

## RESEARCH ARTICLE

# Multiregion single-cell sequencing reveals the transcriptional landscape of the immune microenvironment of colorectal cancer

Wei Wang<sup>1,3,4</sup> | Yu Zhong<sup>1,3,4</sup>  | Zhenkun Zhuang<sup>1,3,4</sup> | Jiarui Xie<sup>1,3,4</sup> | Yueer Lu<sup>1</sup> | Chengzhi Huang<sup>2</sup> | Yan Sun<sup>3,4</sup>  | Liang Wu<sup>3,4</sup> | Jianhua Yin<sup>3,4</sup> | Hang Yu<sup>1</sup> | Zhiqiang Jiang<sup>1</sup> | Shanshan Wang<sup>3,4</sup> | Chunqing Wang<sup>3,4</sup> | Yuanhang Zhang<sup>3,4</sup> | Yilin Huang<sup>1</sup> | Chongyin Han<sup>1</sup> | Zhenggang Zhong<sup>1</sup> | Jialin Hu<sup>1</sup> | Ying Ouyang<sup>1</sup> | Huisheng Liu<sup>1</sup> | Mengya Yu<sup>2</sup> | Xiaochan Wei<sup>3</sup> | Dandan Chen<sup>3</sup> | Lizhen Huang<sup>1</sup> | Yong Hou<sup>3,4,5</sup> | Zhanglin Lin<sup>1</sup> | Shiping Liu<sup>3,4,5,6</sup> | Fei Ling<sup>1</sup> | Xueqing Yao<sup>2</sup> 

<sup>1</sup> School of Biology and Biological Engineering, South China University of Technology, Guangzhou, Guangdong, China

<sup>2</sup> Department of General Surgery, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, School of Medicine, South China University of Technology, Guangzhou, Guangdong, China

<sup>3</sup> BGI-Shenzhen, Shenzhen, China

<sup>4</sup> China National GeneBank, BGI-Shenzhen, Shenzhen, China

<sup>5</sup> Shenzhen Key Laboratory of Single-Cell Omics, Shenzhen, China

<sup>6</sup> The Guangdong-Hong Kong Joint Laboratory On Immunological And Genetic Kidney Diseases, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou, Guangdong, China

## Correspondence

Xueqing Yao, Department of General Surgery, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, School of Medicine, South China University of Technology, Guangzhou, Guangdong 510080, China.

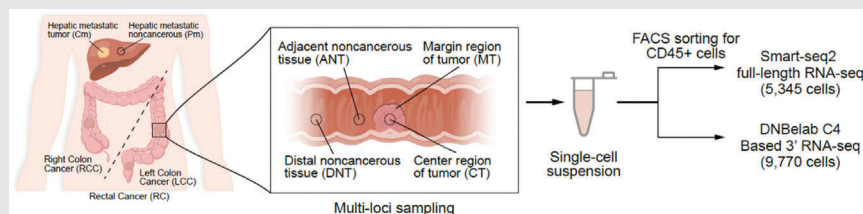
Email: [syaoxueqing@scut.edu.cn](mailto:syaoxueqing@scut.edu.cn)

Fei Ling and Zhanglin Lin, School of Biology and Biological Engineering, South China University of Technology, Guangzhou, Guangdong 510006, China.

Email: [fling@scut.edu.cn](mailto:fling@scut.edu.cn); [zhanglinlin@scut.edu.cn](mailto:zhanglinlin@scut.edu.cn)

Shiping Liu, BGI-Shenzhen, Beishan Industrial Zone, Shenzhen 518083, China.  
Email: [liushiping@genomics.cn](mailto:liushiping@genomics.cn)

## Graphical Abstract



Integrated single-cell RNA sequencing technologies reveal a high-resolution immune landscape of colorectal primary tumors and liver metastasis, identifying major immune cell types and distinct cell functional states of T and B cells as well as predictions of complex cell-cell interactions.

