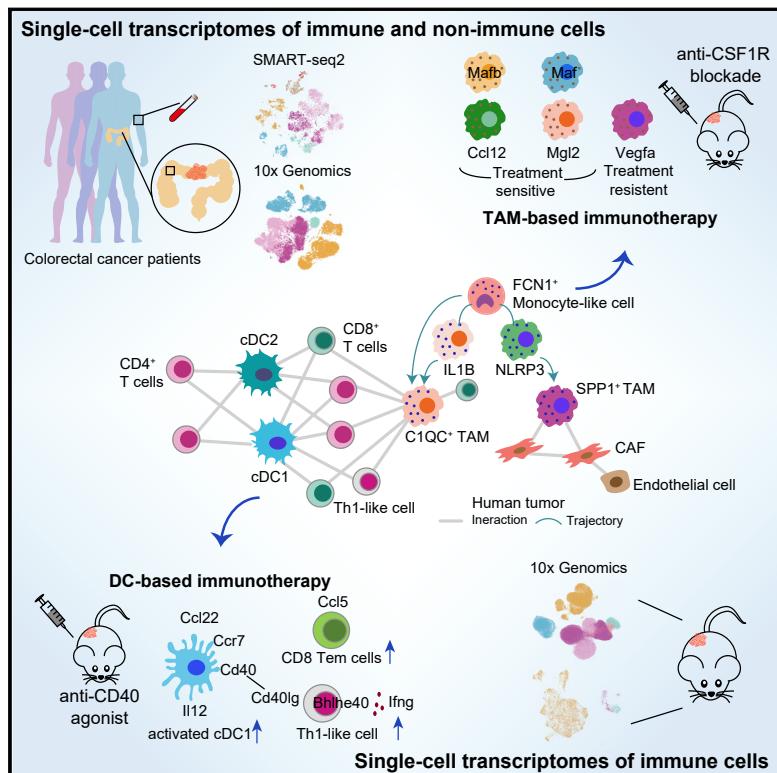


# Single-Cell Analyses Inform Mechanisms of Myeloid-Targeted Therapies in Colon Cancer

## Graphical Abstract



## Authors

Lei Zhang, Ziyi Li,  
Katarzyna M. Skrzypczynska, ...,  
Jackson G. Egen, Zemin Zhang, Xin Yu

## Correspondence

jegen@amgen.com (J.G.E.),  
shenlong1977@163.com (Z.S.),  
xiyu@amgen.com (X.Y.),  
zemin@pku.edu.cn (Z.Z.)

## In Brief

Combined scRNA-seq analyses on the tumor microenvironment in colorectal cancer and murine tumor models identify distinct myeloid populations that convey differential sensitivity to CSF1R blockade and define concerted immune responses involving dendritic cells and T cells upon anti-CD40 treatment.

## Highlights

- scRNA-seq analyses highlight conserved myeloid subsets in human and murine CRC
- Two distinct TAM subsets show inflammatory and angiogenic signatures, respectively
- Two distinct TAM subsets show differential sensitivity to CSF1R blockade
- Anti-CD40 activates specific cDC1s and expands Th1-like and CD8<sup>+</sup> memory T cells



# Single-Cell Analyses Inform Mechanisms of Myeloid-Targeted Therapies in Colon Cancer

Lei Zhang,<sup>1,6</sup> Ziyi Li,<sup>2,6</sup> Katarzyna M. Skrzypczynska,<sup>3,6</sup> Qiao Fang,<sup>1</sup> Wei Zhang,<sup>4</sup> Sarah A. O'Brien,<sup>3</sup> Yao He,<sup>1</sup> Lynn Wang,<sup>3</sup> Qiming Zhang,<sup>2</sup> Aeryon Kim,<sup>3</sup> Ranran Gao,<sup>2</sup> Jessica Orf,<sup>3</sup> Tao Wang,<sup>2</sup> Deepali Sawant,<sup>3</sup> Jiajinlong Kang,<sup>2</sup> Dev Bhatt,<sup>3</sup> Daniel Lu,<sup>3</sup> Chi-Ming Li,<sup>3</sup> Aaron S. Rapaport,<sup>3</sup> Kristy Perez,<sup>3</sup> Yingjiang Ye,<sup>4</sup> Shan Wang,<sup>4</sup> Xueda Hu,<sup>2,5</sup> Xianwen Ren,<sup>2</sup> Wenjun Ouyang,<sup>3</sup> Zhanlong Shen,<sup>4,\*</sup> Jackson G. Egen,<sup>3,\*</sup> Zemin Zhang,<sup>1,2,7,\*</sup> and Xin Yu<sup>3,\*</sup>

<sup>1</sup>Beijing Advanced Innovation Center for Genomics, Peking-Tsinghua Center for Life Sciences, Peking University, Beijing 100871, China

<sup>2</sup>BIOPIC and School of Life Sciences, Peking University, Beijing 100871, China

<sup>3</sup>Department of Inflammation and Oncology and Genome Analysis Unit, Amgen Research, Amgen Inc., South San Francisco, CA 94080, USA

<sup>4</sup>Department of Gastroenterological Surgery, Peking University People's Hospital, Beijing 100044, China

<sup>5</sup>Analytical Biosciences Limited, Beijing 100084, China

<sup>6</sup>These authors contributed equally to this work

<sup>7</sup>Lead contact

\*Correspondence: shenlong1977@163.com (Z.S.), jegen@amgen.com (J.G.E.), zemin@pku.edu.cn (Z.Z.), xiyu@amgen.com (X.Y.)

<https://doi.org/10.1016/j.cell.2020.03.048>

## SUMMARY

Single-cell RNA sequencing (scRNA-seq) is a powerful tool for defining cellular diversity in tumors, but its application toward dissecting mechanisms underlying immune-modulating therapies is scarce. We performed scRNA-seq analyses on immune and stromal populations from colorectal cancer patients, identifying specific macrophage and conventional dendritic cell (cDC) subsets as key mediators of cellular cross-talk in the tumor microenvironment. Defining comparable myeloid populations in mouse tumors enabled characterization of their response to myeloid-targeted immunotherapy. Treatment with anti-CSF1R preferentially depleted macrophages with an inflammatory signature but spared macrophage populations that in mouse and human expresses pro-angiogenic/tumorigenic genes. Treatment with a CD40 agonist antibody preferentially activated a cDC population and increased *Bhlhe40*<sup>+</sup> Th1-like cells and CD8<sup>+</sup> memory T cells. Our comprehensive analysis of key myeloid subsets in human and mouse identifies critical cellular interactions regulating tumor immunity and defines mechanisms underlying myeloid-targeted immunotherapies currently undergoing clinical testing.

## INTRODUCTION

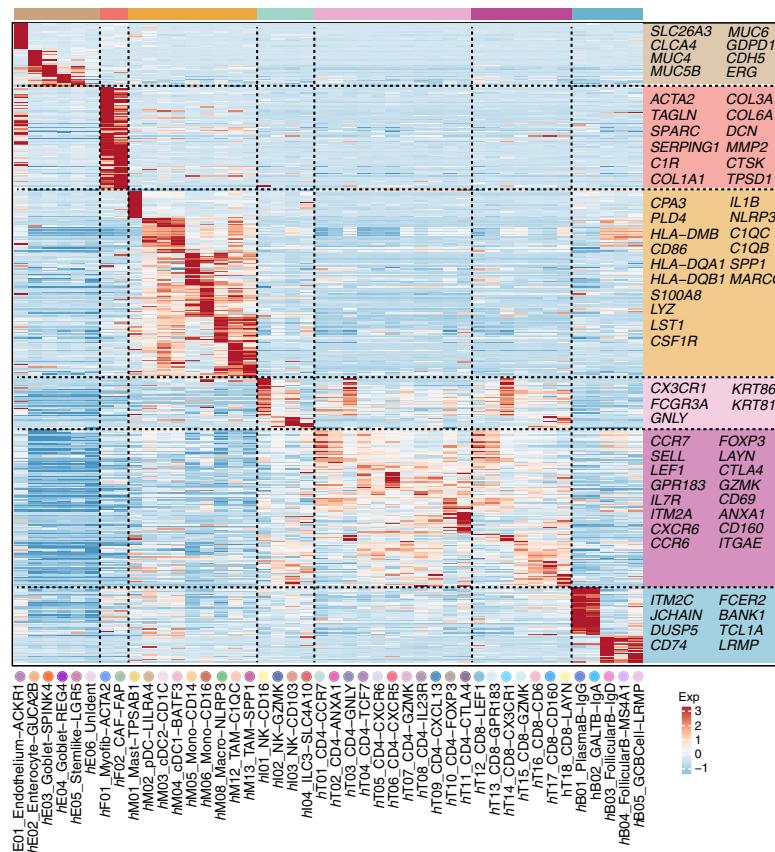
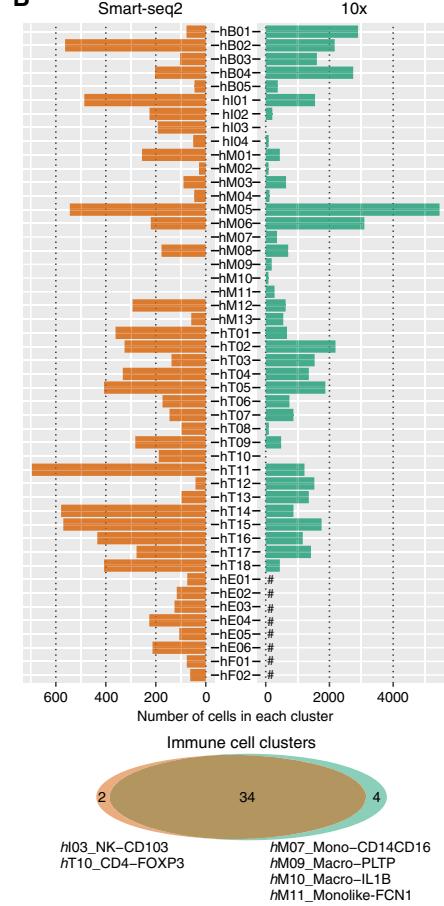
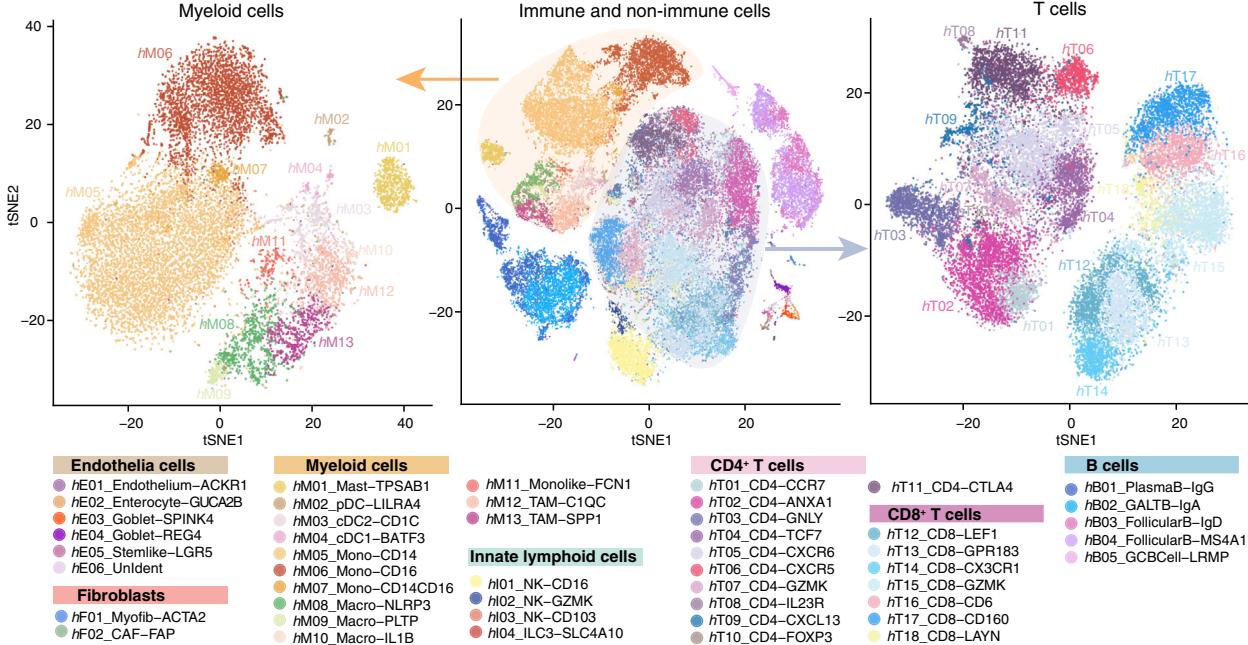
Immune checkpoint blockade (ICB) is capable of disrupting immune surveillance subversion by cancer cells and has dramatically altered cancer treatment paradigms (Ribas and Wolchok, 2018). However, as ICB has shown activity in only a minority of patients with specific cancers (Ribas and Wolchok, 2018; Yarchoan et al., 2017), additional therapeutic strategies to potentiate anti-tumor immunity have been proposed, including deple-

tion of pro-tumorigenic or immune suppressive cells within the tumor microenvironment (TME) and activation of specific immune populations using agonistic antibodies (Mahoney et al., 2015). Unfortunately, our incomplete understanding of the complexity of the TME has led to many of these strategies being indiscriminately advanced to the clinic rather than being selected based on clear mechanistic hypotheses (Binnewies et al., 2018; Galon and Bruni, 2019; Juntila and de Sauvage, 2013).

Single-cell RNA sequencing (scRNA-seq) is a powerful technique for dissecting the complexity of solid tumors, enabling characterization of cell diversity and heterogeneous phenotypic states in unprecedented detail (Papalexis and Satija, 2018; Zhang and Zhang, 2019). scRNA-seq based transcriptome analyses in primary human tumors have not only revealed T cell heterogeneity, but have also begun to elucidate dynamic relationships between T cell populations through integrating analysis of transcriptomes and T cell receptors (Guo et al., 2018; Li et al., 2019; Savas et al., 2018; Tirosh et al., 2016; Zhang et al., 2018; Zheng et al., 2017). Previously, we identified a *BHLHE40*<sup>+</sup> Th1-like cell population that is significantly enriched in tumor samples from colorectal cancer (CRC) patients with high microsatellite instability (MSI), but not microsatellite stability (MSS) (Zhang et al., 2018). Given that only MSI patients respond to ICB (Ganesh et al., 2019; Le et al., 2015), these findings suggest that strategies to enhance the presence and function of these T cells may promote response to ICB.

Recently, single-cell analyses have also revealed the complexity of tumor-infiltrating myeloid cells, including tumor-associated macrophages (TAMs) and dendritic cells (DCs), in multiple different cancer types (Azizi et al., 2018; Chevrier et al., 2017; Lavin et al., 2017; Venteicher et al., 2017; Zhang et al., 2019; Zilionis et al., 2019). TAMs are a heterogeneous cell type that can contribute to malignancy through production of tumor and angiogenic growth factors, extracellular matrix (ECM) remodeling, and immunosuppression (DeNardo and Ruffell, 2019). While immunotherapies aimed at repressing TAM biology have advanced to the clinic, including disruption of macrophage expansion and differentiation by blocking interactions between CSF1R and its ligands CSF1 and IL-34



**A****B****C**

(legend on next page)

(Butowski et al., 2016; Cassier et al., 2015; Lee et al., 2020; Papadopoulos et al., 2017; Ries et al., 2014), minimal monotherapy efficacy against malignant tumors has been observed (Papadopoulos et al., 2017; Ries et al., 2014).

Tumor-associated DCs constitute only a minority of myeloid cells in the tumor but are key orchestrators of anti-tumor T cell responses (Merad et al., 2013). Conventional DCs (cDCs) can be broadly separated into two subsets (cDC1 and cDC2) based on distinct phenotypic markers and functional roles (Haniffa et al., 2013). While the role of cDC2s in the tumor remains poorly defined, cDC1s are critical for the generation of anti-tumor T cell responses based on their ability to cross-present tumor-associated antigens to CD8<sup>+</sup> T cells (Böttcher and Reis E Sousa, 2018; Hildner et al., 2008). Multiple strategies to boost the function of DCs have advanced to the clinic, with approaches to activate the CD40 receptor being widely explored. However, similar to CSF1R inhibitors, CD40 agonists have shown limited monotherapy efficacy (Vonderheide, 2020). Our ability to develop these and other myeloid-targeted therapies is dependent on a comprehensive understanding of myeloid cell heterogeneity in solid tumors and the impact of these therapies on immune cell function and cross-talk within the TME.

While recent studies have attempted to relate the myeloid heterogeneity identified by scRNA-seq in human cancers to cell types found in murine tumor models (Binnewies et al., 2019; Brown et al., 2019; Gubin et al., 2018; Zilionis et al., 2019), the functional consequences of this heterogeneity and the relationships between distinct myeloid populations and the diverse responses to myeloid-targeted immunotherapies remain poorly characterized. Here, we have utilized two scRNA-seq platforms to perform a high-resolution analysis of immune and stromal cell populations in tumors, adjacent normal tissues, and blood from CRC patients. We further constructed a cell-cell interaction network to define key cell populations involved in regulating tumorigenesis and anti-tumor immunity and identified specific populations of TAMs and DCs as central nodes of cellular interaction. To understand the consequences of modulating the function of these cells, we performed additional scRNA-seq to identify analogous immune cell subsets in pre-clinical mouse tumor models in the context of anti-CSF1R and anti-CD40 antibody treatment. By analyzing both human and mouse scRNA-seq data, we reveal the heterogeneity and function of these cells, their role in shaping the TME, and the translatability of myeloid-modulating therapies from pre-clinical models to human cancer.

## RESULTS

### Intratumoral Cell Types in Human CRC Revealed by Combined Plate- and Droplet-based scRNA-seq

We first collected CD45<sup>+</sup> and CD45<sup>-</sup> cells by fluorescence-activated cell sorting (FACS) from tumors, adjacent normal tissues,

and blood of 18 treatment-naive CRC patients (Table S1). The transcriptomes of individual cells were obtained by scRNA-seq via either a droplet-based 10x Genomics platform or a plate-based full-length Smart-seq2 platform (Figure S1A). The data were combined with the full-length transcriptomes of T cells from the same patients we studied previously (Figure S1A) (Zhang et al., 2018). After quality control and filtering, we obtained 43,817 (10x scRNA-seq) and 10,468 (Smart-seq2) transcriptomes of single cells (STAR Methods). As expected, the Smart-seq2 platform captured more genes, including cytokines, CD molecules, ligands/receptors and transcription factors, and exhibited weak batch effects compared to the 10x scRNA-seq platform (Figures S1B–S1F).

To define the major population and subpopulation structure of the tumor-infiltrating leukocytes, we performed graph-based clustering (Butler et al., 2018) on these two datasets separately, identifying the majority of known immune cell types, including myeloid cells, innate lymphoid cells (ILCs), T cells and B cells (Figure S2A). Further unsupervised clustering in each compartment gave rise to a total of 38 and 36 leukocyte cell clusters from the 10x and the Smart-seq2 platform, respectively (Figures 1A and S2B–S2D). We then performed logistic regression modeling (Young et al., 2018) to define the similarities and relationships between clusters from these two datasets (Figure S2E). The leukocyte cell clusters exhibiting high similarities were annotated as the same cell subsets, and the remaining were manually annotated by their specific gene signatures. For lymphocytes, tumor-enriched plasma B cells, gut-associated lymphoid tissue B cells, follicular B cells and germinal center B cells were highly similar between two platforms and could be distinguished by distinct immunoglobulin heavy chain signatures (Figure S2F). Likewise, both platforms captured ILCs, including blood- and tissue-enriched natural killer (NK) cells and normal mucosa-enriched ILC3 cells (Figures S2B and S2C) (Klose and Artis, 2016), as well as tumor-enriched tissue-resident NK cells (hI03) whose expression of inhibitory receptors, cytotoxic molecules, and proliferation-related genes resembled those of CD8<sup>+</sup> exhausted T cells described previously in CRC (Figure S2G) (Zhang et al., 2018). In contrast to lymphocytes, myeloid cells exhibited a greater degree of divergence between these two platforms. Specifically, two intermediate state monocyte subsets (hM07 and hM11) were mainly captured by the 10x platform (Figure 1B), indicating that a larger number of sequenced cells per patient may be needed to identify rare or transitioning cell populations. Ultimately, we obtained 13 myeloid cell clusters, 4 ILC clusters, 18 T cell clusters, and 5 B cell clusters from these two platforms, largely consistent with previous scRNA-seq studies from different cancer types (Azizi et al., 2018; Lambrechts et al., 2018; Zhang et al., 2019). All cell clusters identified from the two datasets could be visualized by a combined t-distributed stochastic neighbor embedding (t-SNE) analysis (Figures 1C and S3A). Each cluster was

**Figure 1. Identification of Intratumoral Cell Types in Human CRC by Smart-seq2 and 10x Genomics scRNA-seq**

(A) Gene expression heatmap analyzed by Smart-seq2 scRNA-seq. Exp, z-score normalized mean expression.

(B) Comparison of clusters generated from Smart-seq2 and 10x platforms. # indicates the lack of identification of cells due to sample collection.

(C) Combined t-SNE plot showing clusters of all immune and non-immune cells from 18 CRC patients analyzed by Smart-seq2 and 10x scRNA-seq.

See also Figures S1, S2, and S3 and Table S1.

comprised of cells derived from multiple different patients and was associated with a different tissue distribution pattern (Figures S3B and S3C).

For non-immune cells from the Smart-seq2 platform, graph-based clustering gave rise to 12 cell clusters (Figure S3D). Malignant cells, defined by inferred copy number variations, exhibited a high degree of heterogeneity in gene expression, thus forming patient-specific clusters (Figures S3E and S3F, STAR Methods). Consistent with previous studies in CRC (Dalerba et al., 2011; Li et al., 2017), we identified several non-malignant cell clusters in normal mucosa and tumors, including vascular endothelial cells, four groups of epithelial cells (enterocytes, stem-like cells and two groups of goblet cells separated by top- and bottom-crypt gene signatures) (Dalerba et al., 2011; Sasaki et al., 2016), and two groups of fibroblasts. Of note, both myofibroblasts and cancer-associated fibroblasts (CAFs) were preferentially found in CRC tumors compared to the normal mucosa (Figures S3G and S3H).

Taken together, while the Smart-seq2 platform captured a greater number of genes allowing a more in-depth analysis of regulatory pathways (Figure S1), the 10x platform effectively obtained more clusters (Figure 1B). Gene depth of Smart-seq2 and cell coverage of 10x scRNA-seq collectively maximized the number of cell types or rare populations and improved the outcome of cell clustering. Thus, we have leveraged the unique advantages of both platforms to define the properties of tumor-infiltrating myeloid cells and their interactions with other cells in CRC.

### Monocyte/Macrophage Subsets Show Tissue-Specific Patterns

We first dissected the gene signatures of all 13 myeloid subsets revealed in this study (Figures 2A–2C, S4A–S4B, and Table S2). Among these clusters, mast cells (hM01) expressed a unique set of genes, such as *TPSAB1/2*, *CPA3*, *MS4A2* and *KIT* (Figure 2B) and, in contrast to lung cancer (Lavin et al., 2017), showed comparable enrichment in both tumors and normal mucosa (Figure 2C), in line with their contribution to the homeostasis of gut bacteria and regulation of inflammation in CRC (Blathner et al., 2010; Krystel-Whittemore et al., 2016).

Three DC subsets (hM02–hM04), plasmacytoid DC (pDC), cDC2, and cDC1 cells, characterized by high expression of *HLA-DRs* and low expression of *CD14* (Figures S4B and S4C), were also identified and were further distinguished by specific expression of *LILRA4/LILRB4*, *CD1C/FCER1A*, and *XCR1/BATF3*, respectively (Figure 2B). These DCs all showed comparable enrichment in both CRC tumors and normal mucosa (Figure 2C). Three blood-enriched clusters (hM05–hM07) were characterized as classical  $CD14^{hi}CD16^-$ , non-classical  $CD14^+CD16^{hi}$ , and intermediate  $CD14^{hi}CD16^+$  monocytes (Figures S4A and S4B), largely consistent with previous reports (Dutertre et al., 2019).

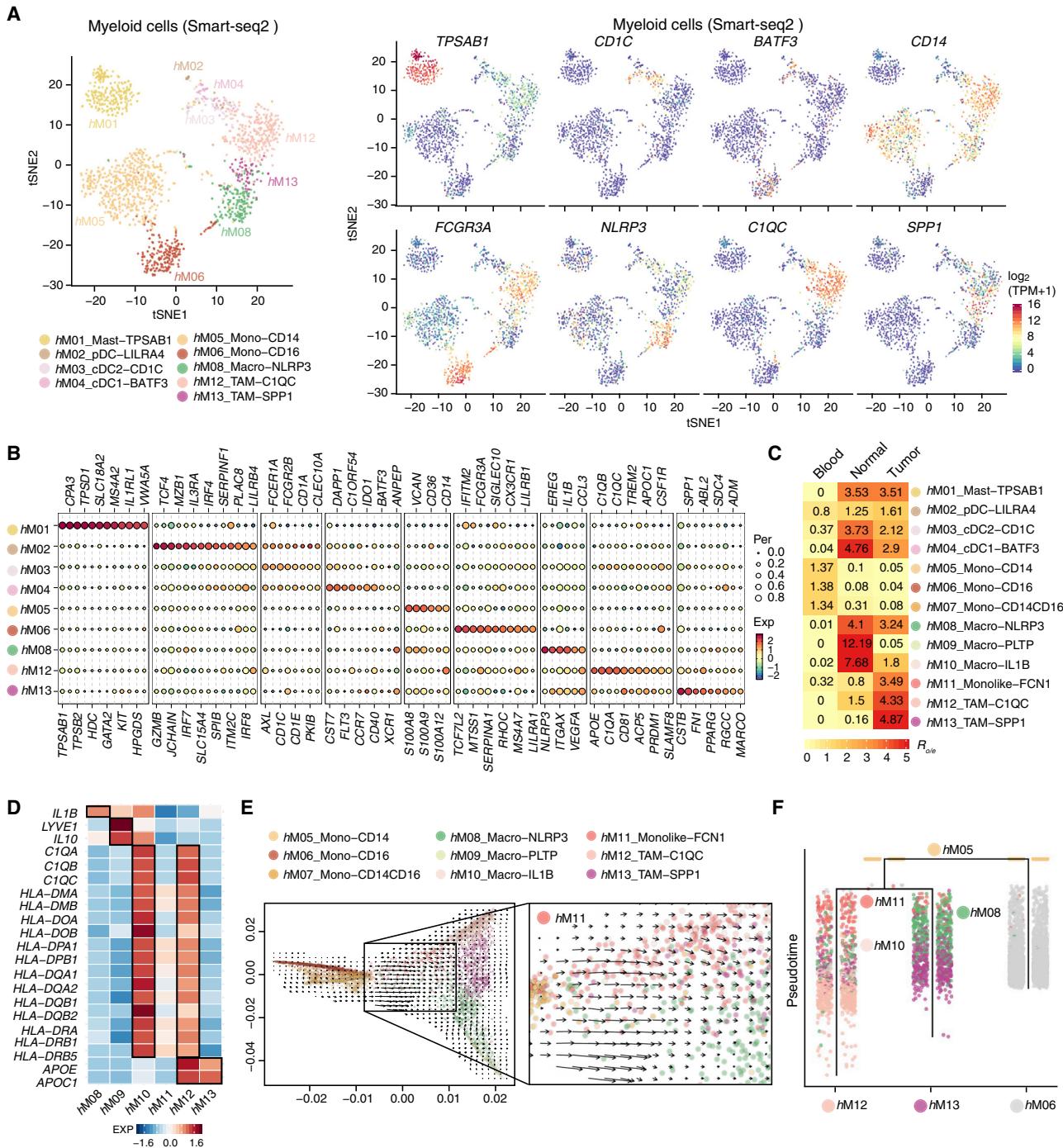
The remaining clusters were identified as macrophages based on their high expression of *CD68*, *CD163*, and *MRC1* (encoding *CD206*). Among them, clusters showing comparable (hM08) or preferential enrichment (hM09 and hM10) in normal mucosa versus tumors were denoted as resident tissue macrophages (RTMs), whereas the remaining tumor-enriched clusters were

denoted as TAMs (hM12–hM13; Figure 2C). Expression of the proinflammatory cytokine gene, *IL1B*, was observed in all RTM populations, with *NLRP3<sup>+</sup>* RTMs showing the highest expression (Figure 2D), consistent with the role of the NLRP3 inflammasome in activation of IL-1 $\beta$  and regulation of intestinal homeostasis (Zaki et al., 2011). While both *NLRP3<sup>+</sup>* and *PLTP<sup>+</sup>* RTMs showed lower expression of the *HLA-DR* gene compared with *IL1B<sup>+</sup>* RTMs, the *PLTP<sup>+</sup>* RTMs exhibited specific expression of *LYVE1* and *IL10* (Figure 2D), resembling recently reported Lyve1 $hi$ MHCII $lo$  monocyte-derived RTMs, which mostly reside alongside blood vessels and possessed a critical role in restraining inflammation and fibrosis (Chakarov et al., 2019).

