TITLE

The cellular genomic diversity, regulatory states and networking of the metastatic colorectal cancer microenvironment

Authors

Anuja Sathe[1], Susan M. Grimes[2], Billy T. Lau[1], Xiangqi Bai[1], Jiamin Chen[1], Carlos Suarez[4], George Poultsides[3], Hanlee P. Ji[1,2]

Institutions

[1] Division of Oncology, Department of Medicine, Stanford University School of Medicine, Stanford, CA, United States

[2] Stanford Genome Technology Center, Stanford University, Palo Alto, CA, United States

[3] Department of Surgery, Stanford University, Stanford, CA 94305, United States

[4] Department of Pathology, Stanford University School of Medicine, Stanford, CA USA

Corresponding author

Hanlee P. Ji

Email: genomics_ji@stanford.edu

Mailing address: CCSR 2245, 269 Campus Drive, Stanford, CA-94305, USA

Running title: Single-cell genomics of metastatic colon cancer

Keywords: Single-cell genomics, tumor microenvironment, colon cancer, metastasis, ex vivo models

ABSTRACT

Metastatic colorectal cancer (**mCRC**) involves tumor cells seeding distant organs. Metastatic CRC genome biology has intrinsic properties related to heterogeneous clonal tumor cells and extrinsic properties of the cellular niche of the tumor microenvironment (**TME**). To characterize mCRC's cell types and states, we used single-cell RNA sequencing (**scRNA-seq**) and DNA sequencing (**scDNA-seq**) on metastases in the liver and omentum, a tissue lining the abdomen. We performed scRNA-seq on 49,637 cells derived from mCRC, paired normal tissue and peripheral blood. We performed whole genome scDNA-seq on 3,683 mCRC cells from the same samples. These metastases had intrinsic heterogeneity, with clone-specific copy number variation and lineage differentiation. For the extrinsic niche, TME macrophages had a cellular state resembling cirrhotic macrophages and foam cells with indicators of increased activity for extracellular matrix (**ECM**) organization. TME fibroblasts had an expression program influencing ECM composition that was not observed in matched normal tissue. Only a few CD8 T cells were present among the liver mCRCs, indicating immune exclusion. Receptor-ligand analysis revealed that TME macrophages and fibroblasts formed an interactome, thus providing intercellular coordination which increased T cell exhaustion and exclusion. For one mCRC, we used in a patient-derived *ex vivo* tissue model that captures the TME components of the original metastatic tumor, applied specific perturbations to the TME immune cells and evaluated cellular state changes with scRNA-seq. Overall, our study identified tumor cell clonal variation, characterized specific TME properties of the CRC metastatic niche and demonstrated an experimental model of perturbing the cellular TME milieu.

INTRODUCTION

Nearly 50% of all patients with colorectal cancer (**CRC**) have metastases, either presenting with distant tumors outside of the colon or developing metastatic recurrence after an initial diagnosis of cancer limited to the colon. The most common site for CRC's distant spread is the liver (Andres et al. 2015). The majority of patients with metastatic colorectal cancer (**mCRC**) receive systemic chemotherapy or targeted therapy. A small number of patients who have a limited number of metastases undergo surgical resection of the tumors from the liver. Regardless of the treatment, mCRC patients have a variable response to any type of therapy with only a small number of patients (17-25%) surviving beyond 10 years (Tomlinson et al. 2007). The cellular genomic diversity, regulatory states and intercellular interactions of any given mCRC contribute to the variation in therapeutic response and impact patient survival. Metastatic tumors are composed of both cancer cells and the surrounding cells of the distal organ site. However, characterizing mCRC's and their genome biology remains a significant challenge given this intermingling of different cell types and their diverse phenotypic characteristics.

Metastasis is a complex process with contributions from the "seed", i.e. migrating tumor cells and the "soil", i.e. the distant organ's cellular microenvironment (Liu et al. 2017). As an intrinsic cancer property, tumor cells have important cellular genomic features that include transcriptional regulatory state, molecular subtype, clonal heterogeneity and specific genomic alterations. These properties enable seeding, colonization and growth at a distant site. As an extrinsic cancer property, the cellular composition of the metastatic site plays a major role. The metastatic process depends on the remodeling of different cells in the tumor microenvironment (**TME**) which has organ-specific architecture. These extrinsic cellular characteristics influence patient prognosis and survival outcomes (Donadon et al. 2018; Pitroda et al. 2018). Citing an example, the liver has a variety of cell types that influence CRC metastasis such as specialized macrophages (i.e. Kupffer cells), endothelial cells and lymphocytes (Williamson et al. 2019). To

gain insights into the biology of mCRC, it is essential to characterize the cellular genomics of both the intrinsic tumor cells and the extrinsic surrounding TME.

Conventional sequencing analysis of cancer collectively averages the transcriptional and genomic DNA information from all of the cells in a heterogenous population such as the TME. For metastatic cancers, this cellular mixing prevents the determination of intrinsic and extrinsic cellular genomic features among the diverse cell types. Addressing the challenges of dissecting the TME cellular milieu, single-cell RNA (**scRNA-seq**) and single-cell DNA sequencing (**scDNA-seq**) provides information from each individual cell and reveals details about the resident populations that would be impossible to discern otherwise (Lim et al. 2020). Single cell gene expression studies have delineated the transcriptional signatures (i.e. cellular state) of resident cells for any given tumor (Lim et al. 2020; Sathe et al. 2020). Single cell DNA sequencing studies have revealed a broad spectrum of genomic alterations such as copy number variation that distinguishes different clonal populations (Leung et al. 2017; Andor et al. 2020; Velazquez-Villarreal et al. 2020).

Several scRNA-seq studies have characterized primary CRCs originating from the colon (Li et al. 2017; Lee et al. 2020; Zhang et al. 2020a). As described by these studies, the primary CRC epithelial cells had specific gene signatures associated with the cellular states of stemness and self-renewal. Moreover, the TME included a variety of different cell types such as fibroblasts, macrophages and other immune cells showing transcriptional differences compared to their counterparts in normal colon. Notably, the intercellular communication networks, as seen by ligand-receptor expression patterns, indicated an immunosuppressive phenotype among these TME cells.

There are only a limited number of studies describing the single cell genomics of metastatic CRCs. For example, Zhang et al. analyzed a single liver metastasis (Zhang et al. 2020b). The authors identified changes in the cellular composition of the TME compared to

normal tissue. There are no published studies using single cell genomics to study peritoneal mCRCs; these are metastatic tumors embedded in the peritoneum tissue which provides a lining of the abdomen. A genomic study using scDNA-seq, exome and targeted sequencing on two patients with mCRC revealed heterogeneity in the tumor epithelium, with evidence of both monoclonal and polyclonal seeding linked to the original progenitor population in the primary tumor (Leung et al. 2017). Heterogeneity in mCRC has also been detected using fluorescent in situ hybridization (Mamlouk et al. 2017) and conventional exome sequencing (Hu et al. 2019).

To characterize the intrinsic and extrinsic properties of metastatic CRC, we used a multi-omics approach with both scRNA-seq and scDNA-seq (**Fig. 1A**). We analyzed a cohort of six mCRC tumors surgically resected from the liver as well as one tumor present in the omentum, another abdominal site. Our study included paired normal tissues and peripheral blood derived cells. Comparing tumor cells to matched normal cells enabled us to identify discrete changes among the individual tumor and TME cells. As an added dimension of genome biology, we studied the functional characteristics of the mCRC TME using a patient-derived *ex vivo* tissue model. Given the cellular diversity of the tumor as found in a patient, we used specific molecules to perturb the mCRC TME followed by scRNA-seq to characterize the transcriptional changes seen among the different immune cell populations. Overall, our study identified specific TME cell properties of the metastatic CRC niche, the extent of tumor cell clonal variation at these sites and potential immunotherapeutic targets in the TME using an experimental *ex vivo* single cell analysis.

RESULTS

Single-cell analysis of colorectal cancer metastasis (mCRC)

We recruited six patients with mCRC, all had liver metastasis. These individuals underwent surgical resection following preoperative chemotherapy with a combination chemotherapy

Figure 1

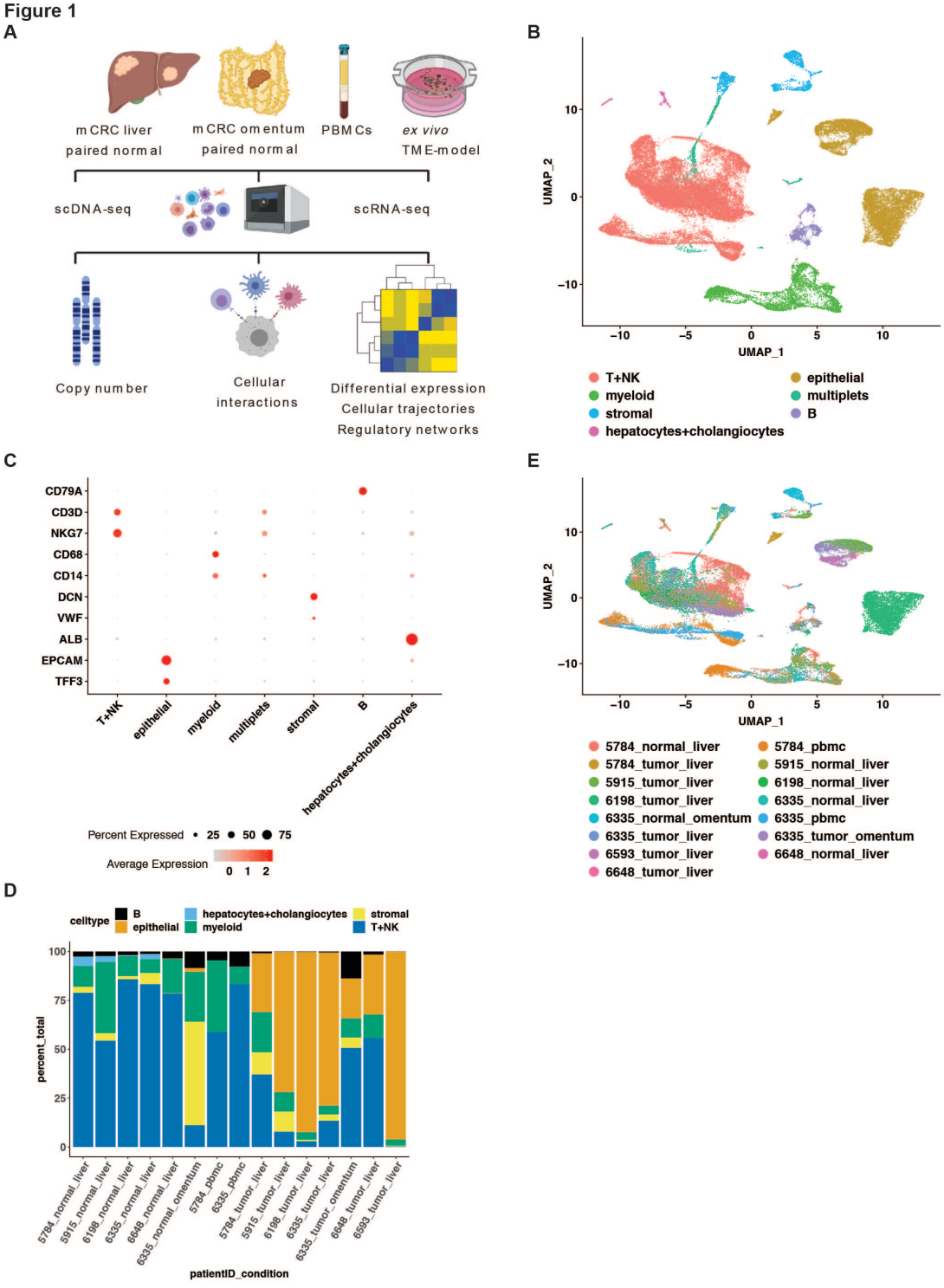

Figure 1: (A) Schematic representation of study design. (B) UMAP representation of dimensionally reduced data following graph-based clustering with marker-based cell type assignments. (C) Dot plot depicting expression levels of specific lineage-based marker genes together with the percentage of cells expressing the marker. (D) Proportion of cell types from each sample detected per cell lineage. (E) UMAP representation annotated with sample information.

involving 5-fluoruracil and oxaliplatin (**Supplemental Table 1**). From surgical resections, we obtained metastatic tumors. In addition, we had paired normal liver tissue from five of these patients and peripheral blood mononuclear cells (**PBMCs**) from two of them. From one patient, we had a second paired synchronous peritoneal metastasis located in the omentum, a tissue lining of the abdomen with matched normal omental tissue. All of these tumors underwent testing for microsatellite instability via immunohistochemistry for the DNA mismatch repair proteins. All of the mCRCs were microsatellite stable (**MSS**).

