

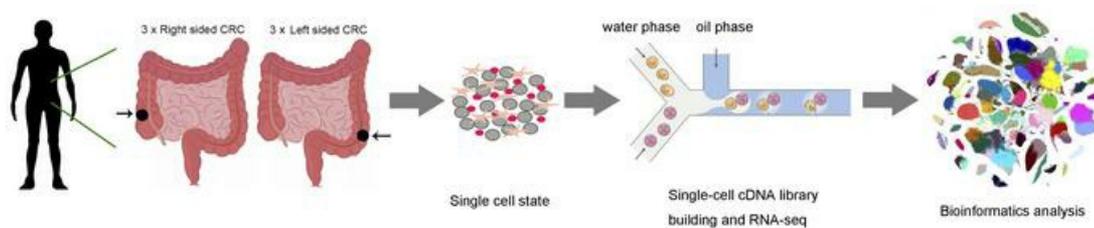
## Resolving the difference between left-sided and right-sided colorectal cancer by single-cell sequencing

Wei Guo, ... , Dawei Chen, Jingxin Li

*JCI Insight*. 2021. <https://doi.org/10.1172/jci.insight.152616>.

Research In-Press Preview Gastroenterology

### Graphical abstract



Find the latest version:

<https://jci.me/152616/pdf>



Resolving the Difference Between Left-sided and Right-sided Colorectal Cancer by Single-cell Sequencing.

Authors: Wei Guo,<sup>1</sup> Cuiyu Zhang,<sup>2</sup> Xia Wang,<sup>2</sup> Dandan Dou,<sup>2</sup> Dawei Chen,<sup>3</sup> \*Jingxin Li<sup>2</sup>

1. Department of Colorectal Surgery, Shandong University Qilu Hospital, Jinan, Shandong, 250012, China

2. Department of Physiology, School of Basic Medical Sciences, Cheeloo College of Medicine, Shandong University, Jinan, Shandong, 250012, China

3. Laboratory of Medical Chemistry, Interdisciplinary Cluster for Applied Genoproteomics (GIGA) Stem Cells, University of Liège, CHU, Sart-Tilman, 4000 Liège, Belgium

Correspondences:

Jingxin Li, Ph.D., Department of Physiology, School of Basic Medical Sciences, Cheeloo College of Medicine, Shandong University, Jinan, Shandong, 250012, China.

e-mail: [ljingxin@sdu.edu.cn](mailto:ljingxin@sdu.edu.cn); Telephone: +8618668983267; Fax: +8653188382502.

Running title: Difference between left-sided and right-sided CRC by scRNA-seq

Word count: 7745

Abstract count: 151

Authors' contributions:

WG, CZ, XW, and DC performed the experiments; WG and CZ analyzed data; WG provided the samples; WG wrote the paper; JL commented on the study and revised the paper; WG and JL designed the research.

Funding:

This work was supported by the National Natural Science Foundation of China (82000779, 31971061); and Taishan Pandeng Scholar Program of Shandong Province (tspd20210321).

**Abstract**

Colorectal cancers (CRCs) exhibit differences in incidence, pathogenesis, molecular pathways and outcome depending on the location of the tumor. The transcriptomes of 27,927 single human CRC cells, from three left-sided and three right-sided CRC patients were profiled by scRNA-seq. Right-sided CRC harbors a significant proportion of exhausted CD8 T cells of a highly migratory nature. One cluster of cells from left-sided CRC exhibiting states preceding exhaustion and a high ratio of “pre-exhausted” to exhausted T cells were favorable prognostic markers. Notably, we identified a novel RBP4<sup>+</sup> NTS<sup>+</sup> subpopulation of cancer cells that exclusively expands in left-sided CRC. Tregs from left-sided CRC showed higher levels of immunotherapy-related genes than those from right-sided CRC, indicating that left-sided CRC may have increased responsiveness to immunotherapy. Antibody-dependent cellular phagocytosis (ADCP) and antibody-dependent cellular cytotoxicity (ADCC) induced by M2-like macrophages were more pronounced in left-sided CRC and correlated with a good prognosis in CRC.

**Keywords:** Colorectal cancer; single-cell sequencing; exhausted CD8 T cells; macrophages.

## Introduction

Often grouped as one disease, right-sided colon cancer (originating from cecum, ascending colon, hepatic flexure) and left-sided colon cancer (originating from splenic flexure, descending colon, sigmoid colon) represent clinically distinct entities with significant differences in their prognosis and treatment outcomes(1, 2). Right-sided colon cancer has a worse prognosis than left-sided colon cancer and rectal cancer(3-5). Extensive sequencing analysis described a characteristic branching pattern of cancer evolution supporting that tumor biology is characterized simultaneously by intratumor heterogeneity and the preservation of ancestral aberrations within the primary tumor and corresponding metastatic sites(6, 7). However, the full spectrum of distinct cell types and their molecular characteristics remain to be well defined in left-sided and right-sided malignant colorectal lesions, which hampers our ability to investigate their differences in colorectal cancer (CRC) pathogenesis.

Advances in single-cell mRNA sequencing (scRNA-seq) have revolutionized our ability to characterize the transcriptional state of thousands of individual cells in-depth. Here, we performed a scRNA-seq survey of 27,927 cells from 6 samples obtained during curative surgery for three left-sided CRCs and three right-sided CRCs, and constructed a single-cell transcriptome atlas for malignant colorectal lesions. We hypothesized that the reasons for the better prognosis of left-side CRC than right-sided CRC might be the number and functional status of different immune cell subpopulations in the tumor microenvironment (TME) in CRC, as well as the level of different signaling pathways of cancer cells themselves and the interaction between cancer cells and tumor microenvironment cells. We used the atlas to construct a network for dissecting the cellular and molecular characteristics of left-sided and right-sided CRC.

## Results

### Single-Cell Atlas of CRC from Left-sided and Right-sided CRC Patients

We generated 27,927 high-quality single-cell transcriptomes from 6 samples obtained during curative surgery for three left-sided and three right-sided CRCs (Figure 1A; Table S1). The quality control (QC) criteria are described in the Materials and Methods.

After performing unsupervised clustering and t-distributed stochastic neighbor (t-SNE) plot analysis (Figure 1B), cluster identities were determined according to the expression of established markers (Figure S1). We also noticed that several cell clusters were enriched in both left-sided and right-sided CRC (Figures 1C). 13,488 single cells originated from left-sided CRC, while 14,439 originated from right-sided CRC (Figure 1C). We utilized the differentially expressed gene (DEG) signatures and attributed clusters to their putative identities and hierarchical similarities (Figure 1D, Figure S1). Figure 2A shows selected DEGs in the form of a heatmap (Figure 2A) and feature plots (Figure 2B). Most cells from right-sided CRC accounted for 71.5% of all cluster 13 cells, while left-sided CRC only accounted for 28.5% (Figure 2C). The proportion of each sample in these clusters was showed in Figure 2D.

These cells were classified into 19 main cell lineages (the last 2 unknown clusters were

incorporated into cluster 19). In addition to cancer cells, we identified 13 immune cell lineages, including B cells, CD4<sup>+</sup> T cells, CD8<sup>+</sup> T cells, regulatory T cells (Tregs), macrophages, neutrophils, natural killer (NK) T cells, dendritic cells (DCs), interleukin (IL)-17-producing helper T (Th17) cells and mast cells, along with 4 nonimmune cell lineages (CD45<sup>-</sup>), including stromal cells, fibroblasts, endothelial cells, and transit amplifying cells (Figure 1B).

### **Cell-Specific Expression Changes in Left-sided and Right-sided CRC**

Projecting the number of DEGs onto the t-SNE plot revealed that cluster 4 cancer cells exhibited the most prominent transcriptomic changes compared to other cell types in the TME of CRC (Figure 3A), which indicated that tumor cell population harbored the most essential transcriptomic differences between left-sided and right-sided CRC.

To understand the biological significance of transcriptional changes between left-sided and right-sided CRC, we performed pathway enrichment analysis with DEGs obtained via unsupervised clustering analysis (Figure 3B). Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis of the DEGs indicated that a handful of genes were associated with neutrophil function (neutrophil-mediated immunity, neutrophil activation, neutrophil activation involved in immune activation, neutrophil degranulation and granulocyte activation), T cell activation, cell adhesion molecule binding and adherens junctions (Figure 3C). Gene ontology (GO) enrichment analysis of the DEGs indicated that a handful of genes were associated with pathways in cancer (Figure 3D).

To identify changes in expression associated with the functional state of different cell types, we distinguished DEGs across cell subsets in CD4 T cells, CD8 T cells, Tregs, M1-like macrophages, 2 were expressed at higher levels in left-sided CRC. In M2-like macrophages, and fibroblasts in the form of volcano plots (Figure 3E).

### **Naive CD4 T Cells are Predominant in Right-sided CRC.**

CD4 T cells were then clustered into 8 subgroups (Figure 4A-B). We examined the specific genes expressed by each CD4 subgroup to identify their functional status. CD4 cluster 4 (CD4-C4) was enriched for CCR7, a specific marker for naive CD4 T cells (Figure 4C). The trajectory was visualized as a t-SNE plot. We noticed that CD4-C4 was present at the start of the differentiation trajectory (Figure 4D). To better understand the trajectories, we defined scores based on previously defined gene signatures(8), and we found that component 1 was highly associated with T cell naiveness, and CD4-C4 exhibited the highest naiveness score (Figure 4E). CD4-C6, representing ~9.7% of all CD4 T cells, was characterized by high expression of PRDM1, suggesting that CD4-C6 was most likely tissue resident memory (Trm) CD4 T cells (Figure S2).

It was reported that tumor-infiltrating Tregs develop not only from recruited Tregs but also from naive T cells in situ in human breast cancer. The abundance of naive CD4<sup>+</sup> T cells and Tregs is closely correlated, and both indicated poor prognosis for breast cancer patients(9). Our data showed that the number of naive CD4 T cells from right-sided CRC was 9-fold

higher than that of those from left-sided CRC. Our CRC data were in line with the above observations from breast cancer. Kaplan-Meier survival curves of overall survival (OS) based on UBE2S (CD4-C3 marker) and FAM177A1 (CD4-C4 marker) expression indicated a poor prognosis (Figure 4F). The similarity network between CD4 T cells and other cell types in our dataset is shown in Figure 4G.

### **Right-sided CRC occupies a Large Proportion of Highly Migratory Exhausted CD8 T Cells**

We applied unsupervised clustering based on t-SNE and identified seven CD8 T cell clusters (Figure 5A-B). Next, we examined the expression of T cell-associated signature genes and known functional markers to define their identities.

CD4-C4 was characterized by high expression of genes associated with naiveness, including CCR7, SELL, LEF1 and TCF7 (Figure 5C). We found that some clusters exhibited distinct expression patterns among all CD8+ T cells in our dataset. Cluster 4 was enriched for the effector T cell marker GNLY and the cytotoxicity-associated gene GZMB. Cluster 4 was characterized by high expression of genes associated with cytotoxicity, including GNLY, PRF1, GZMA and GZMB, also showing high expression of T cell exhaustion markers such as PDCD1, LAG3 and HAVCR2. These data suggested that cluster 4 was most likely exhausted CD8 T cells (Figure 5C). Interestingly, we observed that some cluster of CD8 T cells exhibited states preceding exhaustion. Cluster 6 represented ~10% of all CD8 T cells, and its specific markers included genes associated with cytotoxicity, such as GZMH and GZMK, and chemokines, such as CCL3L3 and CCL4L2. Cluster 6 was characterized by high expression of genes associated with cytotoxicity, including PRF1, GZMA, GZMB, GZMK, IFNG and NKG7 but had low expression of T cell exhaustion markers such as PDCD1, LAG3, TIGIT, CTLA4 and HAVCR2 (Figure 5C).

Right-sided-CRC-derived exhausted CD8 T cells accounted for as many as 67.9% of all CD8-C4 cells, while left-sided-CRC-originated pre-exhausted effector CD8 T cells accounted for 86.8% of all CD8-C6 cells. The ratio of pre-exhausted to exhausted T cells in left-sided CRC was 13.8-fold higher than right-sided CRC (Figure 5D). It has been reported that a high ratio of pre-exhausted to exhausted T cells is associated with a better prognosis than a low ratio in lung adenocarcinoma(8).

CD8-C1 and CD8-C2 were characterized by high expression of PRDM1 and CD69, suggesting these two subgroups were most likely Trm CD8 T cells (Figure 5D, Figure S2).

We applied an unsupervised inference method Monocle to construct the potential developmental trajectories of seven CD8 clusters. CD8-C4 exhaustion cluster and CD8-C6 pre-exhaustion cluster positioned at different ends of the developmental trajectory. A part of the CD8-C2 PRDM1+ Trm CD8 T cell cluster was positioned at the start of the developmental trajectory (Figure 5E). To better understand the trajectories, we defined cytotoxicity scores based on previously defined gene signatures(10, 11), and T cell exhaustion scores based on the average expression of 90 genes highly expressed in

tumor-infiltrating exhausted CD8+ T cells(8). We analyzed the Monocle trajectory in the context of these functional scores, and we found that component 1 was highly associated with T cell exhaustion and cytotoxicity (Figure 5F).