### TAMs Mainly Develop from Unique Tumor-Infiltrating Monocyte-like Precursors

Recent studies in mice have suggested that TAMs can originate both from RTMs and newly recruited monocytes that subsequently differentiate into macrophages (Gubin et al., 2018; Zhu et al., 2017). Using RNA velocity analysis (La Manno et al., 2018) embedded on a diffusion map to infer the future fate of cells, we identified a strong directional flow from CD14-expressing monocytes toward *FCN1<sup>+</sup>* monocyte-like cells and different macrophage populations (Figure 2E). Notably, tumor-enriched *FCN1<sup>+</sup>* monocyte-like cells showed high similarity to blood CD14 $+$  monocytes, likely representing a monocyte population migrating into tumors and harboring a tumor-specific transcriptional program (Figure S4D).

Further deciphering the transcriptional trajectories of macrophages by two orthogonal algorithms, URD (Farrell et al., 2018) and PAGA (Wolf et al., 2019), suggested that *FCN1<sup>+</sup>* monocyte-like cells could subsequently give rise to *C1QC<sup>+</sup>* and *SPP1<sup>+</sup>* TAM populations through different RTMs (Figures 2F and S4E). Interestingly, *C1QC<sup>+</sup>* TAMs were connected to *IL1B<sup>+</sup>* RTMs, and both clusters expressed *C1Q* complement components and *HLA-DR*. Of note, *IL1B<sup>+</sup>* RTMs exhibited lower expression of *APOE* and *APOC1* compared with *C1QC<sup>+</sup>* TAMs (Figure 2D), indicating that these cells could partially resemble the functional phenotypes of *C1QC<sup>+</sup>* TAMs, but have not upregulated transcriptomic programs related to the response to tumors. Conversely, *SPP1<sup>+</sup>* TAMs were primarily connected to *NLRP3<sup>+</sup>* RTMs, and both clusters expressed low-level of *HLA-DRs*. Indeed, a fraction of *NLRP3<sup>+</sup>* RTMs were still present inside the tumor (Figure 2C), suggesting that they could convert to *SPP1<sup>+</sup>* TAMs in the TME. Collectively, our data suggest that both *C1QC<sup>+</sup>* and *SPP1<sup>+</sup>* TAMs develop from tumor-infiltrating monocyte-like precursors, whereas *SPP1<sup>+</sup>* TAMs might also derive from *NLRP3<sup>+</sup>* RTMs (Figure S4F). Furthermore, *in vitro* differentiation studies culturing CD14 $+$  monocytes under normoxic or hypoxic conditions in the presence of cytokines or growth factors that were differentially expressed in TAM populations (Table S3A, STAR Methods) demonstrated that IL-1 $\beta$  and VEGF family members were able to upregulate expression of MARCO, a marker of *SPP1<sup>+</sup>* TAMs, under hypoxic conditions (Figure S4G), consistent with the hypothesis that the TME can promote the development of this TAM population. However, fully elucidating the complex macrophage developmental trajectories occurring within tumors and normal tissues require additional *in vitro* cell differentiation and *in vivo* lineage tracing studies.



**Figure 2. Characterization of Tumor-Infiltrating Myeloid Cells by scRNA-seq in Human CRC**

(A) t-SNE plot showing 9 myeloid clusters (left) and expression levels of selected genes (right) (Smart-seq2 scRNA-seq).

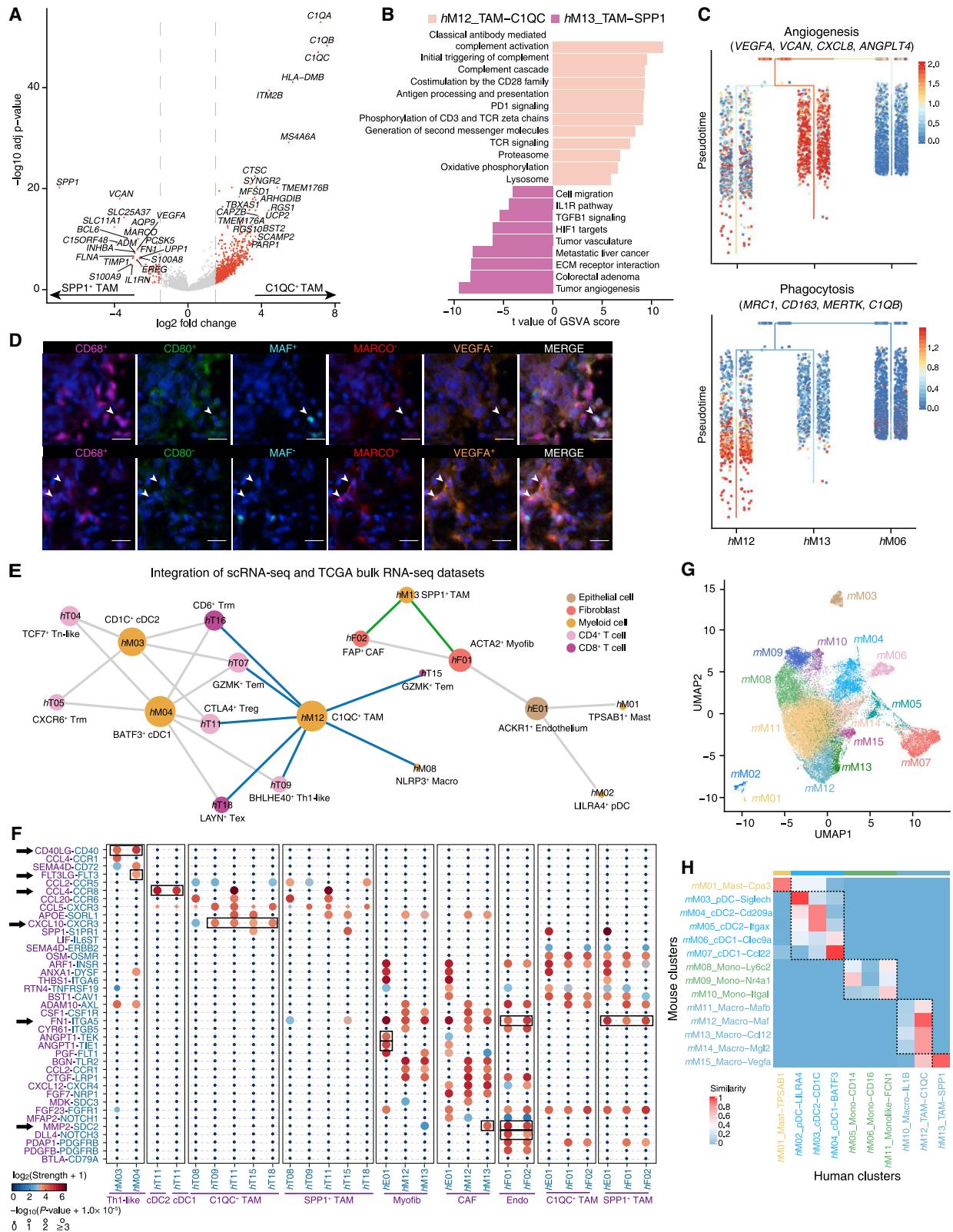
(B) Bubble heatmap showing marker genes across 9 myeloid clusters from (A). Dot size indicates fraction of expressing cells, colored according to z-score normalized expression levels.

(C) Tissue prevalence estimated by Ro/e score (10× scRNA-seq) (STAR Methods).

(D) Heatmap showing expression patterns of selected genes across indicated myeloid clusters (10× scRNA-seq).

(E and F) Inferred developmental trajectory of randomly selected monocytes and macrophages by RNA velocity (E) and monocytes and macrophages enriched in tumors by URD (F) (10× scRNA-seq). Cells are colored according to their cluster origins.

See also Figure S4 and Table S2.



(legend on next page)

### Dichotomous Functional Phenotypes of TAMs in Human CRC

In contrast to breast cancer and lung cancer TAMs that showed a continuous spectrum of phenotypes in the TME (Azizi et al., 2018; Lambrechts et al., 2018), the TAMs in CRC exhibited a remarkable dichotomy (Figure 2F), suggestive of defined cellular differentiation pathways. Importantly, both TAMs expressed distinct transcription factors, with *MAF/MAFB* and *FOS/JUN* in *C1QC<sup>+</sup>* TAMs and *CEBPB* and *ZEB2* in *SPP1<sup>+</sup>* TAMs (Figure S4H), consistent with the development and function of these two subsets of TAMs being stably controlled by transcriptional networks. Expression analyses based on genes associated with “classically activated” (M1) and “alternatively activated” (M2) macrophages in TAMs (Azizi et al., 2018) did not explain the dichotomy of *C1QC<sup>+</sup>* TAM and *SPP1<sup>+</sup>* TAM identified in CRC (Figure S4I, STAR Methods).

Further comparing the differentially expressed genes between TAM populations revealed that *C1QC<sup>+</sup>* TAMs exhibit high expression of complement C1Q genes, *TREM2*, *MERTK*, and *CD80* (Figure 3A and Table S3B). In contrast, the *SPP1<sup>+</sup>* TAMs showed specific expression of *SPP1*, *MARCO*, and *VEGFA* (Figure 3A). Evaluating known pathway expression in both TAM populations using gene set variation analysis (GSVA) revealed a strong enrichment of tumor angiogenesis, ECM receptor interaction, and tumor vasculature pathways in *SPP1<sup>+</sup>* TAMs, while the complement activation and antigen processing and presentation pathways were significantly increased in *C1QC<sup>+</sup>* TAMs (Figure 3B). Strikingly, *SPP1<sup>+</sup>* TAMs additionally exhibited specific enrichment of colorectal adenoma and metastatic liver cancer pathways (Figures 3B and 3C), suggesting a pro-tumorigenic/pro-metastatic role in CRC. Multicolor imaging data also confirmed the existence of these two subsets of TAMs by co-expression of *CD68*, *CD80*, *MAF* and *CD68*, *MARCO*, *VEGFA*, respectively (Figure 3D). Moreover, only *C1QC<sup>+</sup>* TAMs could be identified in the colon mucosa of ulcerative colitis (UC) and healthy individuals (Smillie et al., 2019), whereas *SPP1<sup>+</sup>* TAMs were largely absent in non-cancer tissues (Figures S4J and S4K, STAR Methods), suggesting their dichotomous functional phenotypes in the CRC TME.

### TAM and cDC Subsets Comprise the Core of a Predicted Cell-Cell Interaction Network

To interrogate global cell-cell interactions in CRC, we performed computational modeling by combining scRNA-seq and TCGA

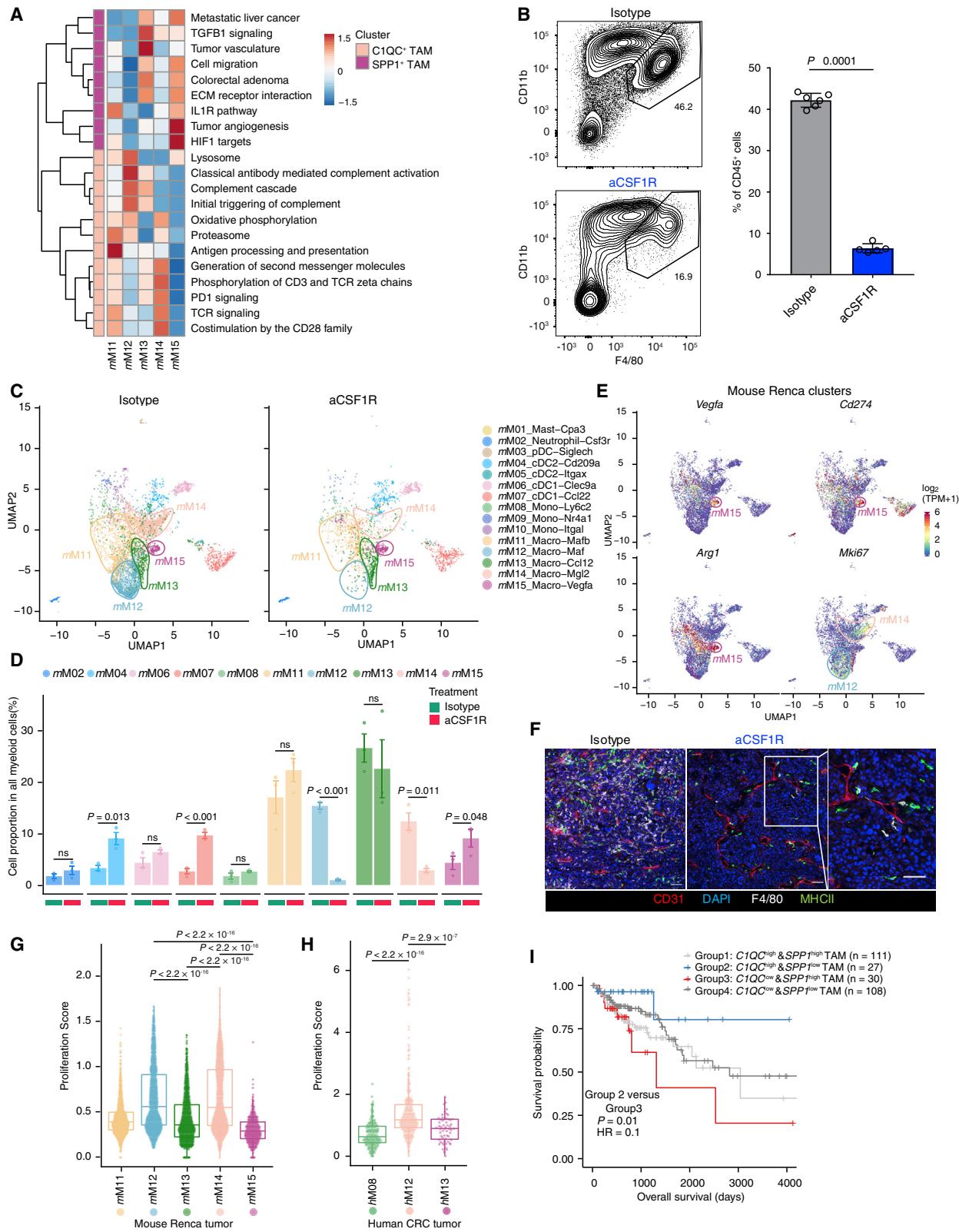
bulk RNA-seq datasets, a method previously used in cancer studies (Jerby-Arnon et al., 2018; Puram et al., 2017; Tirosh et al., 2016). In addition to identifying genes exhibiting strong correlations with the high abundance of a particular cluster (Figure S5A, Table S4A, STAR Methods), we also took those highly correlated genes generated from TCGA and matched them back to our scRNA-seq clusters, thus inferring the co-occurrence of specific cell types (Figure S5B). By applying computational modeling to all cell subsets enriched in tumor, we established a correlative cell-cell interaction network in CRC. Similar analysis was also performed using the normal tissue dataset from GTEx (Figure S5C). These analyses identified known interactions between fibroblasts and endothelial cells in the tumor and between follicular B cells and T follicular helper (Tfh) cells in the normal mucosa (Figures 3E and S5C), suggesting its potential to identify biologically meaningful interactions between less well-characterized cell types.

In the adjacent normal mucosa, interactions were mainly identified between B cells, T cells and DCs, likely reflecting crosstalk in lymphoid follicles of the colon (Figure S5C). In contrast, distinct interactions were identified in tumors, with TAMs and cDCs acting as the core of the predicted network, harboring the most connections with other cell types (Figure 3E). While *SPP1<sup>+</sup>* TAMs showed interactions with CAFs and myofibroblasts, *C1QC<sup>+</sup>* TAMs and two groups of cDCs mainly interacted with other immune cells, especially T cell subsets (Figure 3E), suggesting a function in regulating anti-tumor T cell responses.

To infer molecular interactions mediating cell-cell interactions, we calculated the attraction strengths of ligand-receptor pairs in our scRNA-seq dataset and identified hundreds of interaction pairs displaying significant cell population specificity using a simulation analysis similar to previous methods (STAR Methods) (Ramilowski et al., 2015; Vento-Tormo et al., 2018). Of the ligand-receptor pairs pertaining to myeloid and T cells, CXCL10-CXCR3 was significantly enriched in *C1QC<sup>+</sup>* TAMs (Figures 3F, S5D, and S5E), implicating a potential role of *C1QC<sup>+</sup>* TAMs in recruiting or activating T cells. Compared with *C1QC<sup>+</sup>* TAMs and other leukocytes, *SPP1<sup>+</sup>* TAMs showed preferential expression of *SDC2*, which binds to *MMP2*, a gene highly expressed by CAFs and endothelial cells (Figures 3F, S5D, and S5F) and has been reported to be associated with tumor growth and metastasis in multiple cancer types (Kenny et al., 2008; Zucker and Vacirca, 2004). Other ligand-receptor pairs involving interactions

**Figure 3. Dichotomy of Functional Phenotypes and Predicted Interactions of TAMs in Human CRC**

- (A) Volcano plot showing differentially expressed genes between *C1QC<sup>+</sup>* and *SPP1<sup>+</sup>* TAMs (Smart-seq2 scRNA-seq). *P*-value < 0.05, Two-sided unpaired limma-modulated t test; log2(fold change)  $\geq 1.5$ .
  - (B) Differential pathway enriched in *C1QC<sup>+</sup>* and *SPP1<sup>+</sup>* TAMs (Smart-seq2 scRNA-seq) by GSVA. Two-sided unpaired limma-modulated t test.
  - (C) Projection of TAM clusters by expression of angiogenesis- and phagocytosis-related genes on URD map in Figure 2F.
  - (D) Representative example of a CRC tumor stained by multi-colored IHC. The upper panel indicates *CD68<sup>+</sup>* *CD80<sup>+</sup>* *MAF<sup>+</sup>* TAM (hM12\_TAM-C1QC), and lower *CD68<sup>+</sup>* *MARCO<sup>+</sup>* *VEGFA<sup>+</sup>* TAM (hM13\_TAM-SPP1). Original magnification,  $\times 20$ ; scale bar, 20  $\mu$ m.
  - (E) Correlative network showing interactions of cell subsets identified in CRC tumors based on scRNA-seq and bulk TCGA expression data. Circle size indicates the number of connections with other cell types (STAR Methods).
  - (F) Bubble heatmap showing the mean attraction strength for selected ligand-receptor pairs. Dot size indicates *P*-value generated by permutation test, colored by attraction strength levels (STAR Methods).
  - (G) UMAP plot showing integrated myeloid clusters from MC38 and Renca tumor models.
  - (H) Similarity analysis of myeloid clusters from two mouse models and human CRC (10 $\times$  scRNA-seq).
- See also Figures S4, S5, S6, S8 and Tables S3, S4, S5.



(legend on next page)

between *SPP1*<sup>+</sup> TAMs and other cell subsets included *SPP1-IT-GAV* and *FN1-ITGA5*, with *SPP1*<sup>+</sup> TAMs expressing higher levels of *SPP1* and *FN1* (Figures 3F, S5D, and S5F). Our analyses suggest that *SPP1* and *FN1* may interact with certain integrins to promote tumorigenesis in CRC. Collectively, our scRNA-seq data suggest that TAMs and cDCs comprise the core of a cell-cell interaction network and that TAMs may play dichotomous functions through interacting with distinct immune and stromal cells in the CRC TME.

### Major Tumor-associated Myeloid Cell Populations Are Shared Between Human and Mouse

Next, we sought to study the effect of manipulating distinct myeloid populations on the anti-tumor immune response in mouse tumor models by scRNA-seq and relate these data to our findings on myeloid cell heterogeneity in human. We focused on two therapeutic strategies that have been widely explored in clinical trials: TAM depletion by CSF1R blockade and DC activation by a CD40 agonist. We first evaluated these therapies in multiple syngeneic mouse tumor models, identifying growth of Renca tumors as being sensitive to an anti-CSF1R blocking antibody and growth of MC38 as being sensitive to an anti-CD40 agonist antibody (Figure S6A). Based on these findings, multiple scRNA-seq studies were performed using the 10 $\times$  Genomics platform on immune cells isolated from either Renca tumors following anti-CSF1R treatment (Figure S6B) or MC38 tumors and tumor draining lymph nodes (tdLNs) following anti-CD40 treatment (Figure S6C).

Graph-based clustering was used to identify the major immune cell subsets, including T cell and myeloid populations, in both Renca and MC38 scRNA-seq datasets (Figure S6D and Table S5). To better compare tumor-associated murine myeloid cell subsets with those from human, we next integrated and clustered myeloid cells from both models, identifying 15 discrete cell clusters (Figure 3G). Each cluster was associated with a unique gene expression pattern indicative of a distinct cell type (Figure S6E), and clusters were largely identified in both MC38 and Renca tumors (Figure S6F). To identify relationships between human and mouse populations, we performed a systematic similarity analysis, identifying multiple corresponding myeloid populations across species (Figure 3H).

Focusing first on cDC populations, we identified two murine cDC1 subsets (mM06 and mM07) (Figure S6G), both of which were mapped to a single cDC1 cluster (hM04) in human CRC

(Figure 3H). Further evaluation of heterogeneity within the human *BATF3*<sup>+</sup> cDC1 cluster by re-clustering these cells found evidence for both phenotypes of cDC1 cells within human (Figure S6H). Notably, these two cDC1 phenotypes resemble known immature and activated states of cDC1 cells (Ardouin et al., 2016) and recently described “activated” *LAMP3*<sup>+</sup>*CCR7*<sup>+</sup> DCs in both human and mouse tumors (Figure S6I) (Zhang et al., 2019; Zilionis et al., 2019). We also identified two cDC2 clusters in murine tumors (mM04 and mM05) that mapped to a single cDC2 population in human CRC (hM03) (Figure 3H). Of note, the *Itgax*<sup>+</sup> cDC2 cluster likely represents a classical cDC2 subset, whereas *Cd209a*<sup>+</sup> cDC2 cluster shares features with monocyte-derived DCs that may differentiate from newly infiltrated blood monocytes (Cheong et al., 2010).

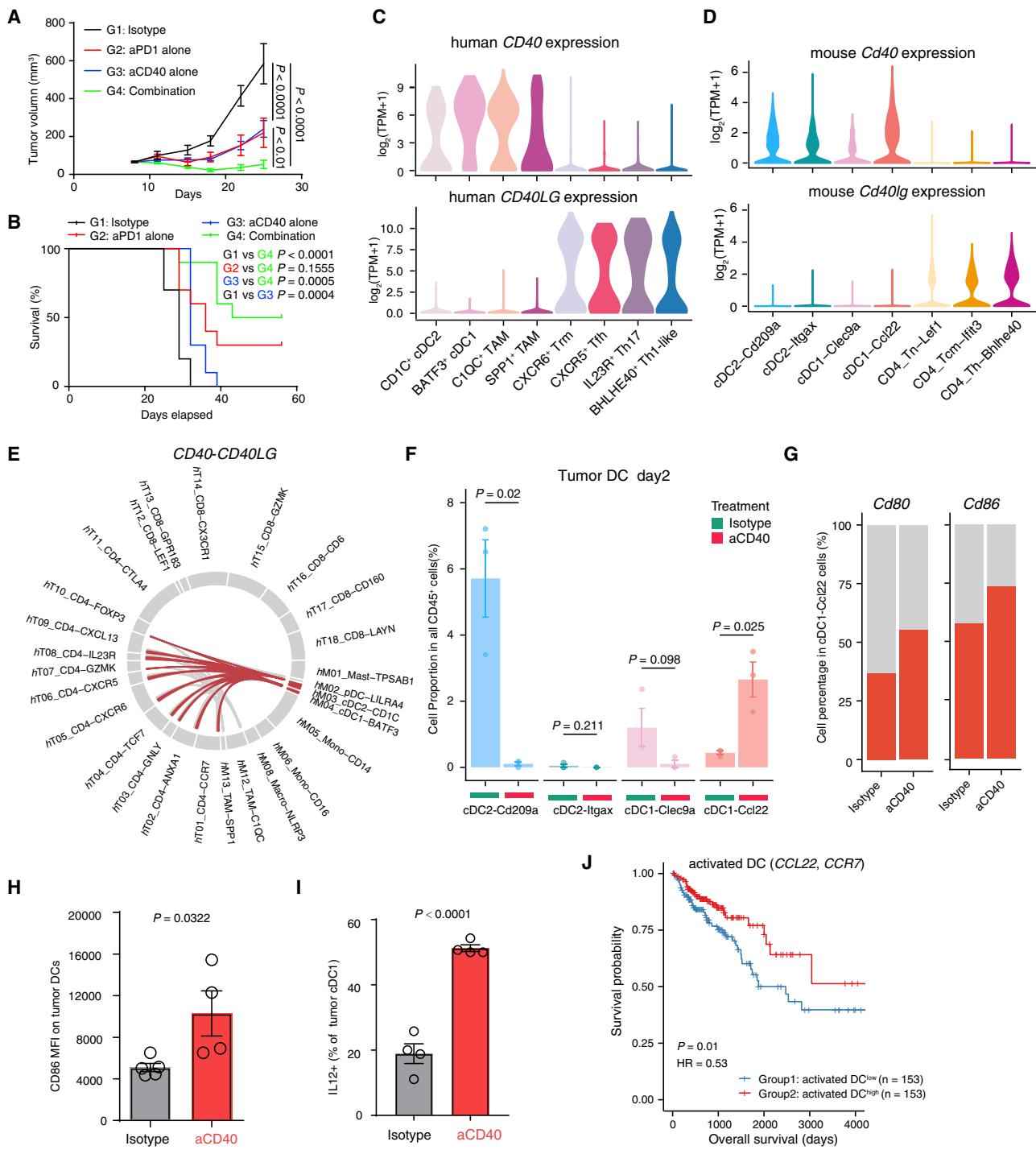
In contrast to DC populations, TAMs showed a greater degree of phenotypic heterogeneity between mouse and human, in line with recent observations in lung cancer (Zilionis et al., 2019). Nevertheless, mouse macrophage clusters (mM11-mM14) showed the highest similarity to human *C1QC*<sup>+</sup> TAMs and minimal overlapping gene expression with human *SPP1*<sup>+</sup> TAMs. In contrast, mouse macrophage cluster mM15\_Macro-Vegfa showed the greatest similarity with human *SPP1*<sup>+</sup> TAMs (Figure 3H). Performing a similar pathway analysis on mouse TAM subsets as was used for human TAM clusters (Figure 3B), we found that mouse TAM populations also segregate on the basis of their angiogenic, hypoxic, and T cell interaction gene signatures (Figure 4A). These data suggest the presence of functionally analogous TAM populations between human CRC patients and mouse tumor models.

