Transcriptional subtype-specific microenvironmental crosstalk and tumor cell plasticity in metastatic pancreatic cancer

Srivatsan Raghavan[1,2,3,4*], Peter S. Winter[1,2,5,6*#], Andrew W. Navia[1,2,5,6,7*], Hannah L. Williams[1,3*], Alan DenAdel[8,9], Radha L. Kalekar[1,2], Jennyfer Galvez-Reyes[2,5,6], Kristen E. Lowder[1,2], Nolawit Mulugeta[2,5,6], Manisha S. Raghavan[1,2], Ashir A. Borah[2], Sara A. Väyrynen[1], Andressa Dias Costa[1,3], Raymond W.S. Ng[1,2], Junning Wang[1], Emma Reilly[1], Dorisanne Y. Ragon[1], Lauren K. Brais[1], Alex M. Jaeger[6], Liam F. Spurr[1,2], Yvonne Y. Li[1,2], Andrew D. Cherniack[1,2], Isaac Wakiro[1], Asaf Rotem[1,2,10], Bruce E. Johnson[1,3,4,10], James M. McFarland[2], Ewa T. Sicinska[1,3], Tyler E. Jacks[6], Thomas E. Clancy[1,3,11], Kimberly Perez[1,3,4], Douglas A. Rubinson[1,3,4], Kimmie Ng[1,3,4], James M. Cleary[1,3,4], Lorin Crawford[8,12], Scott R. Manalis[6,13], Jonathan A. Nowak[3,14,15], Brian M. Wolpin[1,3,4π], William C. Hahn[1,2,3,4π], Andrew J. Aguirre[1,2,3,4π#], Alex K. Shalek[2,3,5,6,7,16π#]

[1]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA
[2]Broad Institute of MIT and Harvard, Cambridge, MA
[3]Harvard Medical School, Boston, MA
[4]Department of Medicine, Brigham and Women's Hospital, Boston, MA
[5]Institute for Medical Engineering & Science, Massachusetts Institute of Technology, Cambridge, MA
[6]Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA
[7]Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA
[8]Center for Computational Molecular Biology, Brown University, Providence, RI
[9]Division of Applied Mathematics, Brown University, Providence, RI
[10]Center for Cancer Genomics, Dana-Farber Cancer Institute, Boston, MA
[11]Department of Surgery, Brigham and Women's Hospital, Boston, MA
[12]Department of Biostatistics, Brown University, Providence, RI
[13]Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA
[14]Department of Oncologic Pathology, Dana-Farber Cancer Institute, Boston, MA
[15]Department of Pathology, Brigham and Women's Hospital, Boston, MA
[16]Ragon Institute of MGH, MIT, and Harvard, Cambridge, MA

*Equal first author contribution
πEqual senior contribution
#To whom correspondence should be addressed:
Peter S. Winter, Ph.D., pswinter@mit.edu
Andrew J. Aguirre, M.D., Ph.D., andrew_aguirre@dfci.harvard.edu
Alex K. Shalek, Ph.D., shalek@mit.edu

ABSTRACT

In pancreatic ductal adenocarcinoma (PDAC), the basal-like and classical transcriptional subtypes are associated with differential chemotherapy sensitivity and patient survival. These phenotypes have been defined using bulk transcriptional profiling, which can mask underlying cellular heterogeneity and the biologic mechanisms that distinguish these subtypes. Furthermore, few studies have interrogated metastases, which are the cause of mortality in most patients with this highly lethal disease. Using single-cell RNA-sequencing of metastatic needle biopsies and matched organoid models, we demonstrate intra-tumoral subtype heterogeneity at the single-cell level and define a continuum for the basal-like and classical phenotypes that includes hybrid cells that co-express features of both states. Basal-like tumors show enrichment of mesenchymal and stem-like programs, and demonstrate immune exclusion and tumor cell crosstalk with specific macrophage subsets. Conversely, classical tumors harbor greater immune infiltration and a relatively pro-angiogenic microenvironment. Matched organoid models exhibit a strong bias against the growth of basal-like cells in standard organoid media, but modification of culture conditions can rescue the basal-like phenotype. This study reframes the transcriptional taxonomy of PDAC, demonstrates how divergent transcriptional subtypes associate with unique tumor microenvironments, and highlights the importance of evaluating both genotype and transcriptional phenotype to establish high-fidelity patient-derived cancer models.

MAIN

While classification of human malignancies by genotype has provided critical structure for understanding tumor biology, tumors can also harbor clinically relevant variation in transcriptional phenotypes.[1] Indeed, for several malignancies such as pancreatic ductal adenocarcinoma (PDAC), classification based on RNA expression has emerged as a genotype-independent predictor of chemotherapy sensitivity and patient survival.[2-6] In PDAC, bulk transcriptional profiling has defined two major transcriptional subtypes, basal-like/squamous (hereafter referred to as “basal”) and classical, where the former is associated with worse prognosis and greater treatment resistance.[3-5,7-14] However, classification based on bulk expression profiling can obscure clinically relevant cellular attributes because it reduces signals from multiple cell types to a single, whole sample average. In reality, PDAC tumors, like many other cancers, are complex multicellular ecosystems shaped by both malignant and microenvironmental features. Unlike in DNA sequencing where mutant and normal reads can be precisely separated, malignant and non-malignant signals in bulk RNA profiles are not easily disentangled, making conclusions about their relationships challenging.

The recent application of single-cell RNA-sequencing (scRNA-seq) to human cancers has revealed that the tumor ecosystem is highly heterogeneous and often consists of continuous phenotypes within both malignant and non-malignant populations.[15-21] The precise cellular characterization this method affords has enabled the re-examination of transcriptional taxonomies and reframed our understanding of the summaries provided by bulk

measurements in multiple cancers.[15,21-26] Such enhanced resolution may be particularly useful in PDAC, where neoplastic cellularity is generally low and stromal content is high. Understanding the distribution and plasticity of malignant and non-malignant states within individual PDAC tumors has important implications for the interpretation of transcriptional subtypes, directing therapy, and monitoring tumor evolution. However, few single-cell studies have been conducted in human PDAC, and these have largely focused on stromal cell types or provided a limited analysis of malignant cells.[11,27-29] We therefore lack a harmonized view of the interplay between malignant transcriptional subtypes and their associated tumor microenvironment (TME).

Our current understanding of PDAC is largely derived from resected primary tumors.[12,13,30] However, the majority of patients with PDAC present with, and succumb to, metastatic disease, which occurs most commonly in the liver.[30] At present, we have little information about the cellular phenotypes and microenvironmental interactions in metastatic lesions. Tissue availability has been a key barrier to enhanced understanding of metastatic disease, as needle biopsies provide an important but cell-limited window into the biology of the metastatic niche.

In conjunction with detailed molecular analysis of patient samples, reliable *ex vivo* models are needed to functionally test clinical and molecular observations. For this purpose, human cancer cell line models are frequently utilized, as is the case in PDAC. However, the methods to generate new cell lines from human tissue are generally inefficient, which limits their utility in personalized medicine.[31] Moreover, once established, cell lines can display significant drift in culture.[32] To address these limitations, several groups have established efficient methods for generating patient-derived organoid cultures from PDAC tissue with the goal of modeling an individual patient's disease.[10,33-35] However, few studies have examined the fidelity and evolution of organoid phenotype and genotype relative to the parental patient tissue.

Here, we developed and employed an optimized translational workflow to perform both high-resolution profiling of patient tissue using scRNA-seq via Seq-Well[36] and derivation of matched organoid models from the same metastatic core needle biopsy. Through this approach we reframe bulk classifications by clarifying the underlying distribution of malignant phenotypes, reveal how microenvironmental heterogeneity is distributed in a transcriptional subtype-dependent manner, and systematically evaluate the *ex vivo* evolution and plasticity of malignant phenotypes.

RESULTS

A clinical pipeline for matched single-cell profiling and organoid model generation

We established a pipeline for collecting needle biopsies from patients with metastatic PDAC (n=23) to generate matched scRNA-seq profiles and organoid models **(Figure 1a, Extended Data Figure 1a, Supplementary Table 1)**. Most samples were obtained from metastatic lesions residing in the liver (19/23), and the majority (21/23) were analyzed by targeted DNA-sequencing which yielded the expected mutational pattern for this disease **(Extended Data Figure 1a)**.[4,12,13] After tissue dissociation, we used 10,000-20,000 viable cells for

111 scRNA-seq via Seq-Well, and the remainder were seeded for organoid culture (**Figure 1a**). This pipeline yielded
112 approximately 1,000 high-quality single cells per biopsy (n=23,042 total cells) and successful early-passage
113 organoid cultures from 70% (16/23) of patient tumor samples **(Extended Data Figure 1a, b**). Dimensionality
114 reduction and shared nearest neighbor (SNN) clustering of the biopsy cells revealed substantial heterogeneity
115 at the single-cell level **(Extended Data Figure 1c; Methods)**. The fractional representation from each biopsy
116 readily split the data into two groups, clusters of admixed cells from multiple patients and distinct patient-specific
117 clusters **(Extended Data Figure 1d)**. This pattern suggested both malignant and non-malignant cells within each
118 biopsy, with patient-specific clusters driven by specific copy number variations (CNVs). To confirm malignant cell
119 identity, we inferred transcriptome-wide CNVs from our single-cell data as previously described.[21,26] CNV
120 alteration scores separated putative malignant and non-malignant cells in each biopsy and demonstrated high
121 concordance with reference targeted DNA-seq (**Figure 1b, c; Extended Data Figure 2a, b).** CNV analysis
122 paired with manual inspection of expression patterns for known markers across single cells supported the
123 identification of malignant cells as well as 11 unique non-malignant cell types **(Extended Figure 1d-f; Figure**
124 **1d, e; Supplementary Table 2).** Thus, we established a robust workflow capable of recovering high quality
125 malignant (n=7,740) and non-malignant (n=15,302) populations from metastatic PDAC needle biopsies with low
126 neoplastic cellularity while also enabling simultaneous generation of matched organoid models.
127
128 **PDAC transcriptional subtypes exist on a continuum and include hybrid expression states**
129
130 We first applied principal component analysis (PCA) to examine major axes of transcriptional variation across
131 malignant cells from all biopsy samples. Notably, we failed to identify canonical driver mutations typically
132 observed in PDAC in one patient sample obtained prior to a pathologically confirmed clinical diagnosis,
133 PANFR0580 **(Extended Data Figure 1a)**; however, we detected a significant fraction of putative malignant cells
134 (n=662) in this biopsy **(Extended Data Figure 2b)**. Principal component 1 (PC1) separated PANFR0580 from
135 all other tumors in our cohort **(Extended Data Figure 3a, top)**. Genes with the strongest negative loading on
136 PC1 were indicative of a neuroendocrine phenotype (*TTR, CHGA, CHGB;* **Extended Data Figure 3a, bottom**)
137 and subsequent pathological evaluation confirmed that this sample was a pancreatic neuroendocrine tumor
138 (PanNET). To focus on transcriptional heterogeneity among PDAC samples, we removed the PanNET cells and
139 performed a new PCA on the remaining 7,078 malignant cells. Inspection of the genes driving the first 3 PCs
140 within PDAC cells revealed separation along previously characterized transcriptional phenotypes
141 (epithelial/mesenchymal transition (EMT)[37], PC1; basal/classical[5], PC2; cell cycle[16], PC3; **Extended Data Figure**
142 **3b,c**), confirming that the main axes of variation in our data align with established transcriptional subtypes.
143 Previous studies using bulk RNA-seq data have converged on two main tumor subtypes, basal and
144 classical.[3-5,11-13] Collapsing the malignant cells from each sample into a pseudo-bulk averaged transcriptome split
145 our cohort into 3 groups: those that exhibit predominately basal character (n=7), those with more classical
146 features (n=4), and those that are intermediate (n=10; **Extended Data Figure 3d).** Examination of basal and
147 classical phenotypes within each biopsy at single-cell resolution suggested that tumors are comprised of a