We sequenced a total of 49,637 single cells from these specimens (22,718 cells from normal liver, 14,553 cells from liver mCRCs, 2212 cells from normal omentum, 3,184 cells from omentum mCRC and 6,970 PBMCs) (**Supplemental Table 2**). For quality control, we filtered cells using quality control measures (11.7% of total cells) (**Supplemental Methods**). One filter involved eliminating cells expressing greater than 30% mitochondrial genes which indicates cell death (Ilicic et al. 2016). We excluded the doublet cells using a computational approach that predicts them based on their similarity with artificial doublets generated using gene expression profiles from randomly selected cells (McGinnis et al. 2019).

To compare the tumor cells with matched normal tissue cells and PBMCs, we aggregated single-cell data across all patients and samples. We normalized the data and performed variance-stabilization transformation by calculating Pearson residuals from a generalized linear model that accounts for technical variation in sequencing depth (Hafemeister and Satija 2019). Following principal component analysis, the top twenty principal components were used to construct a k-nearest neighbors (**KNN**) graph. Afterwards, the cells were iteratively grouped together using the Louvain algorithm (Butler et al. 2018). We performed non-linear dimensionality reduction using Uniform Manifold Approximation and Projection (**UMAP**) (McInnes and Healy 2018) to visualize the resulting clusters.

This procedure identified a total of 35 distinct clusters, each one having cells with similar transcriptional signatures. We identified differentially expressed genes among the clusters using the Wilcoxon rank sum test with genes expressed in greater than 25% of cells in a cluster, having a log fold change greater than 0.25 and a cut-off of $p < 0.05$ following Bonferroni correction. We compared differentially expressed genes in each cluster to literature-based marker genes to identify cell lineages (**Fig. 1B**, **Supplemental Table 3**). Clusters containing markers with multiple lineages were identified as multiplets (i.e. two or more cells in a partition) and were excluded from downstream analysis.

Representative of the organ's cellular architecture, we detected hepatocytes (*ALB*), cholangiocytes (*SOX9, KRT19, DEFB1*), tumor epithelial cells (*TFF3, EPCAM*), stromal cells including endothelial cells (*VWF*), fibroblasts (*DCN*, *COL1A1*) and smooth muscle (*LUM*, *ACTA2*). Representing the immune cell types, we detected myeloid lineage cells including monocytes, macrophages and dendritic cells (*CD14, FCGR3A, CD68, HLA-DP1, CD1C*) and T lymphocytes and NK cells (*CD3D*, *IL7R*, *CD8A*, *GNLY*) (**Figs. 1C, 1D)**. Cells of the same identity clustered together (**Fig. 1E**) regardless of the tissue sample, indicating no prominent batch effects. Next, we performed secondary clustering analysis for each cell lineage.

Malignant epithelial cells exhibit lineage differentiation

Among the tumor epithelial cells, we discovered cellular heterogeneity that involved transcriptional properties related to characteristics of stem cell (i.e. "stemness") and cell differentiation. The tumor epithelial cells formed patient-specific clusters indicative of transcriptional differences among the different mCRCs (**Fig. 2A, 2B**). For three mCRCs (P6198, P6335, P5915), we detected more than one tumor epithelial cell cluster indicative of intratumoral heterogeneity (**ITH**); subpopulations of distinct tumor cells had different gene expression patterns. From patient P6335, we analyzed a synchronous metastasis to the liver and the omentum.

Figure 2

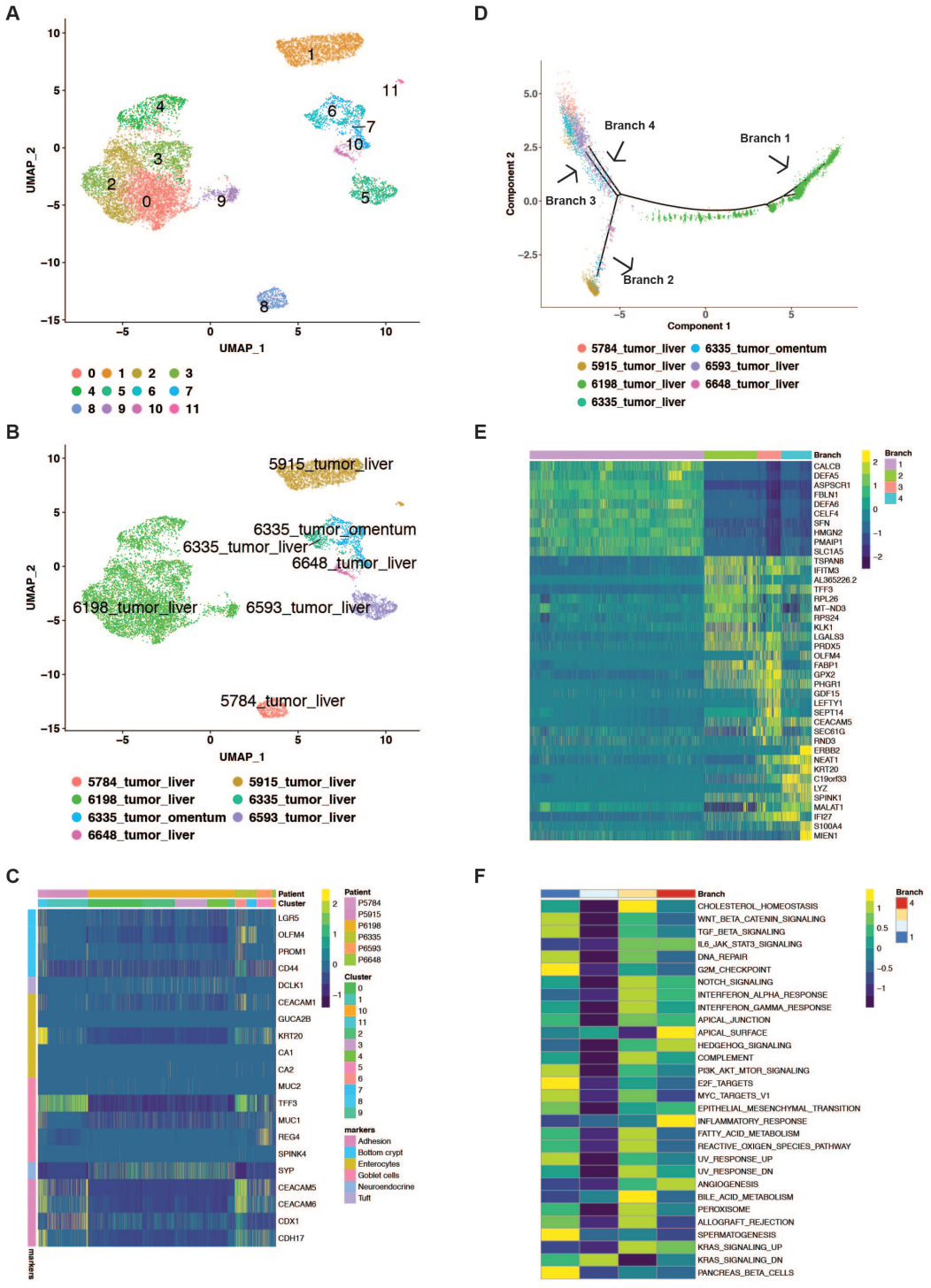

Figure 2: UMAP representation of dimensionally reduced data following graph-based clustering annotated by cluster number (A) or sample (B). (C) Heatmap depicting expression of respective genes per cluster and patient. (D) Trajectory plot of mCRC epithelial cells colored by sample origin and annotated by major trajectory branches 1-4. (E) Heatmap depicting the expression of highest significantly expressed genes per trajectory branch (adjusted p-value < 0.05). (F) Heatmap representing average GSVA enrichment score for respective hallmark pathway for each trajectory branch (ANOVA with Tukey HSD p-value < 0.05).

Cells from both these metastatic sites clustered together indicating similar transcriptional signatures and this result was consistent with a common origin, being the original primary CRC.

We investigated the expression of marker genes for cellular lineages found in the intestinal crypt (**Fig. 2C**). This milieu includes stem cells residing at the base of the crypt, tuft cells, enterocytes or absorptive cells, mucous secreting goblet cells, neuroendocrine cells and adhesion markers associated with CRC (**Supplemental Table 3**) (Li et al. 2017). Across all patients, we detected expression of enterocyte genes (*CEACAM1*, *KRT20*, etc.), goblet cell genes (*TFF3*, *REG4*, *MUC1*, etc.), adhesion genes (*CEACAM5*, *CEACAM6*, etc.) and stemness genes (*LGR5*, *OLFM4*, *PROM1*, *CD44*). Hence, all of these mCRCs had multilineage differentiation patterns observed in normal colon and primary CRCs (Li et al. 2017).

Single cell analysis enabled us to characterize tumor cell populations with different properties, even in the same tumor. Otherwise, these differences may have been overlooked. One of the samples represented a mixture of tumor cell types. Referred to as mixed adeno-neuroendocrine carcinoma (**MANEC**), this colon cancer has both epithelial and neuroendocrine cellular components. P6198's mCRC had higher differential expression of the *SYP* gene, which was also confirmed on clinical immunohistochemistry (**IHC**) staining (**Supplemental Table 1**). Also referred to as synaptophysin, SYP expression is a classic marker for neuroendocrine cancers. However, these tumors cells also expressed epithelial tumor markers such as *MUC1*, *DCLK1* or *LGR5* which are indicative of goblet cell, tuft cell or stemness markers respectively.

Differential gene expression and pathway activation in mCRC tumor epithelium

We performed differential expression analysis at the gene and pathway level to characterize the heterogeneity of tumor epithelial cells. In addition, differences in lineage differentiation and activation of specific signaling pathways were evidence of both interpatient and intrapatient tumor cellular heterogeneity.

Differentially expressed genes in adenocarcinoma mCRCs included the markers *FABP1*, *OLFM4*, *KRT20*, *CEACAM5* and *CEACAM6* (Li et al. 2017) (**Supplemental Fig. 1A**). Expression of these genes is associated with CRC. For example, the product of *CEACAM5* is a carcinoembryonic antigen (**CEA**), which is a protein biomarker in the blood and used for longitudinal monitoring of mCRC. In addition, we detected high expression levels of *TSPAN8* and *HES1*, which are associated with the tumor properties of stemness and invasion (Candy et al. 2013; Zhu et al. 2019). Expression of these genes varied between tumors as well within sub-clusters of the same tumor in P5915 and P6335. Increased *ERBB2* expression was detected in P5784's and P6593's tumors – the clinical immunohistochemistry results from these samples corroborated ERBB2 overexpression (**Supplemental Table 1**). *ERBB2* amplification, encoding the HER2 protein, is a predictive biomarker for the trastuzumab antibody against this receptor in mCRC (De Cuyper et al. 2020). The variation among the tumors was also noted for hallmark pathway activity including for signaling in apical junction complexes and apical surfaces that regulate epithelial cell polarity, Wnt, Notch and PI3K. Moreover, there was variation in gene expression for the angiogenesis pathways and epithelial mesenchymal transition (**Supplemental Fig. 1B**). One of two sub-clusters in both P5915 and P6335 demonstrated a significant downregulation of these pathways compared to normal epithelial cells.

P6198's MANEC metastasis had significantly increased expression of the *DEFA5* and *DEFA6* genes. These genes have been described to be enriched in small intestinal neuroendocrine tumors (Wang and Yu 2020) (**Supplemental Fig. 1A**). P6198 sub-clusters varied in the expression of genes related to differentiation including: *ID3* which is a helix-loop-helix protein involved in transcriptional regulation; transcription factors *JUN*, *JUNB* and *ATF3*. We also observed differential expression of several genes involved in cell cycle regulation such as *MDK*, *CDKN1A*, *HMGB2* and *TOP2A*. Differences in hallmark pathway activity were most evident for

cell cycle related processes (**Supplemental Fig. 1B**) indicative of a sub-cluster with higher rate of proliferation.

Differences in gene expression are subject to confounding effects due to stochasticity. As a solution to this issue, the pseudotime analysis method reconstructs cellular trajectories of cells progressing into stable states representing the final differentiation endpoint (He et al. 2017). We applied pseudotime analysis to our CRC samples and detected four main branches in the trajectories of the mCRC epithelial cells (**Fig. 2D, Supplemental Table 4**). Subsequently, we determined the differential gene expression (**Fig. 2E**) and hallmark pathway activation (**Fig. 2F**).

Branch 1 contained cells from P6198's MANEC metastasis with significant expression of the *DEFA5* and *DEFA6* genes. Branch 4 was dominated by P5784's and P6593's tumors, both of which had an *ERBB2* (HER2) amplification. For these mCRCS, the high expression of *ERBB2* occurred jointly with differential expression of *KRT20*, *SPINK1* and others. This gene expression signature was an indicator of increased apical junction proteins as well as signaling pathways. Branches 2 and 3 had the majority of cells from P5915 and P6335 respectively, albeit with minor contributions from other mCRCs. Cells in Branch 1 shared high expression of several CRC associated genes including *TFF3*, *FABP1, CEACAM5* and others.