### A Pro-Angiogenic Macrophage Population Is Resistant to Therapeutic Treatment with Anti-CSF1R Blockade

We next sought to elucidate the effects of anti-CSF1R on TAM populations in mouse and relate these findings to corresponding populations in human. As expected, treatment of mice bearing Renca tumors with anti-CSF1R decreased the frequency of TAMs relative to treatment with a control antibody (Figure 4B). However, a detectable population of TAMs remained following anti-CSF1R treatment. Further evaluation of this anti-CSF1R resistant population showed a preferential decrease in macrophages expressing high levels of F4/80 (Figures S7A-S7D), suggesting differential sensitivity of distinct macrophage populations to anti-CSF1R treatment. Consistent with these findings, analysis of scRNA-seq data from Renca tumors revealed a

**Figure 4. CSF1R Blockade Leads to Depletion of Select Macrophage Populations from Tumors**

- (A) Heatmap showing different pathways from Figure 3B enriched in the integrated macrophage clusters from mouse tumors by GSVA analysis, colored by z-score transformed mean GSVA scores.
  - (B) Macrophage depletion in Renca tumors from mice dosed with isotype control or anti-CSF1R antibodies assessed by flow cytometry (left). Quantification of macrophages as a percentage of total CD45<sup>+</sup> lymphocytes (right). Two-sided t test.
  - (C) UMAP plots of 7,597 myeloid cells from mice bearing Renca tumors treated with isotype (left) or anti-CSF1R antibody (right).
  - (D) Abundance of mM12\_Macro-Maf and mM14\_Macro-Mgl2 TAM subsets following treatment of mice bearing Renca tumors with anti-CSF1R antibody (Study 2). One-sided t test.
  - (E) UMAP plots showing expression levels of selected genes in different myeloid clusters.
  - (F) Immunofluorescence imaging of Renca tumor sections from mice treated with isotype or anti-CSF1R antibody. Scale bar, 50 $\mu$ m.
  - (G and H) Expression of a proliferation signature across indicated mouse Renca tumor (G) and human CRC (H) myeloid clusters. Two-sided Wilcoxon test.
  - (I) The Kaplan-Meier overall survival curves of TCGA COAD and READ patients grouped by the gene signature expression of *C1QC*<sup>+</sup> TAM and *SPP1*<sup>+</sup> TAM. HR, hazard ratio. Multivariate Cox regression (STAR methods).
- See also Figure S7 and Table S5.



**Figure 5. Anti-CD40 Agonist Treatment Increases the Frequency and Activation State of cDC1 Cells in Tumors**

(A-B) Reduced tumor burden (A, Two-way ANOVA) and improved survival (B, Mantel-Cox test) after treatment with anti-CD40 agonist alone or in combination with anti-PD1.

(C) Violin plot showing expression of *CD40* and *CD40LG* in myeloid and T cell clusters from human CRC (Smart-seq2 scRNA-seq).

(D) Violin plot showing expression of *Cd40* and *Cd40lg* in myeloid and T cell clusters from MC38 tumors.

(E) Chord diagram showing preferential interactions of cDC subsets with various T cell clusters mediated by *CD40-CD40LG* in human CRC.

(F) Anti-CD40 agonist treatment-induced changes in the abundance of tumor-associated cDC1 cells (mM07\_cDC1-Ccl22). One-sided t test.

(G) Increase in the proportion of tumor-infiltrating *Ccl22*<sup>+</sup> cDC1 cells expressing *Cd80* and *Cd86* following treatment of mice bearing MC38 tumors for 2 days.

(legend continued on next page)

near complete loss of clusters mM12 and mM14, but minimal reduction of TAM clusters mM11, mM13, and mM15 following anti-CSF1R treatment (Figures 4C and 4D). Indeed, the frequencies of certain clusters increased in the anti-CSF1R treatment groups (Figure 4D), likely due to the loss of a large macrophage population from the total pool of immune cells. We further examined how differential sensitivity of TAM populations to anti-CSF1R treatment could alter the TME. We found that anti-CSF1R-resistant TAMs preferentially express genes involved in angiogenesis, such as *Vegfa*, and immunosuppression, such as *Cd274* and *Arg1* (Figure 4E). Notably, macrophages remaining in the tumor following anti-CSF1R treatment also appeared to be preferentially associated with tumor vasculature (Figure 4F) (DeNardo et al., 2011), suggestive of a role in regulating angiogenic processes.

Given the role of CSF1R signaling in controlling macrophage homeostasis (Hume et al., 2019), we next evaluated whether TAM proliferation correlated with sensitivity to anti-CSF1R treatment. Anti-CSF1R-sensitive populations (mM12, mM14) had elevated expression of *Mki67* and a higher proliferation score than resistant populations (mM11, mM13, mM15) (Figures 4E and 4G). Notably, in human tumors, C1QC<sup>+</sup> TAMs also have a higher proliferation score than *SPP1*<sup>+</sup> TAMs (Figure 4H). These data suggest that in mice, CSF1R blockade results in preferential depletion of macrophages in cell cycle, leading to differential effects of this treatment across TAM subsets. Extrapolating these findings to humans, CSF1R blockade may preferentially deplete a fraction of the C1QC<sup>+</sup> TAM subset, while sparing *SPP1*<sup>+</sup> TAMs. To further relate our findings in mice to human CRC, we compared patient survival across individuals with different levels of C1QC<sup>+</sup> TAM and *SPP1*<sup>+</sup> TAM gene signatures and found that a low C1QC<sup>+</sup> TAM and high *SPP1*<sup>+</sup> TAM gene signature combination correlated with worse prognostic outcome in CRC patients (Figure 4I). These findings suggest that anti-CSF1R treatment may be insufficient to deplete all macrophage populations with tumor growth-promoting potential, a property that may underlie the poor monotherapy efficacy of anti-CSF1R in the Renca mouse tumor model and in human cancer patients.

### Early Specific Amplification of a cDC1 Population upon Anti-CD40 Treatment

Next, we focused on dissecting the mechanisms of anti-CD40 agonist therapy. Treatment of mice bearing MC38 tumors with anti-CD40 led to reduced tumor growth (Figure S6A), which was further enhanced in combination with PD1 blockade (Figures 5A and 5B). scRNA-seq analyses from both human CRC and mouse MC38 tumor samples revealed that CD40 was expressed on multiple DC and macrophage subsets, especially cDC1 clusters (Figures 5C and 5D). Consistent with these findings and our cell-cell interaction network analysis (Figure 3E) showing the extensive interactions of cDCs with various T cell

subsets (Figure 5E and Table S4B), we found that CD40LG was highly expressed on multiple CD4<sup>+</sup> T cell clusters in human and mouse tumors (Figures 5C and 5D). Although the CD40/CD40LG pathway is known to modulate the interaction between myeloid cells and T cells (Grewal and Flavell, 1998), we aimed to identify specific myeloid and T cell subsets that were preferentially impacted by CD40 agonism.

Analysis of scRNA-seq data from different time points following treatment revealed anti-CD40 agonist-mediated enhancement of *Ccl22*<sup>+</sup> cDC1 cell frequency, while other DC subsets were either decreased or unchanged on day 2 after treatment (Figure 5F). Anti-CD40 antibody also markedly activated these cDC1 cells, as measured by increased CD80 and CD86 expression (Figures 5G and 5H). Further, anti-CD40 increased cDC1 cell production of IL-12 (Figure 5I), a cytokine that enhances Th1 development and IFN $\gamma$  production by CD8<sup>+</sup> T cells (Tait Wojno et al., 2019). While the effects of anti-CD40 antibody on TAMs were less pronounced (Figure S7E), we observed transient reduction of *Mafb*<sup>+</sup> TAMs at day 2 and significant reduction of *Maf*<sup>+</sup> TAMs at day 10 post-treatment (Figure S7E), which may be reflective of widespread changes in tumor-associated immune cell frequency induced by CD40 agonist treatment. Thus, our scRNA-seq analyses identified *Ccl22*<sup>+</sup> cDC1 cells as the primarily myeloid cell type that is activated early after anti-CD40 antibody treatment in the MC38 model. Importantly, signature genes of activated cDC1s positively correlated with favorable overall survival of CRC patients (Figure 5J), suggesting that the ability of anti-CD40 to activate these cells may have relevance to human cancer.

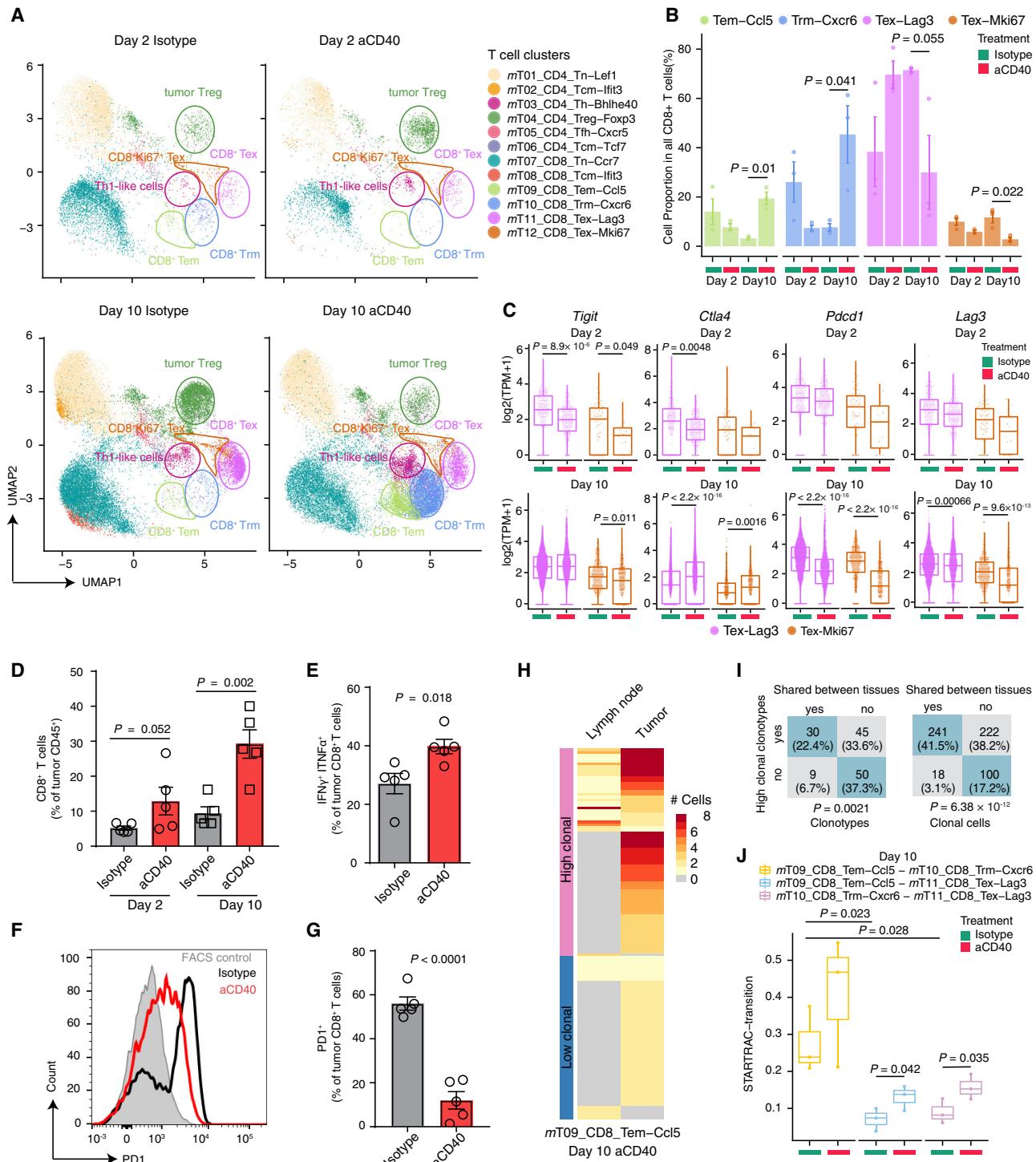
### Anti-CD40 Treatment Increases Effector Memory CD8<sup>+</sup> T Cells and Induces the Activation and Expansion of *Bhlhe40*<sup>+</sup> Th1-like Cells in Tumor

We next investigated the consequences of anti-CD40 treatment on the function of tumor-infiltrating T cells. scRNA-seq analysis of T cells from the MC38 tumor model (Figures S7F and S7G) demonstrated that the frequency of certain CD4<sup>+</sup> and CD8<sup>+</sup> T cell subsets were affected by CD40 agonist treatment (Figure 6A). Since we also captured TCR  $\alpha$ - and  $\beta$ -chain sequences from scRNA-seq (Table S6A), we used the previously developed STARTRAC indices to define the clonal expansion, migration and developmental transition of T cells (Zhang et al., 2018). Anti-CD40 treatment significantly increased the proportion of tumor memory CD8<sup>+</sup> T cell subsets but reduced the frequency of exhausted CD8<sup>+</sup> T cell subsets, especially at 10 days post-treatment (Figure 6B). Moreover, anti-CD40 treatment significantly changed the expression profiles and kinetics of exhaustion markers in *Lag3*<sup>+</sup> Tex and *Mki67*<sup>+</sup> Tex subsets, with expression of *Ctla4* and *Tigit* being rapidly reduced upon treatment and a significant reduction in *Pdcd1* and *Lag3* expression observed at 10 days post-treatment (Figure 6C). Flow cytometry confirmed these scRNA-seq findings, demonstrating that

(H–I) Expression of CD86 (H) and IL-12 (I) protein in MC38 tumor-infiltrating cDC1 cells following treatment with an anti-CD40 agonist antibody, as measured by flow cytometry. Two-sided t test.

(J) Kaplan-Meier overall survival curves of TCGA COAD and READ patients by the mean expression of activated DC gene signature. HR, hazard ratio. Multivariate Cox regression.

See also Figure S7 and Table S4.



**Figure 6. Anti-CD40 Agonist Treatment Increases the Frequency of CD8<sup>+</sup> Memory T Cells in Tumors**

(A) UMAP plot showing alterations of T cell clusters following treatment of MC38 tumors for 2 or 10 days with isotype control or anti-CD40 agonist antibody.

(B) Frequency of CD8<sup>+</sup> memory T cell subsets as a proportion of all CD8<sup>+</sup> T cells. One-sided t test.

(C) Expression of selected genes on *Lag3*<sup>+</sup> Tex and *Mki67*<sup>+</sup> Tex cells. Two-sided Wilcoxon test.

(D-G) Abundance and effector function of CD8<sup>+</sup> T cells (D and E) and frequency of exhausted CD8<sup>+</sup> T cells (F and G) following treatment of MC38 tumors for 2 or 10 days (D) and 7 days (E-G) with an anti-CD40 agonist antibody, as measured by flow cytometry. Two-sided t test.

(legend continued on next page)

anti-CD40 treatment increases effector CD8<sup>+</sup> T cells and reduces PD1<sup>+</sup> exhausted CD8<sup>+</sup> T cells in tumor (Figures 6D–6G). Using STARTRAC analysis, we found that Cc15<sup>+</sup> Tem CD8<sup>+</sup> T cells had a higher migration index than other CD8<sup>+</sup> T cell subsets upon anti-CD40 treatment (Figure S7H), consistent with our earlier findings that CD8<sup>+</sup> Tem cells are more migratory than Trm or exhausted CD8<sup>+</sup> T cells (Zhang et al., 2018). Indeed, anti-CD40 treatment dramatically increased the proportion of Cc15<sup>+</sup> Tem CD8<sup>+</sup> T cells, but not the Cxcr6<sup>+</sup> Trm population in tDLNs (Figure S7I). Further dissecting the TCR clonotypes within the Cc15<sup>+</sup> Tem subset revealed that those cells with higher clonal expansion exhibited more TCR sharing between tumor and tDLN, indicative of more profound mobility of these cells upon anti-CD40 treatment (Figures 6H, 6I, S7J and Table S6B). In addition, the transition index between Cc15<sup>+</sup> Tem and Cxcr6<sup>+</sup> Trm CD8<sup>+</sup> T cells was significantly higher than that of other pairs among CD8<sup>+</sup> T cell subsets at baseline, suggesting that certain Cxcr6<sup>+</sup> Trm cells in tumor may be derived from infiltrated Cc15<sup>+</sup> Tem cells, in a process that was further enhanced by anti-CD40 treatment (Figure 6J). Taken together, our data reveal unique effects of anti-CD40 treatment on the expansion, migration, and transition of tumor-infiltrating Tem and Trm CD8<sup>+</sup> T cells.

We found that CD40 agonist treatment also had a dramatic impact on tumor-infiltrating CD4<sup>+</sup> T cell subsets. Although anti-CD40 treatment significantly amplified Treg cells at 2 days post-treatment, the proportion of Treg cells was reduced at 10 days post-anti-CD40 treatment compared to control-treated groups (Figure 7A). STARTRAC-expansion index confirmed that anti-CD40 treatment enhanced Treg expansion on day 2 and constrained their expansion on day 10 (Figure 7B). In contrast, anti-CD40 treatment specifically increased the proportion of Bhlhe40<sup>+</sup> Th1-like cells on both day 2 and day 10, which was confirmed by anti-CD40 triggered clonal expansion of these cells on both days (Figure 7B). Interestingly, we noticed that anti-CD40 treatment significantly induced the expression of Cd40lg on Bhlhe40<sup>+</sup> Th1-like cells in our scRNA-seq dataset (Figures 7C and 7D), a finding that was confirmed by flow cytometry analysis (Figure 7E). These data suggest that a CD40 agonist might be able to further enhance the crosstalk of tumor-associated BHLHE40<sup>+</sup> Th1-like cells and cDC1 cells.

We previously identified IFNG-expressing BHLHE40<sup>+</sup> Th1-like cells as being enriched in the MSI CRC patient population, which shows favorable response to ICB therapy (Le et al., 2015; Zhang et al., 2018). We confirmed that Bhlhe40<sup>+</sup> CD4<sup>+</sup> T cells isolated from mouse MC38 tumors also expressed higher Bhlhe40 (Figure S7K) and produced significantly more IFN $\gamma$  than other effector CD4<sup>+</sup> T cells (Figures 7F and 7G), expressed the proliferation marker Ki67 by FACS, and were further induced to proliferate by anti-CD40 treatment (Figure 7H). Thus, anti-CD40 activation of tumor cDC1 cells may result in increased frequency of IFN $\gamma$ -producing tumor-infiltrating CD4<sup>+</sup> Th cells.

Consistent with this hypothesis, *in vitro* studies demonstrated that anti-CD40 treatment promoted DC maturation which further increased the expression of Bhlhe40, but not Tbx21, in CD4<sup>+</sup> T cells stimulated with suboptimal concentrations of antigen (Figures 7I and S7L–S7M). Further analyzing the correlation of IFNG-expressing BHLHE40<sup>+</sup> Th1-like cells and cDC cells in human CRC patients revealed that a gene signature of Th1-like cells exhibited preferentially positive correlation with both mature and immature cDC1 cells. These cDC1 cells were also significantly enriched in MSI-H CRC patients (Figure 7J), suggesting a crosstalk between cDC1 cells and BHLHE40<sup>+</sup> Th1-like cells. Importantly, our finding that anti-CD40 increased Bhlhe40<sup>+</sup> Th1-like cells may provide a mechanistic explanation for why CD40 agonist treatment can successfully synergize with anti-PD1 in this model (Figures 5A and 5B).

## DISCUSSION

Here, we have leveraged the advantages of two scRNA-seq approaches to generate an atlas of immune and non-immune cells from human CRC patients. Classifying these cells into specific lineages and deducing a cell-cell interaction network (Figure S8A) revealed specific populations of myeloid cells as central nodes of interaction, leading us to focus on their functions in human and pre-clinical mouse tumor models (Figures S8B and S8C). In contrast to the relative stable phenotypes of lymphocyte, stromal, and DC subsets across different tissues and cancer types, the distinct characteristics of macrophages appeared to be dependent on their tissue/tumor origins (Azizi et al., 2018; Chevrier et al., 2017; Lambrechts et al., 2018; Lavin et al., 2017; Venetiecher et al., 2017; Zilionis et al., 2019). While TAMs in breast and lung cancer patients exhibit a continuous spectrum phenotype (Azizi et al., 2018; Lambrechts et al., 2018), in this study we identified two distinct TAM populations in CRC, consisting of C1QC<sup>+</sup> and SPP1<sup>+</sup> TAMs, both of which may arise from an intermediate state of FCN1<sup>+</sup> monocyte-like cells in the tumor. Consistent with previous studies, neither of these populations fit the M1 and M2 dichotomous phenotypes (Azizi et al., 2018; Müller et al., 2017). Rather, C1QC<sup>+</sup> TAMs showed preferential expression of genes involved in phagocytosis and antigen presentation, while SPP1<sup>+</sup> TAMs were enriched for regulators of angiogenesis. Only SPP1<sup>+</sup> TAMs exhibited preferential enrichment in tumor compared to normal mucosa from healthy donors, suggesting their key roles in the tumorigenesis of CRC.

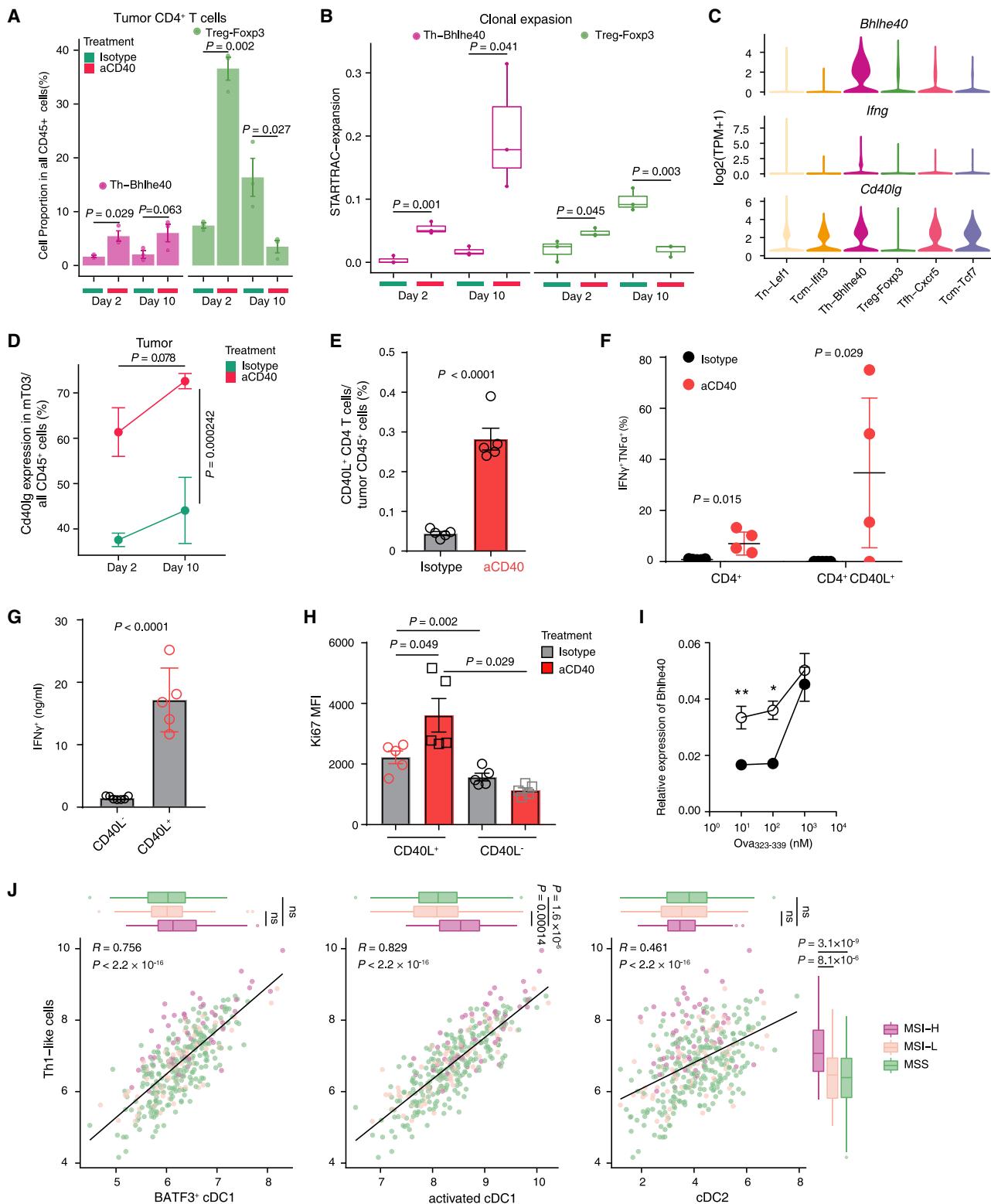
Depleting immunosuppressive TAMs is an attractive therapeutic approach to promote anti-tumor immune responses. However, their heterogeneity, diverse functions, and predicted cell-cell interaction patterns argues for more subset-specific strategies. Indeed, although CSF1R blockade can result in significant TAM depletion, CSF1R inhibition has provided minimal therapeutic benefit in cancer patients as a monotherapy (Papadopoulos et al., 2017; Ries et al., 2014). Recent data suggest

(H) The distribution of clonal clonotypes within the Cc15<sup>+</sup> Tem cluster in lymph node and tumor. Each row represents an individual clonotype. “High clonal” clonotypes had > = 3 cells within Cc15<sup>+</sup> Tem cluster, “low clonal” refers to the remaining clonotypes.

(I) TCR sharing between lymph node and tumor Cc15<sup>+</sup> Tem cells with high versus low clonal clonotypes. Fisher’s exact test.

(J) Developmental transition of CD8<sup>+</sup> T cells quantified by pairwise STARTRAC-transition indices for each mouse. One-sided t test.

See also Figure S7 and Table S6.



**Figure 7. Anti-CD40 Agonist Treatment Expands *Bhlhe40*<sup>+</sup> Th1 Cells in Tumors**

(A) Anti-CD40 agonist treatment-induced changes in the frequency of tumor-infiltrating CD4<sup>+</sup> T cell clusters as a proportion of all CD45<sup>+</sup> cells. One-sided t test.

(B) Clonal expansion levels of CD4<sup>+</sup> T cell clusters quantified by STARTRAC-expansion. One-sided t test.

(C) Violin plot showing expression of selected genes in CD4<sup>+</sup> T cell clusters.

(legend continued on next page)

multiple compensatory mechanisms may limit the anti-tumor activity of CSF1R blockade, including treatment-induced recruitment of *Foxp3*<sup>+</sup> Tregs or myeloid-derived suppressor cell populations (Gyori et al., 2018; Kumar et al., 2017) and the ability of macrophages to rely on alternative signals for recruitment, proliferation and/or survival (Neubert et al., 2018; Pradel et al., 2016). Our data identified murine TAM subsets that are both refractory to depletion by anti-CSF1R and show similarity with human *SPP1*<sup>+</sup> TAMs that have a pro-angiogenic signature and, based on predicted cell-cell interaction network, likely engage in crosstalk with CAFs and endothelial cells. Thus, persistence of this population, and loss of pro-inflammatory *C1QC*<sup>+</sup> TAM populations, may represent a previously unrecognized mechanism of resistance to indiscriminate macrophage depletion via anti-CSF1R therapy. Together with the poor prognosis of patients with high *SPP1*<sup>+</sup> and low *C1QC*<sup>+</sup> TAM signatures, our findings suggest that specific depletion of *SPP1*<sup>+</sup> TAMs could ultimately lead to improved myeloid-targeted immunotherapy or enhanced combination with ICB therapies.

The importance of understanding DC subtypes is similarly demonstrated. Of the cDC1 subsets identified in both MC38 and Renca mouse tumors by scRNA-seq, only one (mM06) expresses the classical cDC1 markers *Clec9a* and *Xcr1*. In contrast, *Ccl22*<sup>+</sup> cDC1s (mM07) bear similarities to the *CD40*<sup>+</sup>*CCR7*<sup>+</sup>*LAMP3*<sup>+</sup> subpopulation of cDC1 cells in human CRC and tumor-associated *LAMP3*<sup>+</sup> DCs recently described by others (Zhang et al., 2019; Zilionis et al., 2019). Previous studies demonstrated that murine cDC1 cells downregulate *Xcr1* and *Clec9a* expression when they mature or become activated (Ardouin et al., 2016; Merad et al., 2013), suggesting *Ccl22*<sup>+</sup> cDC1 cells may be activated cDC1s and not a developmentally distinct subset. Importantly, CD40 agonist treatment significantly increased the frequency *Ccl22*<sup>+</sup> cDC1 cells in MC38 tumors without drastic effects on other myeloid subsets, suggesting that these activated DCs are preferentially targeted by CD40 agonists to promote anti-tumor immune responses.