It was reported that the presence of highly migratory pre-exhausted effector T cells in tumors provides a plausible explanation for the positive response to immunotherapies for non-small-cell lung cancer (NSCLC) patients(8). It was also reported that increased pathogen-specific T cell numbers together with altered migratory patterns can greatly improve immune efficacy(12). Our data showed that CD8-C4 exhausted CD8 T cells from right-sided CRC showed higher levels of components of the focal adhesion, leukocyte transendothelial migration and regulation of the actin cytoskeleton pathways than those from left-sided CRC (Figure 5G). T cell exhaustion is one of the mechanisms by which cancer cells evade the immune system. We concluded that exhausted CD8 T cells from right-sided CRC were more prone to migrate to organs outside the TME, including lymph nodes, liver and lung, leading to a higher tendency of metastasis than left-sided CRC.

The T cell signatures for the co-inhibition program (CTLA4, PDCD1, TIGIT, HAVCR2, LAG3, BTLA, PDPN, CD160, GP49A, LILRB4, CD274, CD200, CD244, PILRA, SIRPB1, LAIR1, CEACAM1, KLRA7, KLRA3, KLRA9, PTGER4, KLRD1, KLRC1, and PROCR) were derived from known markers(8). We observed that most subsets of CD8 T cells induced co-inhibitory programs, and CD8-C1 CD69+ Trm cells, CD8-C4 exhausted CD8 T cells and CD8-C5 FOXP3+ CD8 T cells from right-sided CRC induced stronger co-inhibitory programs than those from left-sided CRC (Figure 5G).

### **Tregs from Left-sided CRC Exhibit Higher Level of Immunotherapy-related Genes**

Tregs suppress the antitumor function of effector T cells and natural killer cells by secreting soluble immunosuppressive factors and expressing inhibitory receptors(13, 14). A high proportion of Tregs in tumor-infiltrating T cells is associated with a poor prognosis in various types of human cancers(15). The prevailing idea is that Tregs are recruited from preexisting circulating Tregs by chemokines or chemokine ligands expressed by tumor cells, stroma or tumor-associated macrophages (TAMs)(16, 17). An alternative possibility is that naive or conventional T cells might be recruited to the tumor and differentiate into Tregs in situ within the immunosuppressive tumor environment(13, 18, 19).

Treg-C1 represented ~20% of all Treg cells, and the specific markers included genes associated with chemokines or chemokine ligands, such as CCL13, CCR7 and CXCR4 (Figure 6A, Figure S2). This indicated that Treg-C1 cells were highly likely recruited from preexisting circulating Tregs by tumor cells, stroma or tumor-associated macrophages (TAMs) from the TME. Treg-C3 was enriched for PRDM1, a specific marker for Trm T cells. We found that Treg-C3 cells shared similar marker gene signatures with CD4-C6 cells (Figure S2) and Trm CD4 T cells, leading us to wonder whether this subgroup of intratumoral Tregs mainly develops from PRDM1+ Trm CD4 T cells. We analyzed the Monocle trajectory for 8 clusters of CD4 T cells and 7 clusters of Tregs, and the results showed that some Treg-C3 cells shared the same position as CD4-C6 cells. Our data suggest that Trm CD4 T cells might differentiate

into Treg-C3 cells in situ within the immunosuppressive tumor environment (Figure 6B).

The number of Treg-C1 cells from right-sided CRC was 9-fold higher than that of those from left-sided CRC. Similarly, the number of Treg-C3 cells from right-sided CRC was 3-fold higher than that of those from left-sided CRC (Figure 6C-D).

We analyzed the Monocle trajectory for CD4 T cells, CD8 T cells and Tregs in the context of functional scores, and we found that component 2 was highly associated with T cell exhaustion, whereas component 1 was positively associated with cytotoxicity (Figure 6E). CD8 T cells showed the highest cytotoxicity level, and Tregs showed the highest T cell exhaustion level.

PDCD1 exhibited high expression in exhausted CD8<sup>+</sup> T cells (CD8-C4) and suppressive tumor Tregs highly expressed CTLA4. This is in line with recent observations that anti-CTLA4 and anti-PD1 therapies target distinct tumor infiltrating lymphocytes (TIL) populations to induce tumor rejection(20). LAG3 was almost exclusively expressed by CD8<sup>+</sup> T cells. HAVCR2 was highly expressed by both suppressive tumor Tregs and exhausted CD8<sup>+</sup> T cells (Figure 6F).

Our data showed that Tregs highly expressed immunotherapy-related genes and that Tregs from left-sided CRC showed higher expression levels than those from right-sided CRC, indicating that left-sided CRC may be more responsive to immunotherapies that function by inducing dysfunctional Tregs (Figure 6G-H). Pathway analysis by gene set variation analysis (GSVA) was showed in Figure 6I.

Survival analysis of the The Cancer Genome Atlas (TCGA) dataset from Kaplan-Meier Plotter (<http://kmplot.com/>) showed that KLF2, a specific marker of Treg-C1, was an unfavorable prognostic marker (Figure 6J). DUSP1, the specific marker of Treg-C2, and RANBP1, the specific marker of Treg-C4, were favorable prognostic markers (Figure 6J). The similarity network between Treg cells and other cell types in our dataset is shown in Figure 6K.

### **The RBP4<sup>+</sup> NTS<sup>+</sup> Cancer Cell Subset Is Unique to Left-sided CRC**

The CRC cells were further divided into 9 subgroups based on t-SNE analysis (Figure 7A). Subpopulation markers were identified across all clusters and lineages, and the top five markers of the main cell lineages were visualized as a bubble chart (Figure 7B). Most subgroups originated from both left-sided and right-sided CRC (Figure 7C).

Strikingly, subgroup 5 was exclusively observed in left-sided CRC patients, which reflects the high tumor heterogeneity between left-sided and right-sided CRC patients (Figure 7D). Subgroup 5 was enriched for the expression of many genes, such as RBP4, NTS, TFF2, REG4, TFF1, SPINK4, GPRC5A, AGR2, AREG, and TFF3 (top 10 specific genes in subgroup 5, Figure 7B). In addition, AGR3 and MUC5AC were also specific markers for subgroup 5 (Figure S3). TFF1, TFF2 and MUC5AC are closely associated with protecting the mucosa from insults by stabilizing the mucus layer, thus promoting the healing process of the colorectal

epithelium(21, 22).

Survival analysis of the TCGA dataset showed that AGR3, MUC5AC, NTS and SPINK4 were favorable prognostic markers (Figure 7E). Moreover, AGR2, AGR3, TFF1, TFF2, MUC5AC and SPINK4 were expressed at higher levels in left-sided than in right-sided CRC, which was verified by IHC (Figure 7F). Survival analysis of GOLGB1, which was specifically expressed in cluster 6, was also favorable prognostic marker. Reduced GOLGB1 expression has been reported to promote the progression of prostate cancer(23). Survival analysis of CLCA1, OLFM4 and PIGR, which were specifically expressed in cluster 8, indicated that they were favorable prognostic markers. It has been reported that increased expression of CLCA1 can suppress CRC aggressiveness(24). Survival analysis of HSPA1A, which was specifically expressed in cluster 4, indicated that it was a poor prognostic factor.

Mapping changes in KEGG pathways during CRC revealed distinctive metabolic characteristics in 9 subgroups of cancer cells. Strikingly, cancer cell subgroup 5, which was exclusively originated from left-sided CRC, presented upregulation of several cancer-associated signaling pathways, including estrogen signaling, ErbB signaling, TNF signaling, HIF-1 signaling and AMPK signaling. The expression of estrogen receptor beta has an inverse relationship with the stage of CRC, and can mediate a protective response by promoting apoptosis(25). Anti-HER2 therapy may have a potentially beneficial role in the treatment of HER2-positive metastatic CRC(26). Upregulation of ErbB signaling indicated the subgroup 5 and left-sided CRC patients may have satisfactory responses to anti-HER2 therapy. Upregulation of TNF signaling in subgroup 5 suggested a good prognosis for left-sided CRC (Figure 8A).

Furthermore, subgroup 5 showed upregulation of the cell death pathway, including apoptosis, necroptosis, autophagy and mitophagy, indicating a good prognosis for left-sided CRC (Figure 8A). The evasion of controlled cell death induction is considered one of the hallmarks of cancer cells(27). Subgroup 5 showed upregulation of lipid metabolism, amino acid metabolism and oxidative phosphorylation (Figure 8A).

Finally, we utilized CellPhoneDB to investigate the interactions between cancer cells and cell subgroups in the TME. As shown in Figure 8B, several collagen-encoding genes secreted by fibroblasts interact with the receptor ( $\alpha 1\beta 1$  complex) expressed on cancer cells, and these ligand-receptor pairs were dramatically upregulated in left-sided CRC compared to right-sided CRC. TNF and TNFSF10 (TRAIL) secreted by M1-like and M2-like macrophages interact with their receptors expressed on cancer cells. These ligand-receptor pairs associated with the TNF signaling pathway were dramatically upregulated in left-sided CRC. TRAIL is a potent anticancer agent owing to its specific targeting of cancerous cells to induce apoptosis while sparing normal cells(28). TGF $\beta$ 1 secreted by the TME interacts with TGF $\beta$  receptor 1 expressed on cancer cells, and this ligand-receptor pair is dramatically upregulated in left-sided CRC. On the other hand, TGF $\beta$ 1 secreted by the TME interacts with TGF $\beta$  receptor 1 expressed on endothelial cells, and this ligand-receptor pair is dramatically upregulated in right-sided CRC. TGF $\beta$  expressed in the colon plays important roles as a

tumor suppressor during colorectal carcinogenesis, while TGF $\beta$  expressed on endothelial cells promotes angiogenesis in CRC(29).

The inferred developmental trajectory suggested a branched structure (Figure S4), with cancer cell-C5 positioned at the beginning of the developmental trajectory, suggesting a possible naive state.

Our data showed that left-sided CRC cancer cells exhibited significantly stronger EGFR signaling, VEGF signaling and ErbB signaling than right-sided CRC cells (Figure 8C), which is consistent with our clinical findings that left-sided CRC is more sensitive to monoclonal antibodies against for EGFR, VEGF and ErbB. The network between cancer cells and other cell types is shown in Figure 8D.

### **Heterogeneity of Macrophages in the TME of CRC**

Two transcriptionally distinct macrophage clusters were revealed. Both clusters expressed macrophage-specific markers CD68 and CD14 (Figure S1). In virtually all tissues, resident macrophages and recruited macrophages exist during acute inflammation and carcinogenesis(30–32). We first examined the expression of MRC1 and CD14, two commonly used markers to distinguish resident macrophages and recruited macrophages. The expression of both genes was observed in both cell clusters but was greater in cluster 6, which was the M2-like macrophage (Figure 9A-C).

We found that M1 marker genes were significantly upregulated in cluster 8 compared with cluster 6. In comparison, M2 marker genes were significantly upregulated in cluster 6 compared with clusters 8. Furthermore, the mean expression across the panel of M1 markers was 2.43-fold higher in cluster 8, whereas the mean expression of the panel of M2 markers was 3.04-fold higher in cluster 6 (Figure 9A-B). This confirms that cluster 6 was M2-like macrophages and cluster 8 was M1-like macrophages.

Our data showed that mean M1 markers expression level of M1-like macrophages from left-sided CRC was 1.43-fold higher than that of right-sided CRC, and mean M2 markers expression level of M2-like macrophages from left-sided CRC was 2.03-fold higher than that of right-sided CRC. We also noticed that M1 markers expression level of M2-like macrophages from left-sided CRC was almost the same as that of M1-like macrophages from right-sided CRC, suggesting that the M2-like macrophage group was a dynamic transitional state of M2 conversion to M1. Our data confirm that the macrophage polarization state is a major determinant of TAM heterogeneity between left-sided and right-sided CRC (Figure 9C).

Pathway analysis of DEGs from cluster 6 and 8 identified pathways known to be important for several vital functions, such as cell death (ferroptosis and necroptosis), macrophage-associated cell function (oxidative phosphorylation, ribosome, spliceosome, proteasome, lysosome, phagosome and RNA degradation), cancer-associated signaling pathway (HIF-1, NOD-like receptor, IL17, TNF, NF-kb, oxytocin signaling) and glutathione

metabolism.

Antibody-dependent cellular phagocytosis (ADCP) represents a significant mechanism in antitumor activity mediated by activated macrophages. Macrophages can kill tumor cells extracellularly via Antibody-dependent cellular cytotoxicity (ADCC)(33). Our data showed that the lysosome and phagosome pathways of M2-like macrophages from left-sided CRC were stronger than those of right-sided CRC, suggesting stronger ADCP and ADCC function and a better prognosis in left-sided CRC (Figure 9D).

The expression of IL1A, IL1B, CCL3, PTGS2, CXCL2/3/8, and CCL3L3 in M1-like macrophages from left-sided CRC was higher than that in M1-like macrophages from right-sided CRC. This suggested higher levels of proinflammatory cytokines and chemokines in M1-like macrophages from left-sided CRC, which further proved the stronger antitumor effect of left-sided-CRC-derived M1 macrophages (Figure 9E).