148 heterogenous mixture of states, likely driving the ambiguous classification of weakly polarized tumors when using
149 hierarchical clustering (**Extended Data Figure 3d,e**). We also observed a significant fraction of malignant cells
150 co-expressing both basal and classical phenotypes, hereafter referred to as "hybrid" cells (~13% of malignant
151 cells, **Figure 2a**, see **Methods**), suggesting that these phenotypes exist on a continuum rather than as discrete
152 states. Classification of each single cell as basal, classical, or hybrid revealed substantial heterogeneity across
153 individual tumors for these phenotypes (**Figure 2b**). These observations underscore the difficulty in assigning
154 intermediate tumors exclusively to basal or classical groups.[11] Thus, where discrete binning was necessary, we
155 employed a basal-classical "score difference" to stratify samples and preserve the polarization for each tumor
156 along this continuum (**Extended Data Figure 3f**).
157 We also used our single-cell data to examine signatures proposed by other bulk RNA sequencing studies
158 to clarify their inter-relationships. Pairwise correlation of all established signatures in malignant cells revealed
159 that many contribute overlapping information and reflect similar underlying biology **(Extended Data Figure 3g**).
160 We observed that cells with higher basal expression were also classified as squamous and quasimesenchymal,
161 while cells with classical signatures were correlated with the pancreatic progenitor subtype **(Extended Data**
162 **Figure 3g,h**).[3,4] By contrast, we did not observe evidence for expression of the immunogenic, ADEX, or exocrine-
163 like transcriptional signatures in malignant cells.[3,4] While the absence of these signatures might represent
164 differences between primary and metastatic disease, these bulk RNA profiles also likely incorporate signals from
165 non-malignant cells in the TME. In support of the latter hypothesis, we find evidence of immunogenic signature
166 expression originating from plasma cells as well as EMT signature expression from both malignant cells and
167 fibroblasts **(Extended Data Figure 3h**). These patterns underscore the need for single-cell resolution to dissect
168 malignant and non-malignant contributions to transcriptional signatures.
169 We next confirmed the presence of basal, classical, and hybrid cells using a novel subtype-specific single-
170 cell multiplexed immunofluorescence (mIF) panel in a cohort of primary resected PDAC (n=15 cases, 46,234
171 cells, **Methods; Figure 2c; Extended Data Figure 4a-c, Supplementary Tables 3, 4**). This orthogonal
172 approach confirmed the intratumoral heterogeneity observed in our scRNA-seq cohort and revealed that PDAC
173 transcriptional subtype diversity occurs on two levels: (i) "mixed" tumors comprised of discrete cells with differing
174 subtype identity, and (ii) hybrid cells which co-express basal and classical programs. These observations indicate
175 that PDAC transcriptional subtypes exist on a continuum, with mixed and hybrid phenotypes occurring even
176 within a single tumor gland (**Figure 2d**).
177
178 **Basal and classical cells exhibit subtype-specific expression programs**
179
180 We next leveraged our single-cell resolution to examine whether specific tumor cell gene expression programs
181 were correlated with either the basal or classical phenotypes. This correlation analysis across malignant cells
182 revealed 1,909 genes significantly associated with either basal or classical expression scores (**Figure 2e;**
183 **Supplementary Tables 5,6; Methods**). Inspection of these genes revealed basal cells are defined by more
184 mesenchymal features and co-express programs associated with transforming growth factor beta (*TGFB2*,

185 *SERPINE1*; TGF-β) signaling, interferon response (*IFI44L, ISG15*; IFN_{Resp}), WNT signaling (*WNT7B, FZD6,*
186 *EPHB2;* WNT), and cell cycle progression (*NASP, TOP2A)*.[37-40] Notably, these patterns are concordant with
187 larger bulk RNA-seq cohorts from primary and metastatic patient samples (**Figure 2e,f; Extended Data Figure**
188 **5a,b**).[12,13] While WNT ligands are included in organoid culture media and thought to be necessary to support
189 tumor cell growth *ex vivo*, we consistently detected only the WNT ligands *WNT7B* and *WNT10A*, which are
190 enriched in malignant basal cells *in vivo* (**Extended Data Figure 5c**).[33,41] Conversely, epithelial and pancreatic
191 progenitor transcriptional programs are enriched in classical PDAC cells (**Figure 2e,f; Extended Data Figure**
192 **5a**). Together, these expression patterns suggest a developmental continuum within PDAC tumors from higher
193 cycling (**Figure 2a**), de-differentiated basal cells to more committed classical epithelial pancreatic progenitors
194 that mirror phenotypes seen in the early developing pancreas.
195
196 **Transcriptional subtypes associate with distinct immune microenvironments**
197
198 Relatively little is known about the structure and composition of the metastatic microenvironment in PDAC, and,
199 more specifically, about how non-malignant heterogeneity associates with the basal to classical continuum. To
200 characterize the cell types in the metastatic niche, we analyzed the non-malignant cells (n=12,830) and refined
201 our broad cell-typing scheme from **Figure 1d** by further subdividing the T/NK cells, monocytes/macrophages,
202 and fibroblasts (**Figure 3a,b**). First, a closer analysis of the T/NK cell cluster revealed 5 cell types—*CD4*+ T,
203 *CD8*+ T, NKT, NK, and *CD16*+ (*FCGR3A*+) NK cells—each expressing the corresponding established markers
204 (**Extended Data Figure 6a-d**). Similarly, an unsupervised examination within the monocyte/macrophage
205 compartment revealed a tumor associated macrophage (TAM) continuum similar to one recently described in
206 colorectal cancer.[42,43] The first two PCs readily identified 3 TAM subsets: "monocyte-like" *FCN1*+, *C1QC+,* and
207 *SPP1*+ macrophages (**Extended Data Figure 6h**). *FCN1*+ "monocyte-like" cells expressed high levels of *IL1B*
208 and *CCR2* and shared some features with CD14+ blood monocytes (*CD300E, S100A8*).[42] *C1QC*+ TAMs
209 resembled a phagocytic phenotype (*CD163*, *MERTK)*, but also demonstrated preferentially high expression of
210 antigen presentation genes (*HLA-DRB1*, *CD74*) and genes described in anti-inflammatory macrophage subsets
211 (*FOLR2*, *CD209*, *AXL, CSF1R*). Conversely, *SPP1*+ TAMs expressed gene programs associated with
212 angiogenesis (*SPP1, FLT1*) and inflammatory response (*CCL2, CCL7, CSF1, CLEC5A*). A fourth subset was
213 positioned as intermediate between these three phenotypes and likely represents a population of actively
214 transitioning/differentiating TAMs (Trans TAM; **Extended Data Figure 6h-j**).[42] Finally, although several scRNA-
215 seq studies in primary resected PDAC have focused on fibroblast phenotypes, we observed few fibroblasts per
216 tumor (**Methods**), with the outliers coming from sampling sites other than the liver (PANFR0637 and
217 PANFR0635) or from a different disease etiology (PANFR0580, PanNET; **Extended Data Figure 6k**).[27-29] Still,
218 in the fibroblasts we recovered we noted evidence of previously identified subtypes including myofibroblastic and
219 inflammatory cancer-associated fibroblasts (myCAFs and iCAFs, respectively) in this metastatic setting
220 (**Extended Data Figure 6l,m**). Taken together, we identified 18 unique cell types/states in the PDAC metastatic
221 microenvironment (**Figure 3a**).

222 We next determined whether the 18 non-malignant cell types/states were represented evenly across the
223 malignant basal-to-classical transcriptional continuum described in **Figure 2**. For this analysis, we computed two
224 quantities: 1) the fractional representation of each non-malignant cell type per biopsy and 2) the correlation of
225 each non-malignant cell type's capture frequency to the average "score difference" (basal/classical polarization;
226 **Extended Data Figure 3f**) derived from the malignant cells in the same biopsy. Cross-correlation of each cell
227 type's fractional representation revealed two distinct patterns that largely diverged by malignant transcriptional
228 subtype association (**Figure 3c**). Overall, cell types traditionally believed to facilitate a more immune-responsive
229 microenvironment were frequently captured together. For example, DC subsets, NK, B, *CD4+* T and
230 inflammatory *FCN1+* TAMs derive from shared microenvironments (hereafter "immune-infiltrated") and tend to
231 associate with more classical tumors (**Figure 3c**). Activated, mature NK cells (*FCGR3A+* NK) were captured
232 most often from these immune-infiltrated biopsies and showed a strong correlation with classical tumors (**Figure**
233 **3d**)**.** Interestingly, *FCGR3A+* NK cells showed the highest expression of cytotoxic markers in our metastatic
234 dataset, even compared to *CD8+* T cells (**Extended Data Figure 6e,f**)**.** Examination of the T cell compartment
235 revealed that *CD4+* T cells were captured more frequently in classical tumors (**Figure 3c,e**), whereas *CD8+* T
236 cells were captured less frequently in immune-infiltrated biopsies and associated more often with an increased
237 basal score. PCA within the *CD8+* compartment revealed a progenitor (*TCF7*, *IL7R*) to differentiated/exhausted
238 (*HAVCR2*, *ENTPD1)* continuum previously associated with differential outcomes to immune checkpoint blockade
239 (**Extended Data Figure 6g**).[20,44] Scoring each *CD8+* T cell over this axis, we observed a progenitor-restricted
240 distribution in most tumors, with only two outlier basal tumors skewing toward more differentiated/exhausted
241 phenotypes (**Figure 3e**). In sum, these findings indicate that much of the cytotoxic activity in the metastatic niche
242 may originate from the innate immune system by way of activated NK cells in the microenvironment of classical
243 tumors.

244 Along with differences in lymphocyte content, the myeloid compartment, specifically TAM phenotypes,
245 showed strong subtype-specific associations. First, we noted selective skewing for the types of TAMs originating
246 from basal versus classical tumors (**Extended data figure 6i**, $P < 2.2\text{x}10^{-16}$, Chi-squared test; **Figure 3c**, *C1QC+*
247 TAM, $r = -0.59$, basal association and *SPP1+* TAM, $r = 0.52$, classical association). Indeed, when examining the
248 monocyte-like to macrophage distribution for TAMs from individual liver biopsies, the most basal-polarized
249 tumors were associated with more macrophage-committed phenotypes (**Figure 3f**). Moreover, by scoring each
250 macrophage using TAM subtype-specific signatures and visualizing them with respect to the likely differentiation
251 trajectory inferred from recent studies (**Supplementary Tables 7-9; Methods**)[42], we confirm a preferential
252 association between *C1QC+* TAMs and basal tumors and, conversely, an enrichment for the inflammatory
253 *FCN1+* monocyte-like and *SPP1+* TAM subsets in tumors with intermediate and classical phenotypes (**Figure**
254 **3g**). In addition to demonstrating that classical tumors are relatively more immune infiltrated, this analysis also
255 identifies distinct microenvironmental phenotypes that co-vary with each PDAC transcriptional subtype and
256 suggests opportunities to direct microenvironmental therapies in a subtype-specific manner.
257

258 **Differential microenvironmental signaling shapes subtype-specific metastatic niches**

Given the striking compositional differences we observed in the immune microenvironment across the basal to classical axis, we next searched for tumor-secreted factors that might influence the structure of the local metastatic niche. Specifically, we analyzed subtype-specific expression patterns for genes detected in malignant cells that were annotated as secreted factors (cytokines, chemokines, growth factors by Gene Ontology; n=218 genes). This analysis nominated 57 basal (orange) and 23 classical-associated (blue) secreted factors (**Figure 4a**). Gene set enrichment analysis (GSEA) demonstrated that basal tumors were enriched for genes associated with growth factor secretion, while classical tumors were enriched for cytokine/chemokine signaling (**Figure 4b**). We observed expression of multiple TGF ligands secreted by basal tumors, consistent with the association of increased TGF-β signaling in basal tumors (**Figure 2f**) and local immune suppression/exclusion. Conversely, several chemokines (*CXCL5*, *CXCL3*) were enriched in classical tumors in agreement with their overall higher degree of immune infiltration and higher fraction of endothelial cells (**Figure 3c**). As such, classical tumors expressed higher levels of *CXCL5* which plays a documented role in enhancing tumor-supportive angiogenesis.[45,46] Consistent with this finding, we observed a strong positive correlation between high average malignant cell expression of *CXCL5* and the fraction of endothelial cells recovered (**Figure 4c**). In basal tumors, we noted increased expression of the ligands *CSF1* and *IL34* (**Figure 4a**) and concomitant expression of their receptor, *CSF1R,* in the basal-associated *C1QC*+ TAMs (**Figure 3c,g; 4d**). Per-tumor analysis revealed a continuum of *C1QC*+ TAM distribution within basal tumors that correlated with high *CSF1R* expression (**Figure 4e,f top**). Malignant cells with strong EMT features (PANFR0545, PANFR0593) expressed the highest levels of *CSF1* and *IL34,* consistent with a role for tumor cells in shaping their local macrophage phenotypes (**Figure 4f bottom**). To extend this finding in larger cohorts, we analyzed bulk RNA-sequencing of primary and metastatic PDAC tumors for markers of transcriptional subtype, TAM, and tumor secretion phenotypes.[12,13] Consistent with our single-cell observations, macrophage markers and the ligands *CSF1* and *IL34* were associated with basal but not classical markers in these samples (n=198, **Figure 4g**). Together, these data provide evidence that subtype-specific intercellular crosstalk shapes distinct niches in the metastatic microenvironment.