Compared to Branch 2 with genes indicating increased apical signaling, Branch 3 was significantly enriched for stemness-related genes such as *OLFM4*, *LEFTY1* and *SEPT14*. This corresponded with an increase in Wnt, TGF-β and Hedgehog signaling in this branch. Overall, the trajectory analysis confirmed that differences in mCRC were accounted for by variation in cancer-related signaling networks.

Influence of copy number variation on tumor heterogeneity

Based on scDNA-seq analysis, tumor heterogeneity in mCRCs was evident based on differences in copy number variation (**CNV**) and aneuploidy changes of chromosomes. To identify clonal CNVs, we performed whole genome single cell sequencing on a subset of four mCRC samples (P5915, P6593, P6335 and P6198). We modeled per-cell read counts per genomic bin as a Poisson distribution dependent on both the GC content and the copy number as previously described (Andor et al. 2020), using Cell Ranger. The GC bias was modeled as a quadratic function with fixed intercept and correction on a cell-by-cell basis was performed. To estimate copy number for each bin, we empirically computed the effect of GC content followed by scaling to generate haploid-scaled copy number calls. To identify candidate breakpoints, we calculated the discontinuity in copy number values among all mappable bins using the log-likelihood ratio statistic. CNVs were generally restricted to regions of the genome (85-90%) where reads could be confidently mapped. We determined single cell copy number profiles with an average median effective coverage of 230 reads per 1 Mb per cell and an average median estimated CNV resolution of 1.38 Mb (**Supplemental Table 5**). Using the CNV segments, we performed hierarchical clustering to determine if there were clonal differences among the epithelial tumor cells (**Fig. 3A**). Importantly, CNVs were not evident in the normal epithelial cells across all samples (**Supplemental Fig. 1C**).

All of the mCRCs had significant levels of aneuploidy and deletions, indicating that these MSS tumors belong to the chromosomal instability (**CIN**) molecular subtype (Guinney et al. 2015). For example, chromosomes 7, 13 and 20 either had trisomy or tetrasomy across all mCRCs. These aneuploidy events are associated with metastasis in CRC (Mamlouk et al. 2017; Hu et al. 2019). We also identified deletions in chromosomes 18 and 21 in all adenocarcinoma mCRCs (P6335, P5915, P6593) and chromosome 8 deletion in P5915. These genomic deletions are common in mCRC (Mamlouk et al. 2017; Hu et al. 2019). Deletions in chromosomes 18 have been identified as a marker for increased risk of metastasis among patients with limited stage

CRCs (Nguyen and Duong 2018). Amplifications in chromosome 13 and 20 have been associated with the development of metastasis in CRC (Haan et al. 2014).

Interestingly, the MANEC metastasis (P6198) had these same events as well as a complete loss of chromosome 8p. While CNV analysis specifically from neuroendocrine tumors has not been extensively described, their mutational landscape is known to resemble colorectal adenocarcinoma (Takizawa et al. 2015). Therefore, our study showed that MANEC tumors with complex differentiation patterns can be linked to genomic instability events characteristic of mCRCs.

Importantly, each mCRC had heterogenous distribution of different CNVs and aneuploidy events. For P6335's mCRC, aneuploidy was evident among 146 of the epithelial tumor cells. Based on different distributions of aneuploid and amplification events on chromosome 3, 11, 13 and 20, there were three major clones and several minor sub-clones (**Fig. 3A**). For P6198's mCRC, 1,862 aneuploid cells contained two major clones that had different cellular distributions for amplification of chromosome 12 with smaller groups of cells varying in chromosome 1, 13 and 19 amplification. For P5915's mCRC, we detected deletions in chromosomes 2, 3, 4, 8, 11 and 17 in a subset of cells. Similarly, P6593's CRC had aneuploid cells that were heterogenous for amplifications in chromosomes 3, 13 and 20. For both these samples, there was a lower number of total cells for scDNA-seq because of the small tissue biopsy and the reduced tumor cell counts as a result of prior chemotherapy (**Supplemental Table 5**).

To understand the impact of this CNV heterogeneity on single-cell gene expression, we inferred the copy number of tumor epithelial cells based on their scRNA-seq derived gene expression using the InferCNV tool (Tickle et al. 2019) (**Fig. 3B**). With transcript counts, one identifies large-scale chromosomal CNV events by using increase or decrease in gene

Figure 3

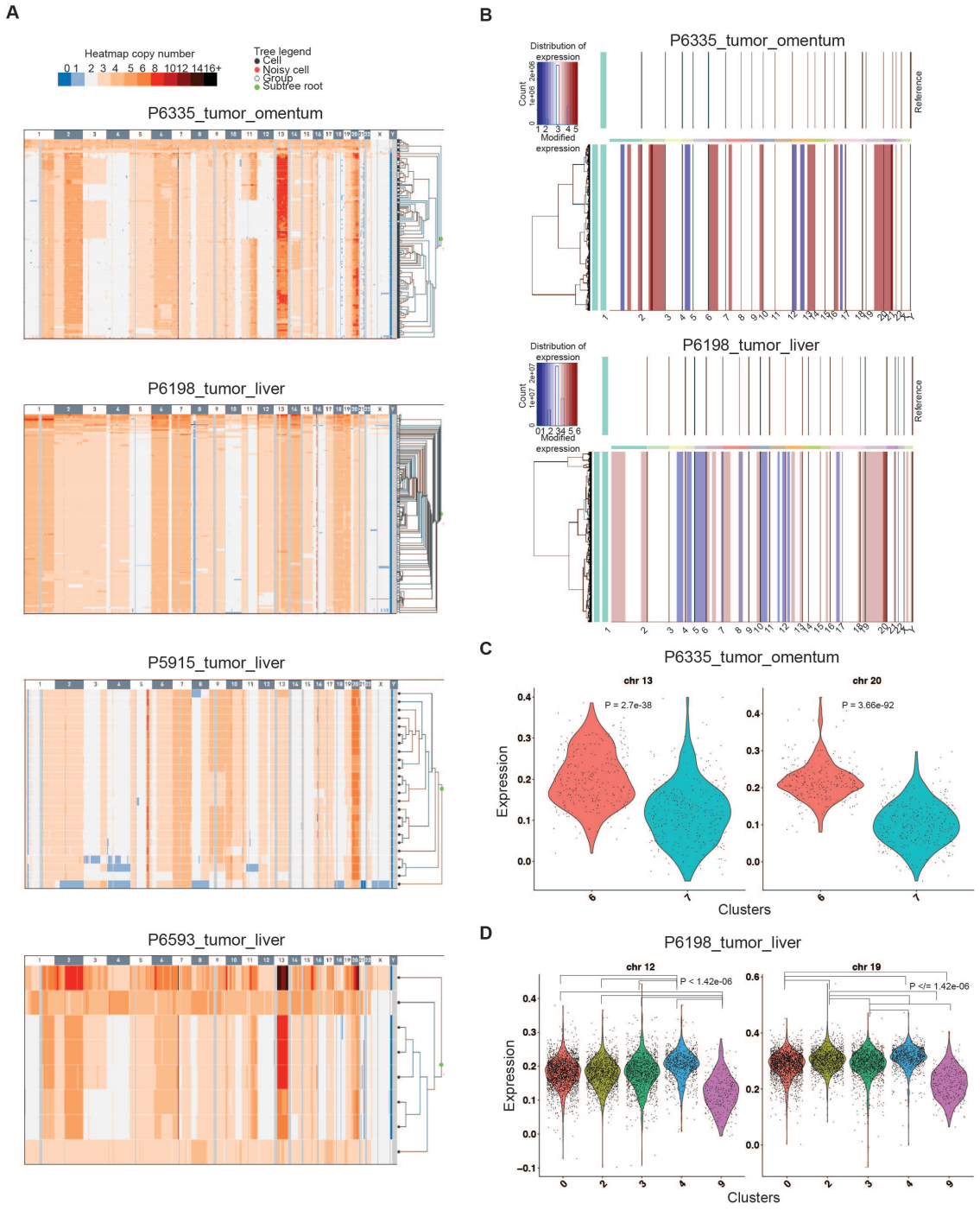

Figure 3: (A) Hierarchical clustering of CNV profiles obtained from scDNA-seq for respective patients. Number of aneuploid cells analyzed per patient: 146 in P6335, 1862 in P6198, 8 in P5915, 24 in P6593 respectively. (B) Inferred copy number profiles for respective patients using scRNA-seq compared to stromal cells from all samples as a diploid reference control. Expression states indicate CNV levels as follows: 1= complete loss, 2 = loss of one copy, 3 = neutral, 4 = addition of one copy, 5 = addition of two copies, 6 = greater than 2 copies. (C-D) Violin plots depicting expression score for genes on respective chromosomes in each epithelial cell sub-cluster in (C) P6335 with T-test p-value and (D) P6198 with ANOVA and Tukey HSD p-value respectively.

expression as a surrogate for gain or loss of copies respectively. Albeit, the resolution of a given CNV is on the order of Mbs and does not provide the resolution of scDNA-seq.

We used stromal cells derived from all patients as a diploid reference control. We calculated an average expression score per chromosome on the basis of genes identified from this analysis (**Supplemental Methods**) (Tirosh et al. 2016). For P6335's tumor, the average expression score for genes on chromosomes 13 and 20 demonstrated an increase in copy number and was significantly different between the two epithelial sub-clusters (**Fig. 3C**). In P6198, several sub-clusters differed significantly in the expression of genes on chromosomes 12 and 19 (**Fig. 3D**). This result identified the importance of gene dosage with some aneuploidy manifesting themselves with a corresponding increase in gene expression.

Interestingly, some DNA-based CNV events were not observed when inferring CNVs from the single cell gene expression data. Thus, gene dosage events as a result of trisomy or higher chromosome number increases did not always lead to increased expression for the affected chromosomes. For P6335's mCRC, we noted this difference for the clonal branch having chromosomes 3 and 11 aneuploidy. For P6198's mCRC, a similar difference was observed for the chromosome 4 aneuploidy. This result is an indicator of additional mechanisms such as epigenetic regulation that alter gene expression and contribute to phenotypic transcriptional heterogeneity (Stamoulis et al. 2019).

Metabolically reprogrammed mCRC macrophages influence ECM re-organization

We discovered that macrophages in the metastatic TME had a single distinct transcriptional state with a gene profile resembling cirrhotic macrophages and foam cells. This metastatic TME state is significantly different that the dichotomous states of macrophages in primary CRC (Zhang et

al. 2020a). Notably, this transcriptional state was not present in normal liver tissue macrophages or peripheral blood monocytes.

To analyze the macrophage single cell data, we performed re-clustering, differential expression analysis and comparisons to lineage marker genes on cells belonging to the myeloid lineage (**Supplemental Fig. 2**). This analysis process enabled the identification of the following cell types: PBMC-dominant *CD14* or *FCGR3A* (CD16) expressing monocyte sub-populations; dendritic cells in the liver (HLA genes, *CD1C*, *CLEC9A*, *IDO1*, etc.); Kupffer cells in the liver (*CD5L*, *MARCO*, *LIPA*, *MAF*, etc.); proliferating cells; mCRC-specific macrophages; additional tissue-monocyte/macrophage clusters detected in normal liver and normal omentum (Villani et al. 2017; MacParland et al. 2018). Cells did not express neutrophil markers (**Supplemental Fig. 2C)**.

Tumor macrophages from mCRCs clustered separately from macrophages in normal tissue and PBMC monocytes (**Fig. 4A**). This profile was different than the classic M1/M2 profiles of macrophage polarization (Martinez et al. 2006; Martinez and Gordon 2014) (**Fig. 4B**). Differential expression analysis revealed significantly higher expression of genes involved in fibrosis including *SPP1*, *TREM2*, *LGALS3* and *CD9*. This transcriptional profile is similar to what is observed among macrophages in liver cirrhosis (Ramachandran et al. 2019) (**Fig. 4C, Supplemental Table 6**). These cells also had high expression of genes such as *APOE* and *APOC* which are involved in cholesterol metabolism. Importantly, the gene expression patterns were similar to what is observed in foam cell macrophages located in atherosclerotic plaques, an obstructive lesion of arterial vessels (Fernandez et al. 2019). Pathway analysis revealed significant enrichment of terms relating to ECM organization and metabolism including HDL-mediated lipid transport, glycosphingolipid and glucose metabolism (**Fig. 4D, Supplemental Table 7**). These metabolic changes impact macrophage function and polarization, a concept referred to as 'immunometabolism' (Viola et al. 2019).