The expression of PDK4, SLC40A1 and TSC22D3 in from right-sided CRC was higher than that in M2-like macrophages left-sided CRC (Figure 9E). PDK4 directly enhances cell proliferation, invasion and chemoresistance in ovarian cancer(34). High expression of SLC40A1 is positively correlated with tumor metastasis and invasion(35). Upregulation of TSC22D3 can subvert therapy-induced anticancer immunosurveillance(36).

Macrophages express ligands for checkpoint molecules, including PD-L1 (CD274) and PD-L2 (PDCD1LG2). Macrophages contribute to the immunosuppression observed in the tumor microenvironment, and macrophage targeting may complement the action of checkpoint blockade inhibitors(33). As shown in Figure 9F, compared to that in other clusters, the expression of immune checkpoints in M2-like macrophages from left-sided CRC was relatively higher. Since all of these immunosuppressive ligands could suppress cytotoxic T lymphocyte function, this evidence proved that the enhanced immunosuppressive properties of macrophages in left-sided CRC and indicated the likelihood of better immunotherapy responses in left-sided CRC.

## **Discussion**

Although the scRNA-seq profiles of intact gastrointestinal organs, including the esophagus, stomach, and colon, have been assessed(37-39), the profiles of CRC, particularly left-sided and right-sided CRC, have not been demonstrated. To our knowledge, this is the first study to define the difference between left-sided or right-sided CRC on the basis of a single-cell atlas. For each cancer, we identified diverse cell types and defined gene expression signatures for these cell types. We also analyzed the transcriptomic changes in some cell types across different lesions. In addition, exhausted CD8 T cells, macrophages and cancer cells, and cellular characteristics related to the responsiveness of left-sided right-sided CRC to checkpoint inhibitor therapy were analyzed in depth to identify cell type-specific markers that are potentially applicable in clinical practice.

Although immunotherapy has dramatically changed the landscape of treatment for many

advanced cancers, the benefit in CRC has thus far been limited to patients with microsatellite instability high (MSI-H); DNA mismatch repair-deficient (dMMR) tumors(40). In our study, it may provide some clues to address the issue left-sided CRC without harboring MSI-H:dMMR may also benefit from immunotherapy as a result of the high expression of checkpoint molecules on M2-like macrophages and T cells.

In summary, we constructed a single-cell transcriptome atlas of left-sided and right-sided CRC. With the atlas, we characterized the expression patterns of diverse cell types in each lesion and analyzed their changes across lesions. Of note, we identified a panel of left-sided CRC cell-specific marker genes, providing a molecular basis for precise prognosis prediction. Our findings provide multidimensional insight into the responsiveness of left-sided and right malignant lesions to checkpoint inhibitor therapy, which may be helpful for predicting the effectiveness of immunotherapy for left-sided and right-sided CRC and may facilitate our understanding of CRC pathogenesis and progression.

## **Methods**

### **Clinical sample collection and preparation.**

Six patients who were pathologically diagnosed with colorectal adenocarcinoma were enrolled in this study. None of the patients had autoimmune disorder or history of prior cancer. None of the patients was treated with chemotherapy, radiation or any other anti-tumor medicines prior to tumor resection. Available clinical characteristics of these patients are summarized in Supplementary Table S1. Clinical samples were collected from Qilu Hospital of Shandong University. This study was approved by Institutional Ethics Committee in Qilu hospital of Shandong University. All patients in this study provided written informed consent for sample collection and data analyses. The processed gene expression data can be accessed from Gene Expression Omnibus database (accession number GSE188711).

### **Single-cell sequencing.**

The protoplast suspension was loaded into Chromium microfluidic chips with 30 v3 chemistry and barcoded with a 10× Chromium Controller (10X Genomics). RNA from the barcoded cells was subsequently reverse-transcribed and sequencing libraries constructed with reagents from a Chromium Single Cell 30 v3 reagent kit (10X Genomics) according to the manufacturer's instructions. Sequencing was performed with Illumina (NovaSeq 6000) according to the manufacturer's instructions (Illumina).

### **Quality control.**

- (1) Remove low-quality reads: scan the read with a 4-base wide sliding window, cutting when the average quality per base drops below 10.
- (2) Remove trailing low quality or N bases (below quality 3).
- (3) Remove adapters.
- (4) Drop reads below the 26 bases long.
- (5) Discard those reads that cannot form paired.

## **Generation and Analysis of Single-Cell Transcriptomes**

Raw reads were demultiplexed and mapped to the reference genome by 10X Genomics Cell Ranger pipeline (<https://support.10xgenomics.com/single-cell-geneexpression/software/pipelines/latest/what-is-cell-ranger>) using default parameters. All downstream single-cell analyses were performed using Cell Ranger and Seurat unless mentioned specifically.

### **Cellranger**

Cellranger reanalyze takes feature-barcode matrices produced by cellranger count or cellranger aggr and reruns the dimensionality reduction, clustering, and gene expression algorithms using cellranger default parameter settings.

### **Seurat**

The Seurat package was used to normalize data, dimensionality reduction, clustering, differential expression. We used Seurat alignment method canonical correlation analysis (CCA)(41) for integrated analysis of datasets. For clustering, highly variable genes were selected and the principal components based on those genes used to build a graph, which was segmented with a resolution of 0.6.

### **Enrichment analysis of marker genes**

Gene Ontology (GO) enrichment analysis of marker genes was implemented by the clusterProfiler R package. We used clusterProfiler R package to test the statistical enrichment of marker genes in KEGG pathways.

### **Definition of exhaustion, naiveness and cytotoxicity scores.**

For exhaustion scores, we first used 90 well-defined T cell exhaustion markers to define the exhaustion score for CD4<sup>+</sup> T cells, CD8<sup>+</sup> T cells and Treg cells after z-score transformation. Similarly, we used the average expression (after z-score transformation) of 4 well-defined naive markers (CCR7, TCF7, LEF1 and SELL) and 12 cytotoxicity associated genes (PRF1, IFNG, GNLY, NKG7, GZMB, GZMA, GZMH, KLRK1, KLRB1, KLRD1, CTSW, CST7) to define the naiveness score and cytotoxicity score for both CD8<sup>+</sup> and CD4<sup>+</sup> T cells, respectively. After delineating the exhaustion, naiveness and cytotoxicity scores of each T cell along the trajectory, we used locally weighted scatterplot smoothing (LOESS) regression to fit the relationships between these scores with Monocle components.

### **Developmental trajectory inference.**

We applied the Monocle (version 2) algorithm with the genes of cell clusters as input to determine the potential lineage differentiation between diverse cell populations. The Monocle function relative2abs was used to convert TPM measurement into messenger RNA cell counts per cell values, and then a CellDataSet object was created with the parameter "expressionFamily = negbinomial". Then the cell differentiation trajectory was inferred with the default parameters of Monocle after dimension reduction and cell ordering.

### **Immunocytochemistry**

For IHC staining, sections were routinely dewaxed and hydrated, then treated with 3% H<sub>2</sub>O<sub>2</sub>

for 10 min to block endogenous peroxidase activities and followed by an overnight incubation at 4 °C with indicated antibodies. Slides were then washed in PBS twice and incubated with goat anti - rabbit/mouse horseradish peroxidase - conjugated secondary antibodies for 30 min at room temperature. Finally, slides were washed and incubated with 3,3' - diami - nobenzidine and counter stained with hematoxylin. The slides were analyzed separately by two pathologists without knowing the patients' clinical information. Anti-AGR2 antibody (ab209224, Abcam) and anti-SPINK4 antibody (ab121257, Abcam) were purchased from Abcam (Cambridge, UK). Anti-AGR3 antibody (PA5-27222, ThermoFisher Scientific) and anti-MUC5AC antibody (PA5-79705, ThermoFisher Scientific) were purchased from ThermoFisher Scientific (Massachusetts, USA). Anti-TFF1 antibody (13734-1-AP, Proteintech) and anti-TFF2 antibody (13681-1-AP, Proteintech) were purchased from Proteintech (Illinois, USA).

### **Pathway analysis.**

The R package limma was used to identify differentially expressed genes with normalized read counts as input and donor as an additional covariate. Genes with Benjamini–Hochberg-adjusted P value < 0.01 and the absolute log<sub>2</sub> (fold change) (log<sub>2</sub> FC) between two groups larger than 1 were used for DAVID (<https://david-d.ncicrf.gov/>) pathway enrichment analysis. In addition, preranked gene set enrichment analysis (GSEA) was also performed, using a Python implementation (package gseapy) was also used for gene set enrichment analysis.

### **Cell-Cell Communication Analysis**

CellPhoneDB is a Python-based computational analysis tool; it enables the analysis of cell-cell communication at the molecular level. A website version was also provided for the analysis of relatively small datasets (<https://www.cellphonedb.org/>). In order to investigate the molecular interaction networks among the cell types or cell clusters, CellPhoneDB was used to analyze major cell types and cell subclusters. Ligand-receptor pairs with p values >0.05, as determined by CellPhoneDB, were filtered, while the others were retained for evaluating the relationship between the different cell clusters.

### **Sduty approval**

The use of human colorectal cancer tissues was approved by Institutional Ethics Committee in Qilu hospital of Shandong University (KYLL-202011-209-01). The study was conducted according to the principles expressed in the Declaration of Helsinki.

### **Authors' contributions**

WG, CZ, XW, and DC performed the experiments; WG and CZ analyzed data; WG provided the samples; WG wrote the paper; JL commented on the study and revised the paper; WG and JL designed the research.

### **Acknowledgements**

Not applicable

## Funding

This work was supported by the National Natural Science Foundation of China (82000779, 31971061); and Taishan Pandeng Scholar Program of Shandong Province (tspd20210321).

## Availability of data and materials

The processed gene expression data can be accessed from Gene Expression Omnibus database (accession number GSE188711).

## Consent for publication

Not required.

## Competing interests

The authors declare that they have no competing interests.

## References

1. Venook AP, Niedzwiecki D, Lenz HJ, Innocenti F, Mahoney MR, O'Neil BH, Shaw JE, Polite BN, Hochster HS, Atkins JN, et al. CALGB/SWOG 80405: Phase III trial of irinotecan/5-FU/leucovorin (FOLFIRI) or oxaliplatin/5-FU/leucovorin (mFOLFOX6) with bevacizumab (RV) or cetuximab (CET) for patients (pts) with KRAS wild-type (wt) untreated metastatic adenocarcinoma of the colon or rectum (MCRC). *J Clin Oncol*. 2014;32(15).
2. Hong TS, Clark JW, and Haigis KM. Cancers of the Colon and Rectum: Identical or Fraternal Twins? *Cancer Discov*. 2012;2(2):117-21.
3. Bufill JA. Colorectal-Cancer - Evidence for Distinct Genetic Categories Based on Proximal Or Distal Tumor Location. *Ann Intern Med*. 1990;113(10):779-88.
4. Distler P, and Holt PR. Are right- and left-sided colon neoplasms distinct tumors? *Digest Dis*. 1997;15(4-5):302-11.
5. Hutchins G, Southward K, Handley K, Magill L, Beaumont C, Stahlschmidt J, Richman S, Chambers P, Seymour M, Kerr D, et al. Value of Mismatch Repair, KRAS, and BRAF Mutations in Predicting Recurrence and Benefits From Chemotherapy in Colorectal Cancer. *J Clin Oncol*. 2011;29(10):1261-70.
6. Birkenkamp-Demtroder K, Olesen SH, Sorensen FB, Laurberg S, Laiho P, Aaltonen LA, and Orntoft TF. Differential gene expression in colon cancer of the caecum versus the sigmoid and rectosigmoid. *Gut*. 2005;54(3):374-84.
7. Gerlinger M, Rowan AJ, Horswell S, Larkin J, Endesfelder D, Gronroos E, Martinez P, Matthews N, Stewart A, Tarpey P, et al. Intratumor Heterogeneity and Branched Evolution Revealed by Multiregion Sequencing. *New Engl J Med*. 2012;366(10):883-92.
8. Guo XY, Zhang YY, Zheng LT, and Zhang ZM. Global characterization of T cells in non-small-cell lung cancer by single-cell sequencing. *Cancer Sci*. 2018;109(6):623-.
9. Su SC, Liao JY, Liu J, Huang D, He CH, Chen F, Yang LB, Wu W, Chen JN, Lin L, et al. Blocking the recruitment of naive CD4(+) T cells reverses immunosuppression in breast cancer. *Cell Res*. 2017;27(4):461-82.
10. Danaher P, Warren S, Dennis L, D'Amico L, White A, Disis ML, Geller MA, Odunsi K, Beechem J, and Fling SP. Gene expression markers of Tumor Infiltrating Leukocytes. *J*

