016).
90. Wang-Gillam, A. et al. A randomized phase II study of cabiralizumab (cabira) + nivolumab (nivo) ± chemotherapy (chemo) in advanced pancreatic ductal adenocarcinoma (PDAC). *J. Clin. Oncol.* **37**, TPS465–TPS465 (2019).
91. Zhang, L. et al. Single-cell analyses inform mechanisms of myeloid-targeted therapies in colon cancer. *Cell* **181**, 442–459.e29 (2020).
92. Srivastava, A., Malik, L., Smith, T., Sudbery, I. & Patro, R. Alevin efficiently estimates accurate gene abundances from dscRNA-seq data. *Genome Biol.* **20**, 65 (2019).
93. Korsunsky, I. et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).
94. Alquicira-Hernandez, J. & Powell, J. E. Nebulosa recovers single-cell gene expression signals by kernel density estimation. *Bioinformatics* **37**, 2485–2487 (2021).
95. Hosein, A. N. et al. Cellular heterogeneity during mouse pancreatic ductal adenocarcinoma progression at single-cell resolution. *JCI Insight* **5**, e129212 (2019).
96. Gabitova-Cornell, L. et al. Cholesterol pathway inhibition induces TGF- $\beta$  signaling to promote basal differentiation in pancreatic cancer. *Cancer Cell* **38**, 567–583.e11 (2020).
97. Jin, S. et al. Inference and analysis of cell-cell communication using CellChat. *Nat. Commun.* **12**, 1–20 (2021).
98. Oh, K., et al. Coordinated single cell tumor microenvironment dynamics reinforce pancreatic cancer subtype, scOh, <https://doi.org/10.5281/zenodo.8066926>, (2023)

## Acknowledgements

This work was supported by the Ruth L. Kirschstein Service Award from the National Institutes of Health to K.O. (CA257489), the Tissue Analytics and Bioinformatics Shared Resources of the Stony Brook's Cancer Center and the Department of Biomedical Informatics. We thank members of the Powers and Shroyer laboratory for the thoughtful discussions on pathological interpretation. We thank Stony Brook Hospital's BioBank for their assistance in acquiring patient tissue samples in this study. Figures were created with BioRender.com.

## Author contributions

R.M. and K.O. conceived the project and interpreted the results. K.O., Y.J.Y., and L.T.H. performed the analysis; K.O., A.S., G.G., J.K., M.G., P.W., M.C., and M.R. acquired, processed, or provided tissues and data. K.O., L.T.H., D.F., and R.M. wrote the manuscript. S.P. and R.M. supervised the project. All authors critically reviewed the manuscript.

## Competing interests

The author declares the following competing interests: RAM holds financial interests in PurIST (Methods and compositions for prognostic and/or diagnostic subtyping of pancreatic cancer US11053550, WO2016061252A1), which has been licensed to GeneCentric Technologies. R.A.M. declares no non-financial competing interests that could be perceived to influence the work reported in this paper. The remaining authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-023-40895-6>.

**Correspondence** and requests for materials should be addressed to Richard A. Moffitt.

**Peer review information** *Nature Communications* thanks Andrew Biankin, Ute Jungwirth, and the other anonymous reviewers for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023