Figure 4

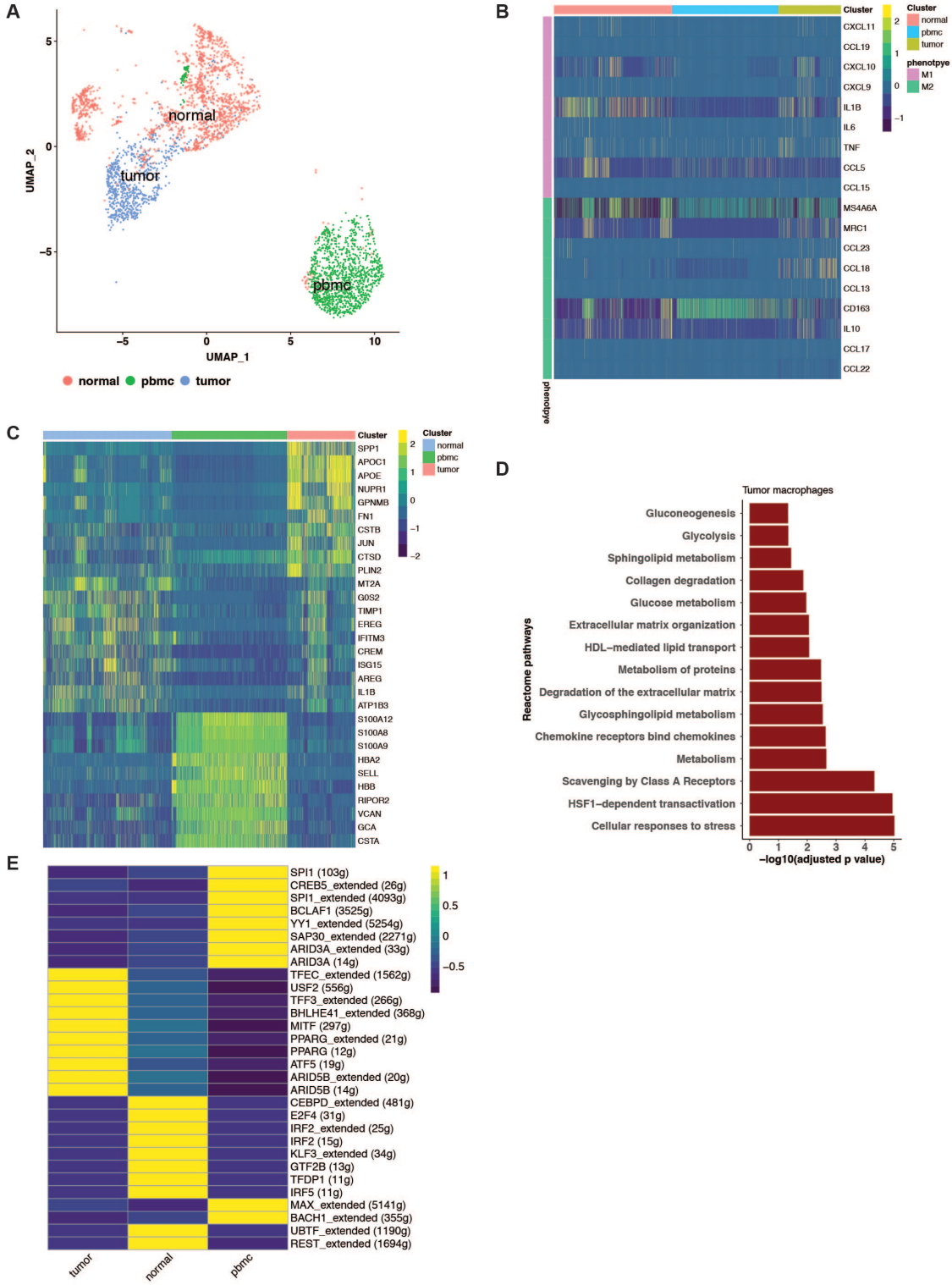

Figure 4: (A) UMAP representation of dimensionally reduced data following graph-based clustering annotated by sample condition. (B) Heatmap depicting expression of respective genes per cluster. (C) Heatmap depicting the expression of highest significantly expressed genes per cluster (adjusted p-value < 0.05). (D) Selected differentially enriched reactome pathways in tumor macrophages. (E) Heatmap depicting average AUC activity of top regulons (ANOVA with Tukey HSD p-value < 0.05) per cluster. Number in parenthesis indicates number of genes in the gene regulatory network.

We inferred the transcriptional regulators of this phenotype by examining the activity of transcription factor and co-factor controlled gene regulatory networks in single cells (Aibar et al. 2017). Tumor macrophages, normal macrophages and PBMC monocytes showed significant differences in the activity of specific gene regulatory networks (**Fig. 4E).** Notably, tumor macrophages were enriched for *PPARG* and *MITF* activity providing a molecular link to the metabolic phenotype that we noted previously using differential gene expression and pathway analysis. *PPARG* encodes a member of the peroxisome proliferator-activated receptor family and regulates lipid pathway signaling and the formation of foam cells in atherosclerosis (Kuznetsova et al. 2020). The *MITF* gene is a transcriptional factor that is known to regulate lipid accumulation in adipose tissue macrophages as well as their inflammatory phenotype (Gabriel et al. 2014).

Metastatic TME has immune exhaustion and exclusion in infiltrating CD8 T cells

The majority of liver mCRCs had relatively few CD8 T cells, suggesting an immune excluded phenotype where T cells are blocked from migrating into the confines of the tumor and surrounding TME (Binnewies et al. 2018). We directly observed this phenomenon with the liver TMEs which contained an average of 133 CD8 T cells. In contrast, paired normal liver contained diverse immune cell types with higher numbers. Interestingly, we detected a rich immune cell infiltrate in the omental metastasis with around 620 CD8 T cells. However, the majority of these cells had characteristics of immune exhaustion. This phenotype is characterized by low cytotoxic potential and the upregulation of multiple immune checkpoints leading to an ineffective anti-tumor response (Wherry and Kurachi 2015).

First, we analyzed all immune cells from mCRCs, PBMCs and normal tissue together revealing 19 clusters (**Supplemental Fig. 3 A, B, C**). Based on their expression of lineage markers, we detected a rich immune infiltrate containing CD4 T, CD8 T, NK, gamma delta T, NK-like, NK-T, MAIT atypical T, plasma and B cells in the normal liver (**Supplemental Fig. 3D**).

Similar to previous findings (MacParland et al. 2018), typical and atypical T and NK cells were mixed in some clusters. We also detected these cell lineages in the normal omentum. As one would expect, T and NK cells in PBMCs clustered separately from the same cell types in the liver and omental tissue; this cluster pattern were indicators of tissue-specific transcriptional differences which we have observed in gastric cancers (Sathe et al. 2020).

We re-clustered tumor-specific lymphocytes from these cells to characterize them in greater detail. The omentum metastasis had a diverse immune infiltrate containing CD4 T cells, heterogenous CD8 T cells, NK cells, regulatory T cells (**Tregs**), plasma and B cells (**Fig. 5 A, B, C**). The CD4 cells expressed *CXCL13* (c10_CD4) and other multiple immune checkpoints (**Supplemental Table 8**). These cells have the potential to regulate TME immune response by affecting the balance between anti-tumor B cells and immunosuppressive Tregs (Gu-Trantien et al. 2017). We also detected the expression of several checkpoint and costimulatory molecules on Tregs (c8_Treg). Tregs have an immunosuppressive role that suppresses any anti-tumor immune cells (Ohue and Nishikawa 2019). Liver mCRCs were comparatively devoid of infiltrating immune cells (**Fig. 5D**), indicating an immune excluded phenotype (Binnewies et al. 2018).

CD8 T cells mediate an effective anti-tumor response. The omentum metastasis (P6335) had with three distinct sub-clusters of CD8 T cells (c1_CD8, c6_CD8, c11_CD8) and a single cluster was detected from liver mCRCs predominantly from P5784 (c2_CD8) (**Fig. 5D**). We examined the expression of known immune checkpoints, costimulatory molecules and cytotoxic effector genes in these cells (**Fig. 5E**). Then we compared these results to previously described gene signatures of cytotoxicity, exhaustion and proliferation (**Fig. 5F**, **Supplemental Methods**). CD8 cells from liver mCRCs had significantly lower cytotoxicity (ANOVA FDR p <

Figure 5

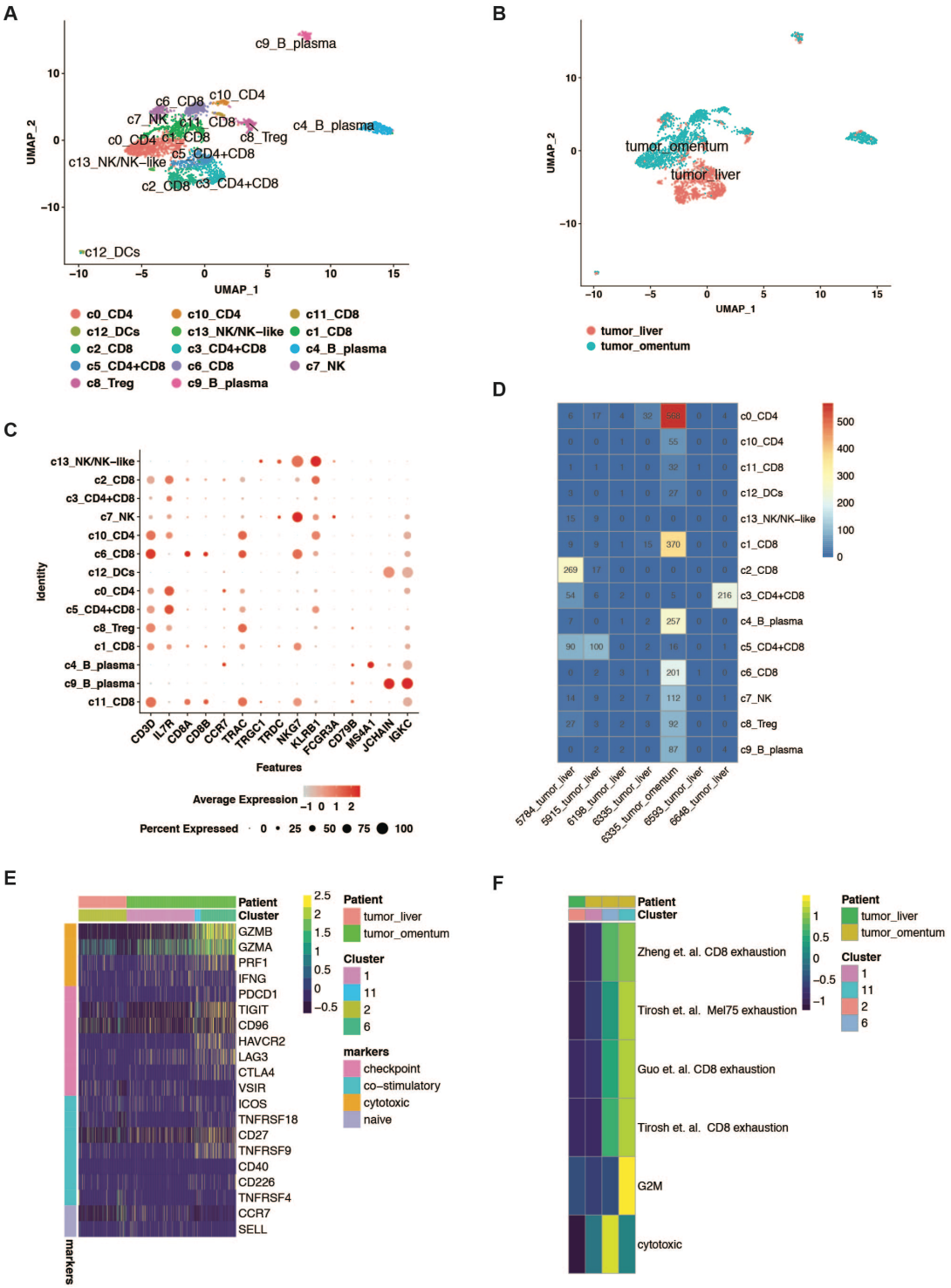

Figure 5: UMAP representation of dimensionally reduced data following graph-based clustering annotated by (A) clusters and (B) sample condition. (C) Dot plot depicting expression levels of specific lineage-based marker genes together with the percentage of cells expressing the marker. (D) Heatmap depicting number of cells from each sample detected in each respective cluster. (E) Heatmap depicting the expression of respective genes in each cluster. (F) Heatmap representing average GSVA enrichment score for respective pathway for each cluster (ANOVA with Tukey HSD p-value < 0.05).

5e-6). The subsets of exhausted T cells in omentum metastasis varied in the extent of enrichment of cytotoxic, exhaustion and proliferation profiles as we have noted previously (Sathe et al. 2020). Exhausted cells expressed multiple immune checkpoints including significant expression of the exhaustion markers *HAVCR2* (TIM3), *TIGIT* and *LAG3* as well as the costimulatory molecule *TNFRSF9* (41-BB) (**Fig. 5E, Supplemental Table 8**). These cells also significantly expressed *CXCL13* and *RBPJ* that we have previously identified in exhausted cells in the gastric cancer TME (Sathe et al. 2020). Thus, we detected a rich infiltration of immune cells with a pronounced T cell exhaustion phenotype in the omentum mCRC and few T cells in the majority of liver mCRCs suggestive of an immune excluded phenotype (Binnewies et al. 2018).

mCRC fibroblasts have a tumor-promoting gene expression program

We discovered transcriptional features of mCRC cancer-associated fibroblasts (**CAFs**) that influence ECM organization and tumor growth, which were absent in the normal tissue counterparts and resembled stromal changes in primary CRC (Lee et al. 2020). After re-clustering and differential expression analysis of stromal cells (**Fig. 6A, 6B, Supplemental Table 9**) we identified the following: omentum-derived mesothelial cells (cluster 10) (Jackson-Jones et al. 2020); omentum-specific mesenchymal cells expressing ECM glycoproteins and proteoglycans (clusters 0, 1,12); mCRC enriched CAFs (clusters 2, 6); endothelial cell subsets (clusters 4, 5, 7, 8, 13); a mixed cluster of hepatic stellate cells and vascular smooth muscle cells (cluster 3).