## RESEARCH ARTICLE

# Multiregion single-cell sequencing reveals the transcriptional landscape of the immune microenvironment of colorectal cancer

Wei Wang<sup>1,3,4</sup> | Yu Zhong<sup>1,3,4</sup>  | Zhenkun Zhuang<sup>1,3,4</sup> | Jiarui Xie<sup>1,3,4</sup> | Yueer Lu<sup>1</sup> | Chengzhi Huang<sup>2</sup> | Yan Sun<sup>3,4</sup>  | Liang Wu<sup>3,4</sup> | Jianhua Yin<sup>3,4</sup> | Hang Yu<sup>1</sup> | Zhiqiang Jiang<sup>1</sup> | Shanshan Wang<sup>3,4</sup> | Chunqing Wang<sup>3,4</sup> | Yuanhang Zhang<sup>3,4</sup> | Yilin Huang<sup>1</sup> | Chongyin Han<sup>1</sup> | Zhenggang Zhong<sup>1</sup> | Jialin Hu<sup>1</sup> | Ying Ouyang<sup>1</sup> | Huisheng Liu<sup>1</sup> | Mengya Yu<sup>2</sup> | Xiaochan Wei<sup>3</sup> | Dandan Chen<sup>3</sup> | Lizhen Huang<sup>1</sup> | Yong Hou<sup>3,4,5</sup> | Zhanglin Lin<sup>1</sup> | Shiping Liu<sup>3,4,5,6</sup> | Fei Ling<sup>1</sup> | Xueqing Yao<sup>2</sup> 

<sup>1</sup> School of Biology and Biological Engineering, South China University of Technology, Guangzhou, Guangdong, China

<sup>2</sup> Department of General Surgery, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, School of Medicine, South China University of Technology, Guangzhou, Guangdong, China

<sup>3</sup> BGI-Shenzhen, Shenzhen, China

<sup>4</sup> China National GeneBank, BGI-Shenzhen, Shenzhen, China

<sup>5</sup> Shenzhen Key Laboratory of Single-Cell Omics, Shenzhen, China

<sup>6</sup> The Guangdong-Hong Kong Joint Laboratory On Immunological And Genetic Kidney Diseases, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou, Guangdong, China

## Correspondence

Xueqing Yao, Department of General Surgery, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, School of Medicine, South China University of Technology, Guangzhou, Guangdong 510080, China.

Email: [syaoxueqing@scut.edu.cn](mailto:syaoxueqing@scut.edu.cn)

Fei Ling and Zhanglin Lin, School of Biology and Biological Engineering, South China University of Technology, Guangzhou, Guangdong 510006, China.

Email: [fling@scut.edu.cn](mailto:fling@scut.edu.cn); [zhanglinlin@scut.edu.cn](mailto:zhanglinlin@scut.edu.cn)

Shiping Liu, BGI-Shenzhen, Beishan Industrial Zone, Shenzhen 518083, China.

Email: [liushiping@genomics.cn](mailto:liushiping@genomics.cn)

Wei Wang, Yu Zhong, Zhenkun Zhuang, Jiarui Xie, Yueer Lu, Chengzhi Huang, and Xueqing Yao contributed equally.

## Abstract

The tumor microenvironment is a complex ecosystem formed by distinct and interacting cell populations, and its composition is related to cancer prognosis and response to clinical treatment. In this study, we have taken the advantage of two single-cell RNA sequencing technologies (Smart-seq2 and DNBe-lab C4) to generate an atlas of 15,115 immune and nonimmune cells from primary tumors and hepatic metastases of 18 colorectal cancer (CRC) patients. We observed extensive changes in the proportions and functional states of T cells and B cells in tumor tissues, compared to those of paired non-tumor tissues. Importantly, we found that B cells from early CRC tumor were identified to be pre-B like expressing tumor suppressors, whereas B cells from advanced CRC tumors tended to be developed into plasma cells. We also identified the association of IgA<sup>+</sup>IGLC2<sup>+</sup> plasma cells with poor CRC prognosis, and demonstrated a significant interaction between B-cell and myeloid-cell signaling, and found CCL8<sup>+</sup> cycling B cells/CCR5<sup>+</sup> T-cell interactions as a potential antitumoral mechanism in advanced CRC tumors. Our results provide deeper insights

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *Clinical and Translational Medicine* published by John Wiley & Sons Australia, Ltd on behalf of Shanghai Institute of Clinical Bioinformatics

### Funding information

Science and Technology Project of Guangzhou, Grant/Award Number: 201704020077; Science and Technology Planning Project of Guangdong Province, Grant/Award Numbers: 2017A0300223006, 2019A030317013; CSCO-Roche Cancer Research Foundation, Grant/Award Number: Y-2019Roche-190; Science and Technology Planning Project of Guangdong Provincial People's Hospital, Grant/Award Number: DFJH201913; Fundamental Research Funds for the Central Universities of China, Grant/Award Number: X2SWD2172910; Science and Technology Program of Guangzhou, Grant/Award Number: 201804010318; Shenzhen Key Laboratory of Single-Cell Omics, Grant/Award Number: ZDSYS20190902093613831; Science, Technology and Innovation Commission of Shenzhen Municipality, Grant/Award Number: GJHZ20180419190827179; Natural Science Foundation of Guangdong Province, Grant/Award Number: 2018A030313379; Guangdong-Hong Kong-Macao-Joint Labs Program from Guangdong Science and Technology, Grant/Award Number: 2019B121205005

into the immune infiltration within CRC, and a new perspective for the future research in immunotherapies for CRC.