- Immunother Cancer*. 2017;5(
11. Tirosh I, Izar B, Prakadan SM, Wadsworth MH, Treacy D, Trombetta JJ, Rotem A, Rodman C, Lian C, Murphy G, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science*. 2016;352(6282):189-96.
  12. Mueller SN, Gebhardt T, Carbone FR, and Heath WR. Memory T Cell Subsets, Migration Patterns, and Tissue Residence. *Annu Rev Immunol*. 2013;31(137-61).
  13. Zitvogel L, Tanchot C, Granier C, and Tartour E. Following up tumor-specific regulatory T cells in cancer patients. *Oncoimmunology*. 2013;2(7).
  14. Nishikawa H, and Sakaguchi S. Regulatory T cells in cancer immunotherapy. *Curr Opin Immunol*. 2014;27(1-7).
  15. Curiel TJ, Coukos G, Zou LH, Alvarez X, Cheng P, Mottram P, Evdemon-Hogan M, Conejo-Garcia JR, Zhang L, Burow M, et al. Specific recruitment of regulatory T cells in ovarian carcinoma fosters immune privilege and predicts reduced survival. *Nat Med*. 2004;10(9):942-9.
  16. Facciabene A, Peng XH, Hagemann IS, Balint K, Barchetti A, Wang LP, Gimotty PA, Gilks CB, Lal P, Zhang L, et al. Tumour hypoxia promotes tolerance and angiogenesis via CCL28 and T-reg cells. *Nature*. 2011;475(7355):226-U141.
  17. Plitas G, Konopacki C, Wu KM, Bos PD, Morrow M, Putintseva EV, Chudakov DM, and Rudensky AY. Regulatory T Cells Exhibit Distinct Features in Human Breast Cancer. *Immunity*. 2016;45(5):1122-34.
  18. Savage PA, Leventhal DS, and Malchow S. Shaping the repertoire of tumor-infiltrating effector and regulatory T cells. *Immunol Rev*. 2014;259(1):245-58.
  19. Belkaid Y, and Oldenhove G. Tuning microenvironments: Induction of regulatory T cells by dendritic cells. *Immunity*. 2008;29(3):362-71.
  20. Wei S, Levine J, Cogdill A, Zhao Y, Anang NA, Andrews M, Sharma P, Wang J, Wargo J, Pe'er D, et al. Distinct cellular mechanisms underlie anti-CTLA-4 and anti-PD-1 checkpoint blockade. *J Immunother Cancer*. 2017;5(
  21. Pelaseyed T, Bergstrom JH, Gustafsson JK, Ermund A, Birchenough GMH, Schutte A, van der Post S, Svensson F, Rodriguez-Pineiro AM, Nystrom EEL, et al. The mucus and mucins of the goblet cells and enterocytes provide the first defense line of the gastrointestinal tract and interact with the immune system. *Immunol Rev*. 2014;260(1):8-20.
  22. Vandembroucke K, Hans W, Van Huysse J, Neiryneck S, Demetter P, Remaut E, Rottiers P, and Steidler L. Active delivery of trefoil factors by genetically modified *Lactococcus lactis* prevents and heals acute colitis in mice. *Gastroenterology*. 2004;127(2):502-13.
  23. Manca S, Frisbie CP, LaGrange CA, Casey CA, Riethoven JJM, and Petrosyan A. The Role of Alcohol-Induced Golgi Fragmentation for Androgen Receptor Signaling in Prostate Cancer. *Mol Cancer Res*. 2019;17(1):225-37.
  24. Li XF, Hu WX, Zhou JJ, Huang YQ, Peng JP, Yuan Y, Yu JK, and Zheng S. CLCA1 suppresses colorectal cancer aggressiveness via inhibition of the Wnt/beta-catenin signaling pathway. *Cell Commun Signal*. 2017;15(
  25. Williams C, DiLeo A, Niv Y, and Gustafsson JA. Estrogen receptor beta as target for colorectal cancer prevention. *Cancer Lett*. 2016;372(1):48-56.
  26. Greally M, Kelly CM, and Cercek A. HER2: An emerging target in colorectal cancer. *Curr*

- Prob Cancer*. 2018;42(6):560-71.
27. Koehler BC, Jager D, and Schulze-Bergkamen H. Targeting cell death signaling in colorectal cancer: Current strategies and future perspectives. *World journal of gastroenterology*. 2014;20(8):1923-34.
  28. Wong SHM, Kong WY, Fang CM, Loh HS, Chuah LH, Abdullah S, and Ngai SC. The TRAIL to cancer therapy: Hindrances and potential solutions. *Crit Rev Oncol Hemat*. 2019;143(81-94).
  29. Li FZ, Cao YN, Townsend CM, and Ko TC. TGF- $\beta$  signaling in colon cancer cells. *World J Surg*. 2005;29(3):306-11.
  30. Zigmond E, Varol C, Farache J, Elmaliah E, Satpathy AT, Friedlander G, Mack M, Shpigel N, Boneca IG, Murphy KM, et al. Ly6C(hi) Monocytes in the Inflamed Colon Give Rise to Proinflammatory Effector Cells and Migratory Antigen-Presenting Cells. *Immunity*. 2012;37(6):1076-90.
  31. Morias Y, Abels C, Laoui D, Van Overmeire E, Guilliams M, Schoupe E, Tacke F, Devries CJ, De Baetselier P, and Beschin A. Ly6C(-) Monocytes Regulate Parasite-Induced Liver Inflammation by Inducing the Differentiation of Pathogenic Ly6C(+) Monocytes into Macrophages. *Plos Pathog*. 2015;11(5).
  32. Shechter R, Miller O, Yovel G, Rosenzweig N, London A, Ruckh J, Kim KW, Klein E, Kalchenko V, Bendel P, et al. Recruitment of Beneficial M2 Macrophages to Injured Spinal Cord Is Orchestrated by Remote Brain Choroid Plexus. *Immunity*. 2013;38(3):555-69.
  33. Mantovani A, Marchesi F, Malesci A, Laghi L, and Allavena P. Tumour-associated macrophages as treatment targets in oncology. *Nat Rev Clin Oncol*. 2017;14(7):399-416.
  34. Li ZH, Peng YX, Li JX, Chen ZJ, Chen F, Tu J, Lin SB, and Wang HS. N-6-methyladenosine regulates glycolysis of cancer cells through PDK4. *Nat Commun*. 2020;11(1).
  35. Zhang QM, He Y, Luo N, Patel SJ, Han YJ, Gao RR, Modak M, Carotta S, Haslinger C, Kind D, et al. Landscape and Dynamics of Single Immune Cells in Hepatocellular Carcinoma. *Cell*. 2019;179(4):829-+.
  36. Yang H, Xia L, Chen J, Zhang SQ, Martin V, Li QQ, Lin SQ, Chen JF, Calmette J, Lu M, et al. Stress-glucocorticoid-TSC22D3 axis compromises therapy-induced antitumor immunity. *Nat Med*. 2019;25(9):1428-+.
  37. Gao S, Yan LY, Wang R, Li JY, Yong J, Zhou X, Wei Y, Wu XL, Wang XY, Fan XY, et al. Tracing the temporal-spatial transcriptome landscapes of the human fetal digestive tract using single-cell RNA-sequencing (vol 20, pg 721, 2018). *Nature Cell Biology*. 2018;20(10):1227-.
  38. Haber AL, Biton M, Rogel N, Herbst RH, Shekhar K, Smillie C, Burgin G, Delorey TM, Howitt MR, Katz Y, et al. A single-cell survey of the small intestinal epithelium. *Nature*. 2017;551(7680):333-+.
  39. Zhang P, Yang MR, Zhang YD, Xiao S, Lai XX, Tan AD, Du SY, and Li S. Dissecting the Single-Cell Transcriptome Network Underlying Gastric Premalignant Lesions and Early Gastric Cancer (vol 27, pg 1934, 2019). *Cell Rep*. 2020;30(12):4317-.
  40. Franke AJ, Skelton WP, Starr JS, Parekh H, Lee JJ, Overman MJ, Allegra C, and George TJ. Immunotherapy for Colorectal Cancer: A Review of Current and Novel Therapeutic Approaches. *Jnci-J Natl Cancer I*. 2019;111(11):1131-41.
  41. Butler A, Hoffman P, Smibert P, Papalexis E, and Satija R. Integrating single-cell

transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol.* 2018;36(5):411-+.

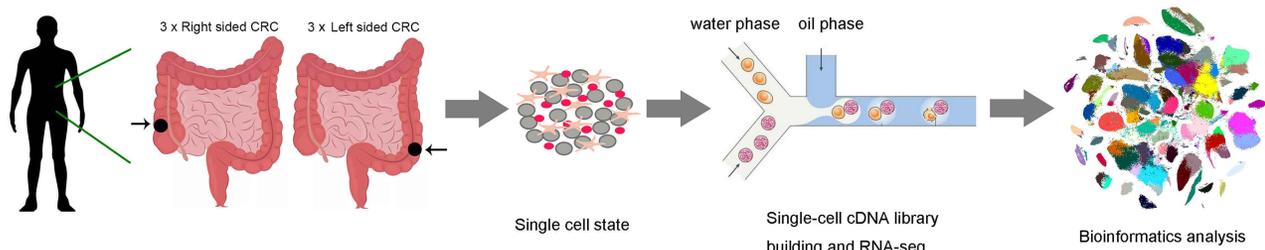
Table S1 Clinical characteristics of six CRC patients enrolled in our study.

Patient No.	Sex	Age	Tumor Location	Tumor Size(cm)	AJCC Stage	LN metastasis	Distant metastasis	MSI status	Metabolic disorders
Left-sided CRC Patient 1	Female	57	Sigmoid	1.5 x 0.5	I	No	No	MSS	No
Left-sided CRC Patient 2	Male	70	Sigmoid	3 x 1	II	No	No	MSS	No
Left-sided CRC Patient 3	Male	65	Sigmoid	2 x 1.5	I	No	No	MSS	No
Right-sided CRC Patient 1	Female	69	Ascending	8 x 1.7	II	No	No	MSS	No
Right-sided CRC Patient 2	Female	80	Ascending	8 x 3.2	II	No	No	MSS	No
Right-sided CRC Patient 3	Male	59	Ascending	10 x 5.5	II	No	No	MSS	No

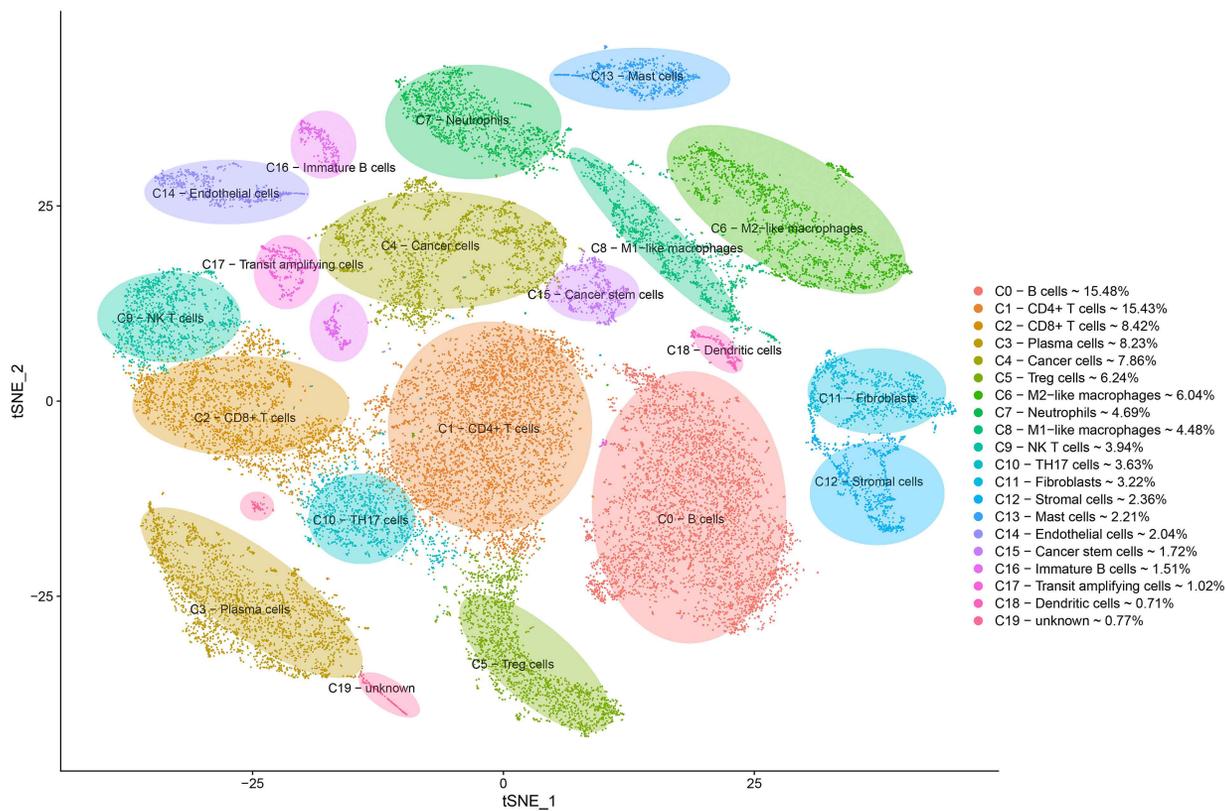
Abbreviation: AJCC, American Joint Committee on Cancer; LN, lymph node; MSI, microsatellite instability.