We compared the differentially expressed genes from the CAFs to the 'matrisome', a set of genes which represents the collection of ECM and ECM-associated proteins (**Supplemental Table 10, Fig. 6C)** (Naba et al. 2012). These cells expressed several collagen genes (*COL1A1*, *COL3A1*, *COL5A1*, etc.) in their core matrisome program together with a variety of ECM

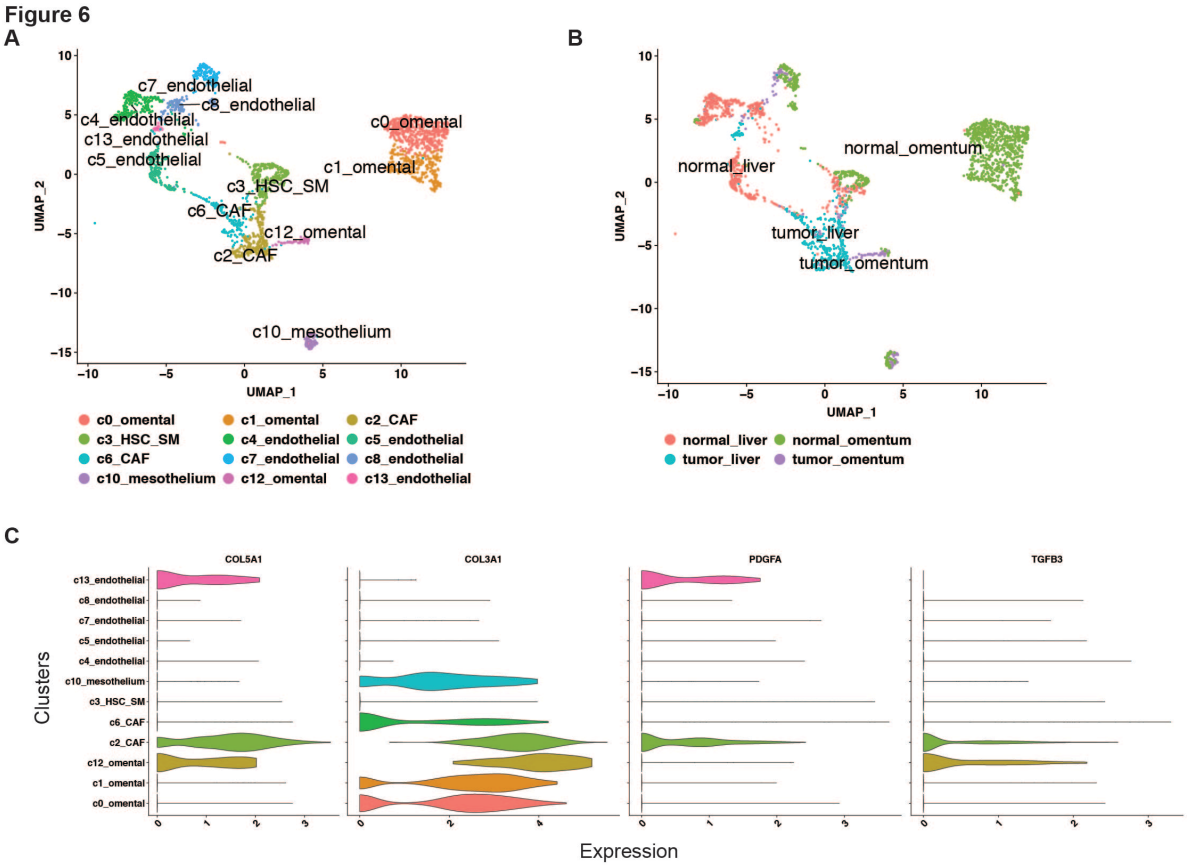

Figure 6: UMAP representation of dimensionally reduced data following graph-based clustering annotated by cluster number (A) or sample condition (B). (C) Violin plots depicting expression level of respective genes in each cluster.

glycoproteins (*FN1*, *TGFBI*, *THBS2*, etc.) and proteoglycans (*BGN*, *VCAN*, etc.). Highly expressed ECM regulators included *MMP11*, *MMP14*, *TIMP1*, *LOXL1* and *LOXL2*. These genes enable ECM remodeling (Liu et al. 2019). The resulting ECM composition influences physical properties such as stiffness and contributes to tumor growth and drug resistance (Shen et al. 2020). Secreted factors included *TGFB1*, *TGFB3*, *VEGFA*, *PDFGFA*, *PDGFC* that can promote tumor growth as well as enable immune evasion (Sahai et al. 2020).

We inferred the gene regulatory networks in CAFs based on the expression of target genes (Aibar et al. 2017). The *ATF3*, *BCLAF1*, *CEPB*, *YY1*, *EGR1*, *GT2F1* and *TAF7* regulons had an average AUC >0.3 revealing the transcription factors that control the CAF phenotype (**Supplemental Table 11**).

Fibroblasts and macrophages form an intercellular signaling network

We discovered a receptor-ligand mediated mutual networking between macrophages and fibroblasts; the two cell types influenced each other's transcriptional phenotype. We also discovered intercellular interactions between these cells that impact both the immunocellular phenotypes of T cell exclusion and exhaustion.

First, we identified cell-type specific receptor-ligand interactions to construct a mCRC TME interactome with a high number of unique interactions mediated by fibroblasts, macrophages and epithelial cells (**Supplemental Fig. 4A, Supplemental Table 13**) (Vento-Tormo et al. 2018). Prominent among these interactions were fibroblast-immune *CXCL12-CXCR4* receptor ligand pair. These genes mediate one of the mechanisms that drives immune suppression in the TME (Chen et al. 2019a) (**Supplemental Fig. 4B**). Namely, the CXCL12 ligand and its co-receptor CXCR4 regulate the mobilization of specific immune cells into tissues. We also detected genes belonging to the Nectin family as well as *LGALS9* that serve as ligands for immune checkpoints *TIGIT* and *HAVCR2* (TIM-3) respectively (Fujihara et al. 2013; Wu et al. 2014; Gorvel and Olive 2020). In addition, we observed differential expression of the macrophage-derived gene *SPP1* that suppresses T cell activation via interaction with *CD44* (Klement et al. 2018).

Next, we identified receptor-ligand interactions that are predicted to regulate specific transcriptional phenotypes in the TME. Our analysis relied on ligand-target and ligand-receptor gene models constructed using published literature (Browaeys et al. 2020). We identified the top ranked ligands that regulate mCRC fibroblasts and its matrisome expression program (**Fig. 7A**). The majority of ligands were secreted by macrophages and fibroblasts (**Fig. 7B**). Confirming the biological validity of this ECM expression program, among the high ranked genes was the ECM regulator *TGFB1*. Multiple studies have established this gene's important role in ECM organization (Hinz 2015). Also, we identified the macrophage secreted ligand *SPP1* which

Figure 7

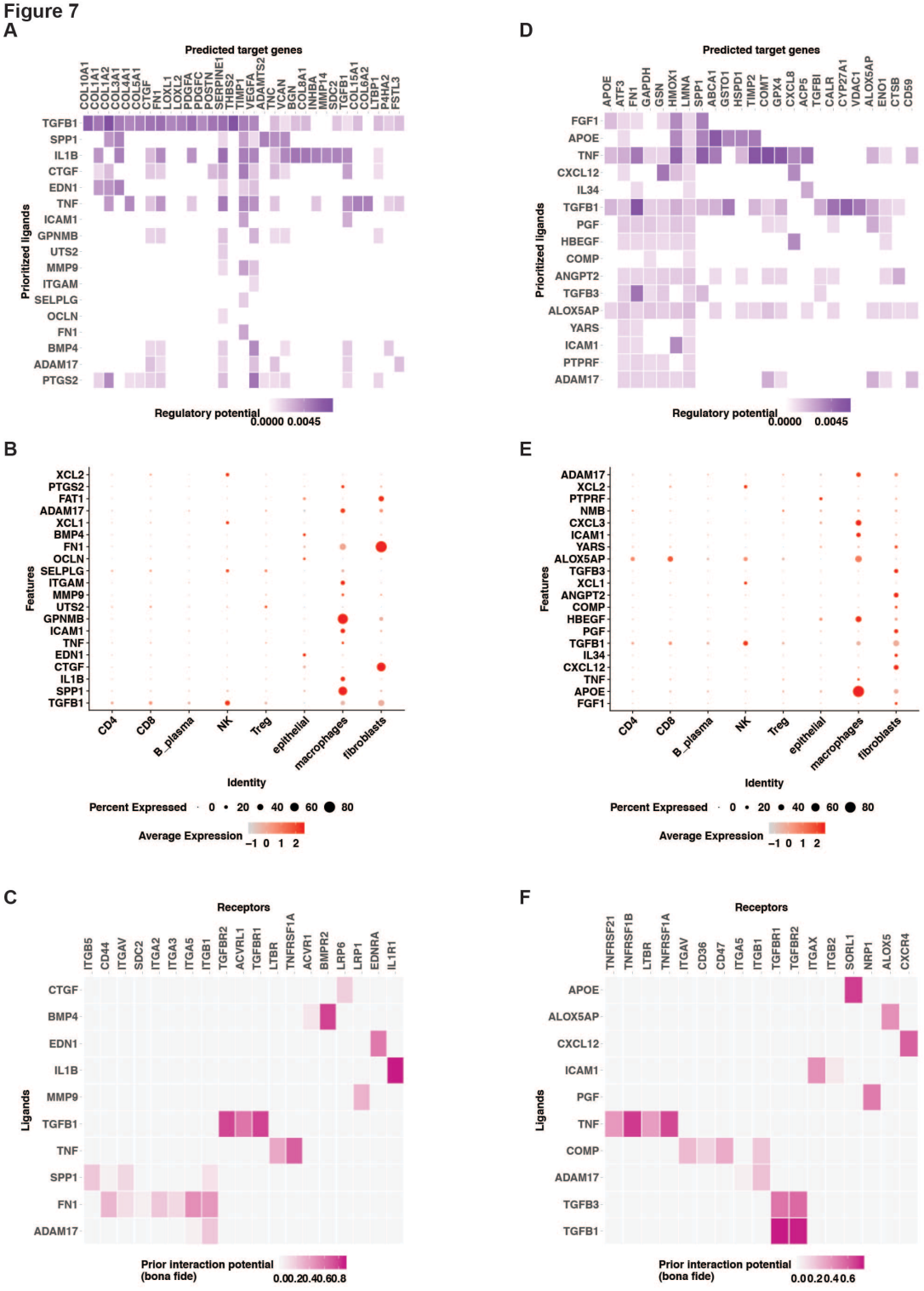

Figure 7: Predicted target genes and their regulatory potential for regulating the (A) fibroblast gene signature or (D) macrophage gene signature. (B, E) Dot plot depicting expression levels of respective ligands together with the percentage of cells expressing them. (C, F) Inferred receptors for the respective ligands with their interaction potential.

regulates the matrisome program in fibroblasts. All of these results confirmed the role of macrophages in regulating ECM organization, expanding the molecular network associated with the fibrogenic macrophage phenotype we previously discovered using differential expression analysis. Corresponding receptors on fibroblasts included members of the integrin, FGF and TGFB family. These genes are additional targets that can be used to experimentally modulate the fibroblast transcriptional phenotype (**Fig**. **7C**).

We investigated the gene expression program in macrophages enriched for reactome pathway terms related to metabolism and extracellular matrix organization. The top ranked ligands were expressed by macrophages and fibroblasts and included *FGF1*, *TNF*, *APOE*, *PGF*, *TGFB1*. (**Fig. 7D, E**). For macrophages, the corresponding receptors include members of the TNF and integrin receptor family and the macrophage immune checkpoint *CD47* (**Fig. 7F**). This analysis thus demonstrates a signaling network between TME macrophages and fibroblasts that influences their transcriptional phenotype via links intercellular communication

We analyzed the expression of the above two fibroblast and macrophage gene signatures in a published dataset generated from conventional RNA-seq of 93 mCRC patients (Pitroda et al. 2018) (**Methods**). The two signatures had a positive Spearman correlation (R2 = 0.59, p-value < 2.2e-16) (**Supplemental Fig. 4C**). This result supports a role for an association between the fibroblast and macrophage transcriptional phenotypes.