## 1 | INTRODUCTION

The tumor microenvironment (TME) consists of a cell component made up of various cell types, including immune cells, inflammatory cells, adipocytes, fibroblasts, and vascular endothelial cells, and noncellular components in and around the tumor.<sup>1</sup> The cellular components within the TME are emerging as key regulators of primary tumor progression, organ-specific metastasis, and therapeutic response.<sup>2</sup> Tumor infiltrating immune cells are the key components,<sup>1,3</sup> including T lymphocytes, B lymphocytes, myeloid cells, mast cells, and natural killer (NK) cells, among others, which form an ecosystem that modulates all aspects of tumor development.<sup>3</sup> Among these, T cells are the most abundant and best-characterized immune cells in the TME of solid tumors.<sup>1,3</sup> The presence of  $CD4^+$  T helper 1 (Th1) and cytotoxic  $CD8^+$  T lymphocytes can prevent tumor growth by targeting antigenic tumor cells, and high densities of activated  $CD8^+$  T cells within the tumor niche are associated with favorable prognoses in various cancers.<sup>3</sup> Deeper understanding of the immune complexity of TME will help to identify advanced biomarkers, and allow for devising novel immunotherapy strategies.<sup>4</sup> Along this line, the single-cell technology has rapidly become a powerful approach

for the analysis of TME, including that of colorectal adenocarcinoma (CRC).<sup>5</sup>

CRC is the fourth leading cause of cancer-related mortality worldwide, and the prognosis for CRC patients who experience recurrence or metastasis is extremely poor.<sup>6,7</sup> Surgery, radiotherapy, and chemotherapy have long been the leading strategies for CRC patients. However, the therapeutic effect has never been satisfying, especially for advanced patients with metastatic lesions. In the past decade, immunotherapy has emerged as a potentially effective systemic treatment for advanced CRC patients. So far a great focus has been placed on tumor-infiltrating T lymphocytes (TILs) as they can directly affect prognosis and the response to immunotherapy.<sup>8</sup> Several studies have demonstrated that the type, location, and density of TILs can be used to predict the overall survival (OS) and progression-free survival (PFS) of CRC patients via an immune score strategy.<sup>9,10</sup> Using single-cell technology, Zhang et al.<sup>8</sup> first analyzed the T-cell subpopulations, and illustrated the distinct TIL cell landscape of CRC. Li et al.<sup>11</sup> and De Vries et al.<sup>12</sup> characterized the immune cells and depicted the immune landscape of primary colorectal tumors and matched normal mucosa, respectively. More recently, Zhang et al.<sup>5</sup> used two single-cell sequencing methods to characterize the immune and stromal

populations of CRC primary tumors in human. However, most of these studies on the TME of CRC have largely focused on T cells or myeloid cells, whereas a comprehensive atlas of tumor-infiltrating B cells is largely lacking.<sup>8,12,13</sup> Since the immune system of the gastrointestinal mucosa is chronically challenged by dietary antigens and bacteria, and immune protection driven by antigens in the gastrointestinal mucosa is mainly attributed to antibody-producing B cells and plasma cells,<sup>14,15</sup> the characterization of B cells in CRC should provide new insights into the immune ecosystem of CRC. Additionally, CRC metastasizes to the liver in over half of patients,<sup>16</sup> compared to that of primary tumors, but the biology of CRC liver metastasis (CLM) is currently poorly characterized.

To better understand the immune ecosystem of CRC primary tumor, and with an emphasis on B cells, we performed full-length scRNA-seq (Smart-seq2) on ~6,000 immune cells from primary tumor (included the center area [CT] and the invasive margin area [MT] of tumor), adjacent normal mucosa tissues (ANT), distinct normal mucosa tissues, hepatic metastatic tumors (Cm), and paired noncancerous tissue (Pm) of eight CRC patients. Furthermore, to gain a comprehensive understanding of the immune cellular landscape of hepatic metastases, we also performed full-length scRNA-seq (Smart-seq2) on ~400 cells from paired Cm and Pm. To verify the full-length scRNA-seq results, we utilized a high-throughput platform recently developed in-house, DNBelab C4, which is a scalable and cost-effective approach for microfluidic droplet-based 3' scRNA-seq technology<sup>17</sup> to obtain high-quality data for 10,000 cells from various tissues in 10 CRC patients. In this study, we found that the discrepancies in previous studies of B cells in tumor may be due to the heterogeneity of B-cell subtypes in cancer and noncancerous tissues. We demonstrated that IgA<sup>+</sup>IGCL2<sup>+</sup> plasma cells were associated with poor prognosis in CRC. Paired ligand-receptor analyses demonstrate that myeloid cells and certain B-cell subsets could regulate multiple T-cell subsets in CRC tumor via various ligand-receptor pairs. This study provides deeper insights into the immune infiltration within CRC.

