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

2

74   measurements in multiple cancers.[15,21-26] Such enhanced resolution may be particularly useful in PDAC, where

75   neoplastic cellularity is generally low and stromal content is high. Understanding the distribution and plasticity of

76   malignant and non-malignant states within individual PDAC tumors has important implications for the

77   interpretation of transcriptional subtypes, directing therapy, and monitoring tumor evolution. However, few single-

78   cell studies have been conducted in human PDAC, and these have largely focused on stromal cell types or

79   provided a limited analysis of malignant cells.[11,27-29] We therefore lack a harmonized view of the interplay between

80   malignant transcriptional subtypes and their associated tumor microenvironment (TME).

81   Our current understanding of PDAC is largely derived from resected primary tumors.[12,13,30] However, the

82   majority of patients with PDAC present with, and succumb to, metastatic disease, which occurs most commonly

83   in the liver.[30] At present, we have little information about the cellular phenotypes and microenvironmental

84   interactions in metastatic lesions. Tissue availability has been a key barrier to enhanced understanding of

85   metastatic disease, as needle biopsies provide an important but cell-limited window into the biology of the

86   metastatic niche.

87   In conjunction with detailed molecular analysis of patient samples, reliable *ex vivo* models are needed to

88   functionally test clinical and molecular observations. For this purpose, human cancer cell line models are

89   frequently utilized, as is the case in PDAC. However, the methods to generate new cell lines from human tissue

90   are generally inefficient, which limits their utility in personalized medicine.[31] Moreover, once established, cell

91   lines can display significant drift in culture.[32] To address these limitations, several groups have established

92   efficient methods for generating patient-derived organoid cultures from PDAC tissue with the goal of modeling

93   an individual patient's disease.[10,33-35] However, few studies have examined the fidelity and evolution of organoid

94   phenotype and genotype relative to the parental patient tissue.

95   Here, we developed and employed an optimized translational workflow to perform both high-resolution

96   profiling of patient tissue using scRNA-seq via Seq-Well[36] and derivation of matched organoid models from the

97   same metastatic core needle biopsy. Through this approach we reframe bulk classifications by clarifying the

98   underlying distribution of malignant phenotypes, reveal how microenvironmental heterogeneity is distributed in

99   a transcriptional subtype-dependent manner, and systematically evaluate the *ex vivo* evolution and plasticity of

100 malignant phenotypes.

101

102 **RESULTS**

103

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

105

106 We established a pipeline for collecting needle biopsies from patients with metastatic PDAC (n=23) to generate

107 matched scRNA-seq profiles and organoid models **(Figure 1a, Extended Data Figure 1a, Supplementary**

108 **Table 1)**. Most samples were obtained from metastatic lesions residing in the liver (19/23), and the majority

109 (21/23) were analyzed by targeted DNA-sequencing which yielded the expected mutational pattern for this

110 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

heterogenous mixture of states, likely driving the ambiguous classification of weakly polarized tumors when using hierarchical clustering (**Extended Data Figure 3d,e**)**.** We also observed a significant fraction of malignant cells co-expressing both basal and classical phenotypes, hereafter referred to as "hybrid" cells (~13% of malignant cells, **Figure 2a**, see **Methods**), suggesting that these phenotypes exist on a continuum rather than as discrete states. Classification of each single cell as basal, classical, or hybrid revealed substantial heterogeneity across individual tumors for these phenotypes (**Figure 2b**)**.** These observations underscore the difficulty in assigning intermediate tumors exclusively to basal or classical groups.[11] Thus, where discrete binning was necessary, we employed a basal-classical "score difference" to stratify samples and preserve the polarization for each tumor along this continuum (**Extended Data Figure 3f**).

We also used our single-cell data to examine signatures proposed by other bulk RNA sequencing studies to clarify their inter-relationships. Pairwise correlation of all established signatures in malignant cells revealed that many contribute overlapping information and reflect similar underlying biology **(Extended Data Figure 3g**). We observed that cells with higher basal expression were also classified as squamous and quasimesenchymal, while cells with classical signatures were correlated with the pancreatic progenitor subtype **(Extended Data Figure 3g,h**).[3,4] By contrast, we did not observe evidence for expression of the immunogenic, ADEX, or exocrine-like transcriptional signatures in malignant cells.[3,4] While the absence of these signatures might represent differences between primary and metastatic disease, these bulk RNA profiles also likely incorporate signals from non-malignant cells in the TME. In support of the latter hypothesis, we find evidence of immunogenic signature expression originating from plasma cells as well as EMT signature expression from both malignant cells and fibroblasts **(Extended Data Figure 3h**). These patterns underscore the need for single-cell resolution to dissect malignant and non-malignant contributions to transcriptional signatures.