Genotype and phenotype evolution of matched patient-derived organoid models

Our observations indicate that basal and classical phenotypes exist along a continuum and exhibit distinct patterns of reciprocal interaction with their local microenvironments. To examine how tumor cell phenotypes adapt and evolve in *ex vivo* microenvironments, we utilized the matched organoid models generated from our metastatic biopsy cohort (**Methods**). For most models, we obtained scRNA-seq samples at the earliest passage possible, typically passage 2 (P2), and again at a later passage (**Extended Data Figure 7a,b**). Notably, only 33% of models derived from basal tumors propagated beyond passage 2, whereas 60% of models derived from classical tumors established long-term cultures (**Extended Data Figure 7b**). Globally, unbiased analysis of malignant biopsy (7,078 cells) and organoid cells (n=14 models, 24,789 cells) revealed that biopsy cells clustered separately from their matched organoid counterparts (**Extended Data Figure 7c,d**). Only two clusters were

296 admixed by donor and originated from early passage organoids (clusters 4 and 32; **Extended Data Figure 7c**).
297 These clusters were defined by expression patterns consistent with fibroblasts (cluster 32) and poorly
298 differentiated epithelial cells (cluster 4), and were not seen in samples from later passages (**Extended Data**
299 **Figure 7e,f**).

300 Comparison of transcriptional phenotypes revealed a striking selection against the basal subtype in
301 organoid culture despite it being the higher cycling subset *in vivo* (**Figure 2a** & **Figure 5a**). To understand the
302 relative contribution of genotype versus phenotype to this bottleneck, we computed the average single-cell
303 genotype (CNV) and phenotype (basal versus classical) correlation distance (*d*) between each biopsy and its
304 matched early passage organoid (**Figure 5b**; **Methods**). Six models, all classical, did not significantly deviate
305 along either the CNV or transcriptional axes outside the expected distance for highly similar samples (intra-
306 biopsy *d* across the cohort; dotted line, $P < 0.05$ for both metrics). Another group, largely basal (right of x-axis
307 dotted line), deviated significantly from their original biopsies along the transcriptional but not the CNV axis.
308 Finally, two basal models (PANFR0545 and PANFR0552) exhibited the strongest deviation from their parent
309 biopsies along both axes (**Figure 5b**, upper right). This analysis demonstrated that early passage organoid
310 models largely maintain genomic features observed in parental tumor tissue, but over half of these models, and
311 in particular models derived from basal tumors, were significantly divergent in phenotype compared to their
312 matched tissue-of-origin.

313 We next examined the subclonal hierarchies within each biopsy-organoid pair. This single-cell
314 comparative analysis identified 4 broad patterns of drift/selection. Pattern 1 consisted of tumors (n=4) where the
315 organoids failed to grow beyond P2; the majority of these were derived from basal tumors (3/4 models), and
316 included the two models (PANFR0545 and PANFR0552) that deviated the most genotypically and phenotypically
317 from their parent biopsies (**Extended Data Figure 8a**). Of the models that propagated beyond P2, Pattern 2
318 models (n=3) showed evidence of selective outgrowth wherein models derived from basal tumors enriched rare
319 subclones tied to more classical or less basal phenotypes (**Extended Data Figure 8b**). In contrast, models within
320 Pattern 3 (n=5) were typified by neutral outgrowth (no overt selection) where the dominant clone(s) in the biopsy
321 grew out in the organoid (**Extended Data Figure 8c**). These models expressed predominantly classical
322 phenotypes and had the least overall deviation from their parent biopsies (**Figure 5b**); none of the models derived
323 from basal tumors displayed this pattern. Finally, Pattern 4 comprised one basal biopsy-organoid pair
324 (PANFR0575) that demonstrated phenotypic plasticity with nearly identical CNVs but a divergent transcriptional
325 phenotype in organoid culture (**Extended Data Figure 8d**). These data illustrate the dramatic adaptation that
326 organoid models undergo *ex vivo* via transcriptional and clonal selection at early passages, especially when
327 derived from basal tumors.

328 When we serially sampled and assessed organoid phenotypes over time, we observed that each model
329 assumed a more classical phenotype regardless of its parent tumor's transcriptional identity, and only the Pattern
330 4 plastic model, PANFR0575, re-acquired its basal phenotype at a later passage (**Figure 5c, d**). Linked genotype
331 and phenotype assessment from iterative passages provided evidence for significant evolution along both CNV
332 and transcriptional axes over time in culture (**Figure 5e**; **Extended Data Figure 8e**). After identifying CNV-

333 defined subclones in the parental biopsy and its associated serial organoid samples (**Figure 5e**; clones A-F;
334 **Methods**), we related cells that were similar in genotype (e.g., all cells within clone A) to their corresponding
335 transcriptional phenotype. In sample PANFR0575 (**Figure 5e**), we observed examples of transcriptional plasticity
336 at early passages within clone A. Cells derived from the parental biopsy were basal, but all other cells in this
337 subclone derived from organoids had classical phenotypes. Interestingly, with successive passaging, several
338 subclones emerged with hybrid and basal phenotypes (clones D and E). While model PANFR0575 is a unique
339 case, it highlights the various ways organoids can evolve in culture, including via transcriptional plasticity (clone
340 A) and the late emergence of rare subclones (clones D and E). In contrast, PANFR0489R was initially basal, but
341 we observed clonal selection and phenotypic drift toward classical states, as seen in most other models (**Figure**
342 **5c; Extended Data Figure 8e**). Together, these findings demonstrate that multiple mechanisms underlie
343 organoid evolution and divergence from the parental tumor, highlight that transcriptional variation is a key
344 contributor to these differences, and emphasize the importance of deep molecular characterization of patient-
345 derived models prior to functional application.
346
347 **Alterations to the *ex vivo* culture environment revive the basal state in organoids**
348
349 Having demonstrated that distinct expression states as well as the local microenvironment co-vary across the
350 basal to classical axis, we reasoned that different conditions may be needed to preserve basal versus classical
351 transcriptional heterogeneity. Comparing bulk RNA expression data from patient tumors (n=219), organoids
352 (n=44) and cell lines (n=49, CCLE) provided evidence that culture conditions can profoundly influence
353 transcriptional state.[12,13,47] Indeed, most organoid models recapitulate the classical phenotype while cell lines
354 mirror basal expression patterns (**Figure 6a**). To isolate the effects of extracellular matrix dimensionality from
355 media formulation, we cultured established 3-dimensional (3D) PDAC organoid models (n=4) as 2-dimensional
356 (2D) cell lines on tissue culture plastic in the same organoid media and noted that this had little effect on
357 transcriptional subtype across the models tested (**Extended Data Figure 9a**). We then hypothesized that
358 multiple components within standard organoid media[10,33], including WNT3A, R-SPONDIN-1, FGF10, and TGF
359 and BMP pathway inhibitors such as NOGGIN and A-8301, may drive tumor cells toward more classical
360 phenotypes in organoid culture. When established organoid models (n = 4) were grown for 1 week in reduced
361 medium without any additives ("stripped" media, containing only Glutamax, anti-microbials, HEPES buffer, and
362 Advanced DMEM/F12 media; **Figure 6b**; see **Methods**), we observed a significant increase in basal gene
363 expression across single cells (**Figure 6c**; $P < 0.0001$), as well as coordinated sample-level shifts to a more
364 basal phenotype in each model, in some cases returning to levels observed in the parental biopsy (**Figure 6d**).
365 This shift was less pronounced in the model derived from the most classical tumor (PANFR0489, pink outline;
366 **Figure 6d; Extended Data Figure 9b**). Although there was an appreciable effect on the fraction of cycling cells
367 in the stripped media (**Figure 6e,** far right), the organoids continued to grow under these conditions (**Extended**
368 **Data Figure 9c**). These responses were unlikely to be driven by acute selection since the CNV profiles between
369 the conditions remained stable within this timeframe (**Figure 6e**). Collectively, these observations provide

370 evidence for significant *ex vivo* tumor cell plasticity in response to microenvironmental cues and suggest that
371 organoid and cell line culture conditions can be further optimized to recapitulate clinically relevant *in vivo* tumor
372 cell states.
373
374 **DISCUSSION**
375
376 This study demonstrates the precision afforded by scRNA-seq for categorizing and phenotyping relevant
377 malignant and non-malignant cell populations in metastatic PDAC. In the malignant compartment, we confirmed
378 the basal-like and classical transcriptional subtyping framework; however, we found that these subtypes exist on
379 a continuum and include a newly identified "hybrid" phenotype. We show at both the RNA and protein level that
380 most tumors are comprised of all three phenotypes and exhibit notable intratumoral heterogeneity in two ways:
381 (i) basal or classical phenotypes in discrete cells but co-occurring in the same tumor, consistent with a recent
382 report[11], and (ii) co-expression of both states in the same single cell (hybrid cells). Importantly, the identification
383 of these hybrid cells in human tumor biopsies suggests that interconversion may be possible between the
384 classical and basal subtypes. Basal tumor cells exhibit mesenchymal and stem-like features, including TGF-β
385 pathway activation and evidence for WNT signaling. In this tissue context, WNT signaling is likely mediated
386 through the expression of WNT7B and/or WNT10A as these were the only ligands consistently expressed in the
387 cells we captured. WNT7B is a key developmental signal for pancreatic progenitor proliferation, normal
388 morphogenesis, and mesenchymal expansion, and its expression evokes the possibility that basal tumor cells
389 may share similarities with a discrete subset of early pancreatic progenitors.[48] Several studies have suggested
390 a role for WNT signaling in supporting proliferation and cell state specification in PDAC models, but more
391 experimentation is needed to clarify its impact.[41,49,50] Given that PDAC transcriptional subtypes have been
392 associated with differential response to chemotherapy[3,7-10], these new insights into PDAC subtype heterogeneity
393 and their associated biologies have important implications for understanding therapy response in clinical trials.
394 In coordination with malignant cell phenotypes, non-malignant cells establish subtype-specific local
395 immune microenvironments within the PDAC metastatic niche. Our observations support a model wherein
396 classical tumors exhibit greater chemokine signaling and concomitant immune infiltration. Although this has been
397 hypothesized previously[2,4,6], the specific cell types and their phenotypes have remained elusive. Our single-cell
398 dataset clarifies this relationship and identifies a classical TME enriched for endothelial cells and specific myeloid
399 and lymphoid cell types. In the lymphoid compartment, surprisingly, we observed cytotoxic signaling that
400 originates primarily from activated NK cells, suggesting a dominant role for innate immune function in the
401 classical metastatic niche. Conversely, the basal microenvironment is optimally tuned for immune
402 suppression/evasion, which may contribute to the overall lower survival seen in this subtype. The relative paucity
403 of CD4+ T cells found in basal tumors suggests exclusion, possibly driven by the higher levels of TGF gene
404 expression in basal contexts. Somewhat unexpectedly, we found evidence of terminally exhausted *CD8+* T cells
405 in only two basal tumors, and, in most cases, both basal and classical tumors exist in a *CD8+* T-cell progenitor-
406 restricted state. Basal tumor cells exhibited higher levels of IFN response gene expression compared with

classical tumors, suggesting exposure to, and potential tolerance of, the presence of activated T cells.[39,40] Basal tumor cells also shape the myeloid compartment by secreting *CSF1* and *IL34*, with concomitant microenvironmental increase in *C1QC*+ TAM populations that skew towards a tumor supportive, anti-inflammatory phenotype. Notably, even within basal tumors, those with the most mesenchymal characteristics possessed the most potent immune-evasive phenotypes, suggestive of additional layers of variation even within the basal subtype.