Our human cell-cell interaction analyses predicted an unexpected interaction between the *BATF3*<sup>+</sup> cDC1 population and both exhausted CD8<sup>+</sup> T cells as well as a distinct population of the *BHLHE40*<sup>+</sup> Th1-like cells. While we did not observe significant frequency changes in exhausted CD8<sup>+</sup> T cells following anti-CD40 treatment, the expression level of several inhibitory receptors such as PD-1 was reduced. Interestingly, a rapid expansion of *Bhlhe40*<sup>+</sup> Th1-like CD4<sup>+</sup> T cells, similar to those cells identified in MSI CRC (Zhang et al., 2018), was observed upon anti-CD40 treatment that preceded an increase in the frequency of memory CD8<sup>+</sup> T cells. Given that CD40-mediated acti-

vation of DCs promoted *Bhlhe40* expression in CD4<sup>+</sup> T cells *in vitro* and the kinetics of CD40-mediated activation of different immune cells *in vivo*, including cDC1 activation, *Bhlhe40*<sup>+</sup> Th1-like CD4<sup>+</sup> T cell expansion, and memory CD8<sup>+</sup> T cell infiltration, we postulate that the expansion of *Bhlhe40*<sup>+</sup> Th1-like CD4<sup>+</sup> T cells may be downstream of *Ccl22*<sup>+</sup> cDC1 activation. These *Bhlhe40*<sup>+</sup> T cells may potentiate cDC1-mediated CD8<sup>+</sup> T cell infiltration, expansion, and anti-tumor function by providing positive feedback signals to cDC1 cells, in part via their high *Cd40lg* expression. Although the pathways regulating *BHLHE40* expression in T cells are not well defined, it has been shown that TCR activation and cytokines such as IL-1 $\beta$  produced by activated myeloid cells may contribute to its induction (Lin et al., 2016), suggesting potential mechanisms by which anti-CD40 may promote the generation of *Bhlhe40*<sup>+</sup> T cells. Targeted cell-depletion studies will help further define the precise role of the intratumoral cellular interactions in CD40 agonist-mediated responses. Nevertheless, our study provides an in-depth understanding of mechanisms regulating CD40 agonist activity and further supports the importance of *Bhlhe40*<sup>+</sup> Th1-like CD4<sup>+</sup> T cells in anti-tumor immunity and immunotherapy.

Taken together, our study reveals previously unappreciated myeloid-T cell and myeloid-stromal connections within CRC, providing mechanistic insights for immunotherapies currently in clinical development, and demonstrates an approach for dissecting the role of specific tumor-associated immune populations through complementary single cell analysis of both human and mouse tumors. Additionally, our datasets, as provided in an interactive portal (<http://crcleukocyte.cancer-pku.cn/>), can be used as a resource for further exploration of tumor-infiltrating immune cells in human and mouse.

## STAR METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- LEAD CONTACT AND MATERIALS AVAILABILITY
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Human specimens
  - *In vivo* mouse models
- METHOD DETAILS
  - Single cell collection for human samples
  - Single cell sorting, reverse transcription, amplification and sequencing for CRC samples (Smart-Seq2)

(D) Abundance of *Cd40lg* transcripts expressed by *Bhlhe40*<sup>+</sup> Th1 cells in CD45<sup>+</sup> cells. Two-way ANOVA.

(E) Frequency of CD4<sup>+</sup> CD40L<sup>+</sup> T cells in tumors from mice bearing MC38 tumors treated with an anti-CD40 agonist antibody, as quantified by flow cytometry. Two-sided t test.

(F) Frequency of tumor-infiltrating IFN $\gamma$ <sup>+</sup> TNF $\alpha$ <sup>+</sup> cells among CD4<sup>+</sup> T cells or CD4<sup>+</sup> CD40L<sup>+</sup> Th1 cells. Two-sided t test.

(G) IFN $\gamma$  production by CD4<sup>+</sup> CD40L<sup>+</sup> T cells compared with CD4<sup>+</sup> CD40L<sup>-</sup> T cells. Two-sided t test.

(H) Expression of the proliferation marker Ki67 by tumor-infiltrating CD4<sup>+</sup> CD40L<sup>+</sup> Th1 cells compared with CD4<sup>+</sup> CD40L<sup>-</sup> T cells. Two-sided t test.

(I) Expression of *Bhlhe40* by OTII CD4<sup>+</sup> T cells stimulated with ova peptide-pulsed splenic DCs in the presence (open circle) or absence (closed circle) of anti-CD40 agonist antibody. Two-sided t test. \*, p < 0.05; \*\*, p < 0.01.

(J) Correlation between gene signatures for Th1-like cells and *BATF3*<sup>+</sup> cDC1 (left), activated cDC1 (middle) and cDC2 (right) on the TCGA COAD and READ datasets. Pearson correlation and linear regression.

See also Figure S7 and Table S6.

- Single cell sorting, library preparation, and sequencing for CRC samples (10 $\times$  Genomics)
- Tumor dissociation for mouse samples
- Single cell sorting, library preparation, and sequencing for mouse samples
- Multi-color immunofluorescence imaging of mouse tissues
- Intracellular cytokine staining
- Bulk DNA isolation and sequencing
- Microsatellite instability testing
- Multi-color immunohistochemistry of human tissues
- Splenic T cell cultures
- Human monocyte culture

## ● QUANTIFICATION AND STATISTICAL ANALYSIS

- Single-cell RNA-Seq data processing
- Evaluation of batch effect with entropy metric
- Batch effect correction and normalization
- Cell doublet detection and removal
- Bulk exome-seq data processing
- TCR analysis
- Unsupervised clustering analysis
- Dimensionality reduction using t-SNE and UMAP
- Identification of signature genes
- Tissue distribution of clusters
- Similarity analysis of clusters from different datasets
- Identification of malignant cells with CNV estimation
- Developmental trajectory inference
- Gene set enrichment analysis
- Define M1/M2 and proliferation phenotypes
- Comparison of myeloid cell subsets with previous datasets
- Correlative cell-cell interactions inferred by combined scRNA-seq and TCGA datasets
- Identification of significant ligand-receptor pairs
- Survival analysis

## ● DATA AND CODE AVAILABILITY

## ● ADDITIONAL RESOURCES

## SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at <https://doi.org/10.1016/j.cell.2020.03.048>.

## ACKNOWLEDGMENTS

We thank Y.S. Li for sample preparation, F. Wang, X. Zhang for assistance with FACS, and B. Belmontes and J. DeVoss for development of mouse tumor models. We thank the Computing Platform of the CLS (Peking University). This project was supported by National Key Research and Development Program of China. (2016YFC0900100), National Natural Science Foundation of China (81573022, 31530036, 91742203, and 81672375) and Amgen Corporation (USA). L.Z was supported by the Postdoctoral Foundation of CLS.

## AUTHOR CONTRIBUTIONS

X.Y., Z.Z., W.O., J.G.E., and L.Z. designed experiments. L.Z., Q.F., Q.Z., and R.G. generated human scRNA-seq data. X.Y., K.M.S., S.A.Ob., L.W., A.K., D.B., D.L., CM. L. A.S.R. performed mouse scRNA-seq, flow cytometry and *in vitro* experiments. J.O., D.S., K.P. generated *in vivo* mouse models. L.Z., Z.Y.L., Q.F., Y.H., and X.R. performed human scRNA-seq analyses. Z.Y.L., L.Z., J.G.E. and X.Y. performed mouse scRNA-seq analyses. W.Z., J.Y.,

S.W. and Z.S. collected clinical samples. X.H coordinated patient sample collections. Z.Y.L., and T.W. constructed the website. L.Z., Z.Y.L., K.M.S., J.G.E., W.O., Z.Z. and X.Y. wrote the manuscript with input from all authors.

## DECLARATION OF INTERESTS

X.Y., J.G.E., W.O., C.M.L., D.L., D.B., D.S., J.O., A.K., L., L.W., S.A.Ob., K.M.S are employees of Amgen Inc. X.H. and Z.Z. are founders of Analytical Biosciences Limited.

Received: July 12, 2019

Revised: January 2, 2020

Accepted: March 20, 2020

Published: April 16, 2020

## REFERENCES

- Angerer, P., Haghverdi, L., Büttner, M., Theis, F.J., Marr, C., and Buettner, F. (2016). *destiny: diffusion maps for large-scale single-cell data in R*. *Bioinformatics* **32**, 1241–1243.
- Ardouin, L., Luche, H., Chelbi, R., Carpentier, S., Shawket, A., Montanana San-chis, F., Santa Maria, C., Grenot, P., Alexandre, Y., Grégoire, C., et al. (2016). Broad and Largely Concordant Molecular Changes Characterize Tolerogenic and Immunogenic Dendritic Cell Maturation in Thymus and Periphery. *Immunity* **45**, 305–318.
- Azizi, E., Carr, A.J., Plitas, G., Cornish, A.E., Konopacki, C., Prabhakaran, S., Nainys, J., Wu, K., Kiseliolas, V., Setty, M., et al. (2018). Single-Cell Map of Diverse Immune Phenotypes in the Breast Tumor Microenvironment. *Cell* **174**, 1293–1308.
- Binnewies, M., Roberts, E.W., Kersten, K., Chan, V., Fearon, D.F., Merad, M., Coussens, L.M., Gabrilovich, D.I., Ostrand-Rosenberg, S., Hedrick, C.C., et al. (2018). Understanding the tumor immune microenvironment (TIME) for effective therapy. *Nat. Med.* **24**, 541–550.
- Binnewies, M., Mujal, A.M., Pollack, J.L., Combes, A.J., Hardison, E.A., Barry, K.C., Tsui, J., Ruhland, M.K., Kersten, K., Abushawish, M.A., et al. (2019). Unleashing Type-2 Dendritic Cells to Drive Protective Antitumor CD4(+) T Cell Immunity. *Cell* **177**, 556–571.e516.
- Blatner, N.R., Bonertz, A., Beckhove, P., Cheon, E.C., Krantz, S.B., Strouch, M., Weitz, J., Koch, M., Halverson, A.L., Bentrem, D.J., and Khazaie, K. (2010). In colorectal cancer mast cells contribute to systemic regulatory T-cell dysfunction. *Proc. Natl. Acad. Sci. USA* **107**, 6430–6435.
- Böttcher, J.P., and Reis E Sousa, C. (2018). The Role of Type 1 Conventional Dendritic Cells in Cancer Immunity. *Trends Cancer* **4**, 784–792.
- Bray, N.L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34**, 525–527.
- Brown, C.C., Gudjonson, H., Pritykin, Y., Deep, D., Lavallee, V.P., Mendoza, A., Fromme, R., Mazutis, L., Ariyan, C., Leslie, C., et al. (2019). Transcriptional Basis of Mouse and Human Dendritic Cell Heterogeneity. *Cell* **179**, 846–863.e824.
- Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420.
- Butowski, N., Colman, H., De Groot, J.F., Omuro, A.M., Nayak, L., Wen, P.Y., Cloughesy, T.F., Marimuthu, A., Haidar, S., Perry, A., et al. (2016). Orally administered colony stimulating factor 1 receptor inhibitor PLX3397 in recurrent glioblastoma: an Ivy Foundation Early Phase Clinical Trials Consortium phase II study. *Neuro-oncol.* **18**, 557–564.
- Cassier, P.A., Italiano, A., Gomez-Roca, C.A., Le Tourneau, C., Touilmonde, M., Cannarile, M.A., Ries, C., Brillouet, A., Müller, C., Jegg, A.M., et al. (2015). CSF1R inhibition with emactuzumab in locally advanced diffuse-type tenosynovial giant cell tumours of the soft tissue: a dose-escalation and dose-expansion phase 1 study. *Lancet Oncol.* **16**, 949–956.
- Chakarov, S., Lim, H.Y., Tan, L., Lim, S.Y., See, P., Lum, J., Zhang, X.M., Foo, S., Nakamizo, S., Duan, K., et al. (2019). Two distinct interstitial macrophage

- populations coexist across tissues in specific subtissular niches. *Science* 363, eaau0964.
- Cheong, C., Matos, I., Choi, J.H., Dandamudi, D.B., Shrestha, E., Longhi, M.P., Jeffrey, K.L., Anthony, R.M., Kluger, C., Nchinda, G., et al. (2010). Microbial stimulation fully differentiates monocytes to DC-SIGN/CD209(+) dendritic cells for immune T cell areas. *Cell* 143, 416–429.
- Chevrier, S., Levine, J.H., Zanotelli, V.R.T., Silina, K., Schulz, D., Bacac, M., Ries, C.H., Ailles, L., Jewett, M.A.S., Moch, H., et al. (2017). An Immune Atlas of Clear Cell Renal Cell Carcinoma. *Cell* 169, 736–749.e718.
- Dalerba, P., Kalisky, T., Sahoo, D., Rajendran, P.S., Rothenberg, M.E., Leyrat, A.A., Sim, S., Okamoto, J., Johnston, D.M., Qian, D., et al. (2011). Single-cell dissection of transcriptional heterogeneity in human colon tumors. *Nat. Biotechnol.* 29, 1120–1127.
- DeNardo, D.G., and Ruffell, B. (2019). Macrophages as regulators of tumour immunity and immunotherapy. *Nat. Rev. Immunol.* 19, 369–382.
- DeNardo, D.G., Brennan, D.J., Rexhepaj, E., Ruffell, B., Shiao, S.L., Madden, S.F., Gallagher, W.M., Wadhwani, N., Keil, S.D., Junaid, S.A., et al. (2011). Leukocyte complexity predicts breast cancer survival and functionally regulates response to chemotherapy. *Cancer Discov.* 1, 54–67.
- Dutertre, C.A., Becht, E., Irac, S.E., Khalilnezhad, A., Narang, V., Khalilnezhad, S., Ng, P.Y., van den Hoogen, L.L., Leong, J.Y., Lee, B., et al. (2019). Single-Cell Analysis of Human Mononuclear Phagocytes Reveals Subset-Defining Markers and Identifies Circulating Inflammatory Dendritic Cells. *Immunity* 51, 573–589.e578.
- Farrell, J.A., Wang, Y., Riesenfeld, S.J., Shekhar, K., Regev, A., and Schier, A.F. (2018). Single-cell reconstruction of developmental trajectories during zebrafish embryogenesis. *Science* 360, eaar3131.
- Galon, J., and Bruni, D. (2019). Approaches to treat immune hot, altered and cold tumours with combination immunotherapies. *Nat. Rev. Drug Discov.* 18, 197–218.
- Ganesh, K., Stadler, Z.K., Cerck, A., Mendelsohn, R.B., Shia, J., Segal, N.H., and Diaz, L.A., Jr. (2019). Immunotherapy in colorectal cancer: rationale, challenges and potential. *Nat. Rev. Gastroenterol. Hepatol.* 16, 361–375.
- Grewal, I.S., and Flavell, R.A. (1998). CD40 and CD154 in cell-mediated immunity. *Annu. Rev. Immunol.* 16, 111–135.
- Gubin, M.M., Esaulova, E., Ward, J.P., Malkova, O.N., Runci, D., Wong, P., Noguchi, T., Arthur, C.D., Meng, W., Alspach, E., et al. (2018). High-Dimensional Analysis Delineates Myeloid and Lymphoid Compartment Remodeling during Successful Immune-Checkpoint Cancer Therapy. *Cell* 175, 1443.
- Guo, X., Zhang, Y., Zheng, L., Zheng, C., Song, J., Zhang, Q., Kang, B., Liu, Z., Jin, L., Xing, R., et al. (2018). Global characterization of T cells in non-small-cell lung cancer by single-cell sequencing. *Nat. Med.* 24, 978–985.
- Gyori, D., Lim, E.L., Grant, F.M., Spensberger, D., Roychoudhuri, R., Shuttleworth, S.J., Okkenhaug, K., Stephens, L.R., and Hawkins, P.T. (2018). Compensation between CSF1R+ macrophages and Foxp3+ Treg cells drives resistance to tumor immunotherapy. *JCI Insight* 3, 120631.
- Haniffa, M., Collin, M., and Ginhoux, F. (2013). Ontogeny and functional specialization of dendritic cells in human and mouse. *Adv. Immunol.* 120, 1–49.
- Hildner, K., Edelson, B.T., Purtha, W.E., Diamond, M., Matsushita, H., Kohyama, M., Calderon, B., Schraml, B.U., Unanue, E.R., Diamond, M.S., et al. (2008). Batf3 deficiency reveals a critical role for CD8alpha+ dendritic cells in cytotoxic T cell immunity. *Science* 322, 1097–1100.
- Hume, D.A., Irvine, K.M., and Pridans, C. (2019). The Mononuclear Phagocyte System: The Relationship between Monocytes and Macrophages. *Trends Immunol.* 40, 98–112.
- Jerby-Arnon, L., Shah, P., Cuoco, M.S., Rodman, C., Su, M.J., Melms, J.C., Leeson, R., Kanodia, A., Mei, S., Lin, J.R., et al. (2018). A Cancer Cell Program Promotes T Cell Exclusion and Resistance to Checkpoint Blockade. *Cell* 175, 984–997.e924.
- Junttila, M.R., and de Sauvage, F.J. (2013). Influence of tumour micro-environment heterogeneity on therapeutic response. *Nature* 501, 346–354.
- Kenny, H.A., Kaur, S., Coussens, L.M., and Lengyel, E. (2008). The initial steps of ovarian cancer cell metastasis are mediated by MMP-2 cleavage of vitronectin and fibronectin. *J. Clin. Invest.* 118, 1367–1379.
- Klose, C.S., and Artis, D. (2016). Innate lymphoid cells as regulators of immunity, inflammation and tissue homeostasis. *Nat. Immunol.* 17, 765–774.
- Krystel-Whittemore, M., Dileepan, K.N., and Wood, J.G. (2016). Mast Cell: A Multi-Functional Master Cell. *Front. Immunol.* 6, 620.
- Kumar, V., Dontireddy, L., Marvel, D., Condamine, T., Wang, F., Lavilla-Alonso, S., Hashimoto, A., Vonteddu, P., Behera, R., Goins, M.A., et al. (2017). Cancer-Associated Fibroblasts Neutralize the Anti-tumor Effect of CSF1 Receptor Blockade by Inducing PMN-MDSC Infiltration of Tumors. *Cancer Cell* 32, 654–668.e655.
- La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastriti, M.E., Lönnérberg, P., Furlan, A., et al. (2018). RNA velocity of single cells. *Nature* 560, 494–498.
- Lambrechts, D., Wauters, E., Boeckx, B., Aibar, S., Nittner, D., Burton, O., Bassez, A., Decaluwe, H., Pircher, A., Van den Eynde, K., et al. (2018). Phenotype molding of stromal cells in the lung tumor microenvironment. *Nat. Med.* 24, 1277–1289.
- Lavin, Y., Kobayashi, S., Leader, A., Amir, E.D., Elefant, N., Bigenwald, C., Remark, R., Sweeney, R., Becker, C.D., Levine, J.H., et al. (2017). Innate Immune Landscape in Early Lung Adenocarcinoma by Paired Single-Cell Analyses. *Cell* 169, 750–765.e717.
- Lee, D.T., Uram, J.N., Wang, H., Bartlett, B.R., Kemberling, H., Eyring, A.D., Skora, A.D., Luber, B.S., Azad, N.S., Laheru, D., et al. (2015). PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *N. Engl. J. Med.* 372, 2509–2520.
- Lee, J.H., Chen, T.W., Hsu, C.H., Yen, Y.H., Yang, J.C., Cheng, A.L., Sasaki, S.I., Chiu, L.L., Sugihara, M., Ishizuka, T., et al. (2020). A phase I study of pexidartinib, a colony-stimulating factor 1 receptor inhibitor, in Asian patients with advanced solid tumors. *Invest. New Drugs* 38, 99–110.
- Li, H., Courtois, E.T., Sengupta, D., Tan, Y., Chen, K.H., Goh, J.J.L., Kong, S.L., Chua, C., Hon, L.K., Tan, W.S., et al. (2017). Reference component analysis of single-cell transcriptomes elucidates cellular heterogeneity in human colorectal tumors. *Nat. Genet.* 49, 708–718.
- Li, H., van der Leun, A.M., Yofe, I., Lubling, Y., Gelbard-Solodkin, D., van Akkooi, A.C.J., van den Braber, M., Rozeman, E.A., Haanen, J., Blank, C.U., et al. (2019). Dysfunctional CD8 T Cells Form a Proliferative, Dynamically Regulated Compartment within Human Melanoma. *Cell* 176, 775–789.e718.
- Lin, C.C., Bradstreet, T.R., Schwarzkopf, E.A., Jarjour, N.N., Chou, C., Archambault, A.S., Sim, J., Zinselmeyer, B.H., Carrero, J.A., Wu, G.F., et al. (2016). IL-1-induced Blhle40 identifies pathogenic T helper cells in a model of autoimmune neuroinflammation. *J. Exp. Med.* 213, 251–271.
- MacDonald, K.P., Palmer, J.S., Cronau, S., Seppanen, E., Olver, S., Raffelt, N.C., Kuns, R., Pettit, A.R., Clouston, A., Wainwright, B., et al. (2010). An antibody against the colony-stimulating factor 1 receptor depletes the resident subset of monocytes and tissue- and tumor-associated macrophages but does not inhibit inflammation. *Blood* 116, 3955–3963.
- Mahoney, K.M., Rennert, P.D., and Freeman, G.J. (2015). Combination cancer immunotherapy and new immunomodulatory targets. *Nat. Rev. Drug Discov.* 14, 561–584.
- Mantovani, A., Cassatella, M.A., Costantini, C., and Jaillon, S. (2011). Neutrophils in the activation and regulation of innate and adaptive immunity. *Nat. Rev. Immunol.* 11, 519–531.
- Merad, M., Sathe, P., Helft, J., Miller, J., and Mortha, A. (2013). The dendritic cell lineage: ontogeny and function of dendritic cells and their subsets in the steady state and the inflamed setting. *Annu. Rev. Immunol.* 31, 563–604.
- Müller, S., Kohanbash, G., Liu, S.J., Alvarado, B., Carrera, D., Bhaduri, A., Watchmaker, P.B., Yagnik, G., Di Lullo, E., Malatesta, M., et al. (2017). Single-cell profiling of human gliomas reveals macrophage ontogeny as a basis for regional differences in macrophage activation in the tumor microenvironment. *Genome Biol.* 18, 234.
- Neubert, N.J., Schmittnaegel, M., Bordry, N., Nassiri, S., Wald, N., Martignier, C., Tille, L., Homicsko, K., Damsky, W., Maby-El Hajjami, H., et al. (2018). T

- cell-induced CSF1 promotes melanoma resistance to PD1 blockade. *Sci Transl Med* 10, eaan3311.
- Papadopoulos, K.P., Gluck, L., Martin, L.P., Olszanski, A.J., Tolcher, A.W., Ngarmchamnanirth, G., Rasmussen, E., Amore, B.M., Nagoren, D., Hill, J.S., and Stephenson, J., Jr. (2017). First-in-Human Study of AMG 820, a Monoclonal Anti-Colony-Stimulating Factor 1 Receptor Antibody, in Patients with Advanced Solid Tumors. *Clin. Cancer Res.* 23, 5703–5710.
- Papalexis, E., and Satija, R. (2018). Single-cell RNA sequencing to explore immune cell heterogeneity. *Nat. Rev. Immunol.* 18, 35–45.
- Patel, A.P., Tirosh, I., Trombetta, J.J., Shalek, A.K., Gillespie, S.M., Wakimoto, H., Cahill, D.P., Nahed, B.V., Curry, W.T., Martuza, R.L., et al. (2014). Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* 344, 1396–1401.
- Pradel, L.P., Ooi, C.H., Romagnoli, S., Cannarile, M.A., Sade, H., Rüttinger, D., and Ries, C.H. (2016). Macrophage Susceptibility to Emactuzumab (RG7155) Treatment. *Mol. Cancer Ther.* 15, 3077–3086.
- Puram, S.V., Tirosh, I., Parikh, A.S., Patel, A.P., Yizhak, K., Gillespie, S., Rodman, C., Luo, C.L., Mroz, E.A., Emerick, K.S., et al. (2017). Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer. *Cell* 171, 1611–1624.e1624.
- Ramilowski, J.A., Goldberg, T., Harshbarger, J., Kloppmann, E., Lizio, M., Sa-tagopam, V.P., Itoh, M., Kawaji, H., Carninci, P., Rost, B., and Forrest, A.R. (2015). A draft network of ligand-receptor-mediated multicellular signalling in human. *Nat. Commun.* 6, 7866.
- Ribas, A., and Wolchok, J.D. (2018). Cancer immunotherapy using checkpoint blockade. *Science* 359, 1350–1355.
- Ries, C.H., Cannarile, M.A., Hoves, S., Benz, J., Wartha, K., Runza, V., Rey-Giraud, F., Pradel, L.P., Feuerhake, F., Klaman, I., et al. (2014). Targeting tumor-associated macrophages with anti-CSF-1R antibody reveals a strategy for cancer therapy. *Cancer Cell* 25, 846–859.
- Sasaki, N., Sachs, N., Wiebrands, K., Ellenbroek, S.I., Fumagalli, A., Lyubimova, A., Begthel, H., van den Born, M., van Es, J.H., Karthaus, W.R., et al. (2016). Reg4+ deep crypt secretory cells function as epithelial niche for Lgr5+ stem cells in colon. *Proc. Natl. Acad. Sci. USA* 113, E5399–E5407.
- Savas, P., Virassamy, B., Ye, C., Salim, A., Mintoff, C.P., Caramia, F., Salgado, R., Byrne, D.J., Teo, Z.L., Dushyanthen, S., et al.; Kathleen Cunningham Foundation Consortium for Research into Familial Breast Cancer (kConFab) (2018). Single-cell profiling of breast cancer T cells reveals a tissue-resident memory subset associated with improved prognosis. *Nat. Med.* 24, 986–993.
- Schelker, M., Feau, S., Du, J., Ranu, N., Klipp, E., MacBeath, G., Schoeberl, B., and Raue, A. (2017). Estimation of immune cell content in tumour tissue using single-cell RNA-seq data. *Nat. Commun.* 8, 2032.
- Smillie, C.S., Biton, M., Ordovas-Montanes, J., Sullivan, K.M., Burgin, G., Graham, D.B., Herbst, R.H., Rogel, N., Slyper, M., Waldman, J., et al. (2019). Intra- and Inter-cellular Rewiring of the Human Colon during Ulcerative Colitis. *Cell* 178, 714–730.e722.
- Tait Wojno, E.D., Hunter, C.A., and Stumhofer, J.S. (2019). The Immunobiology of the Interleukin-12 Family: Room for Discovery. *Immunity* 50, 851–870.
- Tirosh, I., Izar, B., Prakadan, S.M., Wadsworth, M.H., 2nd, Treacy, D., Trombetta, J.J., 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.
- Venteicher, A.S., Tirosh, I., Hebert, C., Yizhak, K., Neftel, C., Filbin, M.G., Hovestadt, V., Escalante, L.E., Shaw, M.L., Rodman, C., et al. (2017). Decoupling genetics, lineages, and microenvironment in IDH-mutant gliomas by single-cell RNA-seq. *Science* 355, eaai8478.
- Vento-Tormo, R., Efremova, M., Botting, R.A., Turco, M.Y., Vento-Tormo, M., Meyer, K.B., Park, J.E., Stephenson, E., Polański, K., Goncalves, A., et al. (2018). Single-cell reconstruction of the early maternal-fetal interface in humans. *Nature* 563, 347–353.
- Vonderheide, R.H. (2020). CD40 Agonist Antibodies in Cancer Immunotherapy. *Annu. Rev. Med.* 71, 47–58.
- Whitfield, M.L., George, L.K., Grant, G.D., and Perou, C.M. (2006). Common markers of proliferation. *Nat. Rev. Cancer* 6, 99–106.
- Wolf, F.A., Angerer, P., and Theis, F.J. (2018). SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* 19, 15.
- Wolf, F.A., Hamey, F.K., Plass, M., Solana, J., Dahlin, J.S., Göttgens, B., Rajewsky, N., Simon, L., and Theis, F.J. (2019). PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. *Genome Biol.* 20, 59.
- Wolock, S.L., Lopez, R., and Klein, A.M. (2019). Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. *Cell Syst* 8, 281–291.e289.
- Wu, Y.E., Pan, L., Zuo, Y., Li, X., and Hong, W. (2017). Detecting Activated Cell Populations Using Single-Cell RNA-Seq. *Neuron* 96, 313–329.e316.
- Yarchoan, M., Hopkins, A., and Jaffee, E.M. (2017). Tumor Mutational Burden and Response Rate to PD-1 Inhibition. *N. Engl. J. Med.* 377, 2500–2501.
- Young, M.D., Mitchell, T.J., Vieira Braga, F.A., Tran, M.G.B., Stewart, B.J., Ferdinand, J.R., Collord, G., Botting, R.A., Popescu, D.M., Loudon, K.W., et al. (2018). Single-cell transcriptomes from human kidneys reveal the cellular identity of renal tumors. *Science* 361, 594–599.
- Zaki, M.H., Lamkanfi, M., and Kanneganti, T.D. (2011). The Nlrp3 inflammasome: contributions to intestinal homeostasis. *Trends Immunol.* 32, 171–179.
- Zhang, L., and Zhang, Z. (2019). Recharacterizing Tumor-Infiltrating Lymphocytes by Single-Cell RNA Sequencing. *Cancer Immunol. Res.* 7, 1040–1046.
- Zhang, L., Yu, X., Zheng, L., Zhang, Y., Li, Y., Fang, Q., Gao, R., Kang, B., Zhang, Q., Huang, J.Y., et al. (2018). Lineage tracking reveals dynamic relationships of T cells in colorectal cancer. *Nature* 564, 268–272.
- Zhang, Q., He, Y., Luo, N., Patel, S.J., Han, Y., Gao, R., Modak, M., Carotta, S., Haslinger, C., Kind, D., et al. (2019). Landscape and Dynamics of Single Immune Cells in Hepatocellular Carcinoma. *Cell* 179, 829–845.e820.
- Zheng, C., Zheng, L., Yoo, J.K., Guo, H., Zhang, Y., Guo, X., Kang, B., Hu, R., Huang, J.Y., Zhang, Q., et al. (2017). Landscape of Infiltrating T Cells in Liver Cancer Revealed by Single-Cell Sequencing. *Cell* 169, 1342–1356.e1316.
- Zhu, Y., Herndon, J.M., Sojka, D.K., Kim, K.W., Knolhoff, B.L., Zuo, C., Cullinan, D.R., Luo, J., Bearden, A.R., Levine, K.J., et al. (2017). Tissue-Resident Macrophages in Pancreatic Ductal Adenocarcinoma Originate from Embryonic Hematopoiesis and Promote Tumor Progression. *Immunity* 47, 323–338.e326.
- Zilionis, R., Engblom, C., Pfirschke, C., Savova, V., Zemmour, D., Saatcioglu, H.D., Krishnan, I., Maroni, G., Meyerovitz, C.V., Kerwin, C.M., et al. (2019). Single-Cell Transcriptomics of Human and Mouse Lung Cancers Reveals Conserved Myeloid Populations across Individuals and Species. *Immunity* 50, 1317–1334.e1310.
- Zucker, S., and Vacirca, J. (2004). Role of matrix metalloproteinases (MMPs) in colorectal cancer. *Cancer Metastasis Rev.* 23, 101–117.