We next confirmed the presence of basal, classical, and hybrid cells using a novel subtype-specific single-cell multiplexed immunofluorescence (mIF) panel in a cohort of primary resected PDAC (n=15 cases, 46,234 cells, **Methods; Figure 2c; Extended Data Figure 4a-c, Supplementary Tables 3, 4**). This orthogonal approach confirmed the intratumoral heterogeneity observed in our scRNA-seq cohort and revealed that PDAC transcriptional subtype diversity occurs on two levels: (i) "mixed" tumors comprised of discrete cells with differing subtype identity, and (ii) hybrid cells which co-express basal and classical programs. These observations indicate that PDAC transcriptional subtypes exist on a continuum, with mixed and hybrid phenotypes occurring even within a single tumor gland (**Figure 2d**)**.**

Basal and classical cells exhibit subtype-specific expression programs

We next leveraged our single-cell resolution to examine whether specific tumor cell gene expression programs were correlated with either the basal or classical phenotypes. This correlation analysis across malignant cells revealed 1,909 genes significantly associated with either basal or classical expression scores (**Figure 2e; Supplementary Tables 5,6; Methods**). Inspection of these genes revealed basal cells are defined by more mesenchymal features and co-express programs associated with transforming growth factor beta (*TGFB2*,

5

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**).

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

Along with differences in lymphocyte content, the myeloid compartment, specifically TAM phenotypes, showed strong subtype-specific associations. First, we noted selective skewing for the types of TAMs originating from basal versus classical tumors (**Extended data figure 6i**, *P* < 2.2x10⁻¹⁶, Chi-squared test; **Figure 3c**, *C1QC*+ TAM, *r* = - 0.59, basal association and *SPP1*+ TAM, *r* = 0.52, classical association). Indeed, when examining the monocyte-like to macrophage distribution for TAMs from individual liver biopsies, the most basal-polarized tumors were associated with more macrophage-committed phenotypes (**Figure 3f**). Moreover, by scoring each macrophage using TAM subtype-specific signatures and visualizing them with respect to the likely differentiation trajectory inferred from recent studies (**Supplementary Tables 7-9; Methods**)[42], we confirm a preferential association between *C1QC*+ TAMs and basal tumors and, conversely, an enrichment for the inflammatory *FCN1*+ monocyte-like and *SPP1*+ TAM subsets in tumors with intermediate and classical phenotypes (**Figure 3g**). In addition to demonstrating that classical tumors are relatively more immune infiltrated, this analysis also identifies distinct microenvironmental phenotypes that co-vary with each PDAC transcriptional subtype and suggests opportunities to direct microenvironmental therapies in a subtype-specific manner.

Differential microenvironmental signaling shapes subtype-specific metastatic niches

7

259

260 Given the striking compositional differences we observed in the immune microenvironment across the basal to

261 classical axis, we next searched for tumor-secreted factors that might influence the structure of the local

262 metastatic niche. Specifically, we analyzed subtype-specific expression patterns for genes detected in malignant

263 cells that were annotated as secreted factors (cytokines, chemokines, growth factors by Gene Ontology; n=218

264 genes). This analysis nominated 57 basal (orange) and 23 classical-associated (blue) secreted factors (**Figure**

265 **4a**). Gene set enrichment analysis (GSEA) demonstrated that basal tumors were enriched for genes associated

266 with growth factor secretion, while classical tumors were enriched for cytokine/chemokine signaling (**Figure 4b**).

267 We observed expression of multiple TGF ligands secreted by basal tumors, consistent with the association of

268 increased TGF-β signaling in basal tumors (**Figure 2f**) and local immune suppression/exclusion. Conversely,

269 several chemokines (*CXCL5*, *CXCL3*) were enriched in classical tumors in agreement with their overall higher

270 degree of immune infiltration and higher fraction of endothelial cells (**Figure 3c**). As such, classical tumors

271 expressed higher levels of *CXCL5* which plays a documented role in enhancing tumor-supportive

272 angiogenesis.[45,46] Consistent with this finding, we observed a strong positive correlation between high average

273 malignant cell expression of *CXCL5* and the fraction of endothelial cells recovered (**Figure 4c**). In basal tumors,

274 we noted increased expression of the ligands *CSF1* and *IL34* (**Figure 4a**) and concomitant expression of their

275 receptor, *CSF1R,* in the basal-associated *C1QC*+ TAMs (**Figure 3c,g; 4d**). Per-tumor analysis revealed a

276 continuum of *C1QC*+ TAM distribution within basal tumors that correlated with high *CSF1R* expression (**Figure**

277 **4e,f top**). Malignant cells with strong EMT features (PANFR0545, PANFR0593) expressed the highest levels of

278 *CSF1* and *IL34,* consistent with a role for tumor cells in shaping their local macrophage phenotypes (**Figure 4f**

279 **bottom**). To extend this finding in larger cohorts, we analyzed bulk RNA-sequencing of primary and metastatic

280 PDAC tumors for markers of transcriptional subtype, TAM, and tumor secretion phenotypes.[12,13] Consistent with

281 our single-cell observations, macrophage markers and the ligands *CSF1* and *IL34* were associated with basal

282 but not classical markers in these samples (n=198, **Figure 4g**). Together, these data provide evidence that

283 subtype-specific intercellular crosstalk shapes distinct niches in the metastatic microenvironment.