## 2 | RESULTS

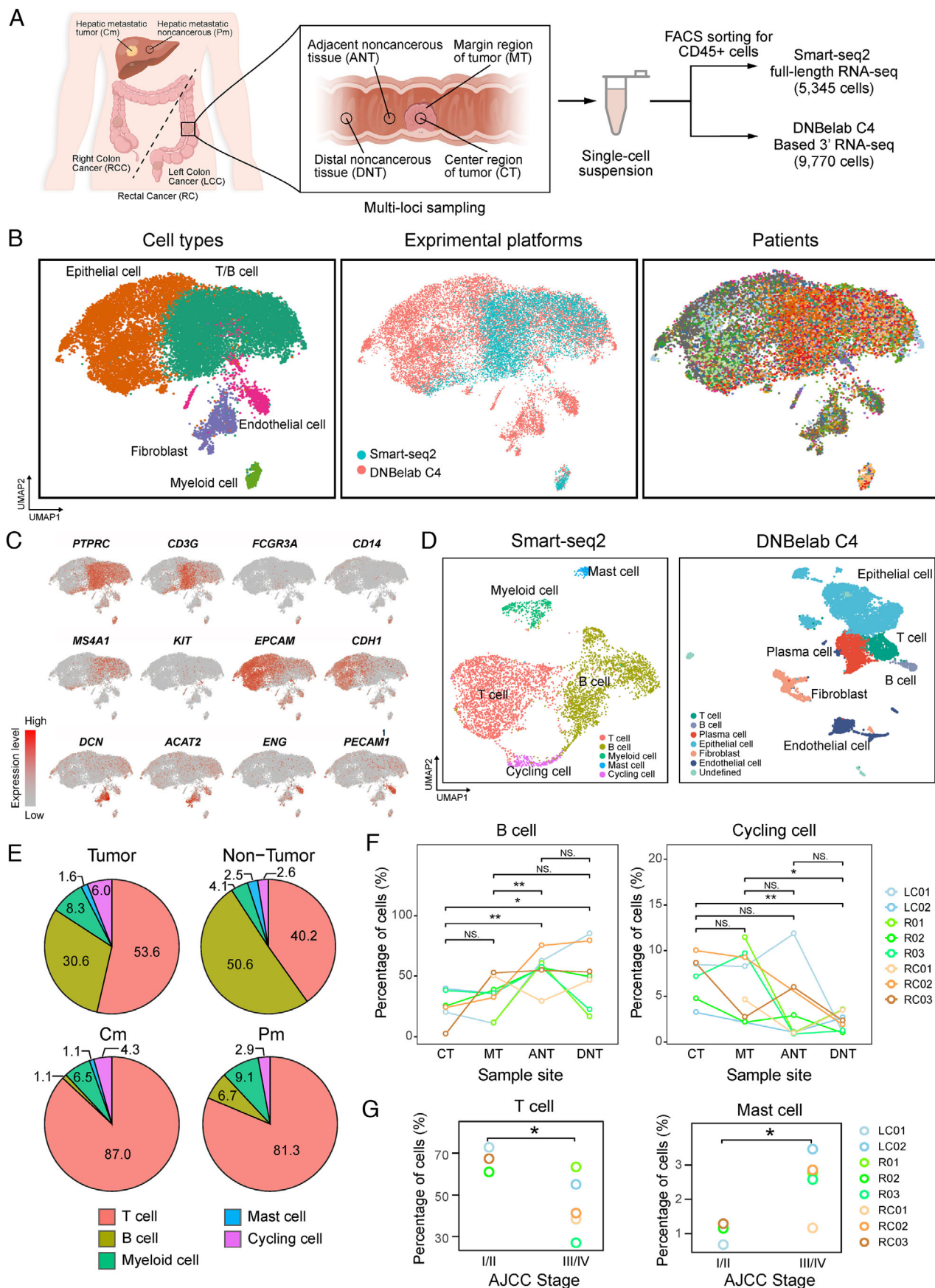
### 2.1 | A single-cell transcription atlas of CRC primary tumors and hepatic metastases