## STAR★METHODS

## KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Antibodies</b>		
Anti-Human CD45 FITC (clone HI30)	Thermo Fisher	Cat# 11-0459-42
Anti-Human CD3 eFluor 450 (clone UCHT1)	Thermo Fisher	Cat# 48-0038-42
Anti-Human CD19 APC (clone HIB19)	Thermo Fisher	Cat# 17-0199-42
Anti-Human CD56 PE (clone TULY56)	Thermo Fisher	Cat# 12-0566-42
Anti-Human CD79A PE/Cy7 (clone HM47)	Thermo Fisher	Cat# 333509
Anti-Human CD14 APC/eFluor 780 (clone 61D3)	Thermo Fisher	Cat# 47-0149-41
Anti-mouse CD45 BV786 (clone 30-F11)	BD Biosciences	Cat# 564225
Anti-mouse CD45 FITC (clone 30-F11)	BD Biosciences	Cat# 553079
Anti-mouse TCR $\beta$ BV421 (clone H57-597)	BD Biosciences	Cat# 562839
Anti-mouse TCR $\beta$ FITC (clone H57-597)	BD Biosciences	Cat# 553171
Anti-mouse Thy1.2 APC-Cy7 (clone 53-2.1)	Thermo Fisher	Cat# 47-0902-82
Anti-mouse CD4 BUV395 (clone GK1.5)	BD Biosciences	Cat# 563790
Anti-mouse CD4 Pacific Blue (clone GK1.5)	BioLegend	Cat# 100428
Anti-mouse CD8 BV650 (clone 53-6.7)	BD Biosciences	Cat# 563234
Anti-mouse CD3 Pacific Blue (clone 17A2)	Biolegend	Cat# 100213
Anti-mouse CD11c APC (clone HL3)	BD Biosciences	Cat# 550261
Anti-mouse CD11c BV786 (clone HL3)	BD Biosciences	Cat# 563735
Anti-mouse B220 PE-Cy7 (clone RA3-6B2)	BD Biosciences	Cat# 552772
Anti-mouse B220 PerCP-Cy5.5 (clone RA3-62B)	BioLegend	Cat#103236
Anti-mouse Ly6G and Ly6C PE-Cy7 (clone RB6-8C5)	BD Biosciences	Cat# 565033
Anti-mouse Ly6C BV711 (clone HK1.4; 1:1000 dilution)	BioLegend	Cat#128037
Anti-mouse Ly6G APC-Cy7 (clone 1A1; 1:100 dilution)	BioLegend	Cat#127624
Anti-mouse CD11b BUV395 (clone M1/70)	BD Biosciences	Cat# 563553
Anti-mouse CD11b BUV737 (clone M1/70)	BD Biosciences	Cat#564443
Anti-mouse CD86 PE (clone GL1)	BD Biosciences	Cat# 553692
Anti-mouse CD86 APC (clone GL1)	BD Biosciences	Cat# 558703
Anti-mouse CD80 BV421 (clone 16-10A1)	Biolegend	Cat# 104725
Anti-mouse Sirp $\alpha$ BV605 (clone P84)	BD Biosciences	Cat# 740390
Anti-mouse XCR1 BV650 (clone ZET)	Biolegend	Cat# 148220
Anti-mouse CD103 FITC (clone M290)	BD Biosciences	Cat# 557494
Anti-mouse IL12 PE (clone C17.8)	Thermo Fisher	Cat# 12-7123-82
Anti-mouse F4/80 PE CY7 (clone BM8)	Biolegend	Cat# 123114
Anti-mouse CD11c APC-Cy7 (N418)	Biolegend	Cat# 117324
Anti-mouse Ki67 PerCP-eFluor710 (clone SolA15)	Thermo Fisher	Cat# 46-5698-82
Anti-mouse IFN $\gamma$ PE (clone XMG1.2)	Thermo Fisher	Cat# 12-7311-82
Anti-mouse TNF $\alpha$ FITC (clone MP6-XT22)	BD Biosciences	Cat# 554418
Anti-mouse PD1 BV605 (clone J43)	BD Biosciences	Cat# 563059
Anti-mouse CD24 BV711 (clone M1/69)	BD Biosciences	Cat# 563450
Anti-mouse MHCII Alexa Fluor700 (clone M5/114.15.2)	Thermo Fisher	Cat# 56-5321-82
Anti-mouse MHCII BV510 (clone M5/114.15.2, 1:1000 dilution)	BioLegend	Cat# 107635
Anti-mouse CD40L APC (clone MR1)	Thermo Fisher	Cat# 17-1541-82
Anti-mouse CD40 mlgG1 (clone FGK45)	Amgen	N/A

(Continued on next page)

***Continued***

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Mouse IgG1 (clone MOPC-21)	BioXCell	Cat# BE0083
Anti-mouse CD40 rat IgG2a (clone FGK45)	BioXCell	Cat# BE0016-2
Rat IgG2a (clone 2A3)	BioXCell	Cat# BE0089
Anti-mouse CD40 mouse IgG1 (clone FGK45)	Amgen	N/A
Anti-mouse PD1 rat IgG2a (clone 29F1A.12)	BioXCell	Cat# BE0273
Anti-mouse CSF1R mouse IgG1 (clone M279)	Amgen	N/A
Anti-Human CD68 (clone KP1; 1:100 dilution)	Abcam	Cat# ab955
Anti-Human CD80 (clone EPR1157(2); 1:200 dilution)	Abcam	Cat# ab134120
Anti-Human c-Maf (clone EPR16484; 1:200 dilution)	Abcam	Cat# ab19942
Anti-Human MARCO (1:100 dilution)	Abcam	Cat# ab231046
Anti-Human MARCO (clone PLK-1)	Invitrogen	Cat# 12-5447-42
Anti-Human VEGFA (clone VG-1; 1:200 dilution)	Abcam	Cat# ab1316
Anti-Mouse CD31 Alexa Fluor 594 (clone MEC13.1; 1:100 dilution)	BioLegend	Cat# 102520
Anti-Mouse F4/80 Alexa Fluor 647 (clone BM8; 1:100 dilution)	BioLegend	Cat# 123122
Anti-Mouse F4/80 BV650 (clone BM8; 1:100 dilution)	BioLegend	Cat# 123149
Anti-Mouse CD68 PE-Cy7 (clone FA-11; 1:100 dilution)	BioLegend	Cat# 137016
Anti-Mouse MHC II Alexa Fluor 488 (clone M5/114.15.2; 1:50 dilution)	BioLegend	Cat# 107616
<b>Biological Samples</b>		
Human primary CRC samples	Peking University People's Hospital	See Table S1 for details
Human peripheral blood from CRC patients	Peking University People's Hospital	See Table S1 for details
Human adjacent normal tissues from CRC patients	Peking University People's Hospital	See Table S1 for details
Human peripheral blood from normal donors	Amgen	N/A
Mouse tumors from Renca tumor-bearing mice	Amgen	N/A
Mouse lymph nodes and tumors from MC38 tumor-bearing mice	Amgen	N/A
<b>Chemicals, Peptides, and Recombinant Proteins</b>		
RPMI-1640	Invitrogen	Cat# 11875-093
Fetal bovine serum (FBS)	Thermo Fisher	Cat# 10099141
HISTOPAQUE-1077	Sigma-Aldrich	Cat# 10771
Ova (323-339)	Anaspec	Cat# AS-27024
dNTP mix (10 mM)	Thermo Fisher	Cat# R0193
Triton X-100	Sigma-Aldrich	Cat# T9284
RNase Inhibitor (40 U/μl)	Takara	Cat# 2313A
External RNA Controls Consortium	Thermo Fisher	Cat# 4456740
Agencourt XP DNA beads	Beckman	Cat# A63882
Actinomycin D	Sigma	Cat# A1410
DNase I	Roche	Cat# 04716728001
Collagenase D	Roche	Cat# 11088858001
Brefeldin A (BFA)	Biolegend	Cat# 420601
Brilliant Stain buffer	BD Biosciences	Cat# 566349
Cell Stimulation cocktail (plus protein transport inhibitors)	Thermo Fisher	Cat# 00-4975-93
Liberase TL	Roche	Cat# 05401020001
LIVE/DEAD Fixable Blue Dead Cell Stain	Molecular Probes	Cat# L34962
DAPI (1:2000 dilution)	Molecular Probes	Cat# D1306
7-AAD Viability Staining Solution	Thermo Fisher	Cat# 00-6993-50

(Continued on next page)

**Continued**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Fixable Viability Dye eFluor506	Thermo Fisher	Cat# 65-0866-14
X-VIVO 10 Serum-free medium	Lonza	Cat# 04743Q
Human recombinant M-CSF	PeproTech	Cat# 300-25
Human recombinant GM-CSF	PeproTech	Cat# 300-03
Human recombinant VEGFA	PeproTech	Cat# 100-20
Human recombinant VEGFB	PeproTech	Cat# 100-20B
Human recombinant VEGFC	PeproTech	Cat# 100-30CD
Human recombinant PIGF	PeproTech	Cat# 100-56A
Human recombinant SPP1	PeproTech	Cat# 120-35
Human recombinant TNF $\alpha$	PeproTech	Cat# 300-01A
Human recombinant IL-1 $\beta$	PeproTech	Cat# 200-1B
Human recombinant IFNG	PeproTech	Cat# 300-02
Human recombinant MIF	PeproTech	Cat# 300-69
Human recombinant RETN	PeproTech	Cat# 450-19
<b>Critical Commercial Assays</b>		
MACS Tumor Dissociation Kit	Miltenyi Biotec	Cat# 130-095-929
MACS CD4+ T cell isolation kit, mouse	Miltenyi Biotec	Cat# 130-104-454
Classical Monocyte Isolation Kit, human	Miltenyi	Cat# 130-117-337
Qubit HsDNA kits	Thermo Fisher	Cat# Q32854
TruePrep DNA Library Prep Kit V2 for Illumina	Vazyme Biotech	Cat# TD503
Chromium Single Cell 3' Library and Bead Kit	10x Genomics	Cat# PN-120237
Chromium Single Cell 5' Library and Bead Kit	10x Genomics	Cat# 1000006
Chromium Single Cell 5' Library Construction Kit	10x Genomics	Cat# 1000020
Mouse IFN $\gamma$ ELISA kit II	BD Biosciences	Cat# 558258
Foxp3/Transcription factor staining buffer set	Thermo Fisher	Cat# 00-5523-00
NEBNext Ultra RNA Library Prep Kit for Illumina Paired-end	NEB	Cat# E7530
Multiplexed Sequencing Library		
SureSelectXT Target Enrichment System for Illumina Paired-End Multiplexed Sequencing Library kit	Agilent	Cat# G9701
Hiseq 3000/4000 SBS kit	Illumina	Cat# FC-410-1003
Hiseq 3000/4000 PE cluster Kit	Illumina	Cat# PE-410-1001
Bhlhe40 Taqman Gene Expression Assay	Thermo Fisher	Cat# Mm00478593_m1
Rpl19 Taqman Gene Expression Assay	Thermo Fisher	Cat# Mm02601633_g1
Tbx21 Taqman Gene Expression Assay	Thermo Fisher	Cat# Mm00450960_m1
TaqMan 2 $\times$ Universal PCR master mix	Applied Biosystems	Cat# 4304437
RNeasy Mini Kit	QIAGEN	Cat# 74104
SuperScript IV First-Strand Synthesis System	Invitrogen	Cat# 18091050
Opal 7-Color Manual IHC Kit	PerkinElmer	Cat# NEL811001KT
<b>Deposited Data</b>		
Data files for human scRNA-seq	This paper	GEO: GSE146771
Data files for mouse scRNA-seq	This paper	ENA: PRJEB34105 ArrayExpress: E-MTAB-8832
Human UC scRNA-seq dataset	Smillie et al., 2019	SCP: SCP259
Human NSCLC scRNA-seq dataset	Zilionis et al., 2019	GEO: GSE127465
Human HCC scRNA-seq dataset	Zhang et al., 2019	EGA: S00001003449
FANTOM5 receptor-ligand database	Ramiliowski et al., 2015	<a href="https://fantom.gsc.riken.jp/5">https://fantom.gsc.riken.jp/5</a>
CellPhoneDB receptor-ligand database	Vento-Tormo et al., 2018	<a href="https://www.cellphonedb.org/">https://www.cellphonedb.org/</a>

(Continued on next page)

**Continued**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental Models: Cell Lines		
MC38	NCI/NIH	ENH204-EP
Renca	ATCC	CRL-2947
Experimental Models: Organisms/Strains		
C57BL/6 inbred mice	Charles River Laboratories	C57BL/6NCrl
BALB/c inbred mice	Charles River Laboratories	BALB/cAnNCrl
OTII transgenic mice	The Jackson Laboratory	B6.Cg-Tg (TcrαTcrβ) 425Cbn/J
Oligonucleotides		
Primer: GAPDH Forward: TTGGCTACAGCAACAGGGTG	This paper	N/A
Primer: GAPDH Reverse: TCTACATGGCAACTGTGAGGAG	This paper	N/A
Software and Algorithms		
FlowJo 10.3	BD Biosciences	<a href="https://www.flowjo.com/">https://www.flowjo.com/</a>
InForm Advanced Image Analysis Software 2.3	PerkinElmer	<a href="https://www.perkinelmer.com/">https://www.perkinelmer.com/</a>
kallisto 0.43.1	Bray et al., 2016	<a href="https://pachterlab.github.io/kallisto">https://pachterlab.github.io/kallisto</a>
Cell Ranger 2.1.0	10x Genomics	<a href="http://10xgenomics.com/">http://10xgenomics.com/</a>
STARTRAC	Zhang et al., 2018	<a href="https://github.com/Japrin/STARTRAC">https://github.com/Japrin/STARTRAC</a>
Seurat 2.3.0/3.0	Butler et al., 2018	<a href="https://satijalab.org/seurat">https://satijalab.org/seurat</a>
Scanpy 1.4.4	Wolf et al., 2018	<a href="https://scanpy.readthedocs.io/en/latest/">https://scanpy.readthedocs.io/en/latest/</a>
velocyto 0.17	La Manno et al., 2018	<a href="http://velocyto.org/velocyto.py/">http://velocyto.org/velocyto.py/</a>
destiny 2.11.3	Angerer et al., 2016	<a href="https://github.com/theislab/destiny">https://github.com/theislab/destiny</a>
URD 1.0.1	Farrell et al., 2018	<a href="https://github.com/farrellja/URD">https://github.com/farrellja/URD</a>
Other		
Interactive explorer of human and mouse total cells	This paper	<a href="http://crcleukocyte.cancer-pku.cn/">http://crcleukocyte.cancer-pku.cn/</a>

**LEAD CONTACT AND MATERIALS AVAILABILITY**

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Z.Z. ([zemin@pku.edu.cn](mailto:zemin@pku.edu.cn)) or to X.Y. ([xiyu@amgen.com](mailto:xiyu@amgen.com)). This study did not generate new unique reagents.

**EXPERIMENTAL MODEL AND SUBJECT DETAILS****Human specimens**

Eighteen colorectal cancer patients, including fourteen females and four males, were enrolled and pathologically diagnosed with colorectal adenocarcinoma at Peking University People's Hospital. Written informed consent was provided by all patients. This study was approved by the Research and Ethical Committee of Peking University People's Hospital and complied with all relevant ethical regulations. Fresh tumor and adjacent normal tissue samples (at least 2 cm from matched tumor tissues) were surgically resected from the above-described patients. All patients, except P0309, P0411, P0720, P1025, and P1212, had peripheral blood and paired tumor and adjacent normal tissues obtained. Patients P0309, P0411, and P1212 had fresh tumor tissues and matched peripheral blood, while patients P0720 and P1025 had fresh tumor tissues and matched adjacent normal tissues. Their ages ranged from 40 to 89 with a median of 68.5. None of them were treated with chemotherapy or radiation prior to tumor resection. The stages of these patients were classified according to the guidance of AJCC version 8. Among these patients, one was diagnosed at stage I, eleven at stage II, and six at stage III. While we did not purposely exclude stage IV patients, none of these patients had distal metastasis, as evidence by the enhanced CT results for abdomen, chest, and pelvic areas before surgery. The available clinical characteristics are summarized in [Table S1](#).

***In vivo* mouse models**

All animal experiments were performed under protocols approved by the Institutional Animal Care and Use Committee (IACUC) of Amgen, Inc. All mice were female and sourced from Charles River Laboratories (Hollister, CA site) or The Jackson Laboratory (Sacramento, CA site) and were provided water and chow *ad libitum* and maintained in a pathogen-free facility. Mice used in syngeneic

neic tumor experiments were 6-8 weeks of age at the time implant. MC38 or Renca cells in serum-free RPMI were inoculated in the right flank at  $3 \times 10^5$  cells per implant and allowed to grow for 8-10 days (MC38) or 12 days (Renca). The mice were then randomized by tumor volume (50-100mm<sup>3</sup>) and treated with the described antibodies.

For MC38 studies, isotype control treatment antibodies were rat IgG2a from BioXCell (2A3) and Amgen in-house generated mouse IgG1. Anti-CD40 antibodies were clone FGK45 (rat IgG2a) from BioXCell and in-house generated FGK45 (mouse IgG1). Anti-PD1 antibody was the 29F1A.12 clone (rat IgG2a) from BioXCell. All anti-CD40 treatments and corresponding isotype controls were diluted in PBS and administered intraperitoneally (i.p.) at 5mg/kg. Anti-PD1 treatment and the corresponding isotype control was diluted in PBS and administered i.p. at 15mg/kg. For efficacy studies, tumors were measured twice a week with digital calipers and tracked using StudyLog until the end of study (twice the median survival of the isotype treated groups). For pharmacodynamic studies, tumors were taken down 24-48 h and 6 days post-initiation of anti-CD40 treatments for tissue collection ([Figure S6A](#))

For Renca studies, isotype control treatment antibodies were mouse IgG1 from BioXCell (MOPC-21). Rat Anti-mouse CSF1R M279 was generated at Amgen ([MacDonald et al., 2010](#)) and converted to a mouse IgG1 antibody backbone. All anti-CSF1R treatments and corresponding isotype controls were diluted in PBS and administered i.p. at 400 µg/mouse three times a week for a total of four doses. For efficacy and pharmacodynamic studies, tumors were measured as above and analyzed at 24 h after the last dose on day 8 post-initiation of anti-CSF1R treatment ([Figure S6A](#)).

## METHOD DETAILS

### Single cell collection for human samples

Single cells were collected from tumor and adjacent normal tissues as described previously ([Zhang et al., 2018](#)). Briefly, tumors and adjacent normal tissues were cut into approximately 1 mm<sup>3</sup> pieces in the RPMI-1640 medium (Invitrogen) with 10% fetal bovine serum (FBS, ScienCell), and enzymatically digested with MACS tumor Dissociation Kit (Miltenyi Biotec) for 30 min on a rotor at 37°C, according to manufacturer's instruction. The dissociated cells were subsequently passed through a 70 µm Cell-Strainer (BD) and centrifuged at 400 g for 10 min. After the supernatant was removed, the pelleted cells were suspended in red blood cell lysis buffer (Solarbio) and incubated on ice for 2 min to lyse red blood cells. After washing twice with 1 × PBS (Invitrogen), the cell pellets were re-suspended in sorting buffer (PBS supplemented with 1% FBS).

PBMCs were isolated using HISTOPAQUE-1077 (Sigma-Aldrich) solution as described previously ([Zhang et al., 2018; Zheng et al., 2017](#)). Briefly, 4 mL of fresh peripheral blood was collected prior to surgery in EDTA anticoagulant tubes and subsequently layered onto HISTOPAQUE-1077. After centrifugation, lymphocyte cells remained at the plasma-HISTOPAQUE-1077 interface and were carefully transferred to a new tube and washed twice with 1 × PBS. Red blood cells were removed via the same procedure described above. These lymphocytes were re-suspended in sorting buffer.

### Single cell sorting, reverse transcription, amplification and sequencing for CRC samples (Smart-Seq2)

Single cell suspensions collected from CRC samples were stained with antibodies against CD45 for FACS sorting, performed on a BD Aria III instrument. To enrich myeloid cells, single cells were further enriched either by gating CD3<sup>-</sup>, CD3<sup>-</sup>CD19<sup>-</sup>/CD79A<sup>-</sup> or CD3<sup>-</sup>CD19<sup>-</sup>CD56<sup>-</sup> cells to exclude lymphoid cells. For B cells and NK cells, single cells were enriched by gating 7AAD<sup>-</sup>CD45<sup>+</sup>CD3<sup>-</sup>CD19<sup>+</sup>CD56<sup>-</sup> and 7AAD<sup>-</sup>CD45<sup>+</sup>CD3<sup>-</sup>CD19<sup>-</sup>CD56<sup>+</sup>. For non-immune cells, single cells were enriched by gating 7AAD<sup>-</sup>CD45<sup>-</sup>. For Patient P0104, P0305, P0309, P0411, P0720, P0728, P1212, P1228, P0413, and P0825, single-cell RNA-seq were performed by the SMART-seq2 platform.

Based on FACS analysis, single cells were sorted into 96-well plates (Axygen) chilled to 4°C, prepared with lysis buffer with 1 µl 10 mM dNTP mix (Invitrogen), 1 µl 10 µM Oligo dT primer, 1.9 µl 1% Triton X-100 (Sigma), and 0.1 µl 40 U/µl RNase Inhibitor (Takara). The single cell lysates were sealed and stored frozen at -80°C immediately. Single cell transcriptome amplifications were performed according to the Smart-Seq2 protocol ([Zhang et al., 2018; Zheng et al., 2017](#)). The External RNA Controls Consortium (ERCC, Ambion; 1: 4,000,000) was added into each well as the exogenous spike-in control before the reverse transcription. The amplified cDNA products were purified with 1 × Agencourt XP DNA beads (Beckman). A procedure of quality control (QC) was performed following the first round of purification, which included the detection of GAPDH by qPCR and fragment analysis by analyzer AATI. For those single cell samples with high quality after QC (cycle threshold < 30), the DNA products were further purified with 0.5 × Agencourt XP DNA beads, and the concentration of each sample was quantified by Qubit HsDNA kits (Invitrogen). Multiplex (384-plex) libraries were constructed and amplified using the TruePrep DNA Library Prep Kit V2 for Illumina (Vazyme Biotech). The libraries were then purified with Agencourt XP DNA beads and pooled for quality assessment by fragment analyzer. For all the 10 patients, purified libraries were analyzed by an Illumina Hiseq 4000 sequencer with 150-bp paired-end reads.

### Single cell sorting, library preparation, and sequencing for CRC samples (10× Genomics)

Single cell suspensions collected from CRC samples were stained with antibodies against CD45 for FACS sorting, performed on a BD Aria III instrument. To enrich myeloid cells, single cells were further enriched either by gating CD3<sup>-</sup> or CD3<sup>-</sup>CD19<sup>-</sup>CD79A<sup>-</sup>CD56<sup>-</sup> cells to exclude lymphoid cells. For T and monocyte-lineage cells, single cells were enriched by gating 7AAD<sup>-</sup>CD45<sup>+</sup>CD3<sup>+</sup> and 7AAD<sup>-</sup>CD45<sup>+</sup>CD3<sup>-</sup>CD14<sup>+</sup> cells, respectively. For Patient P0104, P0123, P0202, P0305, P0323, P0408, P0410, P0613, P1025, P1026, single-cell RNA-seq were performed by the 10× Genomic single cell 3' library platform. Based on FACS analysis, single cells

were sorted into 1.5 mL tubes (Eppendorf) and counted manually under the microscope. The concentration of single cell suspensions was adjusted to 300–350 cells/ul. Cells were loaded between 7,000 and 10,000 cells/chip position using the Chromium Single cell 3' Library, Gel Bead & Multiplex Kit and Chip Kit (10× Genomics, V2 barcoding chemistry) according to the manufacturer's instructions. All the subsequent steps were performed following the standard manufacturer's protocols. Purified libraries were analyzed by an Illumina Hiseq 4000 sequencer with 150-bp paired-end reads.

#### Tumor dissociation for mouse samples

MC38 tissues were mechanically disrupted and digested with Collagenase D (Roche, 0.5 mg/mL) and DNase I (Roche, 40 U/mL). Renca tumor tissues were disrupted in Liberase TL (Roche, 0.2 mg/mL) and DNase I (Roche, 20 U/mL) in RPMI with 2% FBS (Life Technology, Carlsbad, CA, USA). Disruption was performed as follows: 1 × MACS program h\_tumor\_03 (MC38) or h\_tumor\_01 (Renca) using gentleMACS Octo Dissociator (Miltenyi Biotec), followed by incubation at 37°C for 15 min, and further disruption by 2 × MACS program h\_tumor\_02. After disruption, cells were passed through a 70 µm filter and washed with RPMI containing 10% FBS. The single cell suspension was centrifuged at 1,500 rpm for 5 min and resuspended in RPMI with 10% FBS. For sorting cells, MC38 tumor tissues were mechanically disrupted and prepared as above in the absence of Collagenase D and DNase I. For some scRNA-seq experiments, Renca tumor tissues used for scRNA-seq were digested in the presence of 20 µM Actinomycin D (Sigma), and 5µM Actinomycin D was maintained in all downstream staining and sorting steps to minimize artifactual transcriptional changes during tissue processing (Wu et al., 2017).