284

285 **Genotype and phenotype evolution of matched patient-derived organoid models**

286

287 Our observations indicate that basal and classical phenotypes exist along a continuum and exhibit distinct

288 patterns of reciprocal interaction with their local microenvironments. To examine how tumor cell phenotypes

289 adapt and evolve in *ex vivo* microenvironments, we utilized the matched organoid models generated from our

290 metastatic biopsy cohort (**Methods**). For most models, we obtained scRNA-seq samples at the earliest passage

291 possible, typically passage 2 (P2), and again at a later passage (**Extended Data Figure 7a,b**). Notably, only

292 33% of models derived from basal tumors propagated beyond passage 2, whereas 60% of models derived from

293 classical tumors established long-term cultures (**Extended Data Figure 7b**). Globally, unbiased analysis of

294 malignant biopsy (7,078 cells) and organoid cells (n=14 models, 24,789 cells) revealed that biopsy cells clustered

295 separately from their matched organoid counterparts (**Extended Data Figure 7c,d**). Only two clusters were

8

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

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

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

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

9

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

11

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

12

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.

REFERENCES

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

20   Sade-Feldman, M. *et al.* Defining T Cell States Associated with Response to Checkpoint Immunotherapy 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).

540 42   Zhang, L. *et al.* Single-Cell Analyses Inform Mechanisms of Myeloid-Targeted Therapies in Colon
541   Cancer. *Cell* **181**, 442-459 e429, doi:10.1016/j.cell.2020.03.048 (2020).

542 43   Zilionis, R. *et al.* Single-Cell Transcriptomics of Human and Mouse Lung Cancers Reveals Conserved
543   Myeloid Populations across Individuals and Species. *Immunity* **50**, 1317-1334 e1310,
544   doi:10.1016/j.immuni.2019.03.009 (2019).

545 44   Miller, B. C. *et al.* Subsets of exhausted CD8(+) T cells differentially mediate tumor control and respond
546   to checkpoint blockade. *Nat Immunol* **20**, 326-336, doi:10.1038/s41590-019-0312-6 (2019).

547 45   Chen, C. *et al.* CXCL5 induces tumor angiogenesis via enhancing the expression of FOXD1 mediated
548   by the AKT/NF-kappaB pathway in colorectal cancer. *Cell Death Dis* **10**, 178, doi:10.1038/s41419-019-
549   1431-6 (2019).

550 46   Li, A. *et al.* Overexpression of CXCL5 is associated with poor survival in patients with pancreatic cancer.
551   *Am J Pathol* **178**, 1340-1349, doi:10.1016/j.ajpath.2010.11.058 (2011).

552 47   Ghandi, M. *et al.* Next-generation characterization of the Cancer Cell Line Encyclopedia. *Nature* **569**,
553   503-508, doi:10.1038/s41586-019-1186-3 (2019).

554 48   Afelik, S., Pool, B., Schmerr, M., Penton, C. & Jensen, J. Wnt7b is required for epithelial progenitor growth
555   and operates during epithelial-to-mesenchymal signaling in pancreatic development. *Dev Biol* **399**, 204-
556   217, doi:10.1016/j.ydbio.2014.12.031 (2015).

557 49   Arensman, M. D. *et al.* WNT7B mediates autocrine Wnt/beta-catenin signaling and anchorage-
558   independent growth in pancreatic adenocarcinoma. *Oncogene* **33**, 899-908, doi:10.1038/onc.2013.23
559   (2014).

560 50   Steinhart, Z. *et al.* Genome-wide CRISPR screens reveal a Wnt-FZD5 signaling circuit as a druggable
561   vulnerability of RNF43-mutant pancreatic tumors. *Nat Med* **23**, 60-68, doi:10.1038/nm.4219 (2017).

562 51   Mead, B. E. *et al.* Harnessing single-cell genomics to improve the physiological fidelity of organoid-
563   derived cell types. *BMC Biol* **16**, 62, doi:10.1186/s12915-018-0527-2 (2018).

564 52   Ben-David, U., Beroukhim, R. & Golub, T. R. Genomic evolution of cancer models: perils and
565   opportunities. *Nat Rev Cancer* **19**, 97-109, doi:10.1038/s41568-018-0095-3 (2019).

566 53   Balachandran, V. P., Beatty, G. L. & Dougan, S. K. Broadening the Impact of Immunotherapy to
567   Pancreatic Cancer: Challenges and Opportunities. *Gastroenterology* **156**, 2056-2072,
568   doi:10.1053/j.gastro.2018.12.038 (2019).