**Table S1** Clinical characteristics of six CRC patients enrolled in this study.

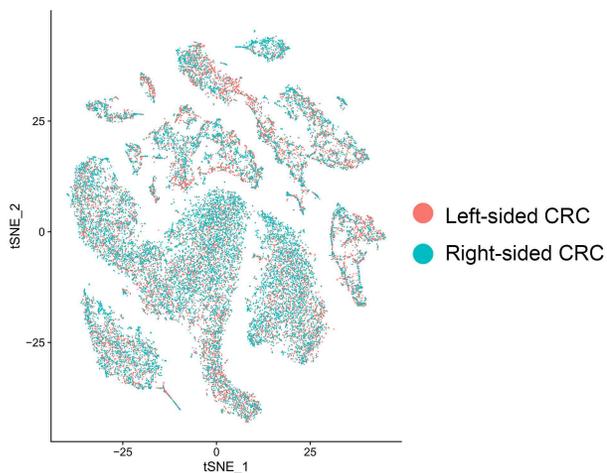
A



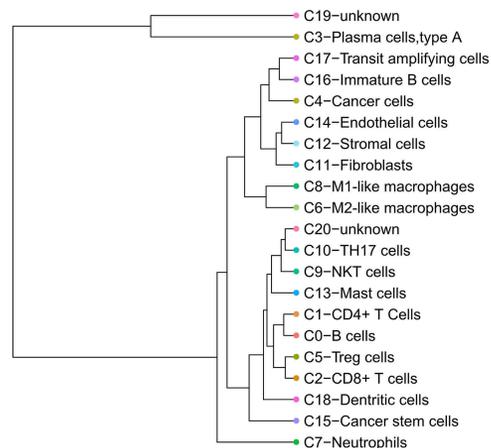
B



C



D



**Figure 1.** Single-Cell Atlas of Colorectal cancer from Left-sided and Right-sided CRC Patients

(A) Schematic diagram highlighting the experimental workflow for the whole study.

(B) The t-SNE plot of 27,927 high-quality cells to visualize cell-type clusters based on the expression of known marker genes.

(C) The t-SNE plot of all types of cells from left-sided CRC samples and right-sided CRC samples.

(D) Unsupervised hierarchical clustering of average gene signatures showing relatedness of cell clusters (correlation distance metric, average linkage).

A

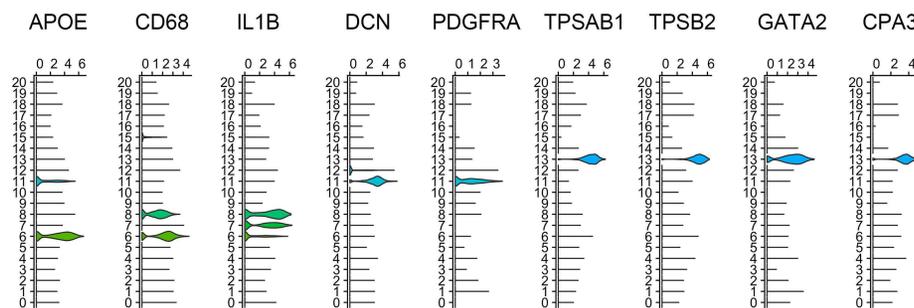
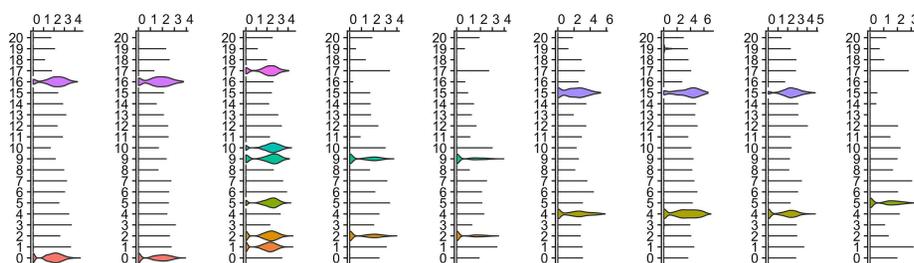
Expression



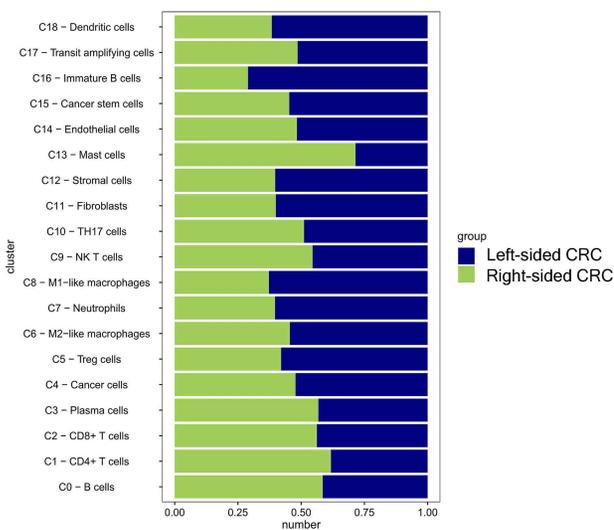
B cells: *MS4A1, CD79A, VPREB3*  
 CD4+ T cells: *IL7R, GIMAP7, SARAF*  
 CD8+ T cells: *GZMK, CCL5, NKG7*  
 Plasma cells: *IGKC, IGH1A, IGLC3*  
 Cancer cells: *TFF3, SPINK4, AGR2*  
 Treg cells: *TNFRSF4, BATF, TNFRSF18*  
 M2-like macrophages: *APOE, C1QA, C1QB*  
 Neutrophils: *S100A8, CXCL8, GOS2*  
 M1-like macrophages: *IL1B, EREG, TIMP1*  
 NK T cells: *GNLY, GZMA, CCL5*  
 TH17 cells: *KLRB1, IL7R, CCL20*  
 Fibroblasts: *COL1A1, COL3A1, COL1A2*  
 Stromal cells: *TAGLN, RGS5, ACTA2*  
 Mast cells: *TPSB2, TPSAB1, CPA3*  
 Endothelial cells: *PLVAP, HSPG2, IGFBP7*  
 Cancer stem cells: *TFF1, PHG1, FABP1*  
 Immature B cells: *TCL1A, RGS13, HIST1H4C*  
 Transit amplifying cells: *STMN1, HMGB2, HIST1H4C*  
 Dendritic cells: *CCL19, CCL22, TXN*  
 unknown: *MZB1, IGHG3, IGHM*  
 unknown: *XCL1, XCL2, AREG*

B

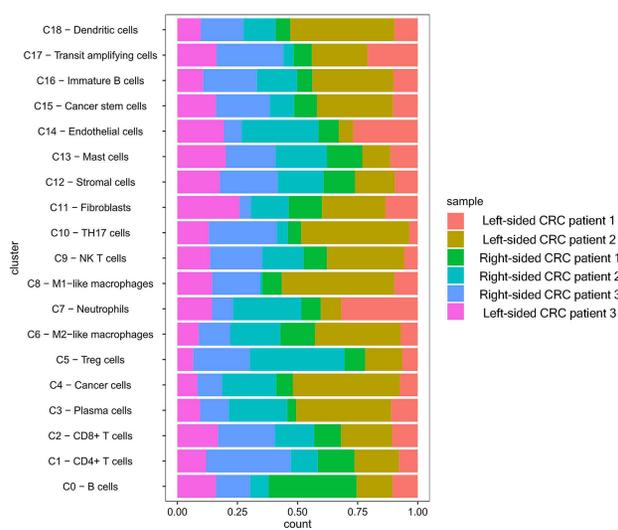
MS4A1 VPREB3 CD3D CD8A CD8B AGR2 TFF3 EPCAM FOXP3



C



D

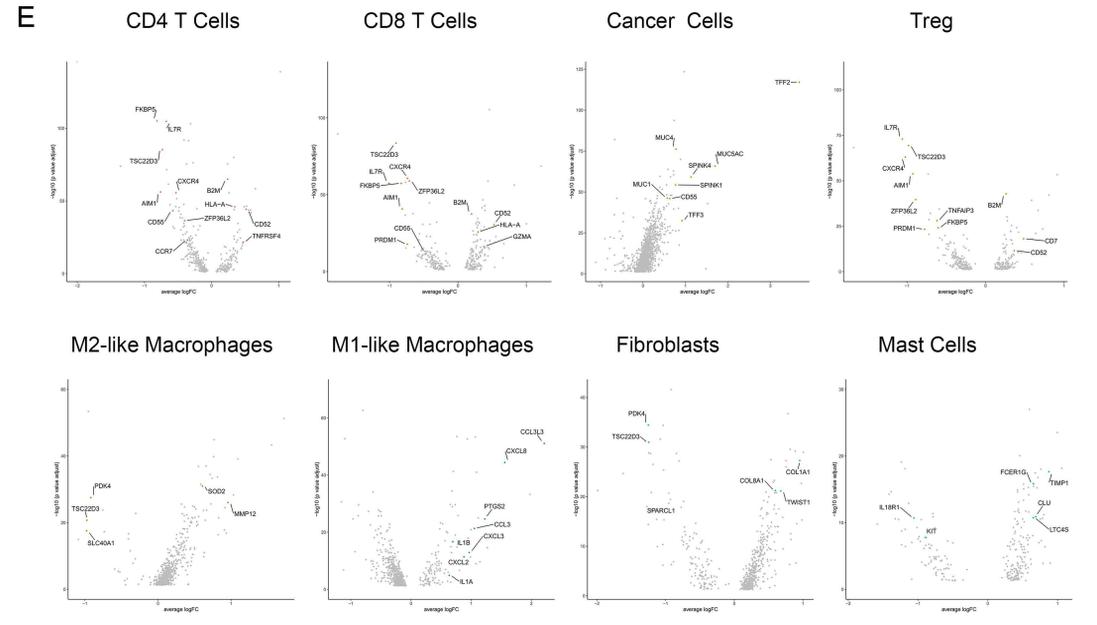
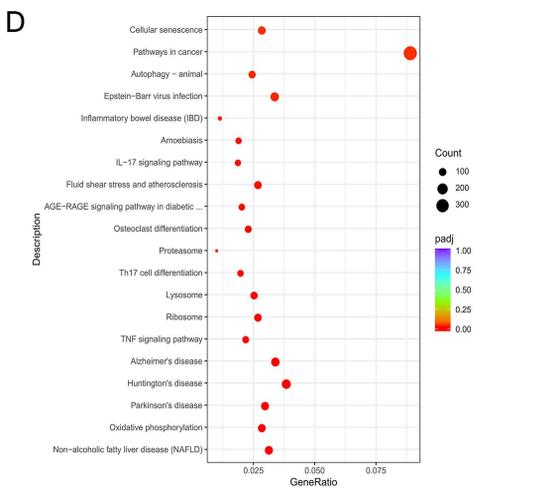
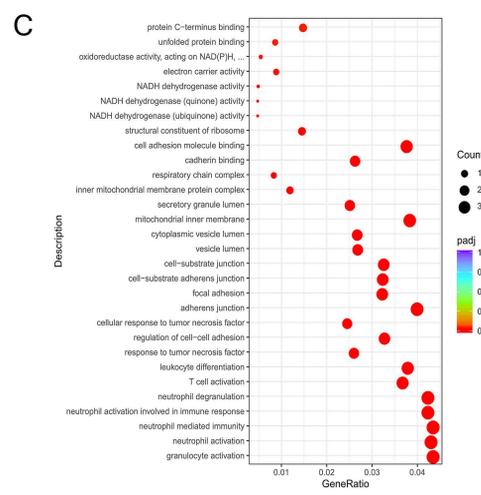
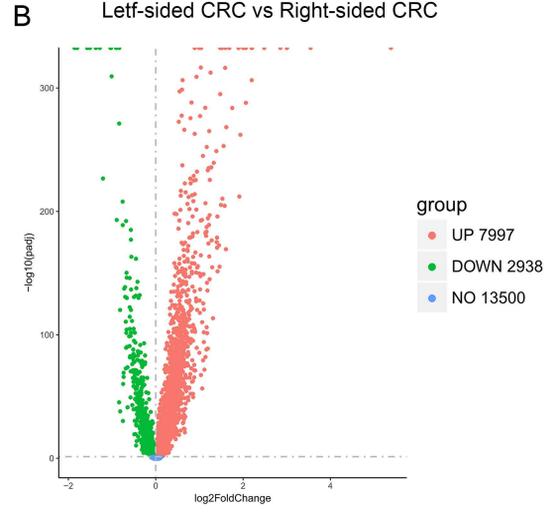
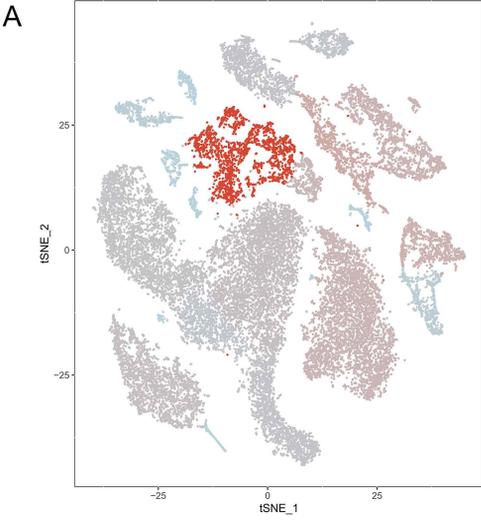


**Figure 2.** Single-Cell Atlas of Colorectal cancer from Left-sided and Right-sided CRC Patients

(A) Heatmap of differentially expressed genes. For each cluster the top 3 genes and their relative expression levels in all CRC cells are shown.