#### Single cell sorting, library preparation, and sequencing for mouse samples

Groups of MC38 tumor-bearing C57BL/6 mice were i.p. injected either with anti-CD40 antibody or isotype control antibody. MC38 tumors and tumor draining lymph nodes (tdLNs) were harvested at day 2 or day 10 post-injection (Figure S6B). Tissues were processed, and single cell suspensions were prepared as described above. For staining of cell surface proteins, cells were incubated with appropriate antibody mixes for 30 min at 4°C. Cells from 3 different tumors and tdLNs were sorted into CD45<sup>+</sup> cells, Thy1.2<sup>+</sup> TCRβ<sup>+</sup> for T cells, CD11c<sup>+</sup>MHCII<sup>+</sup> for DC populations with BD FACSAria cytometer (BD Biosciences). Data acquisition of antibody stained samples was carried out on an LSR-II (BD Biosciences) and analyzed with FlowJo software v10.3 (Tree Star Inc.).

For scRNA-seq analysis of immune cells from Renca tumors, two independent studies were performed for both studies, groups of Renca tumor-bearing BALB/c mice were i.p. injected with either anti-CSF1R antibody or isotype control antibody (Figure S6C). Tissues were processed, and single cell suspensions were prepared as described above. For staining of cell surface proteins, cells were incubated with appropriate antibody mixes for 30 min at 4°C and cells were sorted into bulk CD45<sup>+</sup> cells, CD45<sup>+</sup>CD3<sup>+</sup> T cells, and CD45<sup>+</sup>CD3<sup>-</sup>B220<sup>-</sup>Ly6G<sup>-</sup>Ly6C<sup>-</sup> macrophage/DC enriched populations with a BD FACSMelody TM (BD Biosciences). Sorted cell subsets were combined at a ratio of 75% enriched macrophage/DC cells and 25% bulk CD45<sup>+</sup> cells (Study 1), or 65% bulk CD45<sup>+</sup> cells, 25% enriched macrophage/DC cells, and 10% T cells (Study 2). Study 1 and Study 2 data were combined for all graph-based clustering analysis of immune populations (Figures 3G and 4C). Study 2 data were used for quantification of myeloid cluster frequency in isotype control versus anti-CSF1R antibody treated groups (Figure 4D).

Sorted cell subsets were loaded on the 10× Chromium system and were encapsulated using the Single Cell 5' Library & Gel Bead Kit (10× Genomics). Single-cell gene expression and TCR libraries were generated according to the manufacturer's instructions. Completed libraries were sequenced on HiSeq4000 (Illumina) or NovaSeq (Illumina) platforms at a targeted median read depth of 50,000 reads per cell from total gene expression libraries and 5,000 reads per cell for TCR libraries (cycle specifications 150:8:0:150 [R1:i7;i5:R2]).

#### Multi-color immunofluorescence imaging of mouse tissues

For immunofluorescence imaging of Renca tumors, whole tumors were fixed in 4% PFA/PBS overnight and dehydrated in 30% sucrose/PBS prior to embedding in OCT freezing medium. 16 µm sections were cut on a cryostat, adhered to Superfrost Plus slides (ThermoFisher) and permeabilizedblocked with PBS with 0.3% Triton X-100 (Sigma) and 10% goat serum (Jackson Immunoresearch). Sections were stained with directly conjugated antibodies in PBS with 0.1% Triton X-100 and 5% goat serum. Coverslips were mounted with ProLong Diamond with DAPI (Invitrogen) and images captured on an LSM 510 confocal microscope (Carl Zeiss Microimaging).

#### Intracellular cytokine staining

For intracellular cytokine production of IFNγ and TNFα, cells were stimulated with cell stimulation cocktail plus protein transport inhibitors (ThermoFisher) for 2 h at 37°C with 5% CO2 in RPMI medium supplemented with 10% heat inactivated-FBS (GIBCO). For detecting intracellular IL-12 cytokine, MC38 tumor bearing mice were i.p. injected with anti-CD40 antibody (100ug/mouse) or control isotype. Tumors were harvested 2 days after i.p. injection of antibody and they were incubated with Brefeldin A solution (Biolegend) for 3 h at 37°C with 5% CO2 in RPMI medium containing 10% heat inactivated-FBS. Single cell suspension was obtained after tumor digestion described as above and stained with surface markers followed by staining with IL-12. To determine expression of intracellular proteins, cells were fixed for 30 min at 4°C using Foxp3 Transcription Factor Staining Buffer Set (ThermoFisher) following manufacturer's protocol.

### Bulk DNA isolation and sequencing

Genomic DNA of peripheral blood and tissue samples of CRC patients were extracted using the QIAamp DNA Mini Kit (QIAGEN) according to the manufacturer's specification. The concentrations of DNA were quantified using the Qubit HsDNA Kits (Invitrogen) and the qualities of DNA were evaluated with agarose gel electrophoresis. Exon libraries were constructed using the SureSelectXT Human All Exon V5 capture library (Agilent). Samples were sequenced on the Illumina Hiseq 4000 sequencer with 150-bp paired-end reads.

### Microsatellite instability testing

Microsatellite instability testing was performed as described previously (Zhang et al., 2018). Briefly, DNA purified from tumor tissues of CRC patients using QIAamp DNA Mini Kit (QIAGEN) was subjected to multiplex fluorescent PCR-based assay (Promega) by amplifying 7 loci including 5 mononucleotide repeats (NR21, BAT26, BAT25, NR24, and Mono27) and 2 pentanucleotide repeats (PentaC and PentaD) and was compared with DNA extracted from matched adjacent normal tissues. Multiplex PCR products were analyzed by ABI PRISM 3100 Genetic Analyzer (Applied Biosystems). Patients were defined as MSI-H status by the presence of two or more mononucleotide loci showing instability. MSS was defined as no loci of instability. Among 18 patients in this study, 3 of them were MSI-H (P0123, P0413 and P0825), and the other 15 patients were MSS.

### Multi-color immunohistochemistry of human tissues

Specimens were collected from Peking University People's Hospital within 30 min after the tumor resection and fixed in 10% formalin for 48 h. Dehydration and embedding in paraffin was performed as the following routine methods (Zheng et al., 2017). Paraffin blocks were cut into 5  $\mu$ m sections and adhered to a glass slide heated at 70°C for 1 h, deparaffinized in xylene, and then rehydrated in 100%, 90%, 70% alcohol successively. Presence of TAMs was confirmed using the Opal 7-Color Manual IHC Kit (PerkinElmer, NEL811001KT) according to the protocol of the manufacturer as described previously (Zhang et al., 2018). Briefly, antigen was retrieved by AR9 buffer (pH 6.0, PerkinElmer) and boiled in the oven for 15 min. Samples were blocked at room temperature for 10 min, incubated at room temperature for 1 h with mouse anti-human CD68 (Abcam, clone KP1, 1:100), rabbit anti-human CD80 (Abcam, clone EPR1157(2), 1:200), rabbit anti-human c-Maf (Abcam, clone EPR16484, 1:200), rabbit anti-human Macro (Abcam, 1:100), and mouse anti-human VEGFA (Abcam, clone VG-1, 1:200). A secondary horseradish peroxidase-conjugated antibody (PerkinElmer) was added and incubated at room temperature for 10 min. Signal amplification was performed using TSA working solution diluted at 1:100 in 1  $\times$  amplification diluent (PerkinElmer) and incubated at room temperature for 10 min. The multispectral imaging was collected by Mantra Quantitative Pathology Workstation (PerkinElmer, CLS140089) at 20  $\times$  magnification and analyzed by InForm Advanced Image Analysis Software (PerkinElmer) version 2.3. For each patient, a total of 8-15 high-power fields were taken based on their tumor sizes.

### Splenic T cell cultures

$2 \times 10^6$  OT-II splenocytes were plated in 48-well plates with either 10, 100, or 1000 nM Ova<sub>323-339</sub> peptide (Anaspec). Anti-CD40 (clone FGK45) or mouse IgG1 isotype control (BioXCell) antibodies were added at 10  $\mu$ g/ml, and then cultures were incubated for 3 days. CD4 $^+$  T cells were enriched by magnetic selection (mouse CD4 $^+$  T cell isolation kit, Miltenyi), then subjected to RNA enrichment (RNeasy mini kit, Qiagen) and cDNA first strand synthesis (SuperScript IV, Invitrogen). *Bhlhe40* and *Tbx21* transcript levels were then quantified using a Via 7 RT-PCR system and predesigned TaqMan gene expression arrays (ThermoFisher). Relative expression refers to *Bhlhe40* or *Tbx21* expression normalized to a control gene, *Rpl19*. Immune cell phenotypes were evaluated by flow cytometry and intracellular cytokine staining as described above.

### Human monocyte culture

Human peripheral blood monocytes were prepared by density centrifugation of whole blood diluted 1:3 in Lymphoprep (Accurate Chemical) in 50 mL Leucosep tubes (Greiner Bio-One) followed by enrichment for CD14 $^+$ CD16 $^-$  cells via negative selection using the Classical Monocyte Isolation Kit (Miltenyi). Monocytes were seeded at  $1 \times 10^5$  cells/well in 96-well plates in X-Vivo 10 medium (Lonza) containing 50ng/mL CSF1 or GM-CSF (PeproTech) and the indicated factors (all from PeproTech). Plates were incubated for 6 days under normoxic or hypoxic (5% O<sub>2</sub>) conditions in the presence of cytokines or growth factors that were differentially expressed in TAM populations (Table S3A). The macrophages were analyzed by flow cytometry.

## QUANTIFICATION AND STATISTICAL ANALYSIS

### Single-cell RNA-Seq data processing

The Smart-seq2 data were processed as previously described (Zhang et al., 2018). Low-quality reads were filtered out when: (1) "N" bases accounting for 10% read length; or (2) bases with quality < 5 account for 50% read length; or (3) containing adaptor sequences. The remaining paired-end reads were aligned to the UCSC hg38 human genome reference using *kallisto* (version 0.43.1) (Bray et al., 2016) with parameters "-bootstrap-samples=100-threads=2." The R package *tximport* (version 1.9.12) was used to summarize the transcript-level estimated counts into the matrix of gene-level counts. Low-quality cells were discarded if the cell library size or the number of expressed genes (counts larger than 0) was smaller than pre-defined thresholds, which were the medians of all cells minus

Dataset	Platform	# Genes	# Processed cells
Human	Smart-seq2	15,179	10,468
Human	10× 3' Gene Expression	13,538	43,817
Mouse anti-CD40	10× V(D)J + 5' Gene Expression	15,243	119,698
Mouse anti-CSF1R	10× 3' Gene Expression	13,328	8,578

3 × median absolute deviation. Cells were also removed if their proportions of mitochondrial gene expression were larger than 10% (70% for CD45<sup>-</sup> cells). In the droplet-based 10× data, the *Cell Ranger* toolkit (version 2.1.0) provided by 10× Genomics was applied to align reads and generate the gene-cell unique molecular identifier (UMI) matrix, using the reference genome GRCh38 and GRCm38. For each cell, we quantified the number of genes and UMIs, and kept high quality cells with the detection threshold of 600-4,000 genes and 1,600-25,000 UMIs. We used a relatively high threshold to ensure that we have filtered out the most of barcodes associated with empty partitions or doublet cells. Cells with unusually high detection rate of mitochondrial gene expression were also excluded as described above. The following table shows the number of resulting genes and cells for each dataset collected in this study. For human Smart-seq2 dataset, a total of 5,126 single T cells we generated previously (Zhang et al., 2018) were combined with scRNA-seq data from the same patients in this study.

#### Evaluation of batch effect with entropy metric

We used an entropy-based metric to quantify the batch effect across different patients (Azizi et al., 2018). Genes with the top 1000 standard deviation were used to map the data to a two-dimensional t-SNE by R package *Rtsne* (version 0.15). We constructed a k-NN graph ( $k = 30$ ) based on the Euclidean distance of cells in tSNE coordinates using the R function *kNN* from *dbSCAN* package (version 1.1.2), and defined the mixing of the data across patients as Shannon entropy,

$$H_j = - \sum_{t=1}^T p_j^t \log_2 p_j^t$$

where  $p_j^t$  is the ratio of the number of cells from patient  $t$  in the 30 nearest neighborhood of cell  $j$  and  $\sum_{t=1}^T p_j^t = 1$ .

#### Batch effect correction and normalization

We performed batch correction across different patients in both Smart-seq2 and 10× scRNA-seq data by calculating the batch factor (BF) using the approach in *Pagoda2*. The batch removed counts ( $RC_{nobatch}$ ) was defined as  $C_{ij}/BF_{i,k}$ , where  $C_{ij}$  was the raw count value (or UMI) of gene  $i$  in cell  $j$  and  $BF_{i,k}$  was the batch factor of gene  $i$  in batch  $k$ . In Smart-seq2 data, the transcripts per million (TPM) value was calculated by  $((10^6 \times RC_{nobatch\ ij}) / \text{length of gene } i) / (\sum RC_{nobatch\ ij} / \text{length of gene } i)$ . In 10× data, the expression value was calculated by to create a TPM-like value, and finally computing  $\log_2(TPM + 1)$ . The corrected expression matrices were used as the input for further analysis.

#### Cell doublet detection and removal

We applied *Scrublet* (Wolock et al., 2019) in human 10× dataset to identify artificial libraries generated from two or more cells, where they entered the same microfluidic droplet and were labeled with barcodes. The doublet score for each single cell and the threshold based on the bimodal distribution was calculated using default parameters. We used a cluster-level approach to remove doublet clusters containing large number of potential doublet cells. Specifically, we removed the CD14<sup>+</sup>IL32<sup>+</sup> cluster with a large fraction of potential monocyte-NK cell doublets, which expressed both monocyte signature genes (CD14, S100A8, S100A9, S100A12) and cytotoxic signature genes (IL32, NKG7, GZMA, GZMB) (data not shown).

#### Bulk exome-seq data processing

The bulk exome-seq data were cleaned following the same procedure for scRNA-seq data processing. The cleaned high-quality reads were then aligned to human genome according to the *BWA-PICARD/GATK-strelka* pipeline, and the somatic copy number alterations were obtained by *ADTEx* (version 1.0.4) as we previously described (Zhang et al., 2018).

#### TCR analysis

The TCR-seq data of mouse anti-CD40 treatment dataset was processed using Cell Ranger (version 2.1.0) against the manufacturer-supplied mouse vdj reference genome. In all TCR contigs assembled, we first discarded the low-confidence, non-productive or those UMIs < 2. For cells with two or more  $\alpha$  or  $\beta$  chains assembled, the  $\alpha-\beta$  pair showing the highest expression level (UMI) was defined as the dominant  $\alpha-\beta$  pair in the corresponding cell. If two or more cells had identical dominant  $\alpha-\beta$  pairs, the dominant  $\alpha-\beta$  pair were identified as clonal TCRs, and these T cells were identified as clonal T cells. To integrate TCR results with the gene expression data, the TCR-based analysis was performed only for cells that were identified as T cells. We identified the TCR  $\alpha-\beta$  pairs for

53,499 cells from 33,114 CD4<sup>+</sup> T cells and 37,740 CD8<sup>+</sup> T cells. The detailed TCR  $\alpha$  and  $\beta$  sequences and clonotypes are presented in [Table S6A](#). Among all T cells, 53,499 CD4<sup>+</sup> or CD8<sup>+</sup> T cells with TCR information were used to perform the STARTRAC analysis as we previously described ([Zhang et al., 2018](#)).

### Unsupervised clustering analysis

To cluster single cells by their expression, we used an unsupervised graph-based clustering algorithm implemented in *Seurat v2* (version 2.3.0) ([Butler et al., 2018](#)) for each dataset separately. The highly variable genes were generated with appropriate threshold of the mean expression and dispersion (variance/mean). Principal component analysis (PCA) was performed on about 2000 variable genes. We used the function *FindClusters* on 10-25 PCs with resolution 0.6-2 to perform the first-round cluster and annotated each cluster by known markers. We identified eight major cell types, including 4 immune cell types (T cells, ILCs, B cells and myeloid cells) and 4 non-immune cell types (epithelial cells, endothelial cells, fibroblasts, and malignant cells). Specifically, malignant cell clusters were also characterized by the exceptionally high amounts of expressed genes. The second-round clustering was performed according to the same range of parameters to identify clusters within the major cell types aforementioned. Notably, neutrophils, despite their importance in the tumor microenvironment ([Mantovani et al., 2011](#)), were not sufficiently captured in either Smart-seq2 or 10 $\times$  scRNA-seq platform, which was further confirmed by reference component analysis (RCA) ([Li et al., 2017](#)) and the expression of neutrophil markers (data not shown). The under-representation of these cells was most likely due to the techniques that were employed to purify cells prior to sorting and the known difficulties in capturing adequate RNA from this cell type ([Schelker et al., 2017; Zillionis et al., 2019](#)).

The unsupervised clustering of single cells collected from each mouse model was performed following the same procedures as we did for human scRNA-seq datasets. To further identify the consistent myeloid cell clusters from both mouse datasets, we performed integrative analysis of myeloid cells from these two datasets by *Seurat v3* (version 3.1.0) ([Butler et al., 2018](#)) using the standard integration workflow.

### Dimensionality reduction using t-SNE and UMAP

For visualization, the dimensionality of each dataset was further reduced using either the Barnes-Hut t-Distributed Stochastic Neighbor Embedding (t-SNE) or Uniform Manifold Approximation and Projection (UMAP) with *Seurat* functions *RunTSNE* and *RunUMAP*. The PCs used to calculate the embedding were as the same as those used for clustering. When calculating the UMAP coordinates, the effective scale of embedded points setting (spread) varied from 1 to 2.

To integrate and embed single cells from human CRC generated by two platforms into a shared low-dimension space ([Figure 1C](#)), we utilized integrated analysis (CCA) by the *Seurat v3* function *IntegrateData*. Importantly, the new integrated matrix was obtained and was used only for visualization, instead of clustering or any other downstream analyses.

### Identification of signature genes

We identified differentially expressed genes (DEGs) based on analysis of variance (ANOVA) using the R function *aov*, and the differences between each cluster pairs were tested using Tukey's range test implemented in R function *TukeyHSD* as we previously described ([Guo et al., 2018](#)). In brief, for each cluster, only genes that met these criteria were considered as DEGs: 1) FDR adjusted *P*-value of F test < 0.05; 2) any cluster pair showed significant difference in the HSD test (absolute log<sub>2</sub> fold change > 1 and *P*-value adjusted by Tukey's 'Honest Significant Difference' method < 0.01).

### Tissue distribution of clusters

To quantify the cell type enrichment across tissues, we compared the observed and expected cell numbers in each cluster according to the following formula as we described previously ([Guo et al., 2018; Zhang et al., 2018](#)):

$$R_{o/e} = \frac{\text{Observed}}{\text{Expected}}$$

where the expected cell numbers for each combination of cell clusters and tissues are obtained from the Chi-square test. We assumed that one cluster was enriched in a specific tissue if  $R_{o/e} > 1$ .

### Similarity analysis of clusters from different datasets

To compare the similarities of clusters from different datasets, we trained a logistic regression model using elastic net regularization ([Young et al., 2018](#)). Before training the model, we downsampled each cluster to the minimum size of all clusters in the training dataset to eliminate possible bias attributed to different sizes of training sets. Specifically, for training datasets containing clusters with cell number lower than 50, the downsampled size of these clusters was set as 50. The R function *cv.glmnet* from the *glmnet* package (version 2.0.16) was used to fit a series of *n* binomial logistic regression models with parameters alpha = 0.99, where *n* is the number of clusters in the training data. A 10-fold cross validation was performed in each case. The offset for each model was calculated with  $\log(f / 1 - f)$ , where *f* is the fraction of cells in the cluster being trained.

We used these models to calculate a predicted logit of each cell in the test data for each cluster from the training data with an offset of 0. Predicted logits were then averaged within each cluster and converted to probabilities for visualization, which indicated the similarity of clusters from the test data to those from the training data.

When comparing the similarities between myeloid cells collected from human and two mouse datasets (Figure 3H), we calculated the predicted logits between human 10 $\times$  dataset and anti-CD40 mouse dataset and between human 10 $\times$  dataset and anti-CSF1R dataset separately. Two logits matrices were combined together and then averaged as above.

### Identification of malignant cells with CNV estimation

To distinguish malignant cells from non-malignant cells in human CRC CD45 $^{-}$  cells generated by Smart-seq2, we inferred large-scale chromosomal copy-number variations (CNVs) in each single cell based on a moving averaged expression profiles across chromosomal intervals (Patel et al., 2014; Puram et al., 2017; Tirosh et al., 2016).

We first excluded the genes with low expression levels by the cutoff of  $\log_2(\text{average}(TPM_{i,j}) + 1) < 3.5$  and re-quantified the gene expression as  $E_{i,j} = \log_2(TPM_{i,j} / 10 + 1)$ , where  $TPM_{i,j}$  referred to transcripts per million for gene  $i$  in cell  $j$ . Then the relative expression levels were centered as  $Er_{i,j} = E_{i,j} - \text{mean}(E_{i,j})$ . To avoid considerable impact of any particular gene, we limited the relative expression values to [-3,3] by replacing all values above 3 by 3 and values below -3 by -3. This relative expression matrix was only used for CNV estimation.

Initial CNVs (CNV) were calculated by sorting the analyzed genes by their chromosomal location and applying a moving average to the relative expression values, with a sliding window of 100 genes within each chromosome. The CNV<sub>c</sub> was further centered cell by cell so that the mean CNV<sub>c</sub> of a given cell returned to zero. We have identified and annotated CD45 $^{-}$  sub clusters based on the expression of known maker genes, which helped to redefine CNV estimation using the average patterns of non-malignant cells as a reference. We thus defined multiple baselines as the average CNV<sub>c</sub> of all cells in each non-malignant epithelial cell cluster (hE02-hE06), and the maximal (BaseMax) and minimal (BaseMin) baseline at each gene window across these five clusters. The final CNV estimation CNV<sub>f</sub> of cell  $i$  at position  $j$  was defined as:

$$\text{CNV}_f(i,j) = \begin{cases} \text{CNV}_c(i,j) - \text{BaseMax}(j), & \text{if } \text{CNV}_c(i,j) > \text{BaseMax}(j) + 0.2 \\ \text{CNV}_c(i,j) - \text{BaseMin}(j), & \text{if } \text{CNV}_c(i,j) < \text{BaseMin}(j) - 0.2 \\ 0, & \text{if } \text{BaseMin}(j) - 0.2 < \text{CNV}_c(i,j) < \text{BaseMax}(j) + 0.2 \end{cases}$$

When plotting the heatmap, CNV<sub>f</sub> were limited to [-1,1] and the values between [-0.1,0.1] were replaced by 0 (Figure S3E). To quantitatively distinguish malignant cells from non-malignant cells, we scored each cell with two CNV parameters: (1) CNV score reflected the overall CNV signal, defined as the mean squares of CNV<sub>f</sub> values across the genome; (2) CNV correlation was the Pearson correlation between CNV<sub>f</sub> profile of each cell and the average CNV<sub>f</sub> of all malignant cells identified by unsupervised clustering from the same sample (Figure S3F).

### Developmental trajectory inference

Three different algorithms, including RNA velocity, URD and partition-based graph abstraction (PAGA), were used to infer the developmental trajectories of monocytes and macrophages sequenced by 10 $\times$  Genomics scRNA-seq. When performing the RNA velocity analysis, we first used the R implementation of *destiny* (version 2.11.3) (Angerer et al., 2016) as a nonlinear dimensionality approach to access the diffusion map of up to 1000 random selected cells in each myeloid cluster (hM05-hM13). The *velocityto* python package was then used to recount the spliced reads and unspliced reads based on previously aligned bam files, while the *velocityto* R package was used to calculate RNA velocity values for each gene from each cell and embed RNA velocity vector to the 2-D diffusion map space (La Manno et al., 2018).

To get a clearer pseudo-temporal ordering of all monocytes and macrophages enriched in tumors (hM08, hM10, hM11, hM12, hM13), another algorithm, *URD* (version 1.0.1) (Farrell et al., 2018), was also applied to reconstruct the developmental trajectories into a tree structure (Figure 2F). Monocytes from cluster hM05\_Mono-CD14 were set as root cells, while cells from cluster hM06\_Mono-CD16 were used as control. The analyses were performed based on the following parameters:

```
knn = 200, sigma = 8, divergence.method = 'ks', cells.per.pseudotime.bin = 100, bins.per.pseudotime.window = 10, minimum.visits = 7, p.thresh = 0.001.
```

The third algorithm, PAGA, was calculated by *scanpy* (Wolf et al., 2018) using the top 2000 highly variable genes and 15 PCs (Figure S4E). Based on the calculations of these three algorithms, a model for the developmental trajectories of monocytes and macrophages in CRC was summarized and provided in Figure S4F.

### Gene set enrichment analysis

Gene set variation analysis implemented in the GSVA package (version 1.3.0) was used for gene set enrichment analysis. The gene sets we used were exported using the GSEABase package (version 1.44.0). The differences in pathway activities scored per cell between hM12\_TAM-C1QC and hM13\_TAM-SPP1 cells (Figure 3B) were calculated with LIMMA package (version 3.37.11).

### Define M1/M2 and proliferation phenotypes

The M1/M2 phenotype of each myeloid cell was defined as the mean expression of gene signatures (Azizi et al., 2018). Genes associated with “classically activated” (M1) macrophages include *CCL5*, *CCR7*, *CD40*, *CD86*, *CXCL9*, *CXCL10*, *CXCL11*, *IDO1*, *IL1A*, *IL1B*, *IL6*, *IRF1*, *IRF5* and *KYNU*, while *CCL4*, *CCL13*, *CCL18*, *CCL20*, *CCL22*, *CD276*, *CLEC7A*, *CTSA*, *CTSB*, *CTSC*, *CTSD*, *FN1*, *IL4R*, *IRF4*, *LYVE1*, *MMP9*, *MMP14*, *MMP19*, *MSR1*, *TGFB1*, *TGFB2*, *TGFB3*, *TNFSF8*, *TNFSF12*, *VEGFA*, *VEGFB* and *VEGFC* were used to define the signature of “alternatively activated” (M2) macrophages.

The average expression of known proliferation-related genes was defined as the proliferation score (Whitfield et al., 2006). These proliferation genes include *AURKA*, *BUB1*, *CCNB1*, *CCND1*, *CCNE1*, *DEK*, *E2F1*, *FEN1*, *FOXM1*, *H2AFZ*, *HMGB2*, *MCM2*, *MCM3*, *MCM4*, *MCM5*, *MCM6*, *MKI67*, *MYBL2*, *PCNA*, *PLK1*, *TOP2A*, *TYMS* and *ZWINT*.

### Comparison of myeloid cell subsets with previous datasets

To compare the myeloid cells from our human CRC dataset with those from UC patients and healthy individuals published previously (Smillie et al., 2019), we fetched the UMLs matrix of all myeloid cells and removed the batch effect across different individuals with *bbknn*. After re-clustering of the high-quality cells from UC dataset by *scranpy* (Figure S4J), we used the function *TransferData* from Seurat v3 package to re-classify the myeloid cells from UC dataset based on our CRC 10x dataset. The predicted IDs were obtained and subsequently used to calculate the percentage of cells from UC datasets projected onto different subsets identified in CRC dataset (Figure S4K).

For systemic comparison of cDC1 cells across different human individuals and mouse models, we performed hierarchical clustering of cDC1 cells and activated DC cells from our human CRC, mouse MC38, mouse Renca datasets, recently published UC (Single Cell Portal, SCP259) (Smillie et al., 2019), human and mouse non-small-cell lung cancer (NSCLC) (GEO, GSE127465) (Zillionis et al., 2019), and hepatocellular carcinoma (HCC) datasets (EGA, EGAS00001003449) (Zhang et al., 2019) based on the expression of signature genes related to immature and mature cDC1 (Figure S6I). cDC1 cells were labeled as *mDC1/hDC1* in NSCLC dataset and *DC-c3-CLEC9A* in HCC dataset, while activated DC cells were labeled as *mDC3/hDC3* in NSCLC dataset and *DC-c4-LAMP3* in HCC dataset.