Combining two single-cell RNA sequencing technologies (Smart-seq2 and DNBelab C4), we investigated different cell populations within primary tumors of 18 CRC patients and within matching hepatic metastasis from three of

these patients (Figure 1A, Table S1, see Section 4). Specifically, for patients RC01, RC02, RC03, R01, R02, R03, LC01, and LC02, a modified full-length Smart-seq2 method was used to study the composition of CD45<sup>+</sup> cells isolated from tumors, ANT (proximal to the brim of carcinoma 3-5 cm), and distinct normal tissues (DNT, distal to the brim of carcinoma >10 cm). Besides, paired Cm and adjacent liver (Pm) from patient LC02 were also included. For patients NRC01, NR01, NLC03, NLC04, NLC05, NLC06, NLC07, NLC08, NLC09, and NLC10, a recently developed method based on microfluidic droplet method (DNBelab C4)<sup>17</sup> was used to analyze the cells isolated from tumors, ANT, DNT, Cm, and Pm. After quality control and removing low-quality cells (see Section 4), 5,345 CD45<sup>+</sup> single cells (Smart-seq2) and 9,770 single cells (DNBelab C4) were retained for subsequent analyses. The clinical information and the numbering of cells from each patient are provided in Table S1.

To define the major cell populations within CRC primary and hepatic metastases, we first performed graph-based cell clustering<sup>19</sup> on these two datasets, after removing batch effects among multiple samples (see Section 4). Based on the expression of canonical cell lineage markers (Figure 1C, Figure S1A and S1B), we identified seven major cell types, including four immune cell types (T cells, B cells, myeloid cells, and mast cells) and three nonimmune cell types (epithelial cells, endothelial cells, and fibroblasts) (Figure 1B, Figure S1A, Table S2).

To better define the major population and subpopulation of the tumor-infiltrating leukocytes, we then performed graph-based cell clustering separately on these two datasets. For immune cells from the Smart-seq2 platform, graph-based cell clustering gave rise to 15 cell clusters (Figure 1D). T cells were the predominant immune cell type (49%), followed by B cells (39%), myeloid cells (6%), and mast cells (2%). An additional 4% of cells were positive for proliferation markers, which we annotated as cycling cells (Figure 1B). Each cluster contained cells from a number of different patients, indicating that cell types and expression states in the TME are largely consistent across CRC tumors and do not represent patient-specific subpopulations or batch effects, although the cell-type proportions do vary from person to person (Figure 1B). Not surprisingly, when comparing tumors and matching nontumor tissues, there were significant differences in the proportion of immune cells between tumor (CT and MT) and nonmalignant (ANT and DNT) tissues (Figure 1E,F). T cells and cycling cells were enriched in the tumor, whereas B cells and myeloid cells were depleted; mast cells were detected at similar proportions in both tissue types (Figure 1E,F and Figure S1E). In addition, the infiltration of T cells was significantly increased in both hepatic metastatic tumors and nontumor tissues (87.0% and 81.3%, respectively)



**FIGURE 1** Overview of the immune landscape of colorectal adenocarcinoma (CRC) tumors and nontumor samples. A, Overview of the study design. Colon or rectum tumors and noncancerous samples were collected from 18 patients with CRC. B, Uniform Manifold

(Figure 1E and Figure S1F), and the results were verified by flow cytometry (Figure S1G).

For the immune and nonimmune cells from the DNBe-lab C4 platform, graph-based cell clustering also gave rise to 17 cell clusters (Figure S1B). Based on the expression of known markers, we found that the atlas mainly comprised epithelial cells (*EPCAM*, *CDH1*, *CD24*) and nonepithelial cells (Figure S1C). As shown in Figure 1D, the nonepithelial cell lineages comprised plasma cells (*JCHAIN*, *MZB1*), T cells (*CD3G*, *CD3D*), B cells (*MS4A1*, *CD79A*), fibroblasts (*DCN*, *COL1A2*), and endothelial cells (*GNG11*, *ENG*). The major immune-cell populations, including T cells, B cells, and myeloid-lineage cells for both platforms were qualitatively consistent, demonstrating the validity of our data (Figure S1D).