569 54   Zhu, Y. *et al.* CSF1/CSF1R blockade reprograms tumor-infiltrating macrophages and improves response
570   to T-cell checkpoint immunotherapy in pancreatic cancer models. *Cancer Res* **74**, 5057-5069,
571   doi:10.1158/0008-5472.CAN-13-3723 (2014).

572 55   Candido, J. B. *et al.* CSF1R(+) Macrophages Sustain Pancreatic Tumor Growth through T Cell
573   Suppression and Maintenance of Key Gene Programs that Define the Squamous Subtype. *Cell Rep* **23**,
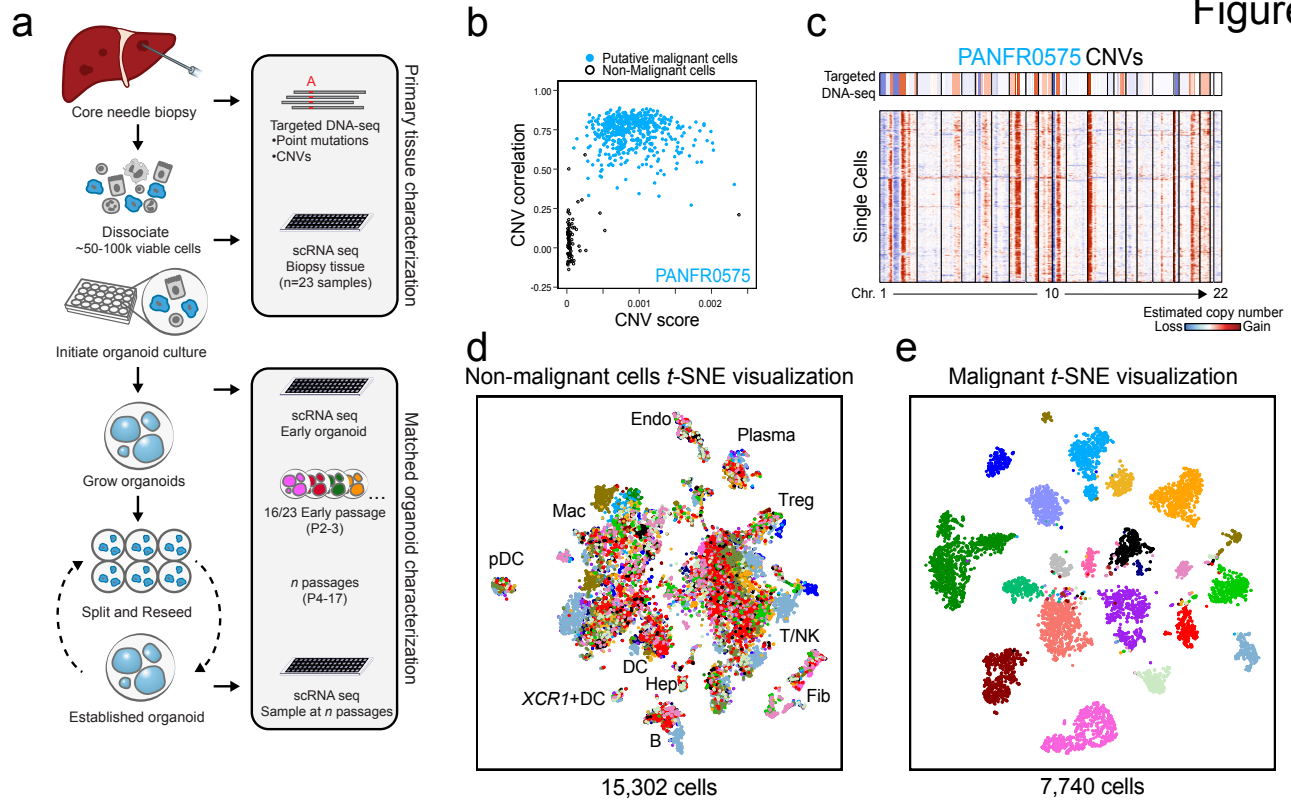574   1448-1460, doi:10.1016/j.celrep.2018.03.131 (2018).