### Correlative cell-cell interactions inferred by combined scRNA-seq and TCGA datasets

#### Cell subtype abundance estimated from bulk expression profiles

For 40 immune cell subtypes identified in human CRC, we assigned them into tumor or normal-enriched group based on their tissue enrichment scores quantified by Ro/e values ( $\text{Ro}/\text{e} > 1$ ; Figures 2C and S3C). Among them, 24 tissue-enriched clusters identified by both Smart-seq2 and 10x Genomics platforms were used in this analysis. To systematically infer the correlative interactions between immune and non-immune cells, we also included three tumor-enriched stromal cell subsets, including fibroblasts (*hF01\_Myofib-ACTA2* and *hF02\_CAF-FAP*), and endothelial cells (*hE01\_Endothelium-ACKR1*). The specific gene signatures for each subtype were derived from Smart-seq2 dataset based on the following criteria: 1) FDR adjusted  $P$ -value of F test  $< 0.05$ ; 2) any cluster pair showed significant difference in the HSD test (absolute  $\log_2$  fold change  $> 1$  and  $P$ -value adjusted by Tukey’s ‘Honest Significant Difference’ method  $< 0.01$ ). 3) AUC  $> 0.65$ ; 4) typical marker genes described in literatures.

For tumor-enriched clusters, the TCGA COAD and READ data were used to evaluate the abundance of each cluster, while the GTEx colon tissue data were used for normal-enriched clusters. The gene expression from these bulk RNA-seq datasets were downloaded from UCSC Xena (<http://xena.ucsc.edu/>). Subsequently, we estimated the relative abundance of each cell subtype by the average expression of z-score normalized log-transformed expression of the cell type specific genes defined above. The lists of cell subtype specific genes were provided in Table S4A.

#### Gene expression and cell subtype abundance correlation matrix

Given the hypothesis that changes in the abundance of cells in the tissue could be influenced by cell-cell interaction (Jerby-Arnon et al., 2018; Puram et al., 2017; Tirosh et al., 2016), we computed the Pearson correlation coefficient between the expression of each gene and the relative abundance of each cell subtype to identify genes that may infer the presence of cells co-occurring with these particular clusters (Figure S8A). Based on the bulk RNA-seq profiles, we obtained a correlation matrix between genes and abundances of cell subtypes. Since the expression levels of certain signature genes in a given cell subtype were highly correlated with the abundance of this specific cell subtype, we defined these genes as “self-expressed genes” based on our Smart-seq2 dataset according to the following criteria: (1) average expression  $> 1$ ; (2) cell frequency of expression  $> 20\%$ . We then filtered these genes and transformed their correlation values to 0 to obtain an adjusted correlation matrix (Figure S8A).

#### Enrichment analysis for highly correlated genes

For a specific cell subtype, to identify which cell subtypes could contribute to the highly correlated non-self-expressed genes, we performed a gene set enrichment analysis based on our Smart-seq2 dataset. In detail, we calculated the mean expression of each gene across each tumor or normal-enriched cell subtype and then performed z-score transformation. Next, the top 13 highly correlated non-self-expressed genes were selected from the adjusted correlation matrix. We then calculated the mean value of z-score transformed expression of these genes in each cell subtype as the enrichment score. Finally, the correlated cell subtype(s) was identified as the z-score transformed enrichment score  $> 1.96$  (Figure S8A). The correlative networks in tumor and normal mucosa were built separately with corresponding correlated cell subtypes (Figures 3E and S5C).

### Identification of significant ligand-receptor pairs

To identify significant ligand-receptor pairs in human CRC using the Smart-seq2 data, we first downloaded the ligand-receptor (LR) pairs from published databases (Ramilowski et al., 2015; Vento-Tormo et al., 2018). We then defined a ligand or receptor as an “expressed” gene in a certain cell type if more than 20% cells had its expression level by the cutoff of  $\log_2(\text{TPM}+1)>3$ , and we set the expression value of unexpressed ligand/receptor to zero. We scored a given LR pair interaction between cell type A and cell type B as the product of average ligand expression across all cells in cell type A and the average receptor expression across all cells in cell type B using the following formula,

$$\text{Int}_{L_i \rightarrow R_i}^{A \rightarrow B} = E_{L_i}^A \times E_{R_i}^B$$

Statistical significance was then assessed by randomly shuffling the cluster labels of all cells and repeating the above steps, which generated a null distribution for each LR pair in each pairwise comparison between two cell types. After running 1,000 times permutations,  $P$ -values were calculated with the normal distribution curve generated from the permuted LR pair interaction scores.

### Survival analysis

The TCGA COAD and READ data were used to evaluate the prognostic performance of individual genes or gene sets derived from specific cell clusters. The gene expression data and the clinical data were downloaded from UCSC Xena (<http://xena.ucsc.edu>). The mean expression of given signatures was grouped into high and low expression groups by the 55<sup>th</sup> and 45<sup>th</sup> quantile values. We used the Cox proportional hazards model implemented in the R package *survival* to correct tumor stage and MSI-state for survival analyses and plotted Kaplan-Meier survival curves using R function *ggsurvplot*.

### DATA AND CODE AVAILABILITY

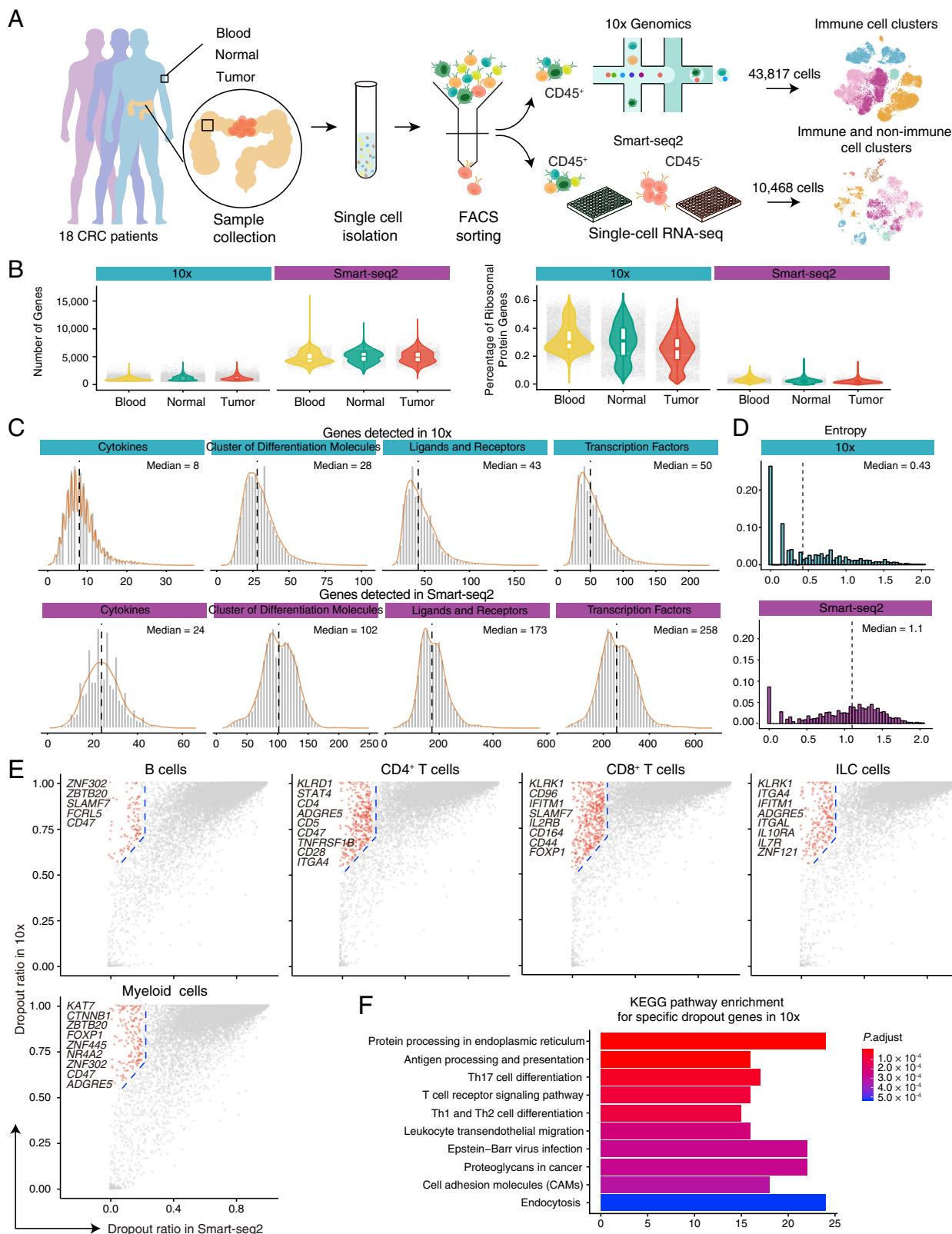
The accession numbers for the raw sequencing data and processed data from mouse samples in this paper are ENA: PRJEB34105 and ArrayExpress: E-MTAB-8832. The accession number for the processed data from human samples in this paper is GEO: GSE146771. The raw FASTQ files are only provided for mouse due to human patient privacy concerns.

### ADDITIONAL RESOURCES

Analysis and visualization of scRNA-seq datasets from both human and mouse can be also performed at <http://crcleukocyte.cancer-pku.cn/>.

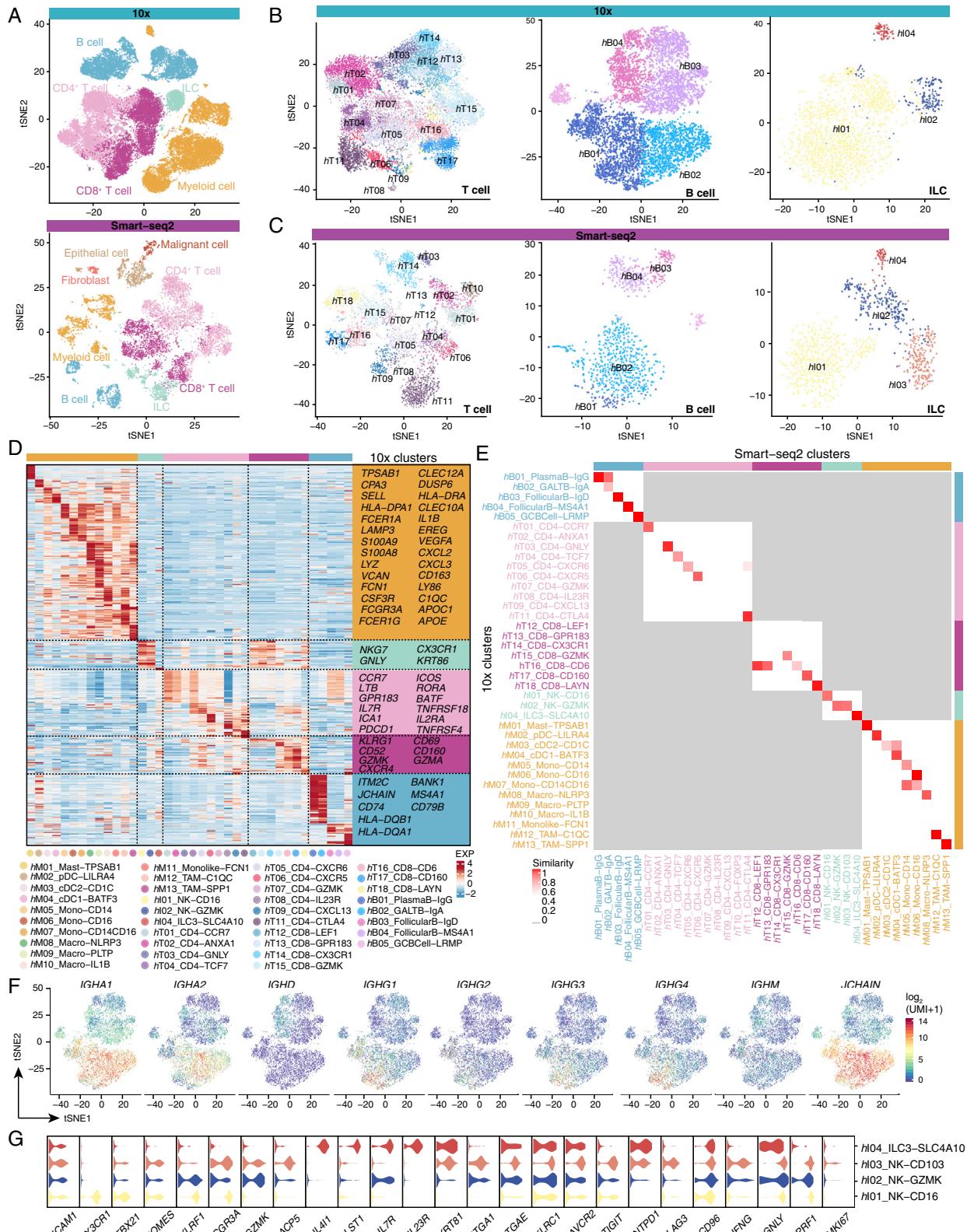
# Supplemental Figures

Cell



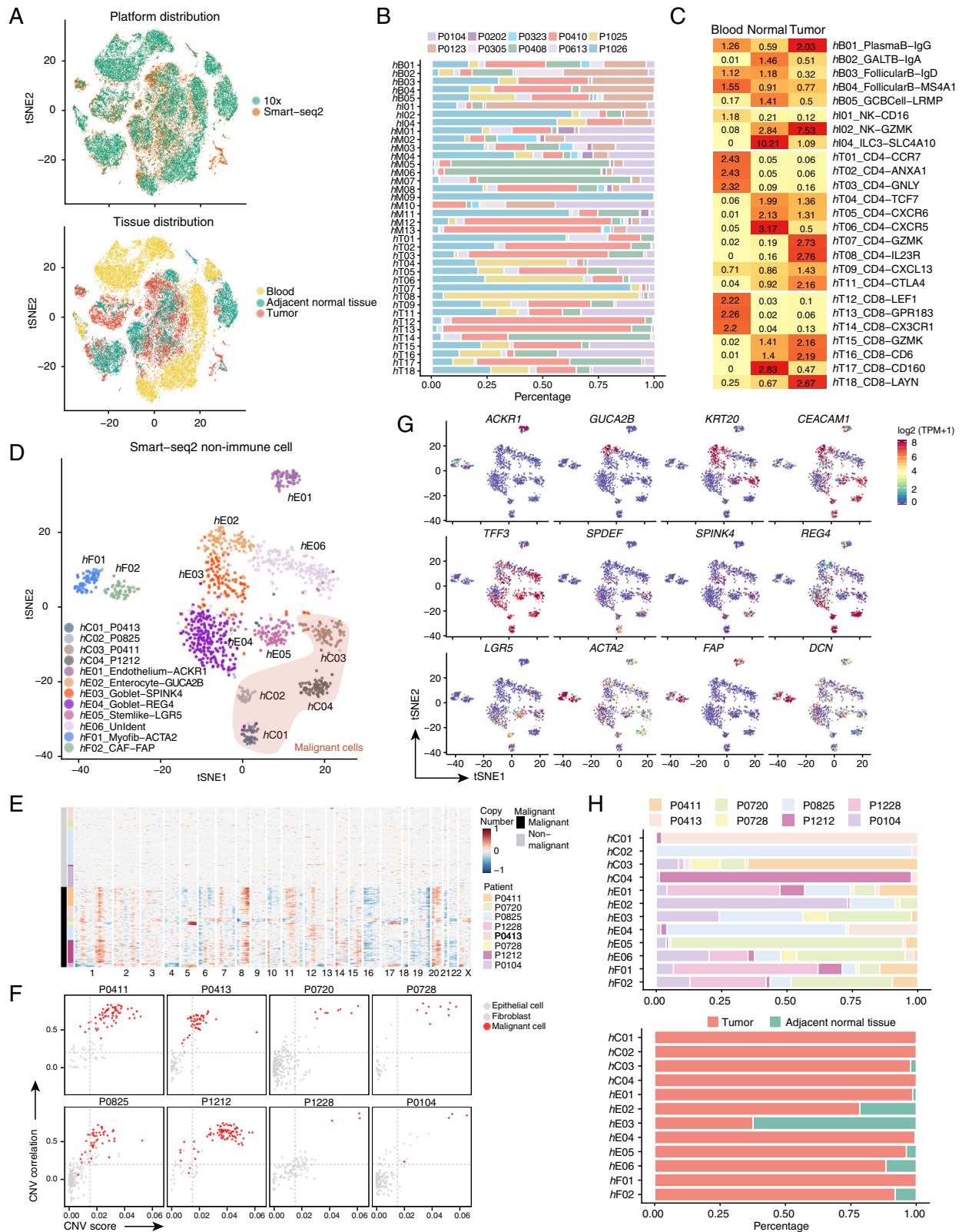
(legend on next page)

- 
- Figure S1. Data Quality of Smart-seq2 and 10 $\times$  Genomics scRNA-seq Datasets from Human CRC, Related to Figure 1**
- (A) Experimental design and analysis of single cells from human CRC samples.
- (B) Number of total genes (left) and percentage of ribosomal protein genes (right) detected in cells from 10 $\times$  and Smart-seq2 scRNA-seq datasets.
- (C) Distribution of genes encoding cytokines, cluster of differentiation molecules, ligands, and receptors, and transcription factors detected in cells from 10 $\times$  (upper) and Smart-seq2 scRNA-seq datasets (lower).
- (D) Histogram showing entropy of patient distribution as a measure of batch effect in 10 $\times$  (upper) and Smart-seq2 scRNA-seq datasets (lower).
- (E) Dropout ratio of genes from different clusters of immune cells detected by 10 $\times$  (y axis) versus Smart-seq2 scRNA-seq (x axis). Points in red indicate specific dropout genes from 10 $\times$  scRNA-seq, which were defined as genes whose dropout ratio in Smart-seq2 is less than 0.2 and in 10 $\times$  are greater than Smart-seq2 plus 0.5.
- (F) Enrichment of different KEGG pathways using specific dropout genes from 10 $\times$  scRNA-seq.



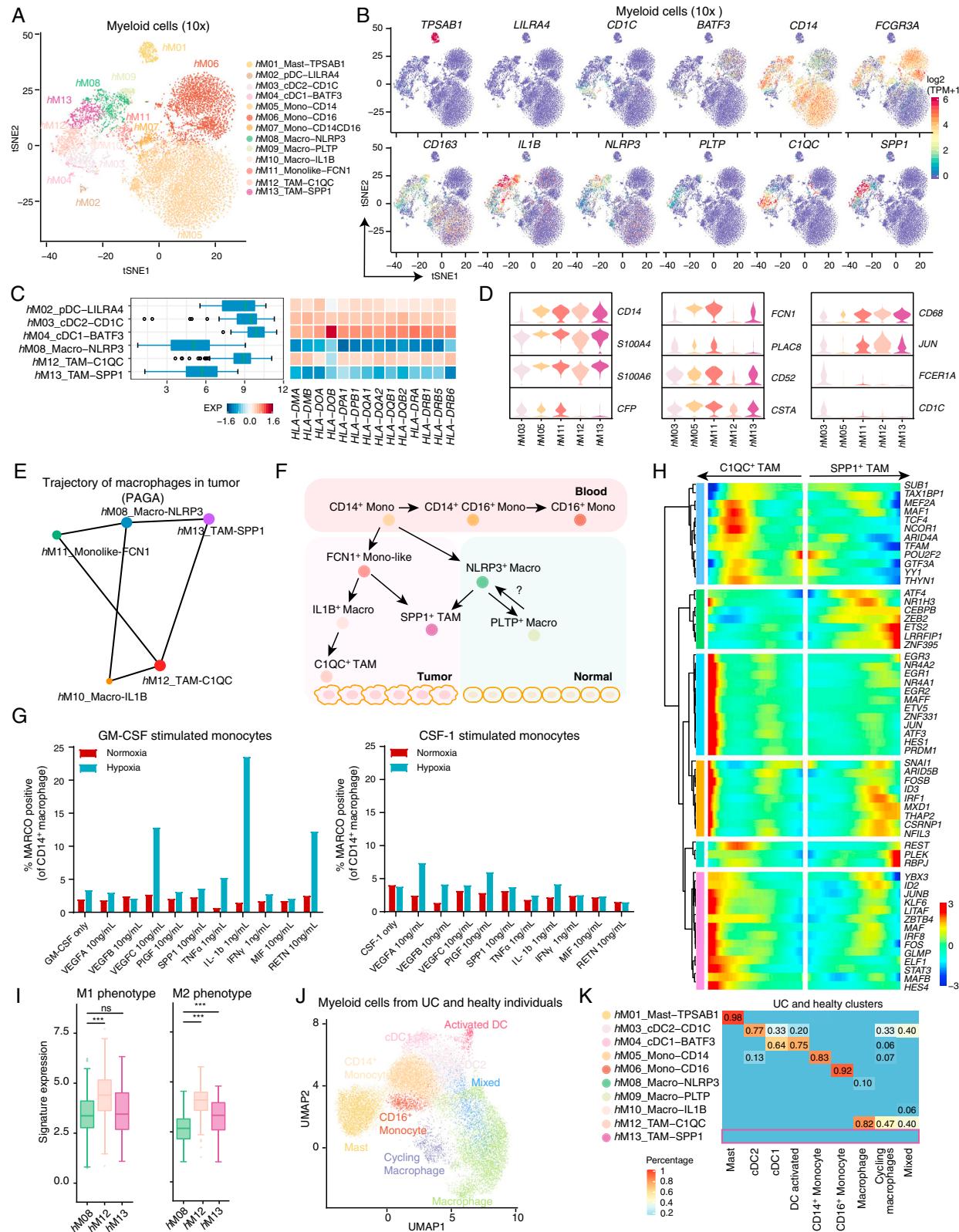
(legend on next page)

- 
- Figure S2. Similarities and Characteristics of Immune Cell Subsets from Human CRC Analyzed by scRNA-seq, Related to Figure 1**
- (A) tSNE plots showing major immune cell subsets identified by 10x Genomics (upper) and Smart-seq2 scRNA-seq (lower).
- (B) t-SNE plots showing 17 T cell clusters (left), 5 B cell clusters (middle) and 3 ILC clusters (right) identified by 10x Genomics scRNA-seq.
- (C) t-SNE plots showing 18 T cell clusters (left), 5 B cell clusters (middle) and 4 ILC clusters (right) identified by Smart-seq2 scRNA-seq.
- (D) Gene expression heatmap of myeloid and lymphoid cell clusters.
- (E) Similarities of lymphoid and myeloid cell clusters from human CRC identified by 10x Genomics (row) and Smart-seq2 (column) scRNA-seq.
- (F) t-SNE plots showing different expression patterns of selective immunoglobulin genes of B cell clusters.
- (G) Violin plots showing expression of selected genes from ILC clusters. Cluster hI03\_NK-CD103 showed co-expression of inhibitory and cytotoxic molecules.



(legend on next page)

- 
- Figure S3. Characteristics of Immune and Non-Immune Cell Subsets from Human CRC Analyzed by scRNA-seq, Related to Figure 1**
- (A) tSNE plots showing the platform distribution (upper) and tissue distribution (lower) of single cells analyzed by Smart-seq2 and 10 $\times$  scRNA-seq platforms. (B) Patient distribution of each myeloid and lymphoid cell cluster.
- (C) Tissue preference of each lymphoid cluster estimated by Ro/e score.
- (D) t-SNE plot showing 12 non-immune cell clusters (4 for malignant and 8 for non-malignant clusters). Malignant cell clusters were classified by inferred copy number variations (CNVs; [STAR Methods](#)).
- (E) Heatmap showing large-scale inferred copy number variations (CNVs) for individual cells (row) from 8 CRC tumors based on the average expression of 100 genes surrounding each chromosomal position (column). CNVs in red indicates amplifications, and blue indicates deletions.
- (F) Classification of malignant cell clusters based on inferred CNV score (x axis) and CNV correlation (y axis) for each tumor ([STAR Methods](#)).
- (G) t-SNE plots showing expression of selected genes from non-immune cell subsets.
- (H) Patient distribution and tissue distribution of each non-immune cell cluster.

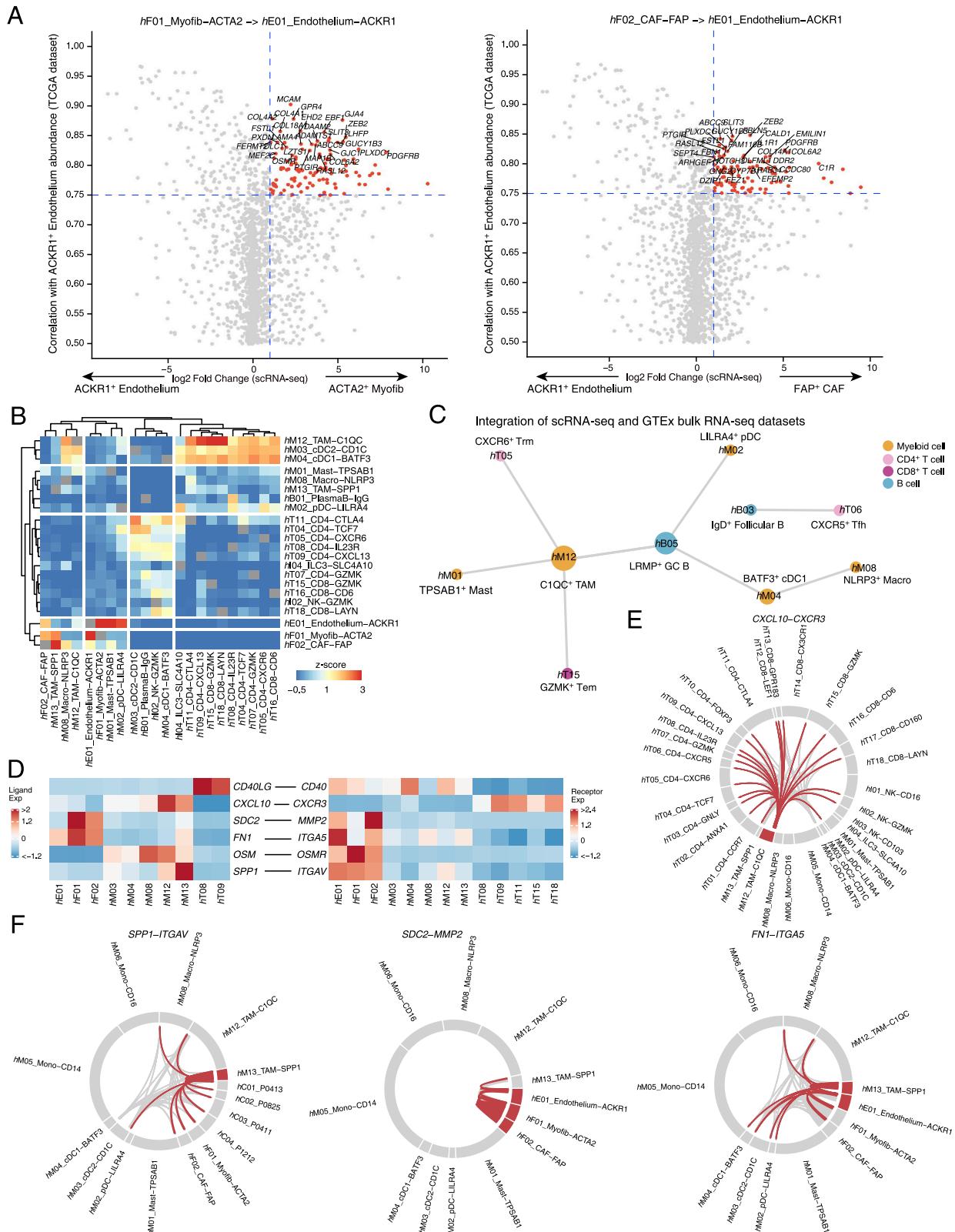


(legend on next page)

---

**Figure S4. Myeloid Cells in Human CRC, Related to Figures 2 and 3**

- (A) t-SNE plot showing 13 myeloid clusters analyzed by 10x scRNA-seq.
- (B) t-SNE plots showing expression of selected genes in myeloid cell clusters from (A).
- (C) Boxplots showing expression of an antigen-presenting signature across indicated myeloid cell clusters (left), calculated by the mean expression of *HLA*-genes, whose expression patterns are also shown in heatmap (right; Smart-seq2 scRNA-seq).
- (D) Violin plots showing expression of selected genes from certain myeloid clusters. Cluster hM11\_Monolike-FCN1 showed expression of signature genes from both *CD14*<sup>+</sup> monocytes and TAMs.
- (E) PAGA graph showing the inferred developmental trajectories for tumor-enriched myeloid clusters.
- (F) Model of the developmental trajectory of monocyte/macrophage lineages in the blood, adjacent normal tissues and tumors from CRC patients.
- (G) Percent MARCO<sup>+</sup> cells of total CD14<sup>+</sup> monocytes following culture under the indicated conditions. Data are from one experiment.
- (H) Heatmap showing the expression of transcription factors identified as varying significantly along the pseudotime trajectory.
- (I) Boxplots showing comparison of M1 (left) and M2 phenotype (right) across indicated macrophage clusters, calculated by the mean expression of corresponding signature genes (Smart-seq2 scRNA-seq). Two-sided Wilcoxon test. \*\*\*, p < 0.001.
- (J) UMAP plot showing the myeloid cell subsets identified in non-cancer tissue from UC patients and healthy individuals.
- (K) Heatmap showing the percentage of myeloid cells from UC dataset projected onto different myeloid cell subsets identified from CRC dataset.