To gain deeper insights into tumor progression and tumor metastasis mechanisms in human CRC, we performed integrative analysis by using the full-length scRNA-seq sequences. The association analysis between tumor, node, and metastasis (TNM) staging and distinct cell type infiltration in CRC tumor revealed that infiltration of T cells was decreased in advanced stage CRC (stage III and IV), in contrast, mast cells were more enriched in advanced stage CRC (Figure 1G). This is consistent with a previous study that found that mast cells were recruited to cancer cells and released pro-tumor factors to promote proliferation of cancer cells.<sup>18</sup>

Taken together, our results suggest that the proportion of cell subpopulations in tumor tissues is significantly different from that of paired normal tissues in CRC, and that liver metastases possess a distinct TME from that of primary CRC.

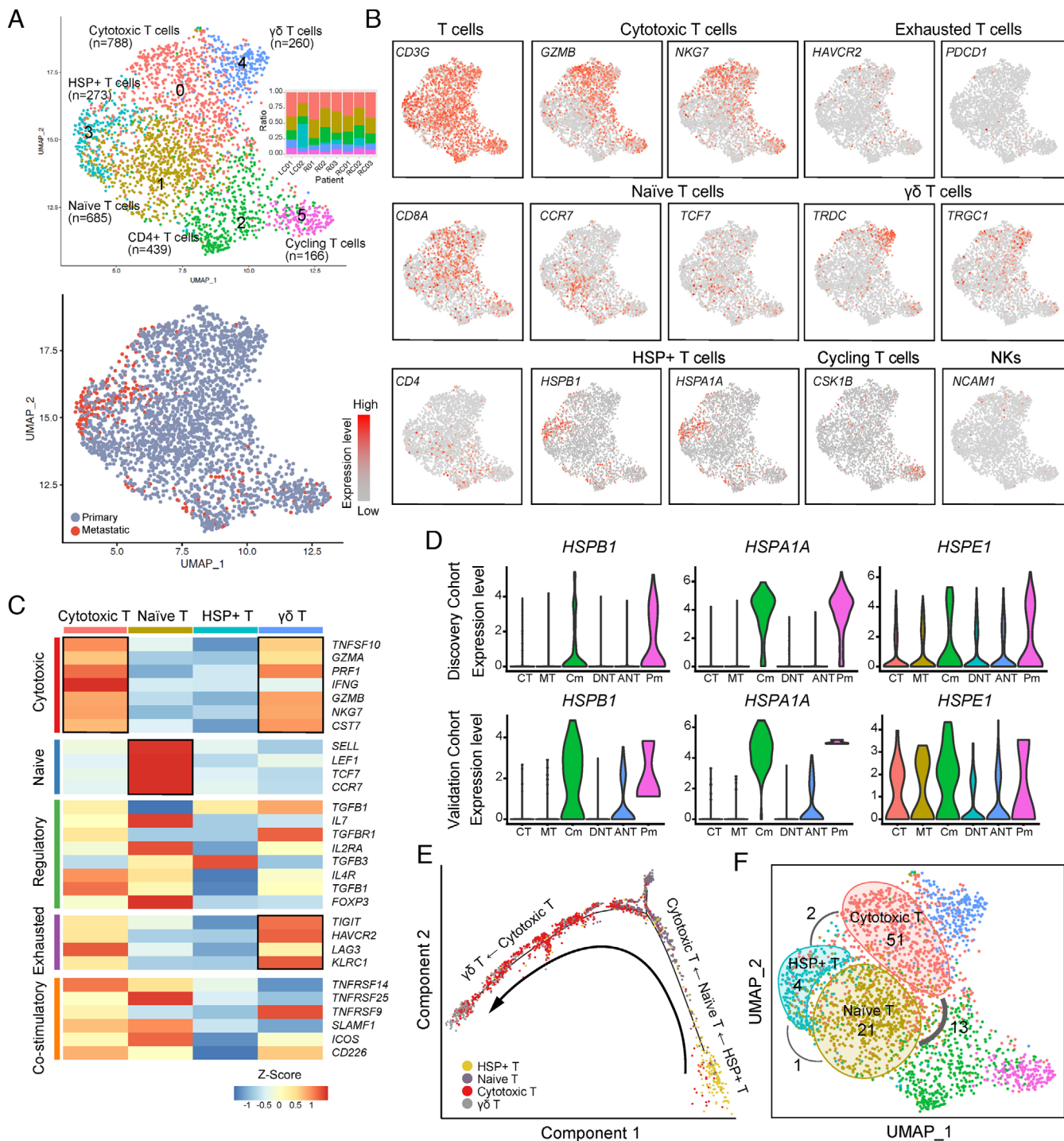
## 2.2 | Heterogeneity and impairment of T cells provide new clinical insights