Comparison of matched biopsies and organoids revealed relative preservation of genomic features in most organoid models, as has previously been demonstrated[10,35], but significant deviation in basal/classical transcriptional state. While classical phenotypes were relatively better preserved, we note strong selection against the basal state under standard organoid media conditions. Serial sampling of organoid models across successive passages demonstrated both phenotypic drift and subclonal outgrowth, such that the dominant clones in some later passage models were only present at low frequencies in the parent tumors. Despite the bias toward classical phenotypes in organoid culture, the rare emergence of basal clones at late passages (PANFR0575; **Figure 5e**) suggests that genotype, in addition to microenvironment, may influence transcriptional plasticity. However, resolution is an important limitation of our clonal tracing, as we cannot comment directly on variation/selection for single mutations. While our findings may explain some of the limitations observed when using PDAC organoid models to predict clinical responses[10,35], they also highlight the significant phenotypic plasticity and adaptability of PDAC cells and, moreover, the utility of primary tissue and matched model comparisons for understanding these features of tumor biology. Interestingly, established PDAC cell lines exhibit predominantly basal phenotypes, but changing matrix dimensionality (2D versus 3D culture) alone did not alter malignant organoid transcriptional phenotypes along the basal-classical axis, implying that variation in adhesive context may affect some but not all biologic behaviors. Encouragingly, the basal phenotype could be recovered by removing exogenous factors from the standard culture media, setting the stage for further optimization of these conditions to adequately support intratumoral heterogeneity and growth.[51] These results highlight that *ex vivo* model growth may not necessarily equate to model fidelity and suggest that experimental conditions, heterogeneity, and plasticity all influence the phenotype of patient-derived organoids.[32,52]

In sum, we show how scRNA-seq can be employed to clarify the structure of the PDAC metastatic niche and uncover formerly unappreciated relationships between tumor transcriptional phenotype and the local immune microenvironment. Although traditionally thought of as a uniformly “immune-cold” tumor, our findings highlight that the immune microenvironment in PDAC harbors a layer of unappreciated complexity closely linked to tumor cell transcriptional subtype that may provide new avenues for therapeutic targeting. Specifically, TAM-directed therapies, such as anti-CSF1R antibodies, could be employed to selectively target transcriptional-subtype-associated populations.[42,53-55] However, while basal tumors associate with a potentially sensitive CSF1R-expressing population (*C1QC*+ TAM), classical tumors harbor TAMs that are resistant to such therapies (*SPP1*+ TAM).[42] Thus, just as we consider combinations to target malignant states, the TME will also likely require tailored combination therapies. These findings provide rationale for future clinical trials to employ high-resolution phenotyping of malignant and non-malignant cells to stratify patients and track tumor evolution in

444 response to therapy. While organoid platforms represent a transformative technology to develop patient-specific
445 tumor models, we demonstrate that some organoid models show a high degree of plasticity and that both their
446 genotype and transcriptional phenotype must be understood to enable their optimal use in personalized
447 medicine. Finally, we provide a framework for relating malignant cells, the TME, and patient-derived model
448 systems that may be applicable in other tumor types with clinically relevant transcriptional variation across the
449 malignant and microenvironmental landscape.

450 **REFERENCES**

451

452 1 Hyman, D. M., Taylor, B. S. & Baselga, J. Implementing Genome-Driven Oncology. *Cell* **168**, 584-599,
453 doi:10.1016/j.cell.2016.12.015 (2017).

454 2 Collisson, E. A., Bailey, P., Chang, D. K. & Biankin, A. V. Molecular subtypes of pancreatic cancer. *Nat*
455 *Rev Gastroenterol Hepatol* **16**, 207-220, doi:10.1038/s41575-019-0109-y (2019).

456 3 Collisson, E. A. *et al.* Subtypes of pancreatic ductal adenocarcinoma and their differing responses to
457 therapy. *Nat Med* **17**, 500-503, doi:10.1038/nm.2344 (2011).

458 4 Bailey, P. *et al.* Genomic analyses identify molecular subtypes of pancreatic cancer. *Nature* **531**, 47-52,
459 doi:10.1038/nature16965 (2016).

460 5 Moffitt, R. A. *et al.* Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of
461 pancreatic ductal adenocarcinoma. *Nat Genet* **47**, 1168-1178, doi:10.1038/ng.3398 (2015).

462 6 Puleo, F. *et al.* Stratification of Pancreatic Ductal Adenocarcinomas Based on Tumor and
463 Microenvironment Features. *Gastroenterology* **155**, 1999-2013 e1993, doi:10.1053/j.gastro.2018.08.033
464 (2018).

465 7 Porter, R. L. *et al.* Epithelial to mesenchymal plasticity and differential response to therapies in pancreatic
466 ductal adenocarcinoma. *Proc Natl Acad Sci U S A*, doi:10.1073/pnas.1914915116 (2019).

467 8 Aung, K. L. *et al.* Genomics-Driven Precision Medicine for Advanced Pancreatic Cancer: Early Results
468 from the COMPASS Trial. *Clin Cancer Res* **24**, 1344-1354, doi:10.1158/1078-0432.CCR-17-2994 (2018).

469 9 O'Kane, G. M. *et al.* GATA6 Expression Distinguishes Classical and Basal-like Subtypes in Advanced
470 Pancreatic Cancer. *Clin Cancer Res*, doi:10.1158/1078-0432.CCR-19-3724 (2020).

471 10 Tiriac, H. *et al.* Organoid Profiling Identifies Common Responders to Chemotherapy in Pancreatic
472 Cancer. *Cancer Discov* **8**, 1112-1129, doi:10.1158/2159-8290.CD-18-0349 (2018).

473 11 Chan-Seng-Yue, M. *et al.* Transcription phenotypes of pancreatic cancer are driven by genomic events
474 during tumor evolution. *Nat Genet* **52**, 231-240, doi:10.1038/s41588-019-0566-9 (2020).

475 12 Cancer Genome Atlas Research Network. Electronic address, a. a. d. h. e. & Cancer Genome Atlas
476 Research, N. Integrated Genomic Characterization of Pancreatic Ductal Adenocarcinoma. *Cancer Cell*
477 **32**, 185-203 e113, doi:10.1016/j.ccell.2017.07.007 (2017).

478 13 Aguirre, A. J. *et al.* Real-time Genomic Characterization of Advanced Pancreatic Cancer to Enable
479 Precision Medicine. *Cancer Discov* **8**, 1096-1111, doi:10.1158/2159-8290.CD-18-0275 (2018).

480 14 Connor, A. A. *et al.* Integration of Genomic and Transcriptional Features in Pancreatic Cancer Reveals
481 Increased Cell Cycle Progression in Metastases. *Cancer Cell* **35**, 267-282 e267,
482 doi:10.1016/j.ccell.2018.12.010 (2019).

483 15 Venteicher, A. S. *et al.* Decoupling genetics, lineages, and microenvironment in IDH-mutant gliomas by
484 single-cell RNA-seq. *Science* **355**, doi:10.1126/science.aai8478 (2017).

485 16 Tirosh, I. *et al.* Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq.
486 *Science* **352**, 189-196, doi:10.1126/science.aad0501 (2016).

487 17 Puram, S. V. *et al.* Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in
488 Head and Neck Cancer. *Cell* **171**, 1611-1624 e1624, doi:10.1016/j.cell.2017.10.044 (2017).

489 18 van Galen, P. *et al.* Single-Cell RNA-Seq Reveals AML Hierarchies Relevant to Disease Progression and
490 Immunity. *Cell* **176**, 1265-1281 e1224, doi:10.1016/j.cell.2019.01.031 (2019).

491 19 Kim, C. *et al.* Chemoresistance Evolution in Triple-Negative Breast Cancer Delineated by Single-Cell
492 Sequencing. *Cell* **173**, 879-893 e813, doi:10.1016/j.cell.2018.03.041 (2018).

493 20 Sade-Feldman, M. *et al.* Defining T Cell States Associated with Response to Checkpoint Immunotherapy
494 in Melanoma. *Cell* **176**, 404, doi:10.1016/j.cell.2018.12.034 (2019).

21 Patel, A. P. *et al.* Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* **344**, 1396-1401, doi:10.1126/science.1254257 (2014).

22 Filbin, M. G. *et al.* Developmental and oncogenic programs in H3K27M gliomas dissected by single-cell RNA-seq. *Science* **360**, 331-335, doi:10.1126/science.aao4750 (2018).

23 Hovestadt, V. *et al.* Resolving medulloblastoma cellular architecture by single-cell genomics. *Nature* **572**, 74-79, doi:10.1038/s41586-019-1434-6 (2019).

24 Neftel, C. *et al.* An Integrative Model of Cellular States, Plasticity, and Genetics for Glioblastoma. *Cell* **178**, 835-849 e821, doi:10.1016/j.cell.2019.06.024 (2019).

25 Tirosh, I. & Suva, M. L. Dissecting human gliomas by single-cell RNA sequencing. *Neuro Oncol* **20**, 37-43, doi:10.1093/neuonc/nox126 (2018).

26 Tirosh, I. *et al.* Single-cell RNA-seq supports a developmental hierarchy in human oligodendroglioma. *Nature* **539**, 309-313, doi:10.1038/nature20123 (2016).

27 Elyada, E. *et al.* Cross-Species Single-Cell Analysis of Pancreatic Ductal Adenocarcinoma Reveals Antigen-Presenting Cancer-Associated Fibroblasts. *Cancer Discov* **9**, 1102-1123, doi:10.1158/2159-8290.CD-19-0094 (2019).

28 Bernard, V. *et al.* Single-Cell Transcriptomics of Pancreatic Cancer Precursors Demonstrates Epithelial and Microenvironmental Heterogeneity as an Early Event in Neoplastic Progression. *Clin Cancer Res* **25**, 2194-2205, doi:10.1158/1078-0432.CCR-18-1955 (2019).

29 Ligorio, M. *et al.* Stromal Microenvironment Shapes the Intratumoral Architecture of Pancreatic Cancer. *Cell* **178**, 160-175 e127, doi:10.1016/j.cell.2019.05.012 (2019).

30 Siegel, R. L., Miller, K. D. & Jemal, A. Cancer statistics, 2020. *CA Cancer J Clin* **70**, 7-30, doi:10.3322/caac.21590 (2020).

31 Drost, J. & Clevers, H. Organoids in cancer research. *Nat Rev Cancer* **18**, 407-418, doi:10.1038/s41568-018-0007-6 (2018).