575

576

577

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
Mac
Treg
pDC
T/NK
DC
Hep
Fib
XCR1+DC
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
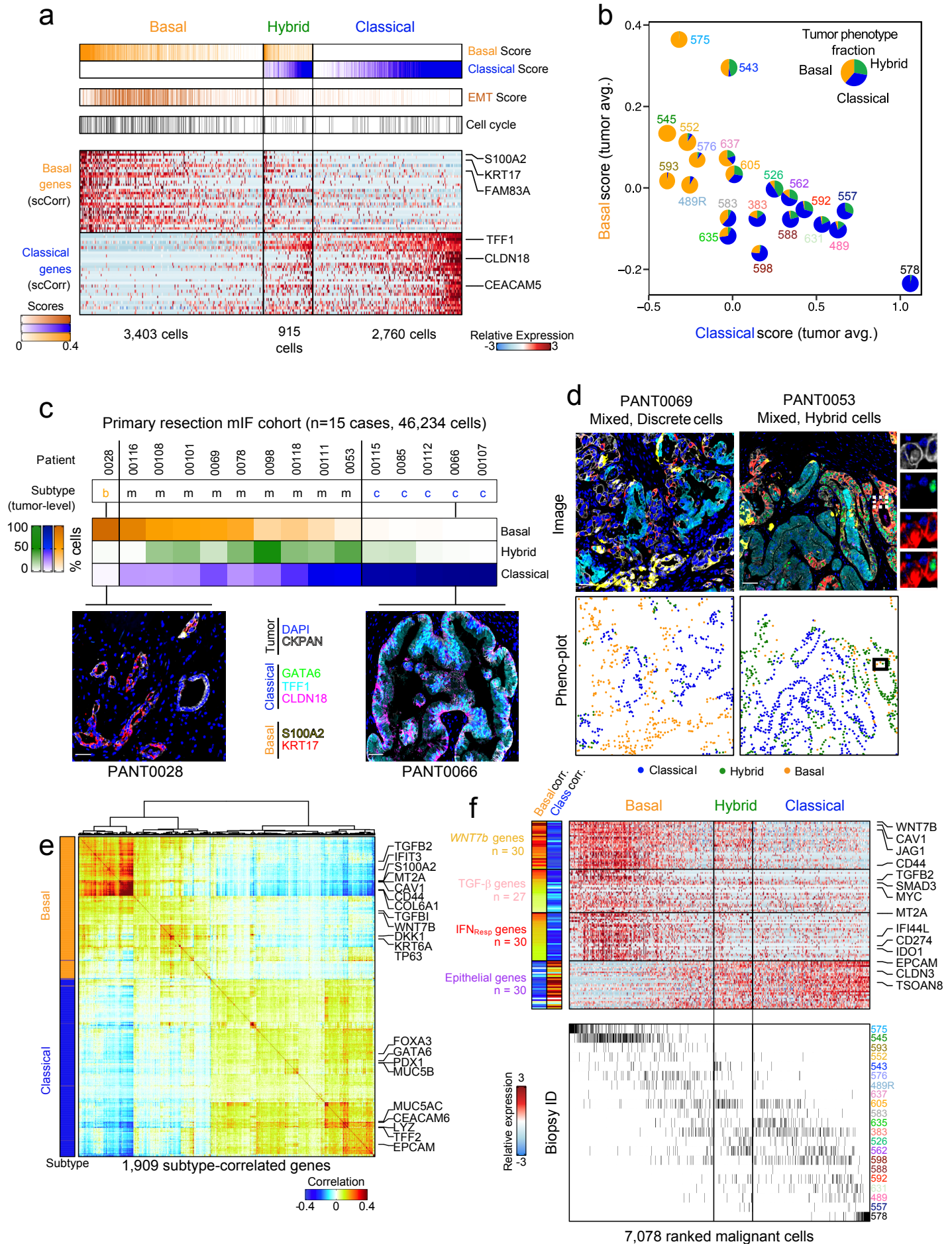
589 **Figure 2. Basal, classical, and hybrid transcriptional states in metastatic PDAC. a**, Heatmap depicts the

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.