T cells in discovery cohort were further reclustered and yielded six distinct clusters (Figure 2A, Table S3). Cluster\_0 cells, highly expressing *GZMB* and *NKG7*, were identified as cytotoxic T cells, which are associated with T-cell activation and cytotoxic function (Figure 2B,C). Cluster\_1 cells, expressing naïve T-cell markers *CCR7* and *TCF7*, were characterized as naïve T cells (Figure 2B). Cluster\_2 cells were identified as CD4<sup>+</sup> T cells due to their expression of *CD4* and *CXCL13* (Figure S2A), which is a marker

of exhausted CD4<sup>+</sup> T helper cells.<sup>19</sup> Cluster\_3, namely HSP<sup>+</sup> T cells, expressing a low level of both cytotoxic markers and naïve markers (Figure 2B,C), were distinct in the high expression level of heat-shock protein (HSP) genes *HSPA1A*, *HSPB1*, and *HSPE1* (Figure S2A). Moreover, HSP<sup>+</sup> T cells were more enriched in metastatic site than primary CRC (Figure 2A). Cluster\_4 cells were identified as  $\gamma\delta$  T cells, as they specifically expressed *TRDC* and *TRGC* (Figure 2B); these cells also highly expressed cytotoxic markers *NKG7* and *GZMB*. Although the exhaustion markers *HAVCR2* and *PDCDI* showed low expression across our dataset, in this cluster the *HAVCR2* (TIM-3) level was relatively higher (Figure 2C), which suggested that the  $\gamma\delta$  T cells possessed cytotoxic functions and partially showed the exhaustion state. Cells in Cluster\_5 were classified as cycling T cells due to their signature expression of cell proliferation genes like *CKS1B* (Figure 2B and Figure S2A).

Surprisingly, the majority of Cluster\_3 cells that expressed HSPs were derived from CLM, with 21.6% and 41.8% in metastatic site Cm and Pm, respectively. HSP-related genes have been shown to be highly expressed in regulatory T cells (Treg) and CD4<sup>+</sup> conventional T (Tconv) cells in breast carcinoma compared to normal breast parenchyma.<sup>20</sup> Moreover, HSP70 has been shown to promote Treg survival and immune suppression.<sup>21</sup> To exclude the possibility that the HSP<sup>+</sup> T cells in our study might just be the artifact results of stress from cell sorting rather than a true state, we examined the expression of *HSPA1A*, *HSPB1*, and *HSPE1* in the validation cohort, and found a similar expression pattern of these HSP genes (Figure 2D). In addition, the presence of HSP<sup>+</sup> T cells in Cm tissues was also confirmed by immunofluorescence (IF) using CD3 and HSP70 antibodies (Figure S2C). Therefore, we supposed these HSP<sup>+</sup> T cells were in their true state. To elucidate whether the HSP<sup>+</sup> T cells were more similar to the naïve state or the cytotoxic state, and their function, a trajectory analysis was performed on the naïve T cells, cytotoxic T cells,  $\gamma\delta$  T cells, and HSP<sup>+</sup> T cells using Monocle2.<sup>22</sup> As shown in Figure 2E, the trajectory began with the HSP<sup>+</sup> T cells, followed by the naïve T cells, and then the cytotoxic T cells, and ended with the  $\gamma\delta$  T cells, suggesting that the HSP<sup>+</sup> cells were more similar to the naïve T cells than the cytotoxic T cells, and with less activity than the naïve T cells. This might be due to the influences of the immune microenvironment in CLM.

Approximation and Projection algorithm (UMAP) plot of all cells from two experimental platforms profiled here. Each cell is color-coded by (left to right): associated cell type, experimental platforms, and patient. C, UMAP plot, color-coded for the expression level (gray to red) of marker genes in each cell type. D, UMAP plots show clustering results of different experimental platforms. E, Cell type distribution of tumor and nontumor tissues as well as paired hepatic metastatic tumor and nontumor tissues in discovery cohort. F, The percentage of B cells and cycling cells in different regions. G, Comparison of the cell distribution of T cells and mast cells and AJCC stage. NS,  $P > 0.05$ ; \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ , as determined by Student's *t*-test.