Perturbing the macrophage signaling network in TMEs

We examined the functional impact of targeting the metastatic niche with a patient-derived *ex vivo* model that preserves all TME lineages. Based on our discovery of a novel transcriptional state of macrophages in the TME, we conducted additional experiments to see if the macrophage signaling could be altered with external perturbations. These experiments provided us with a way

to experimentally challenge the macrophages in the heterogenous TME milieu found in a patient's tumor.

We recapitulated the diverse TME cell lineages and their interactions using an organotypic *ex vivo* model (Jiang et al. 2019) from P8640's liver mCRC resection. We refer to these cultures as 'TME-models'. We sectioned the mCRC tissue into thin 400 µm slices using a vibrating microtome. These slices were overlaid onto a permeable cell-culture insert. Cell culture media was then added onto the insert allowing nutrients to reach the tissue slice. The combination of extremely thin tissue slices and maintained nutrient exchange eliminates the need for specialized ECM support or special growth factors in this system. The native cellular architecture was kept intact with high cell viability. Media acts as a vehicle for added perturbations.

First, we compared the original tissue and cells at the time of surgical resection, referred to as 'T0', and what was maintained ex vivo. Hematoxylin and eosin (**HE**) staining revealed that the morphology of the *ex vivo* TME-models was maintained in comparison to the T0 specimen (**Fig. 8A**). No overt signs of cell death were noted on histology review. With scRNA-seq, we sampled 238 single cells from the T0 resection and an average of 1,292 cells from the TME-models (**Supplemental Table 2**). Based on cell-specific markers, all cell lineages were maintained in the TME-models compared to the T0 tissue sample. The only difference was a reduction in the proportion of the *ex vivo* epithelial cells compared to that in T0 (**Fig. 8B**). Importantly, UMAP clustering revealed that profiles from lineages in the TME-models clustered together with T0 cells indicating that transcriptional phenotypes are maintained *ex vivo* (**Fig. 8C, 8D**). Cells in the *ex vivo* model continued to express lineage markers and genes of interest in various cell types such immune checkpoint expression in T cells, metabolic reprogramming genes in macrophages, ECM re-organization genes in stromal fibroblasts, etc. (**Fig. 8E)**. Overall, these results demonstrate that TME-models successfully recapitulate characteristics of the *in vivo* TME.

Figure 8

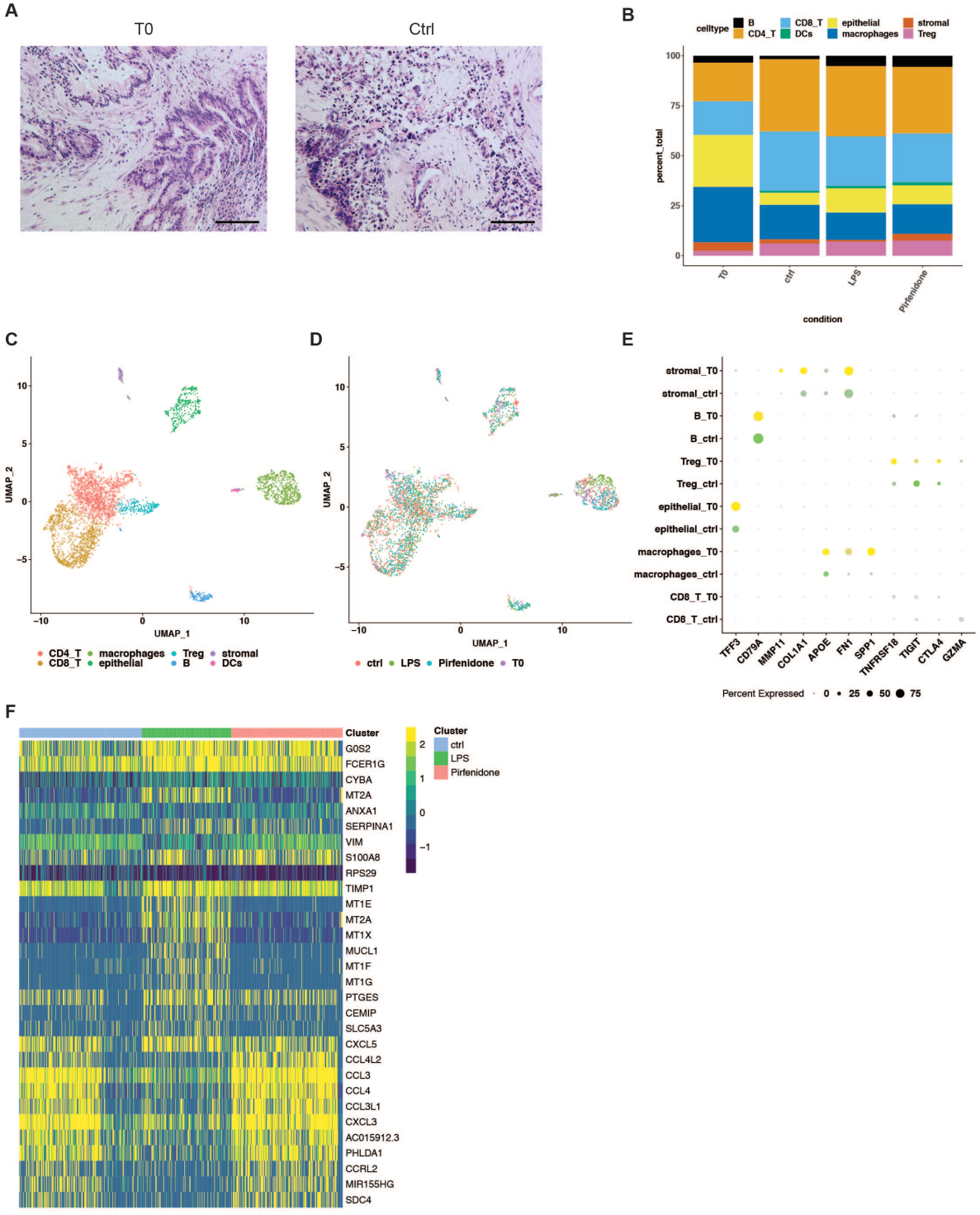

Figure 8: (A) H&E images from representative areas from T0 surgical resection or control (ctrl) ex vivo sample, scale bar = 50 μm. (B) Proportion of cell lineages detected in each sample. UMAP representation annotated with (C) cell lineage and (D) sample information. (E) Dot plot depicting expression levels of respective genes together with the percentage of cells expressing them in each condition. (F) Heatmap depicting the expression of respective genes in each cluster.

We treated the TME-model cells with either control media (**ctrl**) or one of two different perturbations: 1 μg/ml lipopolysaccharide (**LPS**), a known macrophage modulator, or 1 mM pirfenidone which has an impact on fibroblasts. Pirfenidone is an anti-fibrotic and anti-inflammatory agent used in the treatment of pulmonary fibrosis (Lopez-de la Mora et al. 2015). We disaggregated the treated TME-model cells and conducted scRNA-seq and histology at the 24 hour time point after treatment.

We examined the effects of our perturbations on the macrophage sub-population in the TME-models using differential expression analysis (**Fig. 8F, Supplemental Table 11**). LPS led to a significant increase in the expression of metallothionein family genes that have been previously described to mediate its effects on monocytes *in vitro* (Leibbrandt and Koropatnick 1994) as well as *HCK*, a receptor tyrosine kinase that mediates its effects in macrophages (English et al. 1997). LPS reduced the expression of several chemokines including *TNF*, *CCL4L2*, *CCL4*, *CXCL3*, *CXCL2* and mediators of NK-FB pathway such as *NFKBIZ*. Pirfenidone led to a significant increase in in chemokine gene expression that included *CCL4L2*, *CCL3L1*, *CCL4*, *CCL3* and the macrophage regulator *S100A8*. Thus, our *ex vivo* perturbation demonstrated that that TME macrophage phenotype and function could be altered via modulation with small molecules. Critical chemokines that mediate intercellular signaling are altered and their effects can be assessed. Hence, scRNA-seq can successfully be used to investigate the effects of perturbations on specific cellular lineages in *ex vivo* TME-models allowing the investigation of novel TME targets.

DISCUSSION

Our results reveal a complex picture of the intrinsic and extrinsic heterogeneity in mCRC including previously undescribed features of the TME. While a minority of patients with mCRC have excellent prognosis with current treatment strategies, the majority have poor 5-year survival

(Pitroda et al. 2018). We have identified features in the TME including gene expression programs, gene regulatory networks and several inter-cellular interactions that should be explored in future studies as novel targets for mCRC treatment. These targets are of interest also in other cancers to enable the modulation of the immunosuppressive stroma and improve immunotherapy response (Binnewies et al. 2018).

Previous studies have established a role for both monoclonal and polyclonal seeding in mCRC (Leung et al. 2017; Hu et al. 2019). Our results using both scRNA-seq and scDNA-seq revealed that mCRCs contain multiple sub-clones defined at the copy-number level. We also discovered discordance between chromosome arm level events and their gene dosage suggesting that additional regulatory mechanisms determine gene expression, corroborating the findings of other studies (Stamoulis et al. 2019). Experiments that analyze single-cell DNA, RNA and chromatin accessibility performed in parallel can provide further insights into these mechanisms.

Transcriptional level heterogeneity was also confirmed with the activation of various signaling pathways within a single mCRC. We also demonstrated that intratumoral heterogeneity in mCRC is not only characterized by specific genomic DNA events such as allelic imbalances but also by multi-lineage differentiation. It has been proposed that mCRC directly results from an aberrant epithelial mesenchymal transition (EMT) by cancer stem cells in the primary tumor (Pereira et al. 2015). Supporting this concept, we detected the presence of lineage differentiation including stem-like undifferentiated cells, enterocytes and goblet cells in mCRC that have been observed in primary CRC (Li et al. 2017). This underlying cellular complexity suggests that mCRCs contain sub-populations programmed to sustain growth. Thus targeted therapeutics may require multiple strategies that abrogate these different programs to ensure complete tumor eradication.

We discovered that the mCRC TME had significant differences compared to normal tissue including appearance of distinct transcriptional programs in macrophages, fibroblasts and lymphocytes. These changes in the "soil" component of the metastatic cascade could result from the presence of tumor cells in the colonizing metastasis as well as from systemic factors induced by the primary tumor. Investigating these changes in mouse models of metastasis could provide more insights into the origin and regulation of these processes (Liu et al. 2017).

Macrophages had a complex phenotype that does not confirm to the M1/M2 classification as we and others have previously discovered (Martinez and Gordon 2014; Sathe et al. 2020). The *SPP1*, *APOE*, *TREM2*, *CD9*, etc. expressing tumor macrophage phenotype we discovered was enriched for pathways related to ECM reorganization. This expression program has similarities to recently described studies in liver cirrhosis and pulmonary fibrosis where it was demonstrated to be a pro-fibrogenic phenotype (Ramachandran et al. 2019; Reyfman et al. 2019). *SPP1* expressing macrophages have been demonstrated to play a role in promoting primary CRC and also have the potential to influence CD8 function by their role as an immune checkpoint ligand (Zhang et al. 2017; Lee et al. 2020; Zhang et al. 2020a).

This fibrogenic phenotype was accompanied by changes in genes controlling various metabolic pathways including glycolysis, lipid transport and sphingolipid synthesis resembling atherosclerotic foam cells (Fernandez et al. 2019). Macrophage metabolism influences their functional phenotype (Netea-Maier et al. 2018). Our findings provide several metabolic targets including an upstream regulator *PPARG* that can be perturbed to further understand their biology in the context of the TME.

We detected infiltration of various lymphocyte cell types in the omentum mCRC, including exhausted T cells. Peritoneal mCRC is currently treated with surgery and systemic or intraperitoneal chemotherapy (Sanchez-Hidalgo et al. 2019). Our results suggest that these patients could also be targeted with T-cell based immunotherapy. The majority of liver mCRCs

contained few infiltrating lymphocytes suggesting that they exhibit immune desert or immune excluded phenotype characterized by the absence of T cell infiltration in the tumor core (Binnewies et al. 2018). Lack of infiltration was also noted for P6335 liver metastasis although the paired omentum metastasis contained a rich infiltrate, raising the possibility of an organ specific effect.

Fibroblasts and macrophages play a critical role in modulating such an immunosuppressive TME, including the phenotype of T cell exclusion (Binnewies et al. 2018). We discovered fibroblasts specific to the TME with the potential to regulate ECM properties that can in turn promote tumor growth. This fibroblast phenotype could also be influenced by FOLFOX-based neoadjuvant chemotherapy that was administered to all patients in the cohort (Mancini and Sonis 2014).

Single cell characterization of several solid tumors has revealed TME features that are common across all cancers as well as unique to the tissue of origin (Qian et al. 2020). The fibroblast phenotype in both liver and omentum metastases as well as the T cell phenotypes in the omentum metastasis are similar to findings from previous studies in other primary cancers. Unlike most studies in primary cancers, including primary CRC, that have detected multiple macrophage clusters indicative of diverse cellular states (Sathe et al. 2020; Zhang et al. 2020a), we detected a single cluster of macrophages in the mCRC metastatic niche. Further studies should investigate if this transcriptional phenotype is unique to metastasis, indicates a terminal differentiation state in TME macrophages or is a result of organ-specific transcriptional reprogramming.