18
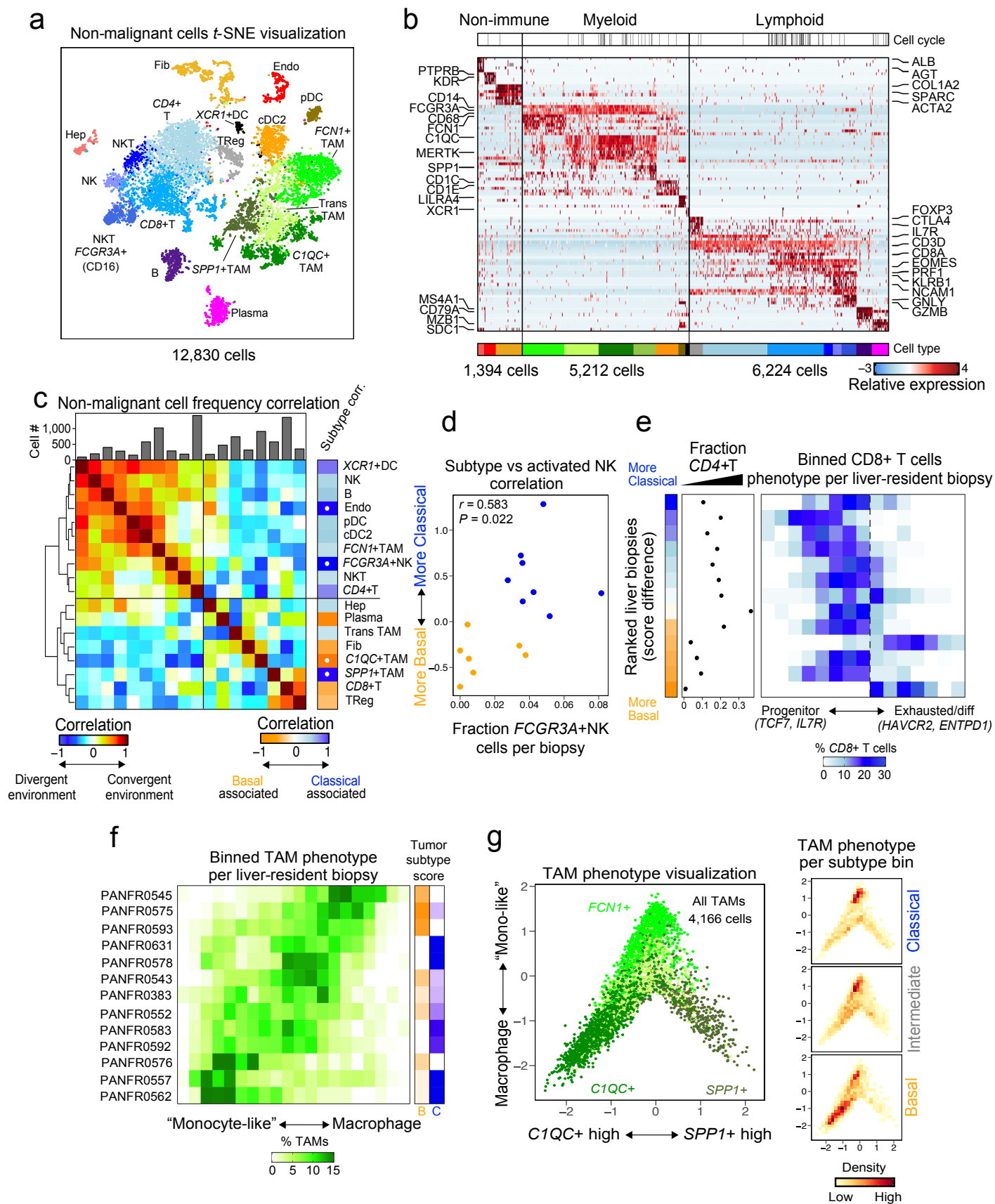
Figure 3

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

19
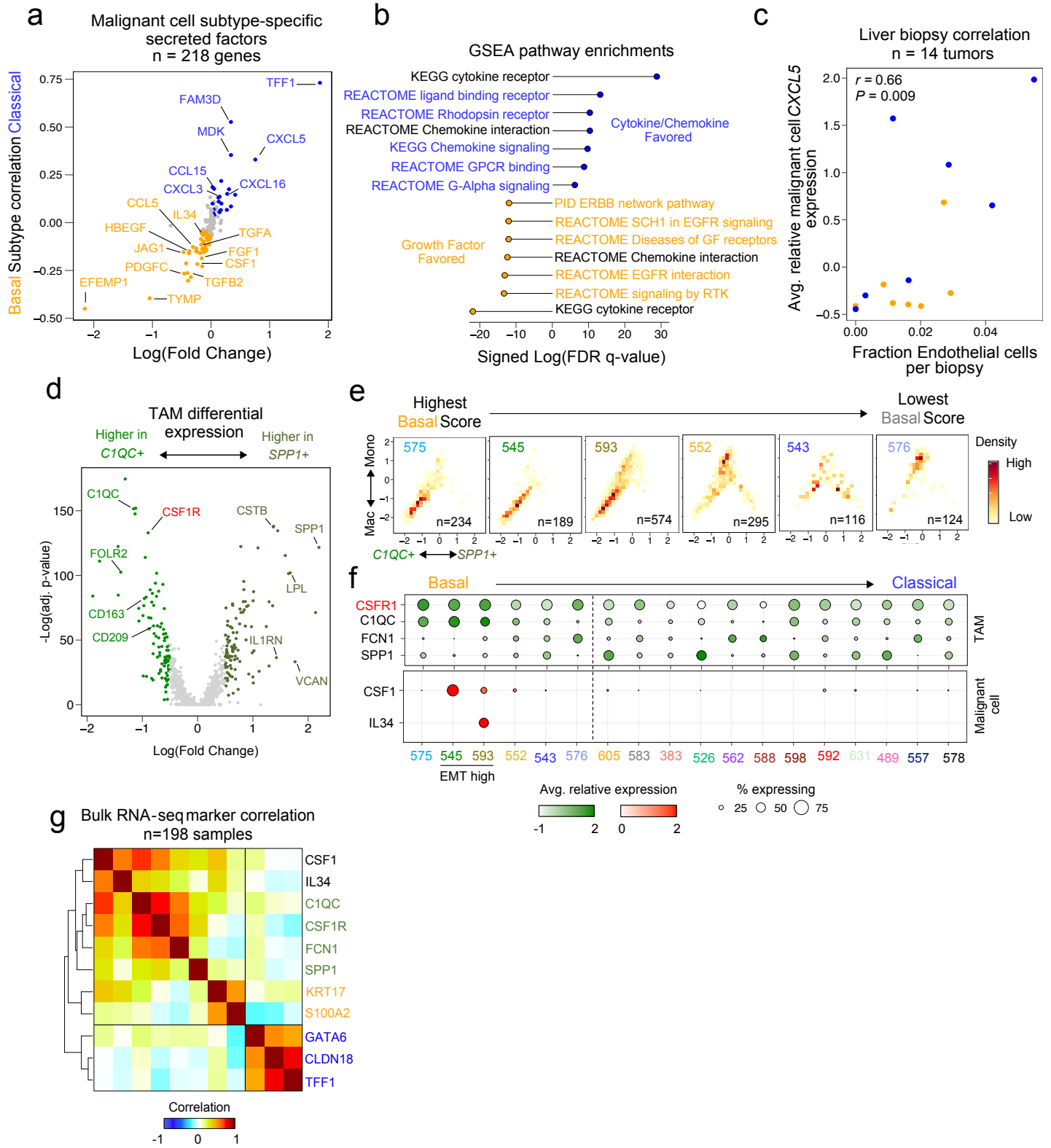
Figure 4

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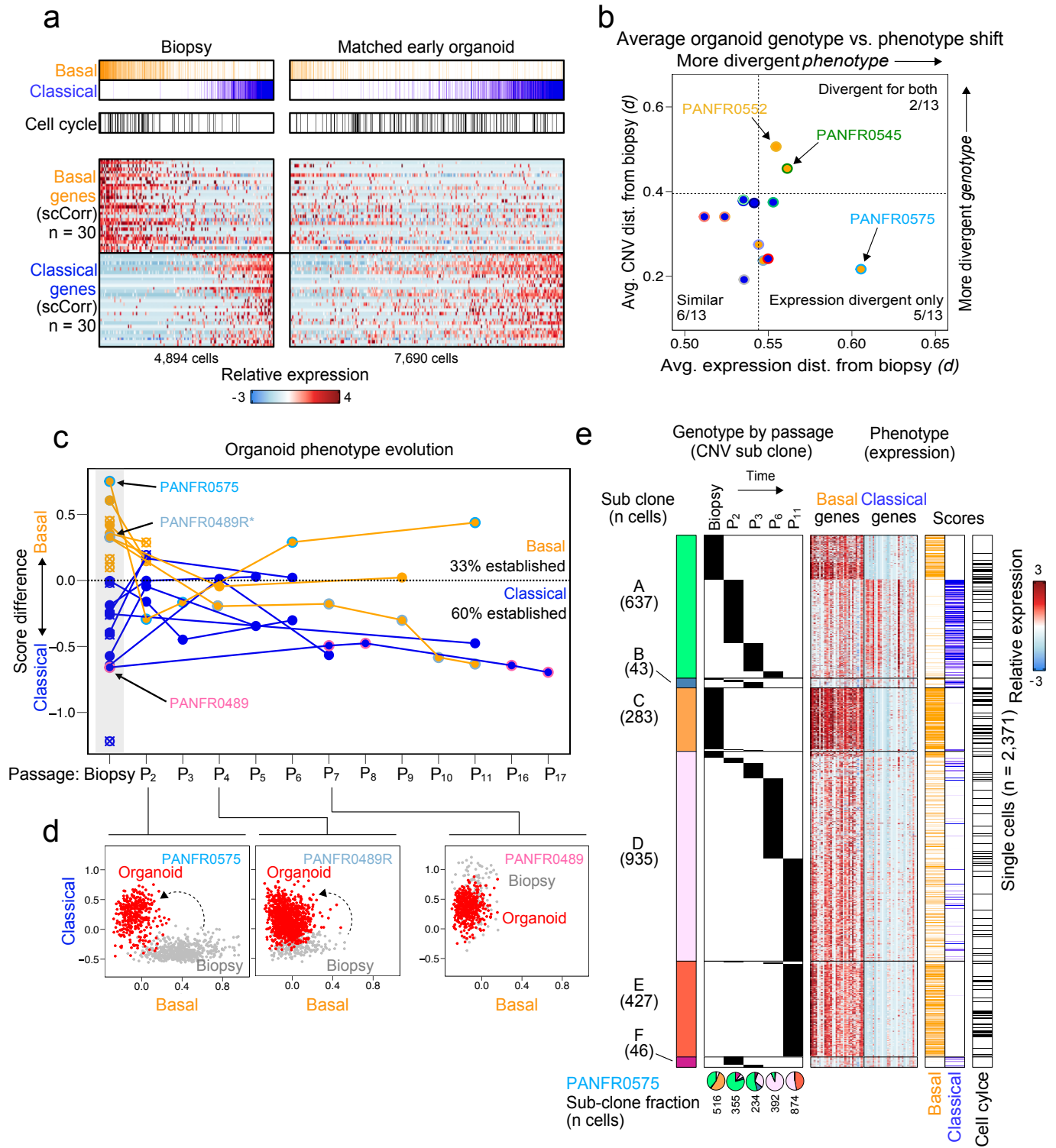
Figure 5. Organoid culture microenvironment selects against the basal state. a, Relative expression for basal and classical genes in biopsy cells (left) and their matched, early passage organoid cells (n=13 models; right). **b**, Scatterplot compares the relative contribution of genotypic drift (CNVs,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