32 Ben-David, U. *et al.* Genetic and transcriptional evolution alters cancer cell line drug response. *Nature* **560**, 325-330, doi:10.1038/s41586-018-0409-3 (2018).

33 Boj, S. F. *et al.* Organoid models of human and mouse ductal pancreatic cancer. *Cell* **160**, 324-338, doi:10.1016/j.cell.2014.12.021 (2015).

34 Huang, L. *et al.* Ductal pancreatic cancer modeling and drug screening using human pluripotent stem cell- and patient-derived tumor organoids. *Nat Med* **21**, 1364-1371, doi:10.1038/nm.3973 (2015).

35 Driehuis, E. *et al.* Pancreatic cancer organoids recapitulate disease and allow personalized drug screening. *Proc Natl Acad Sci U S A*, doi:10.1073/pnas.1911273116 (2019).

36 Gierahn, T. M. *et al.* Seq-Well: portable, low-cost RNA sequencing of single cells at high throughput. *Nat Methods* **14**, 395-398, doi:10.1038/nmeth.4179 (2017).

37 Groger, C. J., Grubinger, M., Waldhor, T., Vierlinger, K. & Mikulits, W. Meta-analysis of gene expression signatures defining the epithelial to mesenchymal transition during cancer progression. *PLoS One* **7**, e51136, doi:10.1371/journal.pone.0051136 (2012).

38 Kim, J. H., Park, S. Y., Jun, Y., Kim, J. Y. & Nam, J. S. Roles of Wnt Target Genes in the Journey of Cancer Stem Cells. *Int J Mol Sci* **18**, doi:10.3390/ijms18081604 (2017).

39 Benci, J. L. *et al.* Tumor Interferon Signaling Regulates a Multigenic Resistance Program to Immune Checkpoint Blockade. *Cell* **167**, 1540-1554 e1512, doi:10.1016/j.cell.2016.11.022 (2016).

40 Benci, J. L. *et al.* Opposing Functions of Interferon Coordinate Adaptive and Innate Immune Responses to Cancer Immune Checkpoint Blockade. *Cell* **178**, 933-948 e914, doi:10.1016/j.cell.2019.07.019 (2019).

41 Seino, T. *et al.* Human Pancreatic Tumor Organoids Reveal Loss of Stem Cell Niche Factor Dependence during Disease Progression. *Cell Stem Cell* **22**, 454-467 e456, doi:10.1016/j.stem.2017.12.009 (2018).

42 Zhang, L. *et al.* Single-Cell Analyses Inform Mechanisms of Myeloid-Targeted Therapies in Colon Cancer. *Cell* **181**, 442-459 e429, doi:10.1016/j.cell.2020.03.048 (2020).

43 Zilionis, R. *et al.* Single-Cell Transcriptomics of Human and Mouse Lung Cancers Reveals Conserved Myeloid Populations across Individuals and Species. *Immunity* **50**, 1317-1334 e1310, doi:10.1016/j.immuni.2019.03.009 (2019).

44 Miller, B. C. *et al.* Subsets of exhausted CD8(+) T cells differentially mediate tumor control and respond to checkpoint blockade. *Nat Immunol* **20**, 326-336, doi:10.1038/s41590-019-0312-6 (2019).

45 Chen, C. *et al.* CXCL5 induces tumor angiogenesis via enhancing the expression of FOXD1 mediated by the AKT/NF-kappaB pathway in colorectal cancer. *Cell Death Dis* **10**, 178, doi:10.1038/s41419-019-1431-6 (2019).

46 Li, A. *et al.* Overexpression of CXCL5 is associated with poor survival in patients with pancreatic cancer. *Am J Pathol* **178**, 1340-1349, doi:10.1016/j.ajpath.2010.11.058 (2011).

47 Ghandi, M. *et al.* Next-generation characterization of the Cancer Cell Line Encyclopedia. *Nature* **569**, 503-508, doi:10.1038/s41586-019-1186-3 (2019).

48 Afelik, S., Pool, B., Schmerr, M., Penton, C. & Jensen, J. Wnt7b is required for epithelial progenitor growth and operates during epithelial-to-mesenchymal signaling in pancreatic development. *Dev Biol* **399**, 204-217, doi:10.1016/j.ydbio.2014.12.031 (2015).

49 Arensman, M. D. *et al.* WNT7B mediates autocrine Wnt/beta-catenin signaling and anchorage-independent growth in pancreatic adenocarcinoma. *Oncogene* **33**, 899-908, doi:10.1038/onc.2013.23 (2014).

50 Steinhart, Z. *et al.* Genome-wide CRISPR screens reveal a Wnt-FZD5 signaling circuit as a druggable vulnerability of RNF43-mutant pancreatic tumors. *Nat Med* **23**, 60-68, doi:10.1038/nm.4219 (2017).

51 Mead, B. E. *et al.* Harnessing single-cell genomics to improve the physiological fidelity of organoid-derived cell types. *BMC Biol* **16**, 62, doi:10.1186/s12915-018-0527-2 (2018).

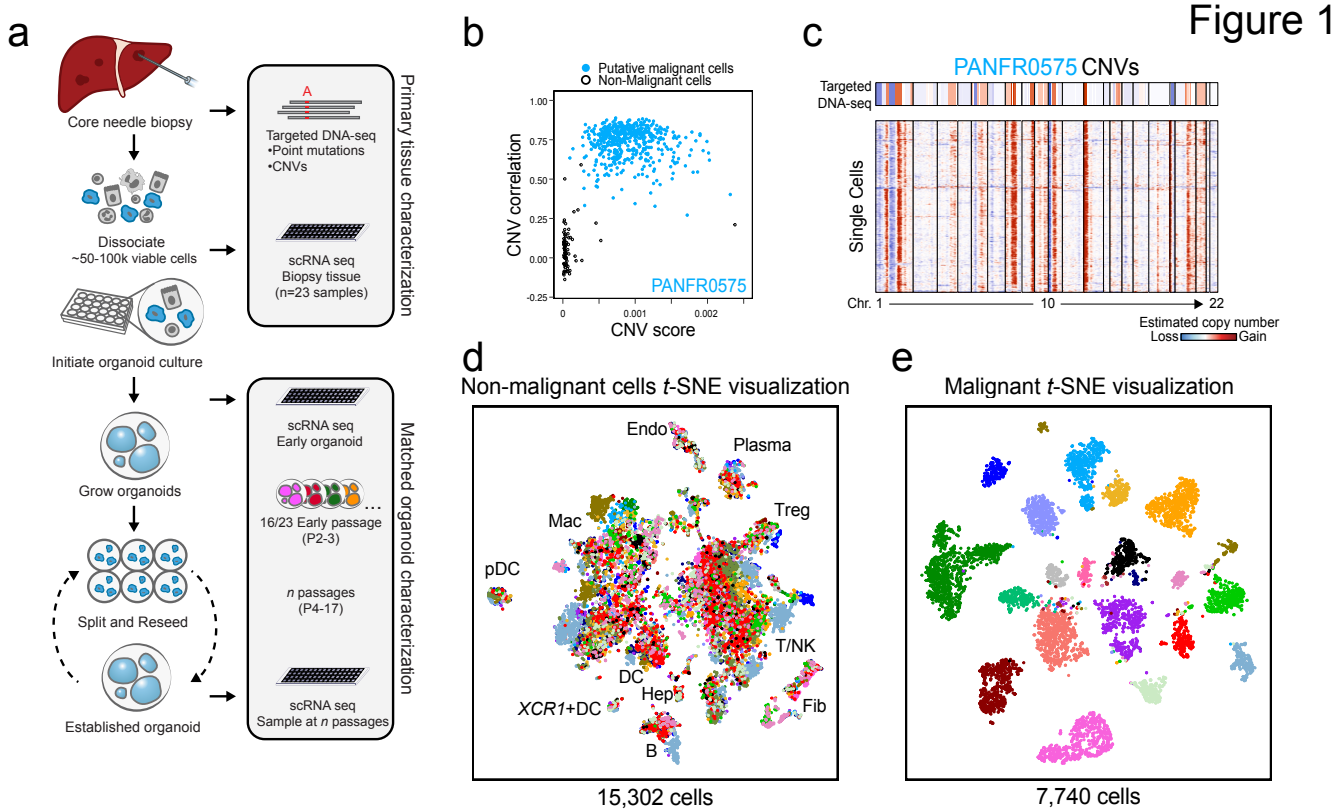
52 Ben-David, U., Beroukhim, R. & Golub, T. R. Genomic evolution of cancer models: perils and opportunities. *Nat Rev Cancer* **19**, 97-109, doi:10.1038/s41568-018-0095-3 (2019).

53 Balachandran, V. P., Beatty, G. L. & Dougan, S. K. Broadening the Impact of Immunotherapy to Pancreatic Cancer: Challenges and Opportunities. *Gastroenterology* **156**, 2056-2072, doi:10.1053/j.gastro.2018.12.038 (2019).

54 Zhu, Y. *et al.* CSF1/CSF1R blockade reprograms tumor-infiltrating macrophages and improves response to T-cell checkpoint immunotherapy in pancreatic cancer models. *Cancer Res* **74**, 5057-5069, doi:10.1158/0008-5472.CAN-13-3723 (2014).

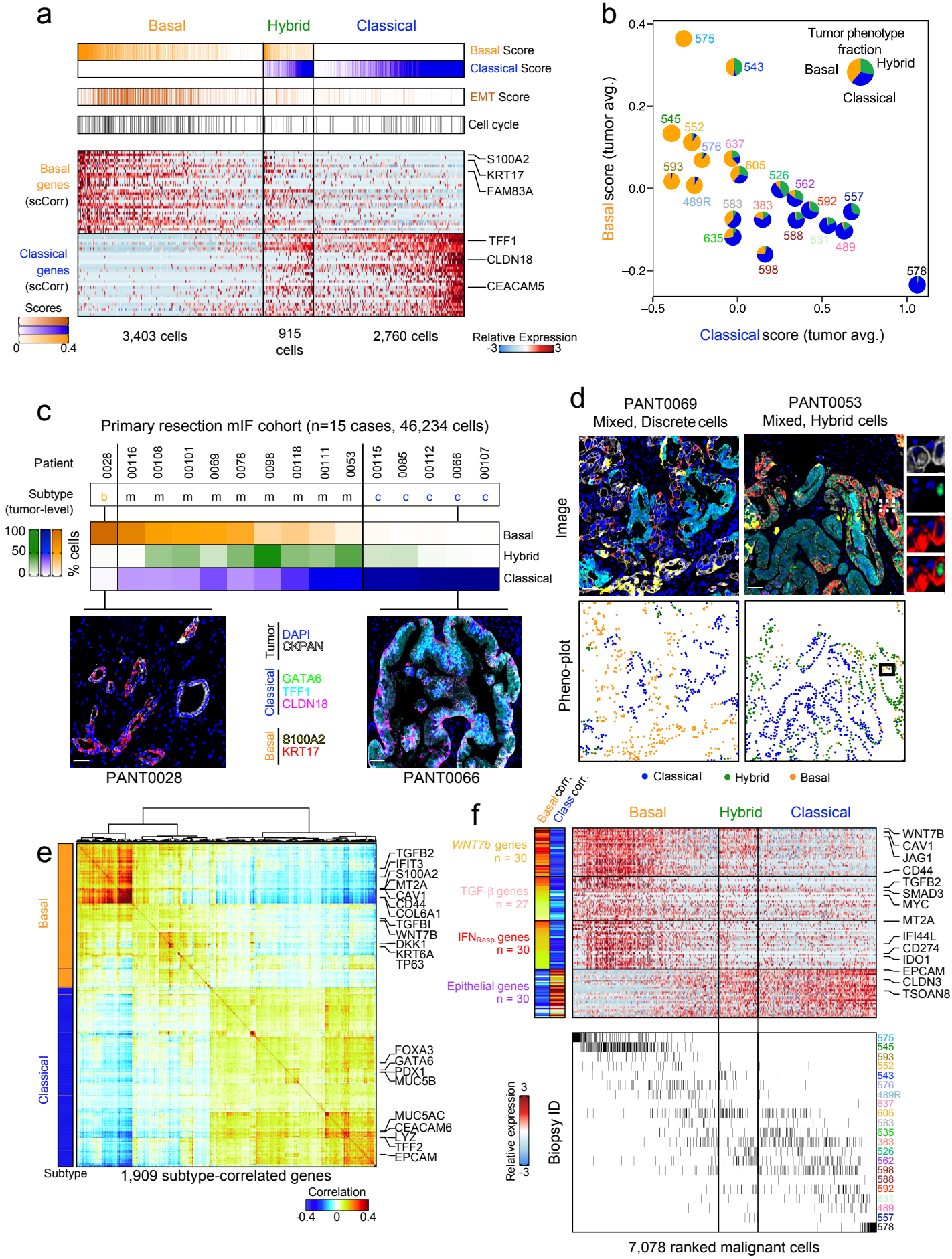
55 Candido, J. B. *et al.* CSF1R(+) Macrophages Sustain Pancreatic Tumor Growth through T Cell Suppression and Maintenance of Key Gene Programs that Define the Squamous Subtype. *Cell Rep* **23**, 1448-1460, doi:10.1016/j.celrep.2018.03.131 (2018).