(B) Violin plots display the distribution of expression of known marker genes across diverse cell types among CRC.

(C-D) For 20 subgroups identified in this profile (left to right): the fraction of cells that originated from left-sided and right-sided CRC samples, and the fraction of cells that originated from each of the six patients.



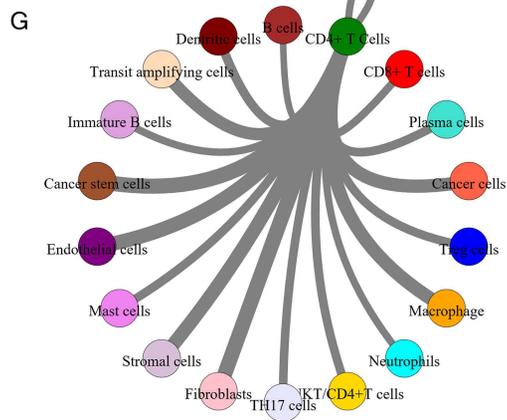
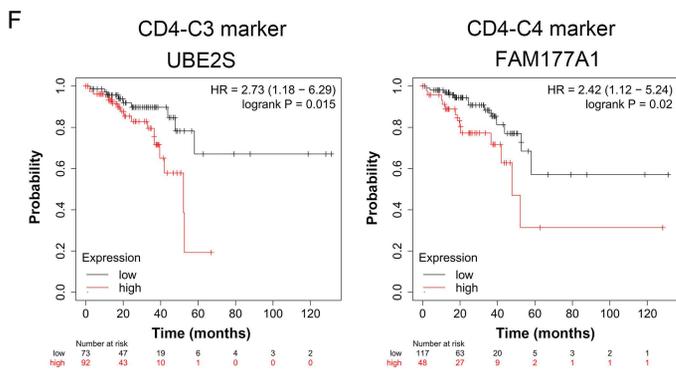
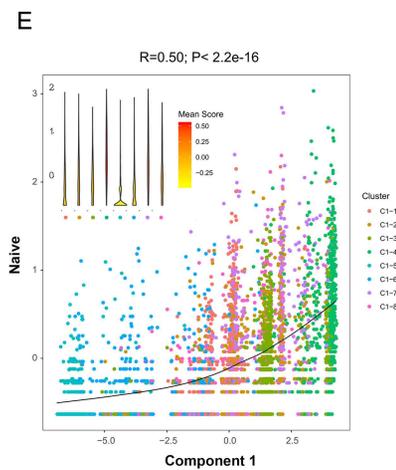
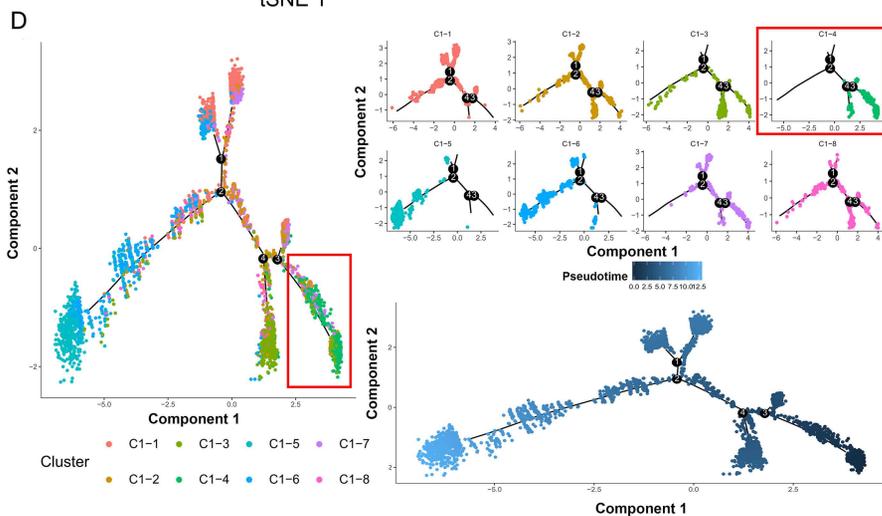
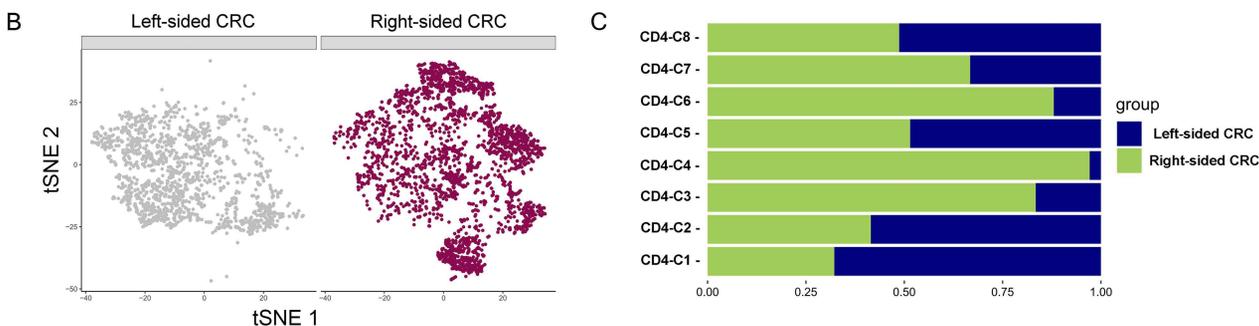
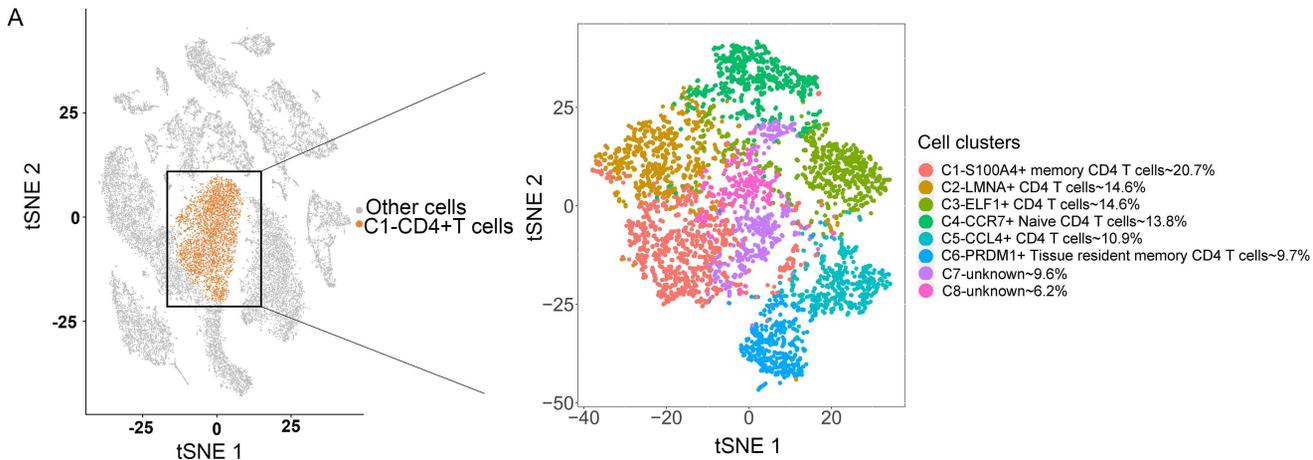
**Figure 3** Cell-Specific Expression Changes in Left-sided and Right-sided CRC

(A) Number of DEGs between Left-sided and Right-sided CRC cells within each cluster projected onto the t-SNE map. DEG:  $|\log \text{ fold change}| > 0.5$ ; adjusted P value  $< 0.05$  was derived by a Wilcoxon rank-sum test.

(B) A volcano plot of DEGs that are upregulated (red) or downregulated (green) between left-sided and right-sided CRC.

(C-D) A cluster profiler identified the enriched gene ontology and Kyoto Encyclopedia of Genes and Genomes processes of DEG.

(E) Unique changes in specific cell subsets between left-sided and Right-sided CRC within CD4 T cells, CD8 T cells, cancer cells, Tregs, M1- and M2-like macrophages, fibroblasts and mast cells compartments. Two-sided paired Student's t-test was used to determine significance.



**Figure 4** Naive CD4 T cells are predominant in Right-sided CRC.

(A) The t-SNE plot that showed the distribution of CD4 T cell lineages (orange, n = 4,310 cells) within the atlas. CD4 T cell populations were re-clustered into eight subclusters (color coding).

(B) Annotation by Left-sided and Right-sided CRC cells.

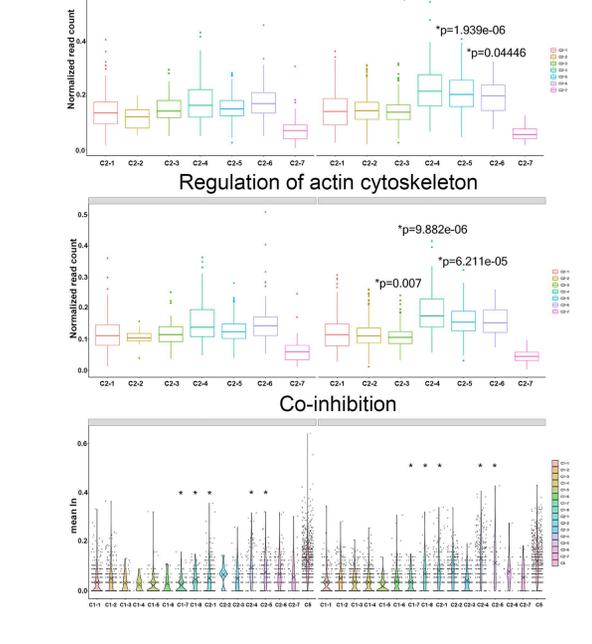
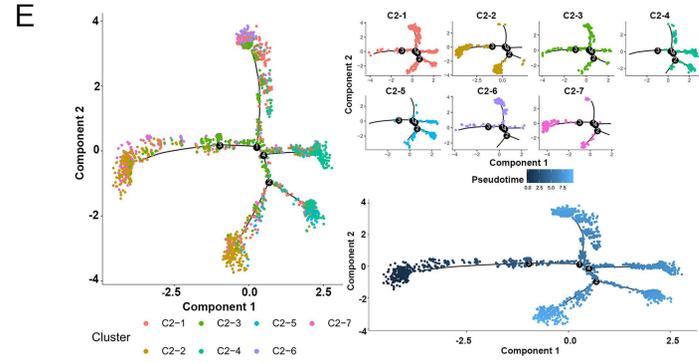
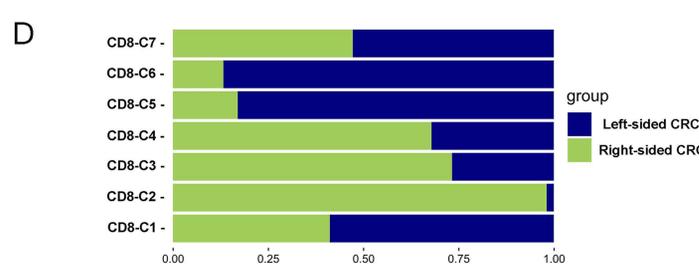
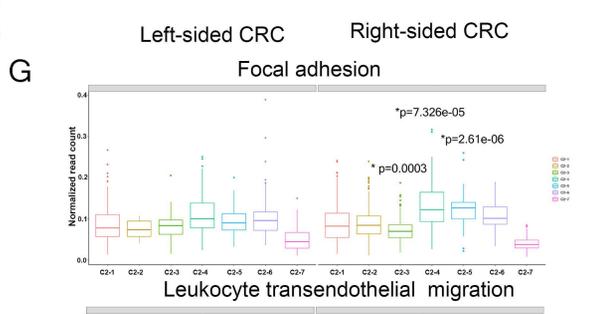
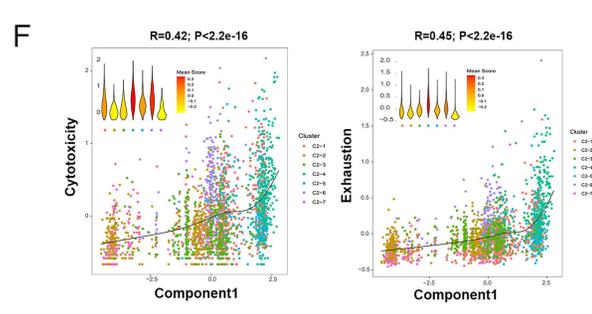
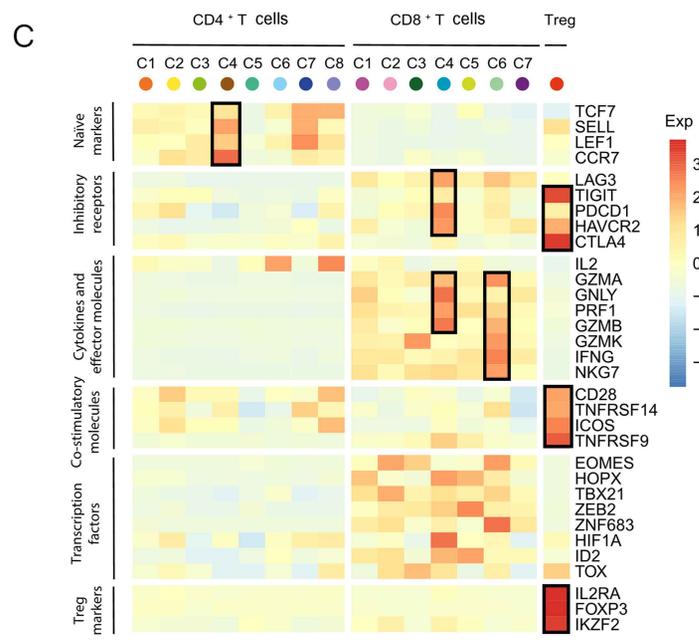
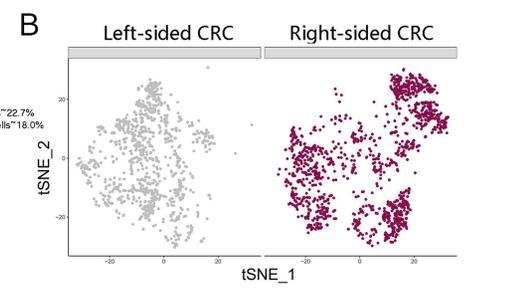
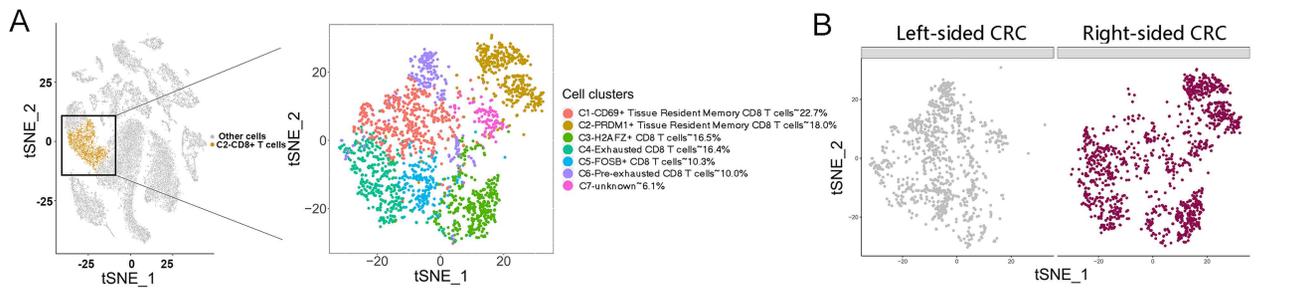
(C) The fraction of cells that originated from left-sided and right-sided CRC samples for 8 subgroups identified in this profile.