The gene expression programs we have discovered can potentially be influenced by tissue dissociation processes. We used the same dissociation protocol for matched normal liver to enable a controlled comparison between tumor and normal microenvironment lineages, unlike

specially developed dissociation protocols for normal liver that ensure adequate dissociation of hepatocytes and stellate cells (MacParland et al. 2018).

We demonstrated that scRNA-seq analysis can be used to understand the effect of perturbations in the TME in a unique patient-derived *ex vivo* model that recapitulates cellular composition and phenotypes of the *in vivo* TME. We detected significant gene expression changes in macrophages exposed to LPS or Pirfenidone. We sampled few fibroblast cells from these samples. This perturbation can be used to understand its effect in modulating the fibroblast phenotype resulting from interactions between macrophages and fibroblasts. Most previous studies even with a known modulator like LPS have been limited to *in vitro* differentiated monocytes or animal models that do not resemble the *in vivo* macrophage phenotype. Our model is based on organotypic slice culture and maintains the native cells in regular cell culture media without the need for either special cytokines or ECM components. Air-liquid interface organoids derived from patients have previously been used to examine the effects of PD-1 immunotherapy using flow cytometry (Neal et al. 2018). The scRNA-seq analysis provides the advantage of an unbiased analysis of the *ex vivo* TME as we have previously demonstrated using a mouse model of gastric cancer (Chen et al. 2019b). Patient-derived TME models in combination with scRNA-seq thus provide a clinically relevant system to investigate novel targets.

METHODS

For the sake of brevity, we provide an overview of the methods with more details reported in the Supplemental Methods.

Sample acquisition

All samples were acquired with informed consent under an approved institutional review board protocol from Stanford University as surgical resections or matched normal tissue from sites displaced at least several centimeters from the tumor.

Single-cell RNA sequencing

The scRNA-seq libraries were generated from cell suspensions using Chromium Single Cell 3′ Library & Gel Bead Kit v2 or Chromium Next GEM Single cell Immune Profiling 5' v1.1 (10X Genomics, Pleasanton, CA, USA) as per manufacturer's protocol and sequenced on Illumina sequencers (Illumina, San Diego, CA) **(Supplemental Table 1)**. 10000 cells were targeted from tissue dissociation suspensions and 3000 for PBMCs with 14 PCR cycles for cDNA and library amplification. A 1% or 2% E-Gel (ThermoFisher Scientific, Waltham, MA, USA) was used for quality control evaluation of intermediate products and sequencing libraries. A Qubit (Thermofisher Scientific) or qPCR with Kapa library quantification kit (Kapa Biosystems, Wilmington, MA) was used to quantify the libraries as per the manufacturer's protocol. Cell Ranger 3 (10X Genomics) 'mkfastq' and 'count' commands were used with default parameters and alignment to GRCh38 to generate matrix of unique molecular identifier **(UMI)** counts per gene and associated cell barcode.

Single-cell DNA sequencing

scDNA-seq libraries were generated using Chromium Single Cell DNA Library & Gel Bead Kit or its beta version (**Supplemental Table 5)** (10X Genomics) as per manufacturer's protocol. 2000 cells were targeted with 14 PCR cycles for sample index PCR. Quality control was performed with a 1% or 2% E-Gel (Thermofisher Scientific) and libraries were quantified using Qubit (Thermofisher Scientific) as per the manufacturer's protocol. Cell Ranger DNA 1.1. (10X Genomics) 'cellranger dna mkfastq' and 'cellranger dna cnv' commands were used to perform reference alignment to

GRCh38, cell calling, copy number estimation and hierarchical clustering. Cell barcodes identified as 'noisy' cells with this pipeline were omitted and outputs were reconstructed from all other cells using the 'cellranger dna reanalyze' command. Resulting hierarchical clustering was visualized in Loupe scDNA Browser (v 1.0.0) (10X Genomics) with 128 nodes. Cells were visualized at a subtree depth suitable to evaluate the majority of aneuploid cells. Subtree depth of 2 was used for P6335 omentum and P6198, 1 for P6593 and 3 for P5915.

Ex vivo slice culture TME-models

Surgical resection was collected in plain RPMI on ice immediately after resection and dissected with iris scissors. VF-310-0Z Compresstome tissue slicer and its accessories (Precisionary, Greenville, NC, USA) were used to generate tissue slices from a piece of the resection. Tissue sample was glued onto the specimen tube base using All Purpose Krazy Glue (Elmer's Products, Inc., Westerville, OH, USA). 3% agarose solution was prepared by diluting UltraPure Low Melting Point Agarose (ThermoFisher Scientific) in water followed by heating in a microwave and cooling for around three-five minutes at room temperature. Tissue sample was allowed to enter the specimen tube and agarose solution was pipetted on top ensuring the sample was completely covered and no bubbles were present. Agarose was solidified by placing pre-chilled chilling block supplied by manufacturer over the specimen tube. Specimen tube was assembled onto compresstome as per manufacturer's instructions and cold PBS was used as a solution in the buffer tank. Slices were generated using advance setting of 3, oscillation of 5 and thickness of 400 μm. Slices were placed onto a 0.4 μm pore size Millicell Cell Culture Insert (Sigma-Aldrich, St. Loius, MO, USA) that was then placed into a 35-mm dish (ThermoFisher Scientific). 1.5 ml of media was placed into the surrounding dish and 0.5 ml was placed onto the slices followed by culture in a cell culture incubator. Media was composed of RPMI, 10% FBS, 1% Antibiotic-Antimycotic (ThermoFisher Scientific) for control slices with either 1 μg/ml lipopolysaccharide

(Sigma-Aldrich) or 1 mM pirfenidone (Tocris Bioscience, Bristol, UK) in the treated conditions. At 24 hours, slice culture TME-models were subjected to fixation for histology and dissociation (**Supplemental Methods**).

DATA ACCESS

Sequencing data deposition is in progress under dbGAP identifier phs001818.v3.p1. Cellranger outputs will be available on https://dna-discovery.stanford.edu/.

AUTHOR CONTRIBUTIONS

AS, GP and HPJ designed the study. AS, BTL and JC developed the methodology and acquired the data. AS, SMG, BTL, XB and HPJ analyzed and interpreted the data. AS and HPJ wrote the manuscript with input from all authors.

ACKNOWLEDGEMENTS

We are grateful to all patients who participated in the study as well as their families. We thank Christine Handy, Christina Wood-Bouwens and Alison Almeda for assistance in sample collection. Figure 1A was created using Biorender.com. This work was supported by US National Institutes of Health grants P01HG000205 (HPJ and SMG), R01HG006137 (HPJ), U01CA217875 (HPJ and AS). HPJ also received support from the American Cancer Society (124571-RSG-13-297-01) and the Clayville Foundation. This work with the Stanford Cancer Institute biobank was supported by a National Cancer Institute Cancer Center Support Grant (P30CA124435). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute. AS received additional support from the Stanford University Translational Research and Applied Medicine (TRAM) pilot grant program.

REFERENCES

Aibar S, Gonzalez-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselmans G, Rambow F, Marine JC, Geurts P, Aerts J et al. 2017. SCENIC: single-cell regulatory network inference and clustering. *Nat Methods* **14**: 1083-1086.

Andor N, Lau BT, Catalanotti C, Sathe A, Kubit M, Chen J, Blaj C, Cherry A, Bangs CD, Grimes SM et al. 2020. Joint single cell DNA-seq and RNA-seq of gastric cancer cell lines reveals rules of in vitro evolution. *NAR Genom Bioinform* **2**: lqaa016.

Andres A, Mentha G, Adam R, Gerstel E, Skipenko OG, Barroso E, Lopez-Ben S, Hubert C, Majno PE, Toso C. 2015. Surgical management of patients with colorectal cancer and simultaneous liver and lung metastases. *Br J Surg* **102**: 691-699.

Binnewies M, Roberts EW, Kersten K, Chan V, Fearon DF, Merad M, Coussens LM, Gabrilovich DI, Ostrand-Rosenberg S, Hedrick CC et al. 2018. Understanding the tumor immune microenvironment (TIME) for effective therapy. *Nat Med* **24**: 541-550.

Browaeys R, Saelens W, Saeys Y. 2020. NicheNet: modeling intercellular communication by linking ligands to target genes. *Nat Methods* **17**: 159-162.

Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. 2018. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* **36**: 411-420.

Candy PA, Phillips MR, Redfern AD, Colley SM, Davidson JA, Stuart LM, Wood BA, Zeps N, Leedman PJ. 2013. Notch-induced transcription factors are predictive of survival and 5-fluorouracil response in colorectal cancer patients. *Br J Cancer* **109**: 1023-1030.

Chen IX, Chauhan VP, Posada J, Ng MR, Wu MW, Adstamongkonkul P, Huang P, Lindeman N, Langer R, Jain RK. 2019a. Blocking CXCR4 alleviates desmoplasia, increases T-lymphocyte infiltration, and improves immunotherapy in metastatic breast cancer. *Proc Natl Acad Sci U S A* **116**: 4558-4566.

Chen J, Lau BT, Andor N, Grimes SM, Handy C, Wood-Bouwens C, Ji HP. 2019b. Single-cell transcriptome analysis identifies distinct cell types and niche signaling in a primary gastric organoid model. *Sci Rep* **9**: 4536.

De Cuyper A, Van Den Eynde M, Machiels JP. 2020. HER2 as a Predictive Biomarker and Treatment Target in Colorectal Cancer. *Clin Colorectal Cancer* **19**: 65-72.

Donadon M, Lleo A, Di Tommaso L, Soldani C, Franceschini B, Roncalli M, Torzilli G. 2018. The Shifting Paradigm of Prognostic Factors of Colorectal Liver Metastases: From Tumor-Centered to Host Immune-Centered Factors. *Front Oncol* **8**: 181.

English BK, Orlicek SL, Mei Z, Meals EA. 1997. Bacterial LPS and IFN-gamma trigger the tyrosine phosphorylation of vav in macrophages: evidence for involvement of the hck tyrosine kinase. *J Leukoc Biol* **62**: 859-864.

Fernandez DM, Rahman AH, Fernandez NF, Chudnovskiy A, Amir ED, Amadori L, Khan NS, Wong CK, Shamailova R, Hill CA et al. 2019. Single-cell immune landscape of human atherosclerotic plaques. *Nat Med* **25**: 1576-1588.

Fujihara S, Mori H, Kobara H, Rafiq K, Niki T, Hirashima M, Masaki T. 2013. Galectin-9 in cancer therapy. *Recent Pat Endocr Metab Immune Drug Discov* **7**: 130-137.

Gabriel TL, Tol MJ, Ottenhof R, van Roomen C, Aten J, Claessen N, Hooibrink B, de Weijer B, Serlie MJ, Argmann C et al. 2014. Lysosomal stress in obese adipose tissue macrophages contributes to MITF-dependent Gpnmb induction. *Diabetes* **63**: 3310-3323.

Gorvel L, Olive D. 2020. Targeting the "PVR-TIGIT axis" with immune checkpoint therapies. *F1000Res* **9**.

Gu-Trantien C, Migliori E, Buisseret L, de Wind A, Brohee S, Garaud S, Noel G, Dang Chi VL, Lodewyckx JN, Naveaux C et al. 2017. CXCL13-producing TFH cells link immune suppression and adaptive memory in human breast cancer. *JCI Insight* **2**.

Guinney J, Dienstmann R, Wang X, de Reynies A, Schlicker A, Soneson C, Marisa L, Roepman P, Nyamundanda G, Angelino P et al. 2015. The consensus molecular subtypes of colorectal cancer. *Nat Med* **21**: 1350-1356.

Haan JC, Labots M, Rausch C, Koopman M, Tol J, Mekenkamp LJ, van de Wiel MA, Israeli D, van Essen HF, van Grieken NC et al. 2014. Genomic landscape of metastatic colorectal cancer. *Nat Commun* **5**: 5457.

Hafemeister C, Satija R. 2019. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol* **20**: 296.

He W, Zhang H, Han F, Chen X, Lin R, Wang W, Qiu H, Zhuang Z, Liao Q, Zhang W et al. 2017. CD155T/TIGIT Signaling Regulates CD8(+) T-cell Metabolism and Promotes Tumor Progression in Human Gastric Cancer. *Cancer Res* **77**: 6375-6388.

Hinz B. 2015. The extracellular matrix and transforming growth factor-beta1: Tale of a strained relationship. *Matrix Biol* **47**: 54-65.

Hu Z, Ding J, Ma Z, Sun R, Seoane JA, Scott Shaffer J, Suarez CJ, Berghoff AS, Cremolini C, Falcone A et al. 2019. Quantitative evidence for early metastatic seeding in colorectal cancer. *Nat Genet* **51**: 1113-1122.

Ilicic T, Kim JK, Kolodziejczyk AA, Bagger FO, McCarthy DJ, Marioni JC, Teichmann SA. 2016. Classification of low quality cells from single-cell RNA-seq data. *Genome Biol* **17**: 29.