Figure 1


a
Core needle biopsy
Dissociate
~50-100k viable cells
Initiate organoid culture
Grow organoids
Split and Reseed
Established organoid
Targeted DNA-seq
•Point mutations
•CNVs
scRNA seq
Biopsy tissue
(n=23 samples)
Primary tissue characterization
scRNA seq
Early organoid
16/23 Early passage
(P2-3)
n passages
(P4-17)
scRNA seq
Sample at n passages
Matched organoid characterization
b
Putative malignant cells
Non-Malignant cells
CNV correlation
CNV score
PANFR0575
c
PANFR0575 CNVs
Targeted
DNA-seq
Single Cells
Chr. 1
10
22
Estimated copy number
Loss
Gain
d
Non-malignant cells t-SNE visualization
Endo
Plasma
Treg
Mac
pDC
T/NK
DC
Hep
XCR1+DC
Fib
B
15,302 cells
e
Malignant t-SNE visualization
7,740 cells

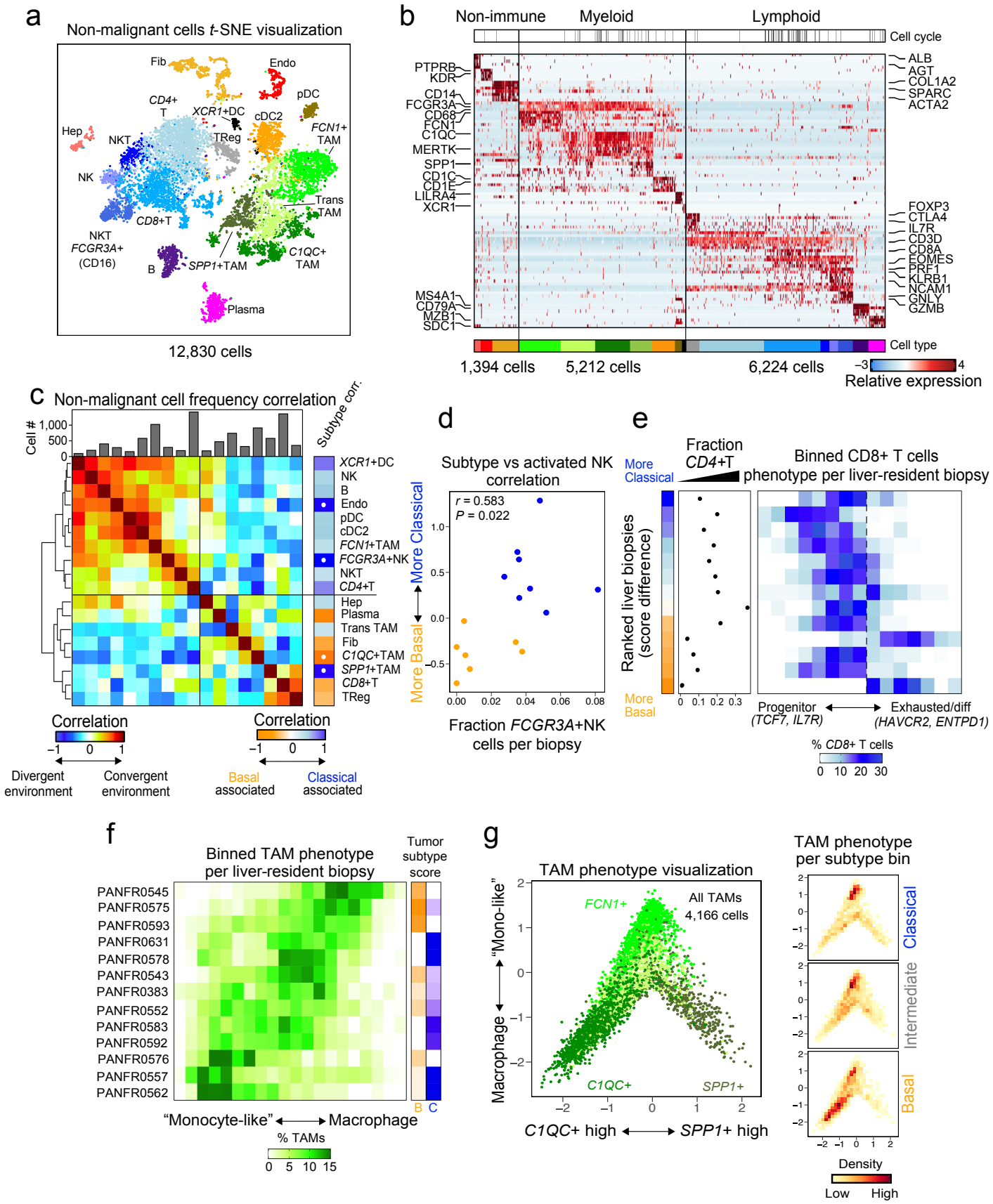
578 **Figure 1. A clinical pipeline for matched single-cell RNA-seq and organoid generation from metastatic**
579 **biopsies. a**, Pipeline for collecting patient samples, dissociation and allocation for scRNA-seq, and parallel
580 organoid development. **b**, CNV correlation (to averaged top 5% of altered cells) versus CNV score (mean square)
581 for each single cell in PANFR0575. Cells are colored by their putative class: malignant (light blue) or non-
582 malignant (empty black circles). **c**, Bulk targeted DNA-seq (top) and single-cell (rows, bottom) CNV profiles
583 arranged by chromosome (columns). **d, e**, *t*-distributed stochastic neighbor embedding (*t*-SNE) visualization for
584 non-malignant (**d**) and malignant (**e**) single cells in the biopsy cohort. Cells are colored by patient. Endo,
585 Endothelial; Fib, Fibroblast; B, B-cell; Hep, Hepatocyte; DC, Dendritic cell; pDC, Plasmacytoid dendritic cell;
586 Mac, Macrophage; T, T-cell; NK, Natural killer cell.
587
588

Figure 2


a
Basal
Hybrid
Classical
Basal Score
Classical Score
EMT Score
Cell cycle
Basal genes (scCorr)
Classical genes (scCorr)
S100A2
KRT17
FAM83A
TFF1
CLDN18
CEACAM5
Scores
0
0.4
3,403 cells
915 cells
2,760 cells
Relative Expression
-3
3
b
Tumor phenotype fraction
Basal score (tumor avg.)
Classical score (tumor avg.)
c
Primary resection mIF cohort (n=15 cases, 46,234 cells)
Patient
Subtype (tumor-level)
% cells
Basal
Hybrid
Classical
Tumor
DAPI
CKPAN
GATA6
TFF1
CLDN18
S100A2
KRT17
PANT0028
PANT0066
d
PANT0069
Mixed, Discrete cells
PANT0053
Mixed, Hybrid cells
Image
Pheno-plot
Classical
Hybrid
Basal
e
Basal
Classical
Subtype
1,909 subtype-correlated genes
Correlation
-0.4
0
0.4
TGFB2
IFIT3
S100A2
MT2A
CAV1
CD44
COL6A1
TGFBI
WNT7B
DKK1
KRT6A
TP63
FOXA3
GATA6
PDX1
MUC5B
MUC5AC
CEACAM6
LYZ
TFF2
EPCAM
f
Basal corr.
Class corr.
WNT7b genes n = 30
TGF-β genes n = 27
IFNResp genes n = 30
Epithelial genes n = 30
Basal
Hybrid
Classical
WNT7B
CAV1
JAG1
CD44
TGFB2
SMAD3
MYC
MT2A
IFI44L
CD274
IDO1
EPCAM
CLDN3
TSOAN8
Relative expression
3
-3
Biopsy ID
7,078 ranked malignant cells

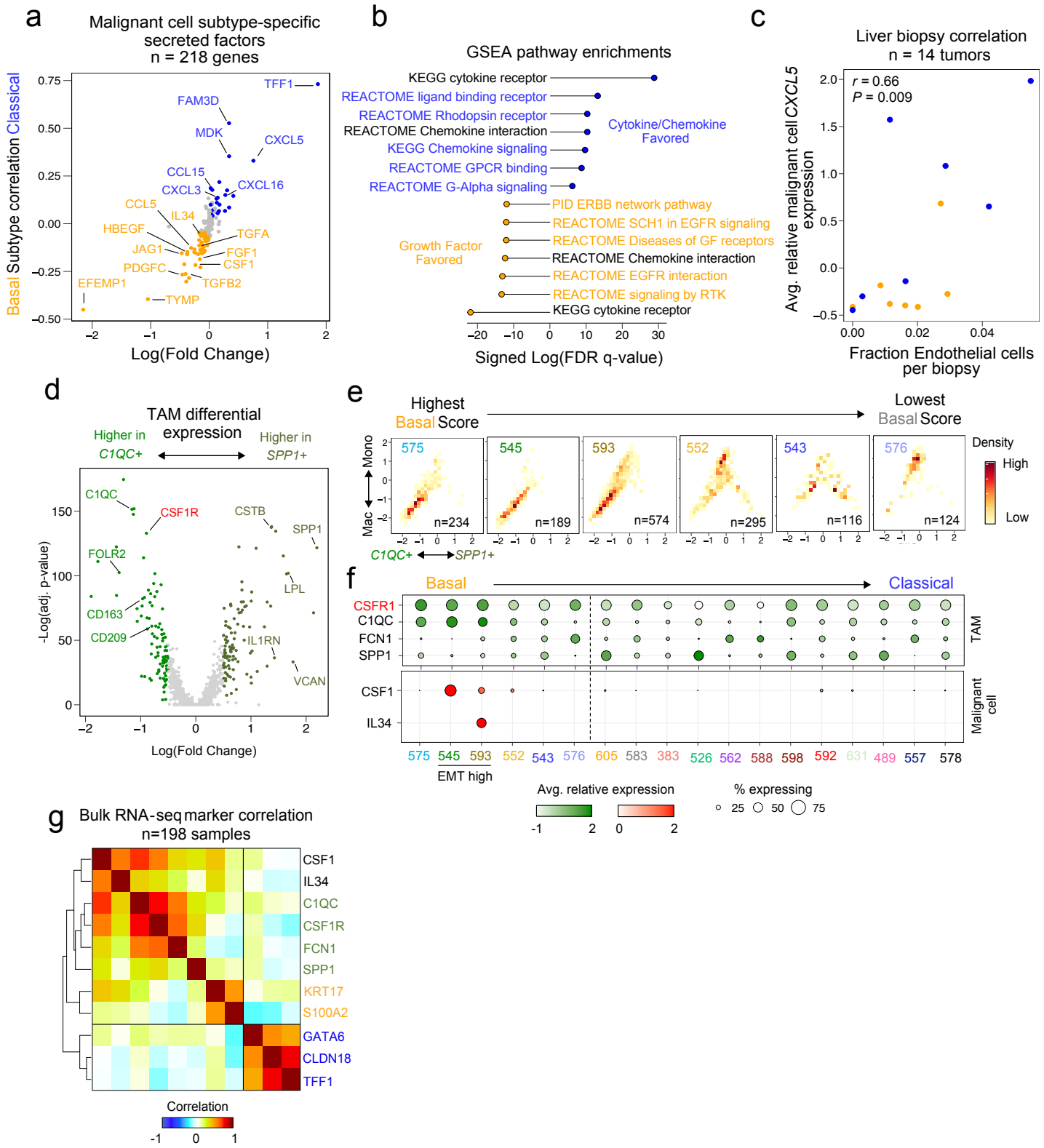
589 **Figure 2. Basal, classical, and hybrid transcriptional states in metastatic PDAC. a**, Heatmap depicts the
590 expression of basal and classical genes (n=30 each, **Methods**) across all malignant cells. EMT, basal, classical,
591 and cell cycle programs are indicated. **b**, PDAC tumors are arranged by their average classical (x-axis) and
592 basal (y-axis) scores. Points are pie charts summarizing the malignant subtype composition within each biopsy.
593 **c**, Composition of each tumor (% cells) across the three expression subtypes in the primary resection cohort
594 (n=15 cases) determined by multiplexed immunofluorescence (b, basal; m, mixed; c, classical). Representative
595 images for strongly polarized tumors are shown (bottom). **d**, Representative mixed tumor images (top) and
596 corresponding pheno-plots (bottom). Pheno-plot points correspond to cells in the image above and are colored
597 by their subtype, marker negative cells are not visualized. Zoom panel on far right (dotted white box, image; solid
598 black box, pheno-plot) shows juxtaposed hybrid and basal cells. **e**, Pairwise correlation of genes significantly
599 associated with basal or classical expression states. Left bar indicates the subtype association of each gene
600 (orange, basal; blue, classical). **f**, Heatmap shows the relative expression of the indicated basal and classical-
601 associated programs, cells are ordered as in **a.** Left heat bar indicates each gene's correlation to either basal or
602 classical subtypes, and the range for these values is the same as in **e**. Bottom plot indicates each single-cell's
603 biopsy of origin.