(legend on next page)

**Figure S5. Cell-Cell Interactions in Human CRC, Related to Figure 3**

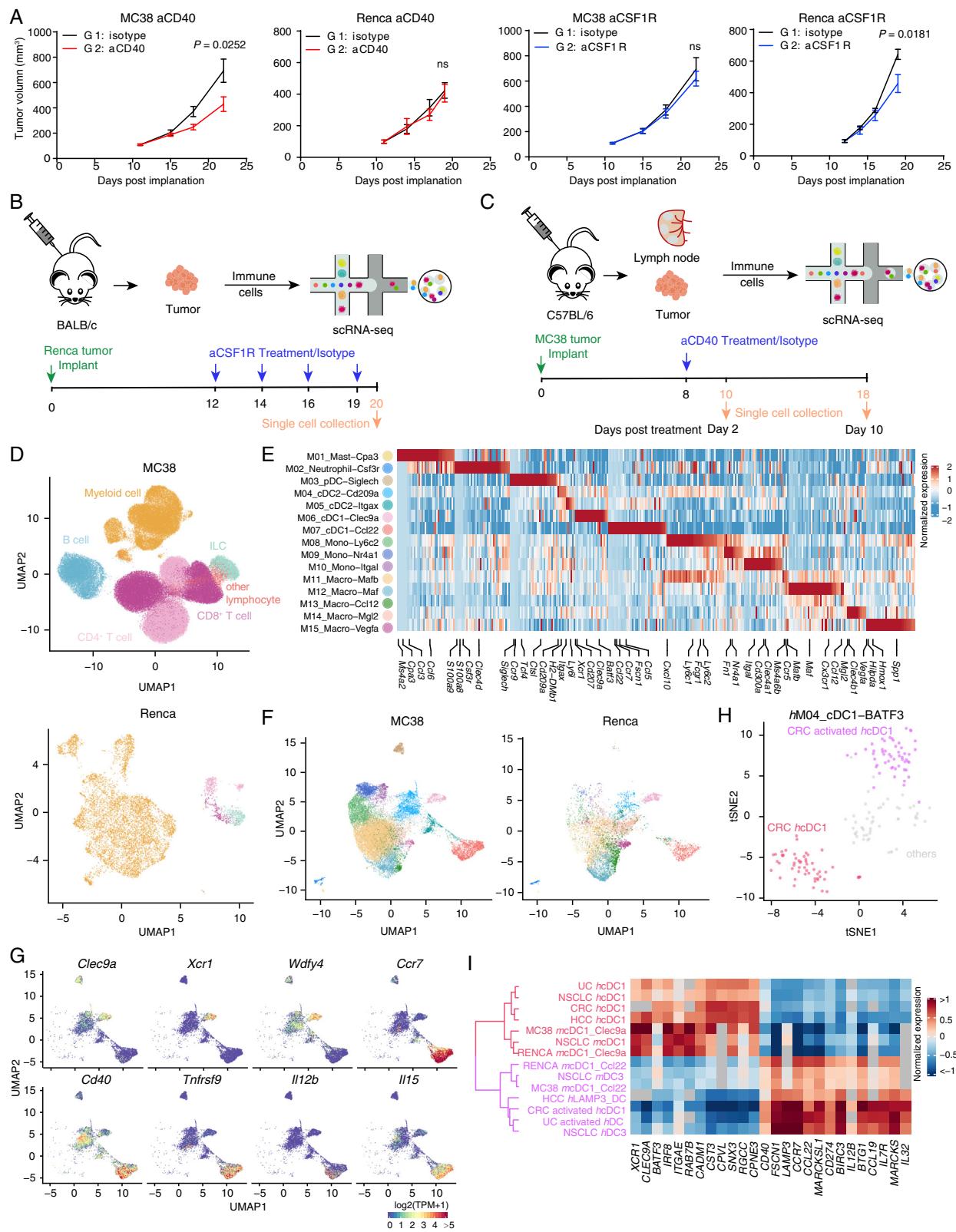
(A) Inferred correlative cell-cell interactions between *ACTA2*<sup>+</sup> myofibroblasts (left) / *FAP*<sup>+</sup> CAF (right) and *ACKR1*<sup>+</sup> endothelium cells based on Smart-seq2 scRNA-seq and TCGA bulk RNA-seq datasets. Specific gene signature for *ACKR1*<sup>+</sup> endothelium cells was subtracted from scRNA-seq and inferred to TCGA COAD and READ dataset as abundance, thereby to identify the correlation of each gene with such abundance across bulk tumors (y axis). The expression of these correlated genes in *ACTA2*<sup>+</sup> myofibroblasts (left) and *FAP*<sup>+</sup> CAF (right) versus *ACKR1*<sup>+</sup> endothelium cells (x axis). Dots in red indicate genes that are specific for *ACTA2*<sup>+</sup> myofibroblasts / *FAP*<sup>+</sup> CAF in the scRNA-seq dataset and also highly associated with the abundance of *ACKR1*<sup>+</sup> endothelium cells in the TCGA dataset.

(B) Heatmap showing enrichment scores for genes highly correlated with the abundance of *C1QC*<sup>+</sup> TAM, *SPP1*<sup>+</sup> TAM and *ACKR1*<sup>+</sup> endothelium cells across TCGA bulk tumors (column), enriched in different immune and non-immune cell clusters (row; [STAR Methods](#)).

(C) Inferred correlative cell-cell interaction network in the normal mucosa calculated based on Smart-seq2 scRNA-seq and GTEx bulk RNA-seq datasets ([STAR Methods](#)). Nodes indicate clusters enriched in normal mucosa, colored according to major cluster origins. Circle size indicates the number of connections with other cell types ([STAR Methods](#)).

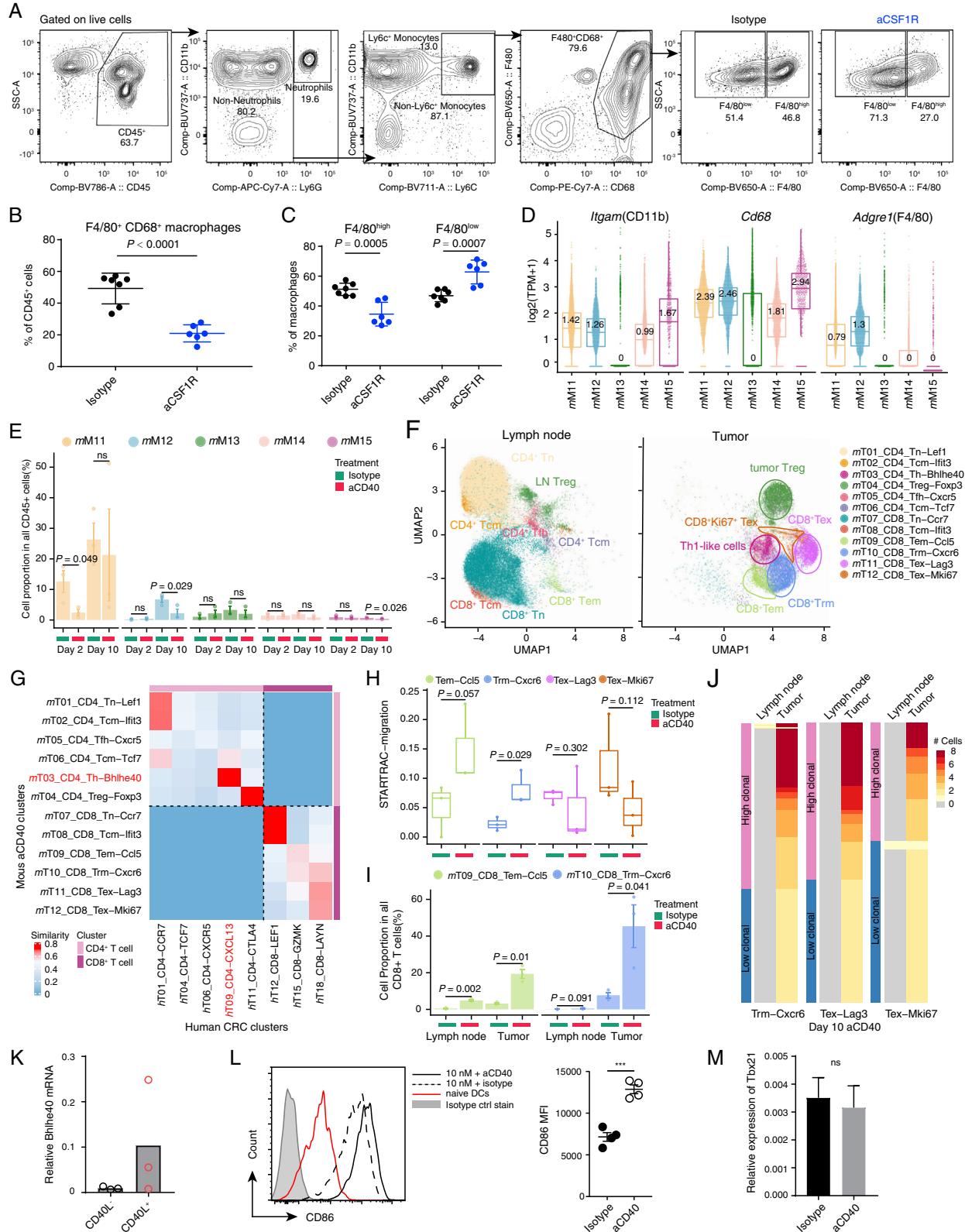
(D) Heatmap showing the expression of ligands and receptors in indicated clusters in human CRC.

(E-F) Chord diagrams showing predicted interactions by *CXCL10-CXCR3*, *SPP1-ITGAV*, *SDC2-MMP2*, and *FN1-ITGA5* in human CRC.



(legend on next page)

- 
- Figure S6. Characteristics of Myeloid Subsets from Mice Following Treatment with CSF1R Blockade or CD40 Agonist, Related to Figure 3**
- (A) Differential responses of MC38 and Renca syngeneic tumor models to treatment with anti-CD40 agonist antibody and anti-CSF1R antibody. Anti-CSF1R dosing of mice bearing Renca and MC38 tumors and anti-CD40 dosing of mice bearing MC38 tumors were performed as described in [STAR Methods](#). Renca tumor bearing mice were treated with anti-CD40 antibody less frequently (100 $\mu$ g of antibody i.p. on days 12 and 19 following tumor implantation) due to anti-CD40 toxicity observed in BALB/c mice. Two-sided t test.
- (B-C) Experimental design for scRNA-seq analysis of immune cells from (B) mice bearing Renca tumors treated with isotype control or anti-CSF1R blocking antibody or (C) mice bearing MC38 tumors treated with isotype control or anti-CD40 agonist antibody.
- (D) UMAP plots showing major immune cell subsets identified in MC38 (upper) and Renca (lower) tumor models.
- (E) Gene expression heatmap of integrated myeloid cell clusters from MC38 and Renca tumor models.
- (F) UMAP plots showing integrated myeloid cell clusters from MC38 (left) and Renca (right) tumors.
- (G) UMAP plots showing expression levels of selected genes in DC clusters from MC38 and Renca tumor models.
- (H) t-SNE plot showing sub-clusters of hM04\_cDC1-BATF3 cells after re-clustering analysis of these cells identified by Smart-seq2 and 10 $\times$  scRNA-seq in human CRC.
- (I) Heatmap showing expression of selected genes across cDC1 and activated DC cells identified from different human and mouse scRNA-seq datasets.  
\*\*\*, p < 0.001.



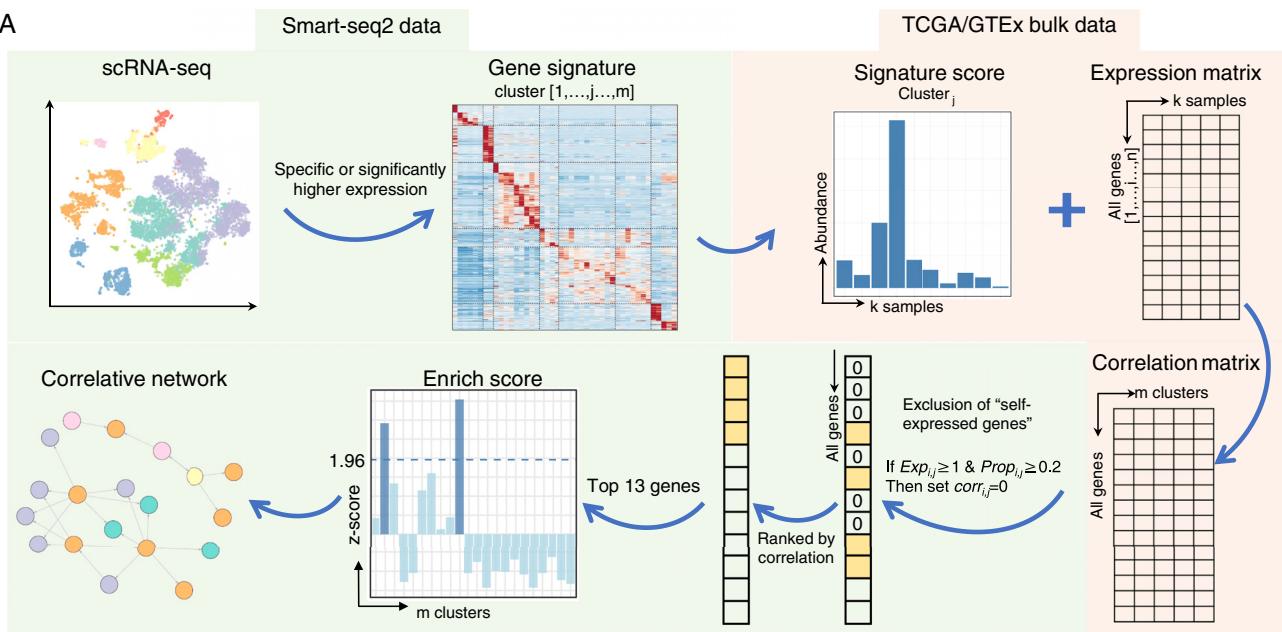
*(legend on next page)*

---

**Figure S7. Alterations of Immune Cell Populations from Mice Following Treatment with CSF1R Blockade or CD40 Agonist, Related to Figures 4–7**

- (A) Alternative gating scheme to examine differential effects of anti-CSF1R treatment on heterogeneous TAM populations following treatment with anti-CSF1R antibody. Two-sided t test.
- (B) Quantification of F4/80<sup>+</sup> CD68<sup>+</sup> macrophages based on the alternative gating strategy from (B) in Renca tumors treated with isotype control or anti-CSF1R antibody. Two-sided t test.
- (C) Preferential depletion of F4/80<sup>high</sup> TAMs in Renca tumors from mice treated with anti-CSF1R antibody.
- (D) Expression levels of *ITGAM* (encoding CD11b), *CD68*, and *ADGRE1* (encoding F4/80) in macrophage populations identified by scRNA-Seq in Renca tumors.
- (E) Changes in the frequency of TAM clusters as a proportion of all CD45<sup>+</sup> cells on days 2 and 10 following anti-CD40 treatment. One-sided t test.
- (F) UMAP plots showing T cell clusters in MC38 tumors and tumor-draining lymph nodes.
- (G) Similarity of T cells from mouse MC38 clusters and human CRC clusters (Smart-seq2 scRNA-seq).
- (H) Migration potential of CD8<sup>+</sup> T cell clusters quantified by STARTRAC-migr indices for each mouse treated with isotype control or anti-CD40 agonist antibody on day 10. One-sided t test.
- (I) Changes in the frequency of cells within different CD8<sup>+</sup> T cell clusters from lymph nodes and tumors as a proportion of total CD8<sup>+</sup> T cells following 10 days of anti-CD40 treatment. One-sided t test.
- (J) Distribution of clonal clonotypes of mT08, mT09, and mT10 clusters in lymph nodes and tumors, related to Figure 6H. Each row represents an individual clonotype. “High clonal” and “low clonal” clonotypes were defined as those in Figure 6H.
- (K) Preferential expression of *Bhlhe40* by sorted CD4<sup>+</sup> CD40L<sup>+</sup> T cells compared to sorted CD4<sup>+</sup> CD40L<sup>-</sup> T cells, as measured by RT-PCR.
- (L) Increase in the expression of CD86 on ova peptide pulsed OTII splenic DCs after anti-CD40 agonist treatment. Two-sided t test. \*\*,  $p < 0.001$ .
- (M) Expression of *Tbx21* in T cells stimulated with splenic DCs pulsed with suboptimal concentrations of ova peptide is not potentiated by anti-CD40 agonist antibody treatment. Two-sided t test.

A



B

Human CRC myeloid cell cluster names	Functional properties	Representative genes		Tissue enrichment P N T
		TFs	Other genes	
<i>hM01_Mast-TPSAB1</i>	Mast	<i>GATA2, EPAS1, PBX1</i>	<i>TPSAB1, CPA3, TPSD1, MS4A2, KIT, SIGLEC6, SIGLEC8, CD22, CSF1, TIMP3</i>	+/- +++ +++
<i>hM02_pDC-LILRA4</i>	pDC	<i>IRF4, IRF7, IRF8, SP1B, SOX4</i>	<i>LILRA4, SLC15A4, PLD4, CDC50, IL3RA, LY9, SELL, GAS6</i>	++ ++ ++
<i>hM03_cDC2-CD1C</i>	cDC2	<i>HMGAA1, PFDN1, IRF4</i>	<i>CD1C, CLEC10A, FCGR2B, S1RPA, ADAM8, FCER1A, AXL, ADAM28, LY86, TIMM13, ARAF</i>	+/- +++ +++
<i>hM04_cDC1-BATF3</i>	cDC1	<i>BATF3, ID2, ETV3</i>	<i>XCR1, CLEC9A, PTDSS1, SCARB1, IL6ST, CD40, TNFRSF10B, IDO1, CST7, CLIC2, NET1, ANXA6</i>	+/- +++ +++
<i>hM05_Mono-CD14</i>	Classical monocyte	<i>CEBPD</i>	<i>FCN1, CD14, CD36, SELL, S100A8, S100A12, CLEC12A, MS4A6A, CXCL14</i>	++ +/- +/-
<i>hM06_Mono-CD16</i>	Nonclassical monocyte	<i>TCFL2, KLF3, IKZF1, FLI1</i>	<i>FCN1, FCGR3A, FCGR3B, LILRA/B1, CX3CR1, IFITM1, ICAM2, MTSS1, CDKN1C, CDH23, SLC44A2</i>	++ +/- +/-
<i>hM07_Mono-CD14CD16</i>	Intermediate monocyte	<i>CEBPD, ZFP36L2</i>	<i>FCN1, CD14, FCGR3A, CD300E, FAM65B</i>	++ +/- +/-
<i>hM08_Macro-NLRP3</i>	NLRP3 <sup>+</sup> RTM	<i>NFKB2</i>	<i>NLRP3, AQP9, METRNL, GPR132, ANPEP, PLAUR, EREG, VEGFA, PTGS2</i>	+/- +++ +++
<i>hM09_Macro-PLTP</i>	PLTP <sup>+</sup> RTM	<i>KLF2, KLF4, EGR1, CREM, HIF1A</i>	<i>PLTP, LYVE1, IL10, STAB1, SEPP1, SNX6, CTSL, THBS1, MARCO, CXCL3, CALR, CD163, CD36</i>	- +++ +/-
<i>hM10_Macro-IL1B</i>	IL1B <sup>+</sup> RTM	<i>HES1, REL, BTG2</i>	<i>IL1B, C1Qα, C1Qβ, C1QB, CCL3L3, HLA-DQA1/2, HLA-DQB1, GPR183, CD83, CLEC10A, MRC1</i>	+/- +++ ++
<i>hM11_Monolike-FCN1</i>	Monocyte-like cell	<i>CEBPD, MAFB</i>	<i>FCN1, HLA-DQA1, HLA-DQB1, HSPA1A, HSPA1B, CXCR4, FCGR2A, FCGR3A</i>	+/- + +++
<i>hM12_TAM-C1QC</i>	C1QC <sup>+</sup> TAM	<i>MAFB, ATF3, MAF, MEF2A, PA2G4</i>	<i>C1Qα/B/C, MERTK, FPR3, TREM2, MS4A4A, SLC20B1, NRP1, SLAMF8, FCGR1A, ENG, SIRPA</i>	- ++ +++
<i>hM13_TAM-SPP1</i>	SPP1 <sup>+</sup> TAM	<i>CEBPB</i>	<i>SPP1, IL1RN, OLR1, CXCL2, VEGFA, EREG, C15orf48, GPNMB, PHLDA1, AQP9, TNF3, NDRG1</i>	- +/- +++

C

Mouse myeloid cell cluster names	Functional properties	Representative genes		aCD40	aCSF1R	
		TFs	Other genes	D2	D10	R
<i>mm01_Mast-Cpa3</i>	Mast	<i>Gata2, Nfe2, Kit7</i>	<i>Cpa3, Ms4a2, Ifitm1, Hdc, Slc41a3, Csf1, Syt3, Hgf, Il6, Il4, Ccl3, Cd9, Cd69, Il18rap</i>	NA	NA	NA
<i>mm02_Neutrophil-Csf3r</i>	Neutrophil	<i>Ets2, Mxd1, Kit2, Egr1, Csrmp1</i>	<i>Csf3r, Cxcr2, Cd24a, Csar1, Trem1, Il1r2, S100a8, S100a9, Cxcl2, Il1rn, Clec4d, Clec4e, Upp1</i>	↑	NA	↑
<i>mm03_pDC-Siglech</i>	pDC	<i>Tcf1, Runx2, Bcl11a, Tsc22d1, If7</i>	<i>Siglech, Ly6d, Cox6a2, Smim5, Kit1, Rppgrip1, Lefty1, Upb1, Ccr9, Cd7, Sell, Cd164, Plip, P2ry14, Ptprs</i>	NA	NA	NA
<i>mm04_cDC2-Cd209a</i>	cDC2	<i>Bcl3, Blhhe40, Cbf2a13</i>	<i>Cd209a, Cd300a, Cd83, Itgb7, Vrk1, Napsa, Ms4a4c</i>	↓ ↓ →	↑ ↑	
<i>mm05_cDC2-Itgax</i>	cDC2	<i>Batf3, Cbf2a13, Pa2g4</i>	<i>Itgax, Cd300a, Adrbk2, Il4i1, Mcemp1, Tmem176a, Slc38a2, Ramp1</i>	NA	NA	NA
<i>mm06_cDC1-Clec9a</i>	cDC1	<i>Batf3, Cbf2a13, Pa2g4</i>	<i>Xcr1, Itgb7, Arsb, Ckb, Fgd2, Naga, Pak1, Rab7b, Wdfy4, Ppm1m, Cd24a, Flt3, Cd83, Slamf7, Slamf8</i>	↓	NA	↑
<i>mm07_cDC1-Ccl22</i>	cDC1	<i>Relb, Etv3, Batf3, Aebp2, Nfk2b</i>	<i>Ccl22, C05, Il15, Ccr7, Il15ra, Plxnc1, Prmp, Cd40, Birc2, Fscn1, Anxa3, Cacnb3, Nudt17, Soc52, Tspan3, Serpinb6b</i>	↑ ↑	NA	↑ ↑
<i>mm08_Mono-Ly6c2</i>	Monocyte	<i>Stat2, Mx1, Bach1, Bcl3</i>	<i>Ly6c2, Mgst1, Cd9, Smox, Ilf3, Ilf205, Ilf2, Isg20, Plaur, Tlr2, Cd14, Cd300lf, Tgm2, Cxcl10, Fn1, F13a1, Ccl2</i>	↓ ↓ ↓ ↓	↑ ↑	↑
<i>mm09_Mono-Nr4a1</i>	Monocyte	<i>Nr4a1, Rara, Kit13, Nfl3, Kit4, Bcl3</i>	<i>Cd300a, Il17ra, Trnsfrf21, Trnsfrf13, Cd302, Cd14, Ifit6, Gstm1, Idh1, Gsr, Adgre5, Fn1, F13a1, Anxa1, Ramp1, C3</i>	↓ ↓ ↑	NA	
<i>mm10_Mono-Itgal</i>	Monocyte	<i>Nr4a1, Pou2l2, Kit4, Ly1</i>	<i>Itgal, Ace, Cd300a, Ceacam1, Spn, Trnsfrf1b, Il17ra, Lrp1, Adgre4, Trem4, Ear2, Sltk10</i>	↑ ↑	NA	
<i>mm11_Macro-Mafb</i>	Macrophage	<i>Mafb, Cebpb</i>	<i>C1qa/b, Axl, Spin1, C3ar1, F11r, Car5, Ccr1, Fgr1, Fgr4, Ly6i, Cfb, Mpnp14, Clec5a, Ier3</i>	↓ ↓	↑	
<i>mm12_Macro-Maf</i>	Macrophage	<i>Maf</i>	<i>C1qa/b/c, Axl, Ccl12, Trem2, Tgfb1, Cd81, Cd72, Cd63, Abhd12, Adgre1, Ms4a7, Nlpcr, Olfml3, Tmem119, Hpgds</i>	NA ↓	↓ ↓	
<i>mm13_Macro-Ccl12</i>	Macrophage	<i>Ir7, Mafl</i>	<i>C1qa/b/c, Ccl12, Sdc3, Lgmn, H2-Aa</i>	↑ ↓	→	
<i>mm14_Macro-Mgl2</i>	Macrophage	<i>Kif4</i>	<i>Mgl2, Lrp1, Fn1, Plip, Axl, Ear2, Tnfp3, Birc5</i>	→ ↓	↓ ↓	
<i>mm15_Macro-Vegfa</i>	Macrophage	<i>Cebpb, Blhhe40, Atf4, Tgif1, Atf3</i>	<i>Spp1, Vegfa, Mmp12, Adam8, Cd274, Cd63, Thbs1, C3ar1, Il1rn, Clec4d, Emp1, Arg1, Ero1, Hlpda, Hmox1, Sgk1</i>	NA NA	↑ ↑	

(legend on next page)

---

**Figure S8. Analysis Framework for Correlative Cell-Cell Interaction Networks and Summary of Myeloid Cell Subsets in Human and Mouse Tumors, Related to Figure 3 and STAR Methods**

(A) Workflow illustrating correlative network analysis. Cell subtype specific genes were identified from the Smart-seq2 scRNA-seq dataset and were used to estimate the relative abundance in TCGA/GTEX bulk RNA-seq datasets. After filtering the self-expressed genes, high correlative genes were calculated. Finally, an enrichment analysis based on Smart-seq2 profile was performed to identify all correlated cell subtypes and build the correlative network ([STAR Methods](#)).

(B) Overview of myeloid cell cluster characteristics from human CRC. +++ indicates  $R_{o/e} > 2$ ; ++,  $1 < R_{o/e} \leq 2$ ; +,  $0.5 < R_{o/e} \leq 1$ ; +/-,  $0 < R_{o/e} \leq 0.5$ ; -,  $R_{o/e} = 0$ .

(C) Overview of myeloid cell cluster characteristics from MC38 and Renca murine tumors. R, Response to treatment.  $\uparrow\uparrow$  indicates  $P\text{-value} < 0.05$  and fold change (FC)  $> 1$ ;  $\uparrow$ ,  $P\text{-value} \geq 0.05$  and FC  $> 1.2$ ;  $\rightarrow$ ,  $P\text{-value} \geq 0.05$  and  $0.83 < \text{FC} \leq 1.2$ ;  $\downarrow$ ,  $P\text{-value} \geq 0.05$  and FC  $\leq 0.83$ ;  $\downarrow\downarrow$ ,  $P\text{-value} < 0.05$  and FC  $< 1$ ; NA, proportions in both isotype and treatment groups were less than 1%.