21
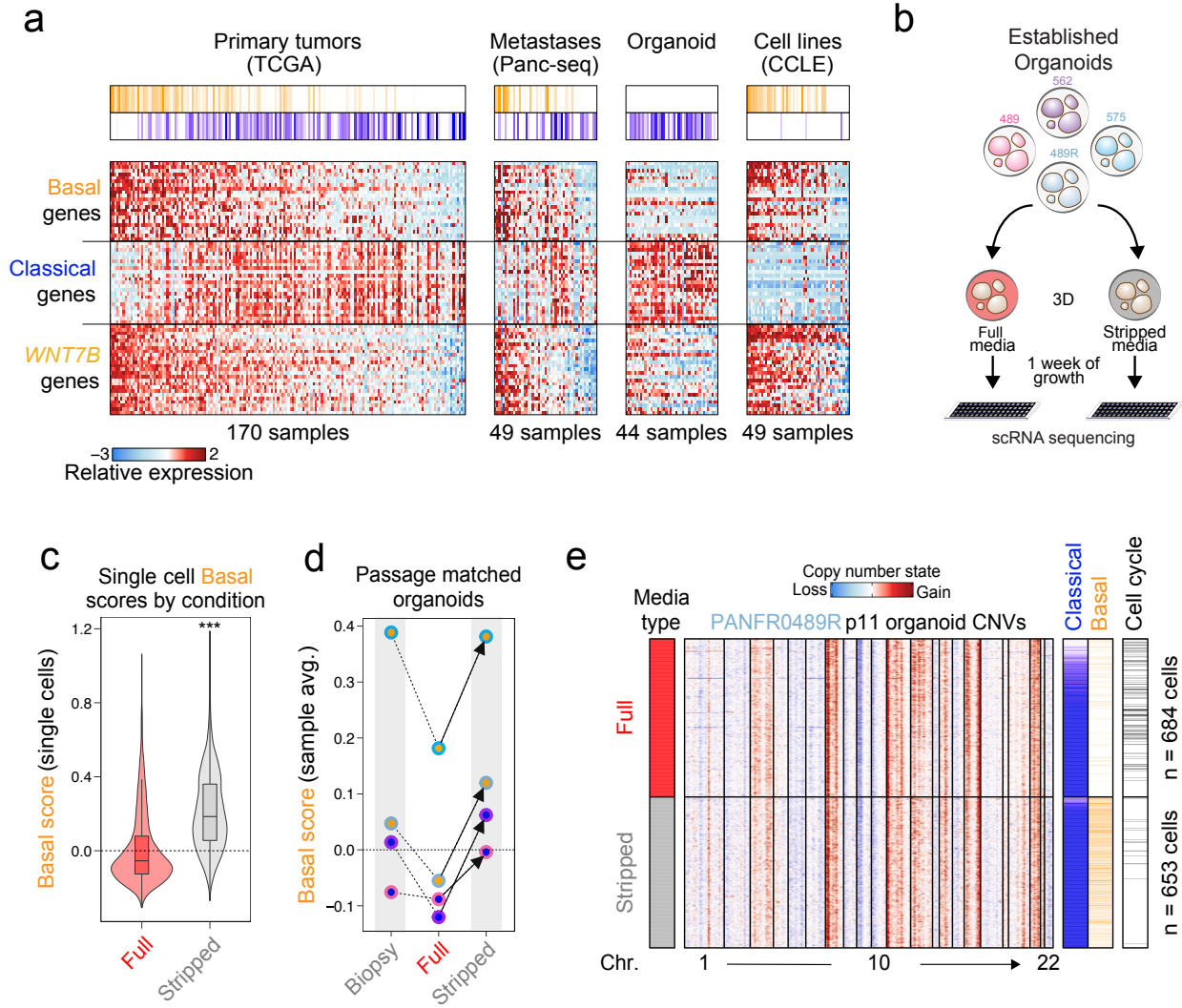
Figure 6

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 = 2.2x10^{-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.

22