Figure 3


a
Non-malignant cells t-SNE visualization
Fib
Endo
CD4+ T
XCR1+DC
pDC
cDC2
Hep
NKT
TReg
FCN1+ TAM
NK
Trans TAM
CD8+T
NKT FCGR3A+ (CD16)
B
SPP1+TAM
C1QC+ TAM
Plasma
12,830 cells
b
Non-immune
Myeloid
Lymphoid
Cell cycle
PTPRB
KDR
CD14
FCGR3A
CD68
FCN1
C1QC
MERTK
SPP1
CD1C
CD1E
LILRA4
XCR1
MS4A1
CD79A
MZB1
SDC1
ALB
AGT
COL1A2
SPARC
ACTA2
FOXP3
CTLA4
IL7R
CD3D
CD8A
EOMES
PRF1
KLRB1
NCAM1
GNLY
GZMB
Cell type
1,394 cells
5,212 cells
6,224 cells
−3
4
Relative expression
c
Non-malignant cell frequency correlation
Cell #
1,000
500
0
Subtype corr.
XCR1+DC
NK
B
Endo
pDC
cDC2
FCN1+TAM
FCGR3A+NK
NKT
CD4+T
Hep
Plasma
Trans TAM
Fib
C1QC+TAM
SPP1+TAM
CD8+T
TReg
Correlation
−1
0
1
Divergent environment
Convergent environment
Correlation
−1
0
1
Basal associated
Classical associated
d
Subtype vs activated NK correlation
r = 0.583
P = 0.022
More Classical
More Basal
1.0
0.5
0.0
−0.5
0.00
0.02
0.04
0.06
0.08
Fraction FCGR3A+NK cells per biopsy
e
Fraction CD4+T
More Classical
More Basal
Binned CD8+ T cells phenotype per liver-resident biopsy
Ranked liver biopsies (score difference)
0
0.1
0.2
0.3
Progenitor (TCF7, IL7R)
Exhausted/diff (HAVCR2, ENTPD1)
% CD8+ T cells
0
10
20
30
f
Binned TAM phenotype per liver-resident biopsy
Tumor subtype score
PANFR0545
PANFR0575
PANFR0593
PANFR0631
PANFR0578
PANFR0543
PANFR0383
PANFR0552
PANFR0583
PANFR0592
PANFR0576
PANFR0557
PANFR0562
B
C
"Monocyte-like"
Macrophage
% TAMs
0
5
10
15
g
TAM phenotype visualization
All TAMs
4,166 cells
FCN1+
C1QC+
SPP1+
"Mono-like"
Macrophage
2
1
0
−1
−2
−2
−1
0
1
2
C1QC+ high
SPP1+ high
TAM phenotype per subtype bin
Classical
Intermediate
Basal
Density
Low
High

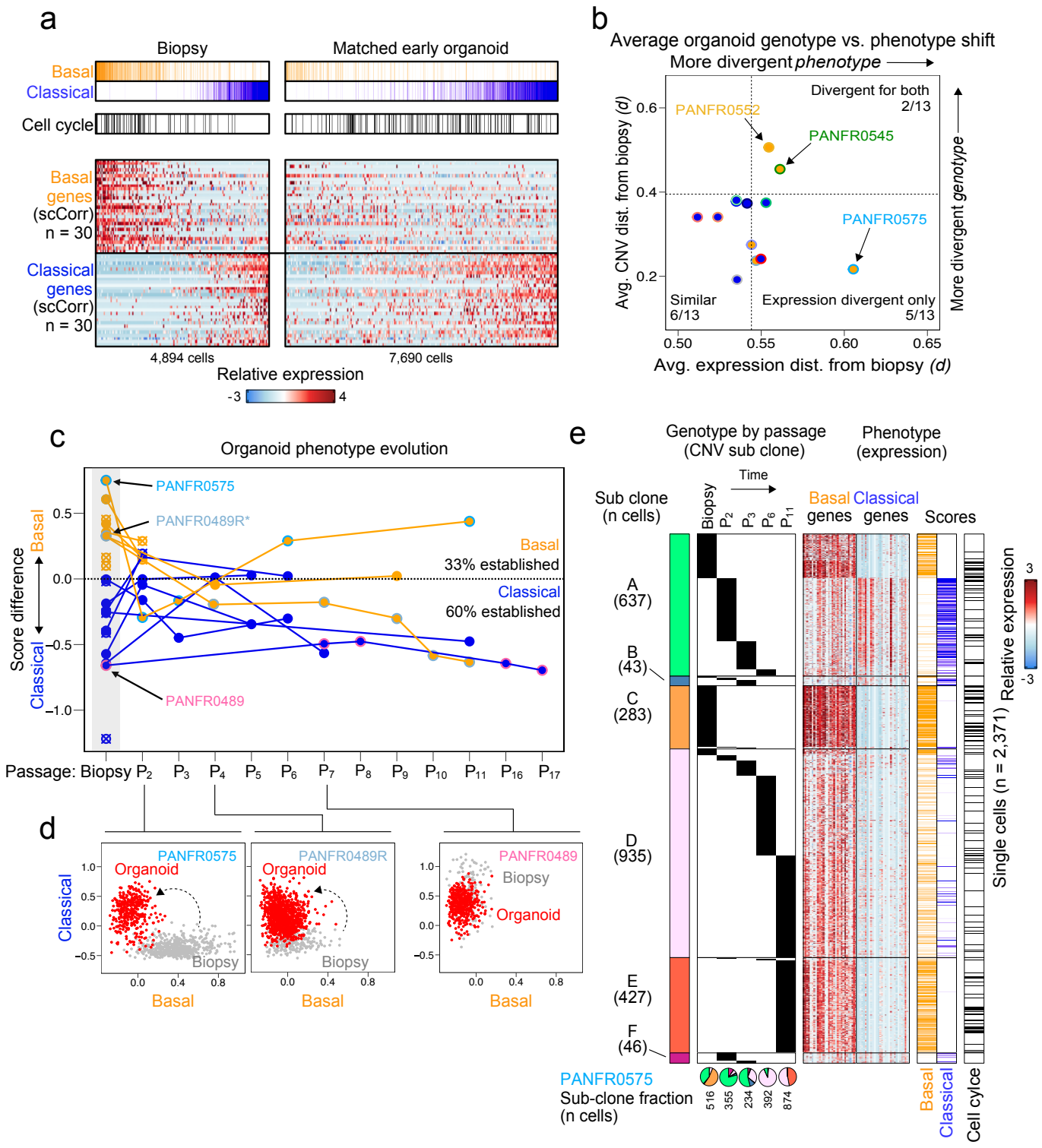
604 **Figure 3. Asymmetric distribution of immune phenotypes across the basal to classical continuum. a**, *t*-
605 SNE visualization of non-malignant cells identified in the metastatic microenvironment, abbreviations are the
606 same as in **Figure 1d** (TAM, tumor associated macrophage; Trans, Transition; NKT, natural killer T cell). **b,**
607 Heatmap shows the relative expression for select cell type markers. Top bar indicates the binarized cell cycle
608 program (black, cycling) and the bottom color bar corresponds to the cell type colors noted in **a**. **c,** Cross-
609 correlational heatmap and hierarchical clustering for similarity in the capture frequency of non-malignant cell
610 types from each biopsy. Rainbow coloration in the main heatmap indicates convergence (yellow to red) or
611 divergence (white to blue) across cell types. Right heat bar indicates preferential association for each cell type
612 with either the basal (orange, negative values) or classical (blue, positive values) malignant transcriptional
613 subtypes. Color ranges for both quantities are Pearson's *r*, white dots indicate $P < 0.05$ for the subtype
614 associations. Top bar chart indicates the total number of cells for each type. **d,** Scatter plot compares each liver
615 biopsy's position on the basal to classical continuum (y-axis, score difference) to the relative abundance of
616 activated NK cells captured from its microenvironment. Points represent individual biopsies and are colored by
617 their discretized transcriptional subtype (n = 15). **e,** Distribution (blue heat) of CD8+ T cell phenotype across the
618 progenitor to exhausted/differentiated continuum in each liver-resident biopsy. Biopsies are sorted by the score
619 difference (far left heat bar). The corresponding fractional capture of CD4+ T cells is indicated left of the main
620 heat map for each sample. **f,** Distribution (green heat) of TAM phenotype for the macrophages captured in each
621 liver biopsy. Biopsies with <100 macrophages were excluded. Heatmap is ranked by average monocyte-like to
622 macrophage skew and both average basal and classical scores are indicated (right). **g**, Phenotypic hierarchy for
623 TAM subsets using the expression scores for each phenotype across all TAMs captured in the dataset. The
624 distribution (density; high = more TAMs) across the phenotypic hierarchy is visualized (right) according to
625 malignant transcriptional phenotype as in **Extended Data Figure 3d**.