(D) Differentiation trajectory of CD4 T cells in CRC, with each color coded for pseudotime and clusters.

(E) Monocle components were correlated with functional features of CD4 T cells (the 4,310 cells as in Figure 4A), including scores of naïveness calculated by the mean expression of gene sets related to this T cell status (see Methods).

(F) Kaplan-Meier survival curves of OS based on UBE2S and FAM177A1 expression using the online bioinformatics tool Kaplan-Meier Plotter.

(G) The similarity network between CD4 T cell and diverse cell types in our dataset. The thickness of edges in the network was denoted as the Pearson correlation coefficient between the centroids of any pair of cell types.



**Figure 5** Right-sided CRC occupies a large proportion of highly migratory exhausted CD8 T cells.

(A) The t-SNE plot that showed the distribution of CD8 T cell lineages (orange,  $n = 2,351$  cells) within the atlas. CD8 T cell populations were re-clustered into seven subclusters (color coding).

(B) Annotation by Left-sided and Right-sided CRC cells.

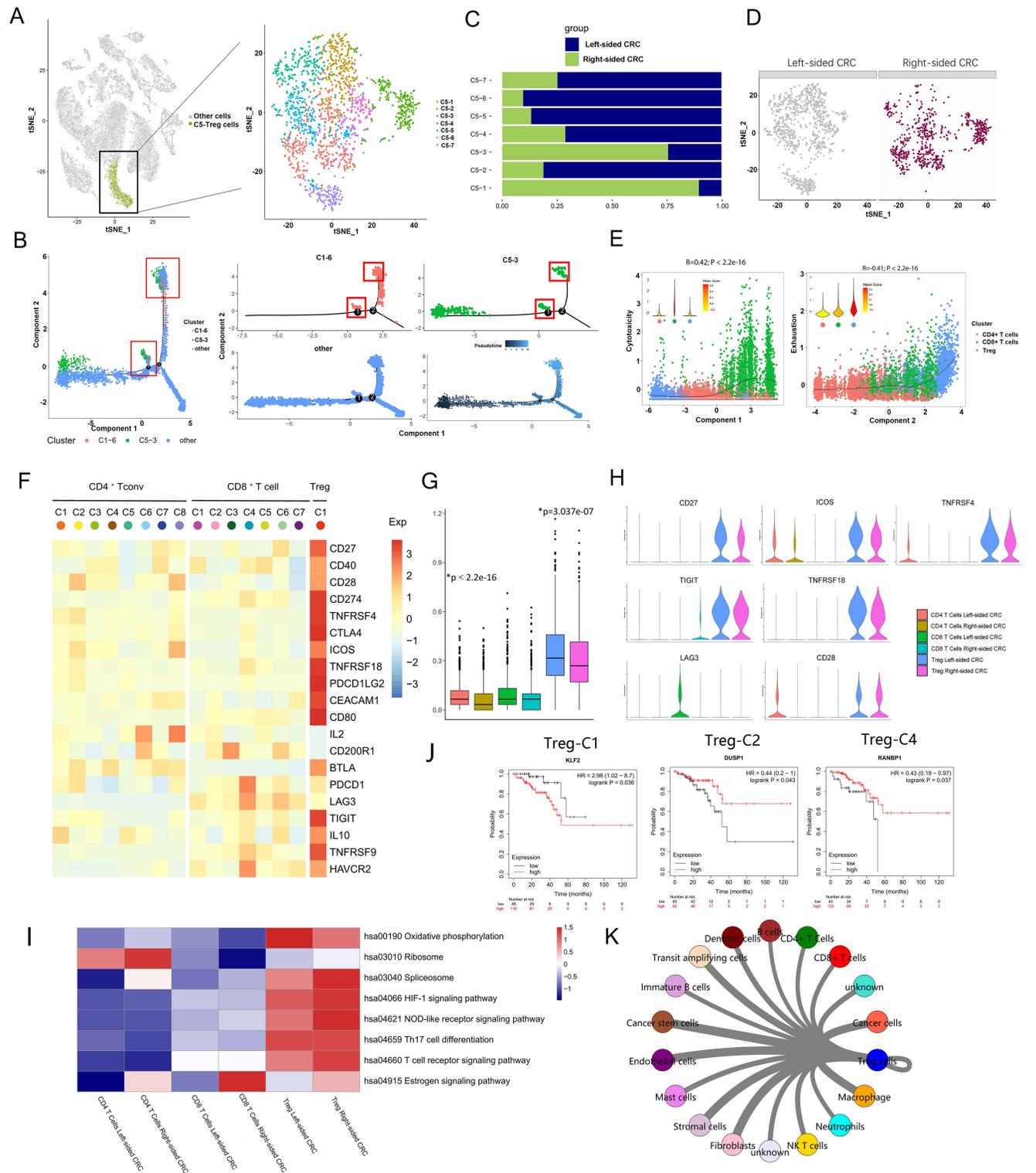
(C) z-score normalized mean expression of selected T cell function-associated genes in each cell cluster. Black boxes highlight the prominent patterns defining known T cell subtypes.

(D) The fraction of cells that originated from left-sided and right-sided CRC samples for 7 subgroups identified in this profile.

(E) Differentiation trajectory of CD8 T cells in CRC, with each color coded for pseudotime and clusters.

(F) Monocle components were correlated with functional features of CD8<sup>+</sup> T cells (the 2,351 cells as in Figure 5A), including scores of exhaustion and cytotoxicity calculated by the mean expression of gene sets related to T cell status.

(G) Box plots of the expressions of Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways and co-inhibition program of all CD4 and CD8 T cell clusters between left-sided and right-sided CRC. \* $P < 0.05$ ; Two-sided paired Student's t-test was used to determine significance.



**Figure 6** Tregs from Left-sided CRC exhibit higher level of immunotherapy-related genes

(A) The t-SNE plot that showed the distribution of Treg cell lineages (green, n = 1,742 cells) within the atlas. Treg cell populations were re-clustered into seven subclusters (color coding).

(B) Differentiation trajectory of CD4 T cells and Treg cells in CRC, with each color coded for CD4-C6 (Tissue Resident Memory CD4 T cells), Treg-C3 and pseudotime.

(C) The fraction of cells that originated from left-sided and right-sided CRC samples for 7 subgroups identified in this profile.

(D) Annotation by Left-sided and Right-sided CRC cells.

(E) Monocle components were correlated with functional features of Treg cells (the 1,742 cells as in Figure 5a), including scores of exhaustion and cytotoxicity calculated by the mean expression of gene sets related to the T cell status.

(F) Heatmap of the expression patterns of genes currently targeted by immunotherapies.

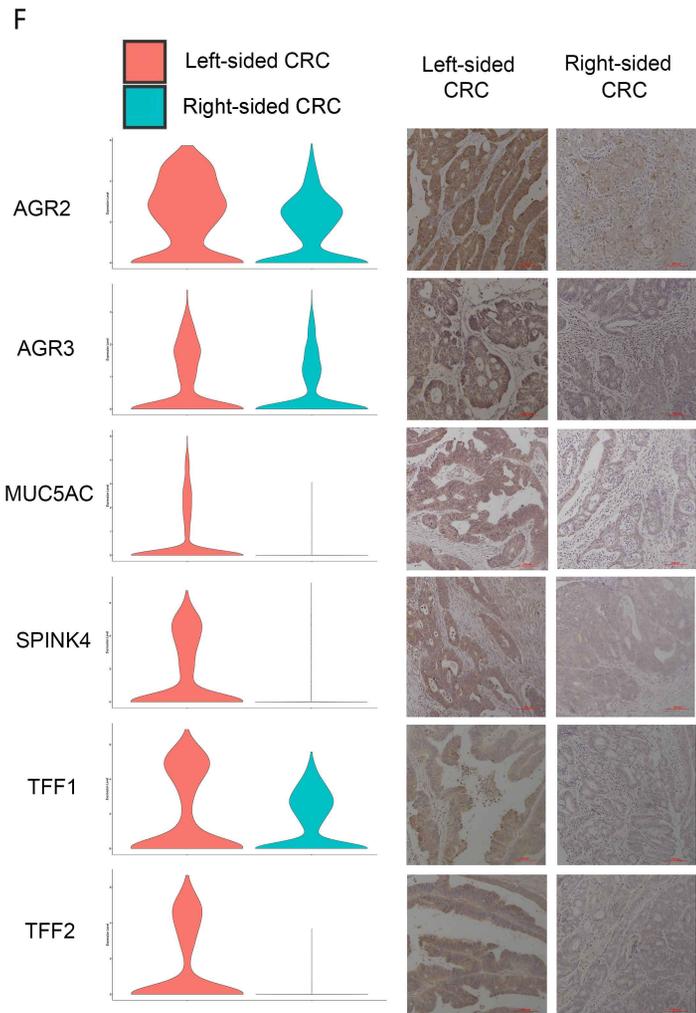
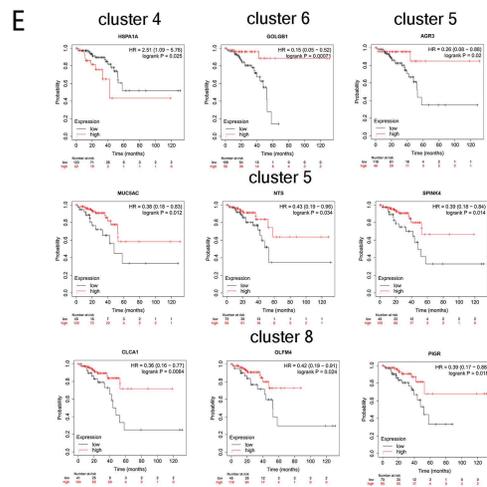
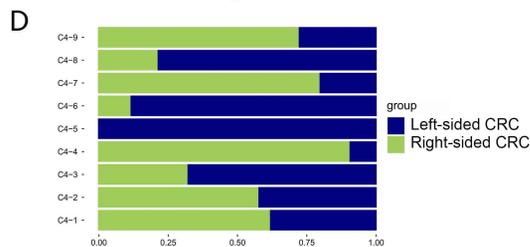
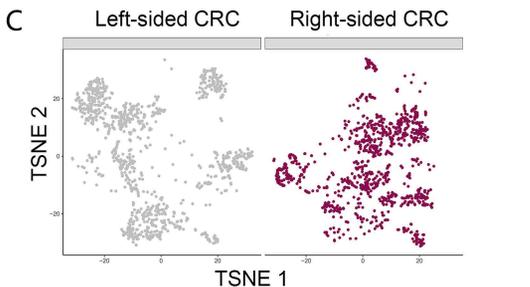
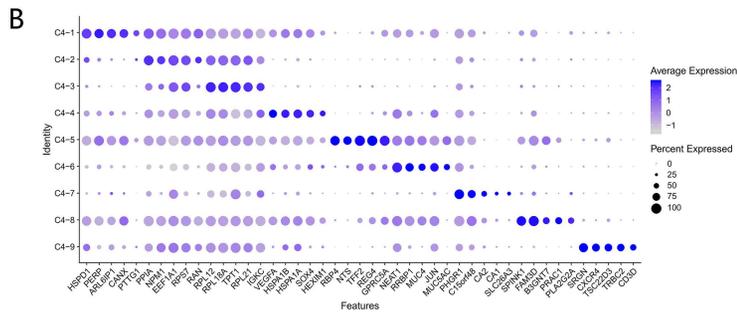
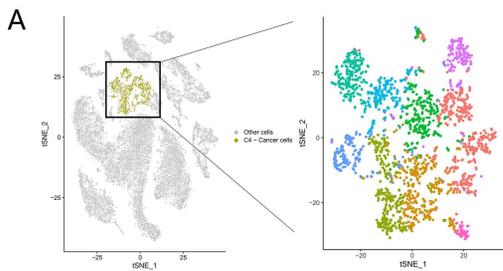
(G) Box plots of mean expressions of genes currently targeted by immunotherapies of all CD4, CD8 T cell and Treg clusters between left-sided and right-sided CRC.