Jackson-Jones LH, Smith P, Portman JR, Magalhaes MS, Mylonas KJ, Vermeren MM, Nixon M, Henderson BEP, Dobie R, Vermeren S et al. 2020. Stromal Cells Covering Omental Fat-Associated Lymphoid Clusters Trigger Formation of Neutrophil Aggregates to Capture Peritoneal Contaminants. *Immunity* **52**: 700-715 e706.

Jiang X, Seo YD, Sullivan KM, Pillarisetty VG. 2019. Establishment of Slice Cultures as a Tool to Study the Cancer Immune Microenvironment. *Methods Mol Biol* **1884**: 283-295.

Klement JD, Paschall AV, Redd PS, Ibrahim ML, Lu C, Yang D, Celis E, Abrams SI, Ozato K, Liu K. 2018. An osteopontin/CD44 immune checkpoint controls CD8+ T cell activation and tumor immune evasion. *J Clin Invest* **128**: 5549-5560.

Kuznetsova T, Prange KHM, Glass CK, de Winther MPJ. 2020. Transcriptional and epigenetic regulation of macrophages in atherosclerosis. *Nat Rev Cardiol* **17**: 216-228.

Lee HO, Hong Y, Etlioglu HE, Cho YB, Pomella V, Van den Bosch B, Vanhecke J, Verbandt S, Hong H, Min JW et al. 2020. Lineage-dependent gene expression programs influence the immune landscape of colorectal cancer. *Nat Genet* **52**: 594-603.

Leibbrandt ME, Koropatnick J. 1994. Activation of human monocytes with lipopolysaccharide induces metallothionein expression and is diminished by zinc. *Toxicol Appl Pharmacol* **124**: 72-81.

Leung ML, Davis A, Gao R, Casasent A, Wang Y, Sei E, Vilar E, Maru D, Kopetz S, Navin NE. 2017. Single-cell DNA sequencing reveals a late-dissemination model in metastatic colorectal cancer. *Genome Res* **27**: 1287-1299.

Li H, Courtois ET, Sengupta D, Tan Y, Chen KH, Goh JJL, Kong SL, Chua C, Hon LK, Tan WS et al. 2017. Reference component analysis of single-cell transcriptomes elucidates cellular heterogeneity in human colorectal tumors. *Nat Genet* **49**: 708-718.

Lim B, Lin Y, Navin N. 2020. Advancing Cancer Research and Medicine with Single-Cell Genomics. *Cancer Cell* **37**: 456-470.

Liu Q, Zhang H, Jiang X, Qian C, Liu Z, Luo D. 2017. Factors involved in cancer metastasis: a better understanding to "seed and soil" hypothesis. *Mol Cancer* **16**: 176.

Liu T, Zhou L, Li D, Andl T, Zhang Y. 2019. Cancer-Associated Fibroblasts Build and Secure the Tumor Microenvironment. *Front Cell Dev Biol* **7**: 60.

Lopez-de la Mora DA, Sanchez-Roque C, Montoya-Buelna M, Sanchez-Enriquez S, Lucano-Landeros S, Macias-Barragan J, Armendariz-Borunda J. 2015. Role and New Insights of Pirfenidone in Fibrotic Diseases. *Int J Med Sci* **12**: 840-847.

MacParland SA, Liu JC, Ma XZ, Innes BT, Bartczak AM, Gage BK, Manuel J, Khuu N, Echeverri J, Linares I et al. 2018. Single cell RNA sequencing of human liver reveals distinct intrahepatic macrophage populations. *Nat Commun* **9**: 4383.

Mamlouk S, Childs LH, Aust D, Heim D, Melching F, Oliveira C, Wolf T, Durek P, Schumacher D, Blaker H et al. 2017. DNA copy number changes define spatial patterns of heterogeneity in colorectal cancer. *Nat Commun* **8**: 14093.

Mancini ML, Sonis ST. 2014. Mechanisms of cellular fibrosis associated with cancer regimen-related toxicities. *Front Pharmacol* **5**: 51.

Martinez FO, Gordon S. 2014. The M1 and M2 paradigm of macrophage activation: time for reassessment. *F1000Prime Rep* **6**: 13.

Martinez FO, Gordon S, Locati M, Mantovani A. 2006. Transcriptional profiling of the human monocyte-to-macrophage differentiation and polarization: new molecules and patterns of gene expression. *J Immunol* **177**: 7303-7311.

McGinnis CS, Murrow LM, Gartner ZJ. 2019. DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. *Cell Syst* **8**: 329-337 e324.

McInnes L, Healy J. 2018. UMAP: uniform manifold approximation and projection for dimension reduction. doi:https://arxiv.org/abs/1802.03426 ArXiv.

Naba A, Clauser KR, Hoersch S, Liu H, Carr SA, Hynes RO. 2012. The matrisome: in silico definition and in vivo characterization by proteomics of normal and tumor extracellular matrices. *Mol Cell Proteomics* **11**: M111 014647.

Neal JT, Li X, Zhu J, Giangarra V, Grzeskowiak CL, Ju J, Liu IH, Chiou SH, Salahudeen AA, Smith AR et al. 2018. Organoid Modeling of the Tumor Immune Microenvironment. *Cell* **175**: 1972-1988 e1916.

Netea-Maier RT, Smit JWA, Netea MG. 2018. Metabolic changes in tumor cells and tumor-associated macrophages: A mutual relationship. *Cancer Lett* **413**: 102-109.

Nguyen HT, Duong HQ. 2018. The molecular characteristics of colorectal cancer: Implications for diagnosis and therapy. *Oncol Lett* **16**: 9-18.

Ohue Y, Nishikawa H. 2019. Regulatory T (Treg) cells in cancer: Can Treg cells be a new therapeutic target? *Cancer Sci* **110**: 2080-2089.

Pereira L, Mariadason JM, Hannan RD, Dhillon AS. 2015. Implications of epithelial-mesenchymal plasticity for heterogeneity in colorectal cancer. *Front Oncol* **5**: 13.

Pitroda SP, Khodarev NN, Huang L, Uppal A, Wightman SC, Ganai S, Joseph N, Pitt J, Brown M, Forde M et al. 2018. Integrated molecular subtyping defines a curable oligometastatic state in colorectal liver metastasis. *Nat Commun* **9**: 1793.

Qian J, Olbrecht S, Boeckx B, Vos H, Laoui D, Etlioglu E, Wauters E, Pomella V, Verbandt S, Busschaert P et al. 2020. A pan-cancer blueprint of the heterogeneous tumor microenvironment revealed by single-cell profiling. *Cell Res* doi:10.1038/s41422-020-0355-0.

Ramachandran P, Dobie R, Wilson-Kanamori JR, Dora EF, Henderson BEP, Luu NT, Portman JR, Matchett KP, Brice M, Marwick JA et al. 2019. Resolving the fibrotic niche of human liver cirrhosis at single-cell level. *Nature* **575**: 512-518.

Reyfman PA, Walter JM, Joshi N, Anekalla KR, McQuattie-Pimentel AC, Chiu S, Fernandez R, Akbarpour M, Chen CI, Ren Z et al. 2019. Single-Cell Transcriptomic Analysis of Human Lung Provides Insights into the Pathobiology of Pulmonary Fibrosis. *Am J Respir Crit Care Med* **199**: 1517-1536.

Sahai E, Astsaturov I, Cukierman E, DeNardo DG, Egeblad M, Evans RM, Fearon D, Greten FR, Hingorani SR, Hunter T et al. 2020. A framework for advancing our understanding of cancer-associated fibroblasts. *Nat Rev Cancer* **20**: 174-186.

Sanchez-Hidalgo JM, Rodriguez-Ortiz L, Arjona-Sanchez A, Rufian-Pena S, Casado-Adam A, Cosano-Alvarez A, Briceno-Delgado J. 2019. Colorectal peritoneal metastases: Optimal management review. *World J Gastroenterol* **25**: 3484-3502.

Sathe A, Grimes SM, Lau BT, Chen J, Suarez C, Huang RJ, Poultsides G, Ji HP. 2020. Single-Cell Genomic Characterization Reveals the Cellular Reprogramming of the Gastric Tumor Microenvironment. *Clin Cancer Res* doi:10.1158/1078-0432.CCR-19-3231.

Shen Y, Wang X, Lu J, Salfenmoser M, Wirsik NM, Schleussner N, Imle A, Freire Valls A, Radhakrishnan P, Liang J et al. 2020. Reduction of Liver Metastasis Stiffness Improves Response to Bevacizumab in Metastatic Colorectal Cancer. *Cancer Cell* **37**: 800-817 e807.

Stamoulis G, Garieri M, Makrythanasis P, Letourneau A, Guipponi M, Panousis N, Sloan-Bena F, Falconnet E, Ribaux P, Borel C et al. 2019. Single cell transcriptome in aneuploidies reveals mechanisms of gene dosage imbalance. *Nat Commun* **10**: 4495.

Takizawa N, Ohishi Y, Hirahashi M, Takahashi S, Nakamura K, Tanaka M, Oki E, Takayanagi R, Oda Y. 2015. Molecular characteristics of colorectal neuroendocrine carcinoma; similarities with adenocarcinoma rather than neuroendocrine tumor. *Hum Pathol* **46**: 1890-1900.

Tickle T, Tirosh I, Georgescu C, Brown M, Haas B. 2019. inferCNV of the Trinity CTAT Project. Vol 2020. Klarman Cell Observatory, Broad Institute of MIT and Harvard, Cambridge, MA, USA.

Tirosh I, Izar B, Prakadan SM, Wadsworth MH, 2nd, Treacy D, Trombetta JJ, Rotem A, Rodman C, Lian C, Murphy G et al. 2016. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* **352**: 189-196.

Tomlinson JS, Jarnagin WR, DeMatteo RP, Fong Y, Kornprat P, Gonen M, Kemeny N, Brennan MF, Blumgart LH, D'Angelica M. 2007. Actual 10-year survival after resection of colorectal liver metastases defines cure. *J Clin Oncol* **25**: 4575-4580.

Velazquez-Villarreal EI, Maheshwari S, Sorenson J, Fiddes IT, Kumar V, Yin Y, Webb MG, Catalanotti C, Grigorova M, Edwards PA et al. 2020. Single-cell sequencing of genomic DNA resolves sub-clonal heterogeneity in a melanoma cell line. *Commun Biol* **3**: 318.

Vento-Tormo R, Efremova M, Botting RA, Turco MY, Vento-Tormo M, Meyer KB, Park JE, Stephenson E, Polanski K, Goncalves A et al. 2018. Single-cell reconstruction of the early maternal-fetal interface in humans. *Nature* **563**: 347-353.

Villani AC, Satija R, Reynolds G, Sarkizova S, Shekhar K, Fletcher J, Griesbeck M, Butler A, Zheng S, Lazo S et al. 2017. Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science* **356**.

Viola A, Munari F, Sanchez-Rodriguez R, Scolaro T, Castegna A. 2019. The Metabolic Signature of Macrophage Responses. *Front Immunol* **10**: 1462.

Wang Q, Yu C. 2020. Expression profiling of small intestinal neuroendocrine tumors identified pathways and gene networks linked to tumorigenesis and metastasis. *Biosci Rep* **40**.

Wherry EJ, Kurachi M. 2015. Molecular and cellular insights into T cell exhaustion. *Nat Rev Immunol* **15**: 486-499.

Williamson T, Sultanpuram N, Sendi H. 2019. The role of liver microenvironment in hepatic metastasis. *Clin Transl Med* **8**: 21.

Wu C, Thalhamer T, Franca RF, Xiao S, Wang C, Hotta C, Zhu C, Hirashima M, Anderson AC, Kuchroo VK. 2014. Galectin-9-CD44 interaction enhances stability and function of adaptive regulatory T cells. *Immunity* **41**: 270-282.

Zhang L, Li Z, Skrzypczynska KM, Fang Q, Zhang W, O'Brien SA, He Y, Wang L, Zhang Q, Kim A et al. 2020a. Single-Cell Analyses Inform Mechanisms of Myeloid-Targeted Therapies in Colon Cancer. *Cell* **181**: 442-459 e429.

Zhang Y, Du W, Chen Z, Xiang C. 2017. Upregulation of PD-L1 by SPP1 mediates macrophage polarization and facilitates immune escape in lung adenocarcinoma. *Exp Cell Res* **359**: 449-457.

Zhang Y, Song J, Zhao Z, Yang M, Chen M, Liu C, Ji J, Zhu D. 2020b. Single-cell transcriptome analysis reveals tumor immune microenvironment heterogenicity and granulocytes enrichment in colorectal cancer liver metastases. *Cancer Lett* **470**: 84-94.

Zhu R, Gires O, Zhu L, Liu J, Li J, Yang H, Ju G, Huang J, Ge W, Chen Y et al. 2019. TSPAN8 promotes cancer cell stemness via activation of sonic Hedgehog signaling. *Nat Commun* **10**: 2863.