Figure 4


a
Malignant cell subtype-specific secreted factors
n = 218 genes
Basal Subtype correlation Classical
Log(Fold Change)
TFF1
FAM3D
MDK
CXCL5
CCL15
CXCL3
CXCL16
CCL5
IL34
HBEGF
TGFA
JAG1
FGF1
CSF1
PDGFC
TGFB2
EFEMP1
TYMP
b
GSEA pathway enrichments
KEGG cytokine receptor
REACTOME ligand binding receptor
REACTOME Rhodopsin receptor
REACTOME Chemokine interaction
KEGG Chemokine signaling
REACTOME GPCR binding
REACTOME G-Alpha signaling
Cytokine/Chemokine Favored
PID ERBB network pathway
REACTOME SCH1 in EGFR signaling
REACTOME Diseases of GF receptors
REACTOME Chemokine interaction
REACTOME EGFR interaction
REACTOME signaling by RTK
KEGG cytokine receptor
Growth Factor Favored
Signed Log(FDR q-value)
c
Liver biopsy correlation
n = 14 tumors
r = 0.66
P = 0.009
Avg. relative malignant cell CXCL5 expression
Fraction Endothelial cells per biopsy
d
TAM differential expression
Higher in C1QC+
Higher in SPP1+
C1QC
CSF1R
CSTB
SPP1
FOLR2
LPL
CD163
CD209
IL1RN
VCAN
-Log(adj. p-value)
Log(Fold Change)
e
Highest Basal Score
Lowest Basal Score
Mono
Mac
C1QC+
SPP1+
575
545
593
552
543
576
n=234
n=189
n=574
n=295
n=116
n=124
Density
High
Low
f
Basal
Classical
CSFR1
C1QC
FCN1
SPP1
TAM
CSF1
IL34
Malignant cell
575 545 593 552 543 576 605 583 383 526 562 588 598 592 631 489 557 578
EMT high
Avg. relative expression
% expressing
25 50 75
g
Bulk RNA-seq marker correlation
n=198 samples
CSF1
IL34
C1QC
CSF1R
FCN1
SPP1
KRT17
S100A2
GATA6
CLDN18
TFF1
Correlation

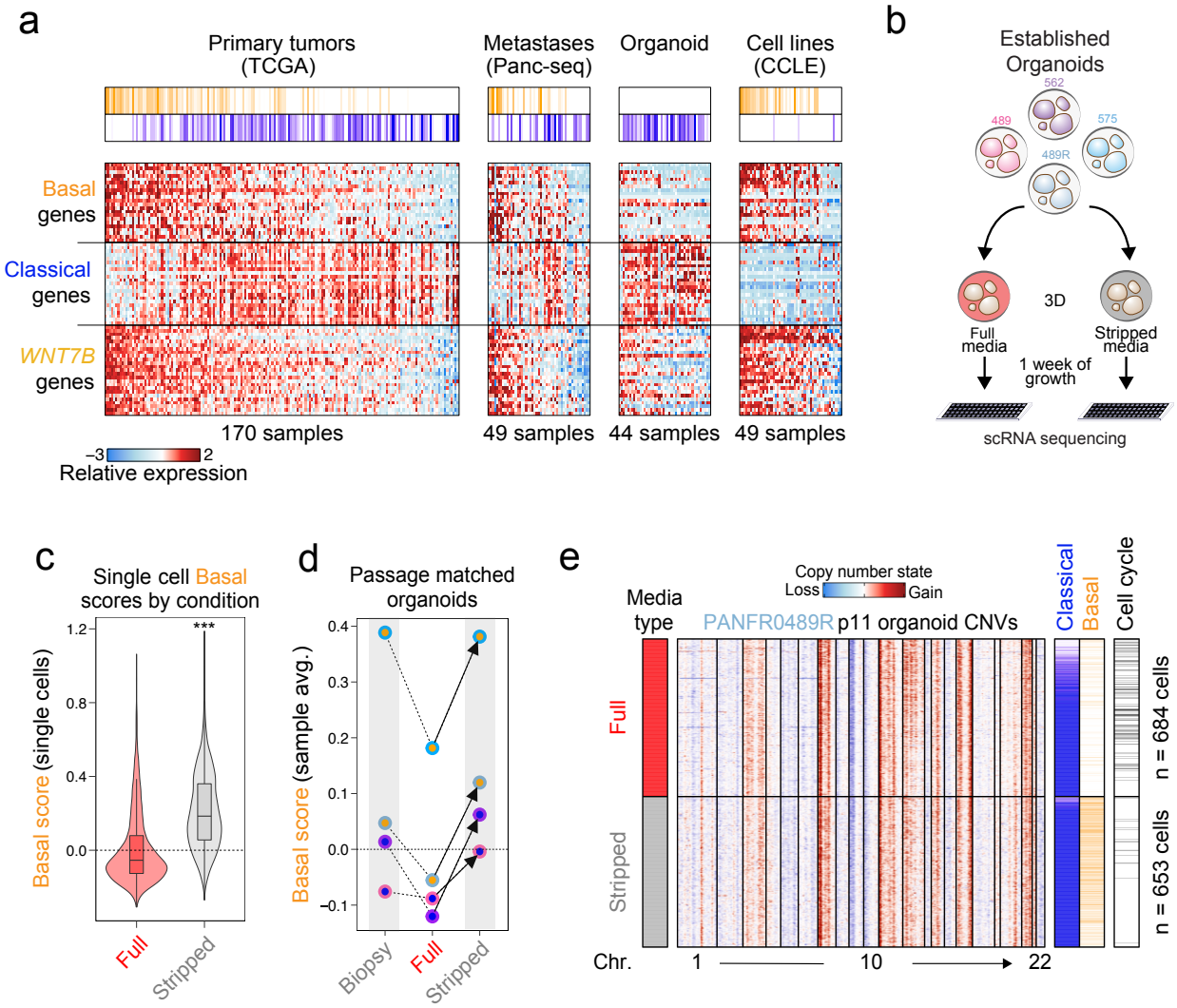
626 **Figure 4. Differential microenvironmental crosstalk shapes subtype-specific metastatic niches. a,**
627 Scatterplot comparing differential expression (x-axis) and subtype correlation coefficient (y-axis) for the 219
628 genes annotated as secreted growth factors, cytokines, or chemokines detected in malignant cells. Genes
629 passing significance are assigned as "subtype specific" for either basal or classical ($P < 0.05$, DE; $P < 0.003$,
630 correlation). **b**, Pathway enrichments for the top genes associated with each subtype. Shared enrichments are
631 in black, orange and blue denote basal and classical-unique terms, respectively. **c,** Scatterplot comparing the
632 fraction of endothelial cells captured and the average expression in malignant cells for *CXCL5.* Each point
633 represents one biopsy (n=14). **d,** Differential expression between the committed TAM subsets *SPP1*+ and
634 *C1QC*+. Genes are colored by their subtype selectivity ($P < 0.05$; Log(Fold Change) > 0.5). **e,** TAM phenotypic
635 hierarchy plots for basal subtype tumors (Figure 3g, bottom), split by biopsy and sorted by decreasing basal
636 score. Heat indicates distribution of total TAMs (bottom right for each plot) per biopsy as in **Figure 3g**. **f,** Dot
637 plots for TAM (top, green fill) and malignant (bottom, red fill) expression of the indicated genes sorted by basal-
638 classical polarization. Size of the dot indicates fraction of cells expressing a given gene. Left of the dotted line
639 are tumors visualized in **e**; basal tumors high for EMT program expression (PANFR0545 and PANFR0593) are
640 indicated. **g**, Cross-correlation for markers of TAM subsets (green), basal (orange) and classical (blue) markers
641 used in mIF studies, and putative malignant secreted factors (black).

Figure 5


a
Biopsy
Matched early organoid
Basal
Classical
Cell cycle
Basal genes (scCorr) n = 30
Classical genes (scCorr) n = 30
4,894 cells
7,690 cells
Relative expression
-3
4
b
Average organoid genotype vs. phenotype shift
More divergent phenotype
Divergent for both 2/13
PANFR0552
PANFR0545
PANFR0575
Similar 6/13
Expression divergent only 5/13
Avg. CNV dist. from biopsy (d)
Avg. expression dist. from biopsy (d)
More divergent genotype
c
Organoid phenotype evolution
PANFR0575
PANFR0489R*
PANFR0489
Basal 33% established
Classical 60% established
Score difference
Basal
Classical
Passage: Biopsy P2 P3 P4 P5 P6 P7 P8 P9 P10 P11 P16 P17
d
PANFR0575
Organoid
Biopsy
PANFR0489R
PANFR0489
Classical
Basal
e
Genotype by passage (CNV sub clone)
Phenotype (expression)
Time
Sub clone (n cells)
Biopsy
Basal genes
Classical genes
Scores
A (637)
B (43)
C (283)
D (935)
E (427)
F (46)
Relative expression
Single cells (n = 2,371)
PANFR0575 Sub-clone fraction (n cells)
516
355
234
392
874
Basal
Classical
Cell cylce

642 **Figure 5. Organoid culture microenvironment selects against the basal state. a**, Relative expression for
643 basal and classical genes in biopsy cells (left) and their matched, early passage organoid cells (n=13 models;
644 right). **b**, Scatterplot compares the relative contribution of genotypic drift (CNVs, y-axis) versus phenotypic drift
645 (basal/classical gene expression, x-axis). Both quantities represent distance ($d = (1-r)/2$); higher value = greater
646 distance) derived from Pearson correlation (r). Each point is one organoid/biopsy pair and summarizes the
647 average d between organoid cells and their matched initial biopsy. Dotted lines are $P < 0.05$ comparing average
648 intra-biopsy (biopsy cells to themselves) d across the cohort for both metrics. Fill colors denote classification of
649 original tumor, point outline color is the biopsy identifier. **c**, Line plot for each biopsy and its successive organoid
650 samples (*see **Methods**). Points represent the sample averaged score at the indicated timepoints, lines tie
651 samples derived from the same initial biopsy. Color indicates if the original biopsy was initially measured as
652 basal (orange) or classical (blue). Colored point outlines denote all samples from the indicated original biopsy.
653 Crossed empty circles indicate when a sample failed to grow. **d**, Representative scatterplots for single-cell basal
654 and classical scores in biopsy (grey) and the indicated organoid passage (red) sample. **e**, Genotype and
655 phenotype evolution in PANFR0575. Cells are sorted first by their subclone (A-F, color bar far left; **Methods**)
656 and then sample of origin (Biopsy or organoid, right of subclone color bar; P_n, Organoid passage number). Each
657 single cell's corresponding phenotype is shown in the center heatmap and far right expression score bars (Cell
658 cycle, black). The fraction of each subclone in each sample is indicated with pie charts at the bottom, cell
659 numbers per sample are below.

Figure 6


a
Primary tumors (TCGA)
Metastases (Panc-seq)
Organoid
Cell lines (CCLE)
Basal genes
Classical genes
WNT7B genes
170 samples
49 samples
44 samples
49 samples
−3
2
Relative expression
b
Established Organoids
562
489
575
489R
3D
Full media
Stripped media
1 week of growth
scRNA sequencing
c
Single cell Basal scores by condition
Basal score (single cells)
1.2
0.8
0.4
0.0
***
Full
Stripped
d
Passage matched organoids
Basal score (sample avg.)
0.4
0.3
0.2
0.1
0.0
−0.1
Biopsy
Full
Stripped
e
Copy number state
Loss
Gain
Media type
PANFR0489R p11 organoid CNVs
Classical
Basal
Cell cycle
Full
Stripped
n = 684 cells
n = 653 cells
Chr.
1
10
22

660 **Figure 6. Recovery of the basal state in altered media conditions. a**, Relative expression for 90 genes
661 representing basal, classical, and *WNT7B* expression programs across bulk RNA-seq samples from primary
662 resections (TCGA) and metastatic biopsies (Panc-Seq), as well as organoid and cell line (CCLE) models.
663 Phenotype scores are indicated and samples are ranked by their score difference. **b**, Schematic for depleted
664 media experiment. **c**, Single-cell violin plots for basal score in passage matched organoids grown in the indicated
665 media conditions (***P = 2.2×10^{-16}). **d**, Dot plot represents the sample average basal score in the indicated
666 conditions. Lines tie samples and color outlines indicate sample identity. Each sample's biopsy basal score is
667 included for reference. **e**, Inferred CNVs, expression scores, and cell cycle status for each cell from either
668 Stripped (grey) or Full (red) organoid media conditions in the PANFR489R experiment.