(H) Violin plots display the distribution of expression of CD27, ICOS, TNFRSF4, TIGIT, TNFRSF18, LAG3, CD28 across CD4 T cell, CD8 T cell and Treg cell among CRC.

(I) Kaplan-Meier survival curves of OS based on KLF2, DUSP1 and RANBP1 expression using the online bioinformatics tool Kaplan-Meier Plotter.

(J) Differences in 8 hallmark pathway activities scored with GSEA software. Shown are t values calculated by a linear model.

(K) The similarity network between Treg cell and diverse cell types in our dataset. The thickness of edges in the network was denoted as the Pearson correlation coefficient between the centroids of any pair of cell types. \*P < 0.05; Two-sided paired Student's t-test was used to determine significance.



**Figure 7** A RBP4+ NTS+ Cancer Cell Subset Is Unique to Left-sided CRC

(A) The t-SNE plot that showed the distribution of cancer cell lineages (yellow, n = 2,196 cells) within the atlas. Cancer cell populations were re-clustered into nine subclusters (color coding).

(B) Top five marker genes of 15 major cell types identified in this profile.

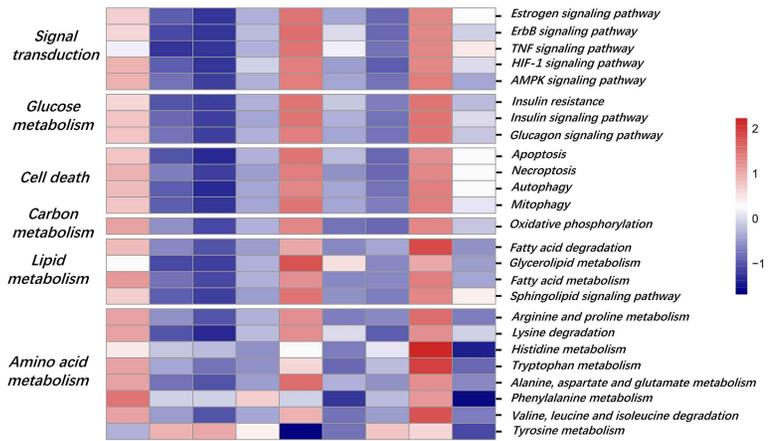
(C) Annotation by Left-sided and Right-sided CRC cells.

(D) The fraction of cells that originated from left-sided and right-sided CRC samples for 9 subgroups identified in this profile.

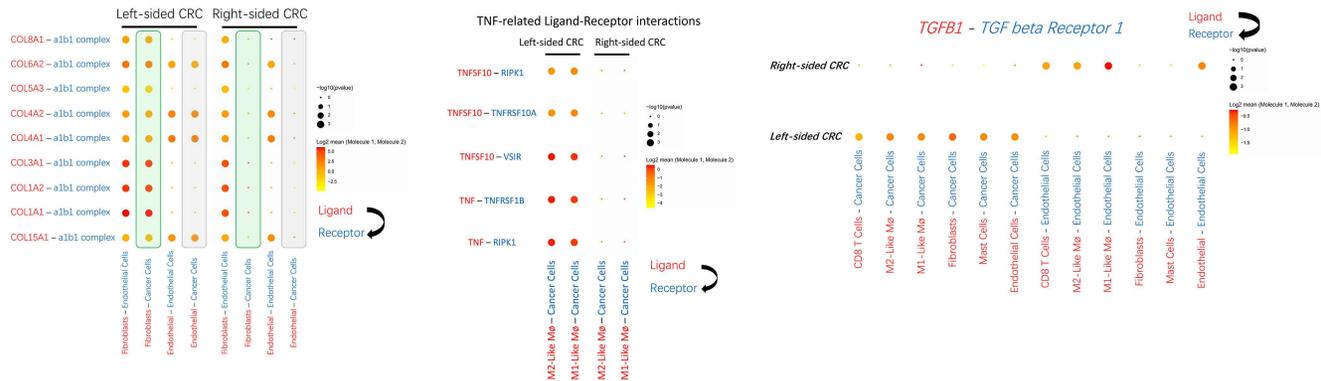
(E) Kaplan-Meier survival curves of OS based on HSPA1A, GOLGB1, AGR3, MUC5AC, NTS, SPINK4, CLCA1, OLFM4 and PIGR expression using the online bioinformatics tool Kaplan-Meier Plotter.

(F) Violin plots and immunohistochemistry display the distribution of expression of AGR2, AGR3, MUC5AC, SPINK4, TFF1, TFF2 between left-sided and right-sided CRC.

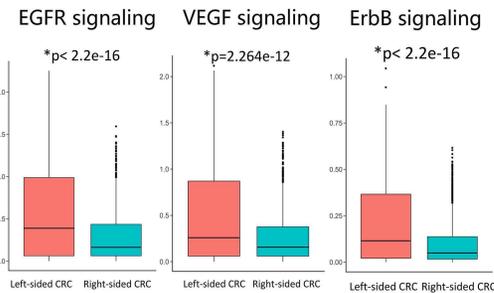
**A**



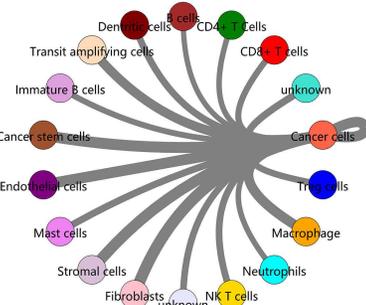
**B**



**C**



**D**



**Figure 8** A RBP4+ NTS+ Cancer Cell Subset Is Unique to Left-sided CRC

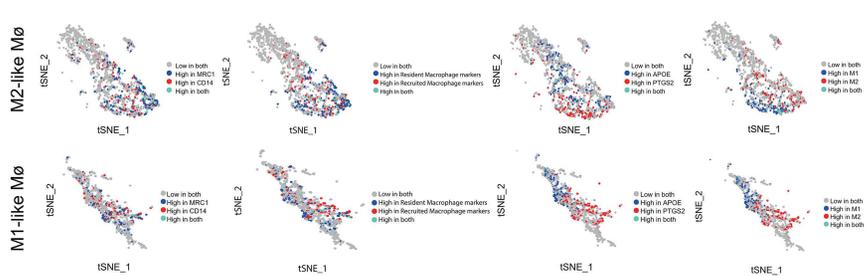
(A) Differences in hallmark pathway activities scored with GSVA software. Shown are t values calculated by a linear model.

(B) Ligand-receptor interaction between cancer cells and TME-infiltrated cell clusters detected by CellPhoneDB 2. Selected ligand-receptor pairs are shown in the bubble plot.

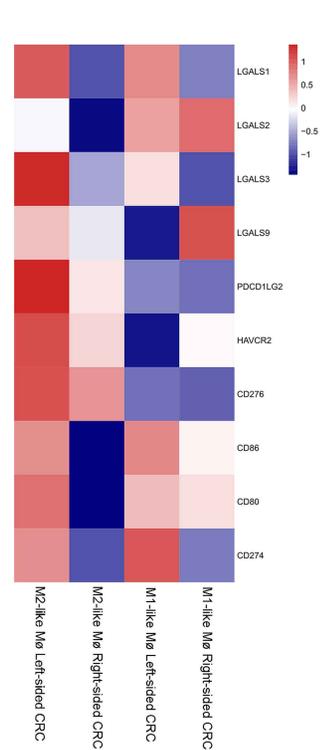
(C) Box plots of the expressions of Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways enriched differentially expressed genes of cancer cell cluster between left-sided and right-sided CRC.

(D) The similarity network between cancer cell and diverse cell types in our dataset. The thickness of edges in the network was denoted as the Pearson correlation coefficient between the centroids of any pair of cell types. \* $P < 0.05$ ; Two-sided paired Student's t-test was used to determine significance.

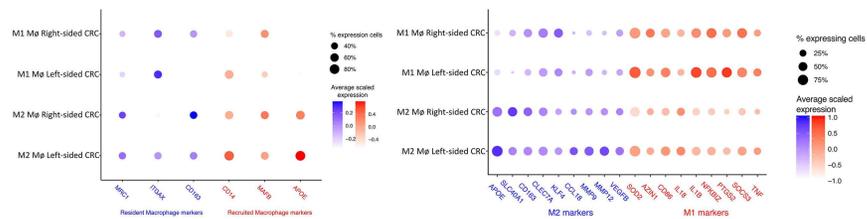
A



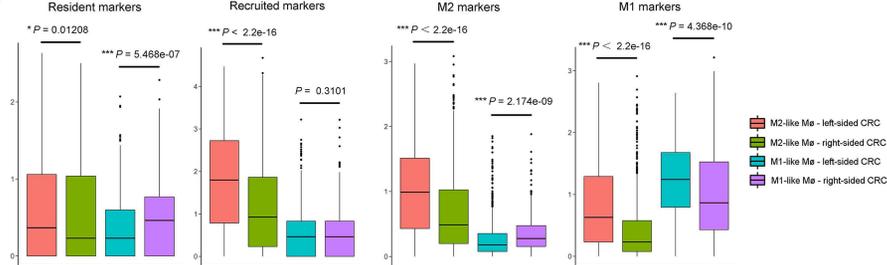
F



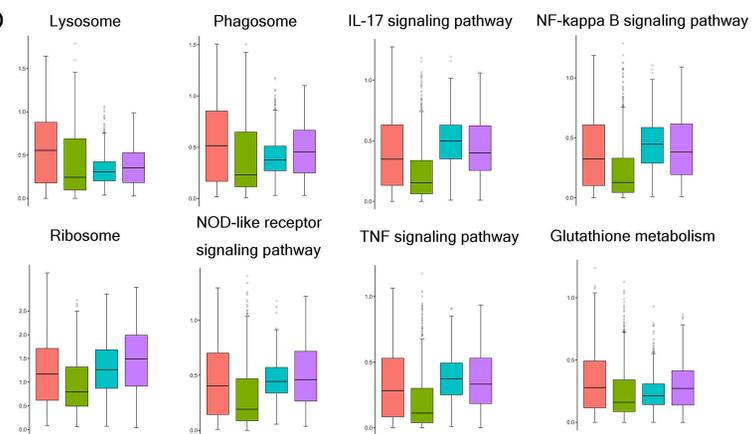
B



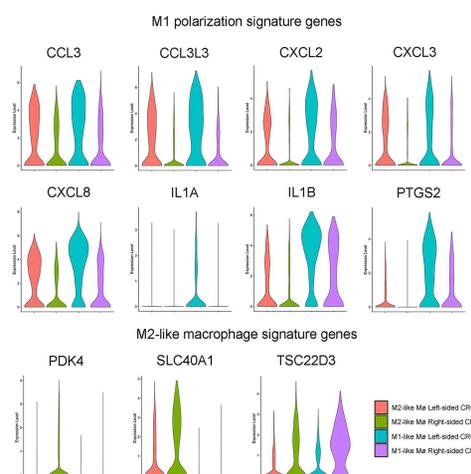
C



D



E



**Figure 9** Heterogeneity of Macrophages in the TME of CRC

(A) Relative expression of Mrc1 and CD14, APOE and PTGS2 overlaid on tSNE plot. Summary expression of 3 resident biomarkers (Mrc1, Itgax, and CD163) and 3 recruited biomarkers (Cd14, Apoe, and Mafb), M1 markers and M2 markers overlaid on tSNE plot.

(B-C) Bubble plot comparing expression of resident (blue) and recruited (red) biomarkers, M2 (blue) and M1 (red) markers across M1-like or M2-like macrophage clusters from left-sided or right-sided CRC.

(D) Mean normalized expression of genes annotated for enriched pathways of M1-like and M2-like macrophage cluster from left-sided and right-sided CRC.

(E) Violin plots display the distribution of expression of M1 and M2 polarization signature genes between left-sided and right-sided CRC.

(F) Heatmap of positive immune checkpoint expression on macrophages. The row Z score was implicated to represent the expression level. \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ ; Two-sided paired Student's t-test was used to determine significance.