**FIGURE 2** Characterization of T-cell subpopulations in colorectal adenocarcinoma (CRC). **A**, Uniform Manifold Approximation and Projection algorithm (UMAP) plot of 2,571 T cells, color-coded by their associated cluster and bar plot shows distribution of cell clusters among patients (top) or sample type (bottom). **B**, UMAP plot, color-coded for the expression level (gray to red) of marker genes in each cell type. **C**, Heatmap showing the z-score of T-cell functional genes for each cluster. **D**, The expression level of heat-shockprotein (HSP) genes in discovery cohort and validation cohort. **E**, The ordering of T cells along pseudotime in a two-dimensional state-space defined by Monocle2. Each point corresponds to a single cell, and each color represents a T-cell cluster. **F**, Clone expansion and sharing among naïve, cytotoxic, and HSP<sup>+</sup> T cells. The number within the circle refers to the expansion clone number in that specific cluster. Lines connecting clusters are based on the degree of TCR sharing, with the thickness of the line representing the number of shared TCR clones.

As previously reported, the diversity of expanded T-cell receptor (TCR) clones is associated with prognosis in tumor and a greater diversity of expanded TCR clones has been observed in activated T cells than in inactivated T cells.<sup>23</sup> To further confirm the activity state of the HSP<sup>+</sup> T cells, we assembled full-length TCR  $\alpha$  and  $\beta$  sequences for 938 T cells using TRACER,<sup>24</sup> among which 84 clones were considered to be expanded clones consisting of more than two cells. We analyzed the clone expansion and clonal sharing between three  $\alpha\beta$  T-cell subpopulations: naïve, cytotoxic, and HSP<sup>+</sup> T cells. The number of expanded clones in these three T subtypes were 21, 51, and 4 (Figure 2F), respectively, along the trajectory path, indicating that the HSP<sup>+</sup> T cells were more inactive than the naïve T cells. The results above suggest that the T cells from metastatic sites were less active with fewer expansion events than those from primary sites, demonstrating heterogeneity between the microenvironment of metastatic and primary sites.

The difference between T cells in tumor tissue and non-malignant tissue can also be observed, much as previously reported.<sup>8</sup> For example, the naïve T cells were significantly enriched in the nonmalignant tissue ( $P < 0.001$ ), while the CD4<sup>+</sup> T helper cells and cycling T cells were enriched in the tumor tissue ( $P = 0.0011$  and  $0.018$ , respectively) (Figure 3A). No significant difference was observed between the center and marginal regions of tumors, nor between adjacent noncancerous and distal noncancerous tissues (Figure 3B).

We further asked whether the same T-cell subpopulations, but derived differently from tumor or normal tissues, had different gene expression patterns. We thus performed a gene set variation analysis (GSVA) for cells from tumor (CT and MT) and noncancerous (ANT and DNT) tissues, separately (see Section 4). For the cytotoxic T cells, tumor-derived T cells showed stronger cholesterol homeostasis activities (Figure 3C) and significantly higher expression levels of *ACAT1*, a key cholesterol esterification enzyme (Figure 3D). This high level of cholesterol may be related to T cells dysfunction. According to Yang et al,<sup>25</sup> the anti-tumor response of CD8<sup>+</sup> T cells is associated with cholesterol metabolism in a skin melanoma mouse model. Inhibition of *ACAT1* enhanced the proliferation of CD8<sup>+</sup> T cells and successfully delayed the growth of tumors in mice.<sup>25</sup> It was also reported that the expression of *ACAT1* could serve as a potential prognostic marker in prostate cancer.<sup>26</sup> To interrogate the association of *ACAT1* level with prognosis in CRC, we performed survival analysis and found that high *ACAT1* expression is related to poor disease-free survival in colon adenocarcinoma and rectum adenocarcinoma cohorts from The Cancer Genome Atlas (TCGA) cohort (Figure 3E). Therefore, *ACAT1* may be a potential prognostic marker for CRC.

For the naïve T cells, the proportion of this cell type in tumors was significantly lower than in nontumor (Figure 3A), but strong proliferative activities (G2M checkpoint and E2F targets), allograft rejection activities, as well as strong IFN- $\gamma$  and IFN- $\alpha$  responses were observed. This indicates that the naïve T cells in the TME may be more similar to activated T cells than the naïve T cells in adjacent normal tissues.

Finally, comparison of gene expression between early and advanced T cells revealed that the cytolytic molecule *GZMA* highly expressed in T cells derived from early CRC tumor (Figure 3F and Figure S2B). Upregulated *GZMA* in CRC (from TCGA) was also found to be associated with longer PFS ( $P = 0.036$ ) (Figure 3G). Therefore, impairment of T-cell-mediated immune response and enrichment of mast cell could be a factor for poor prognosis of CRC.

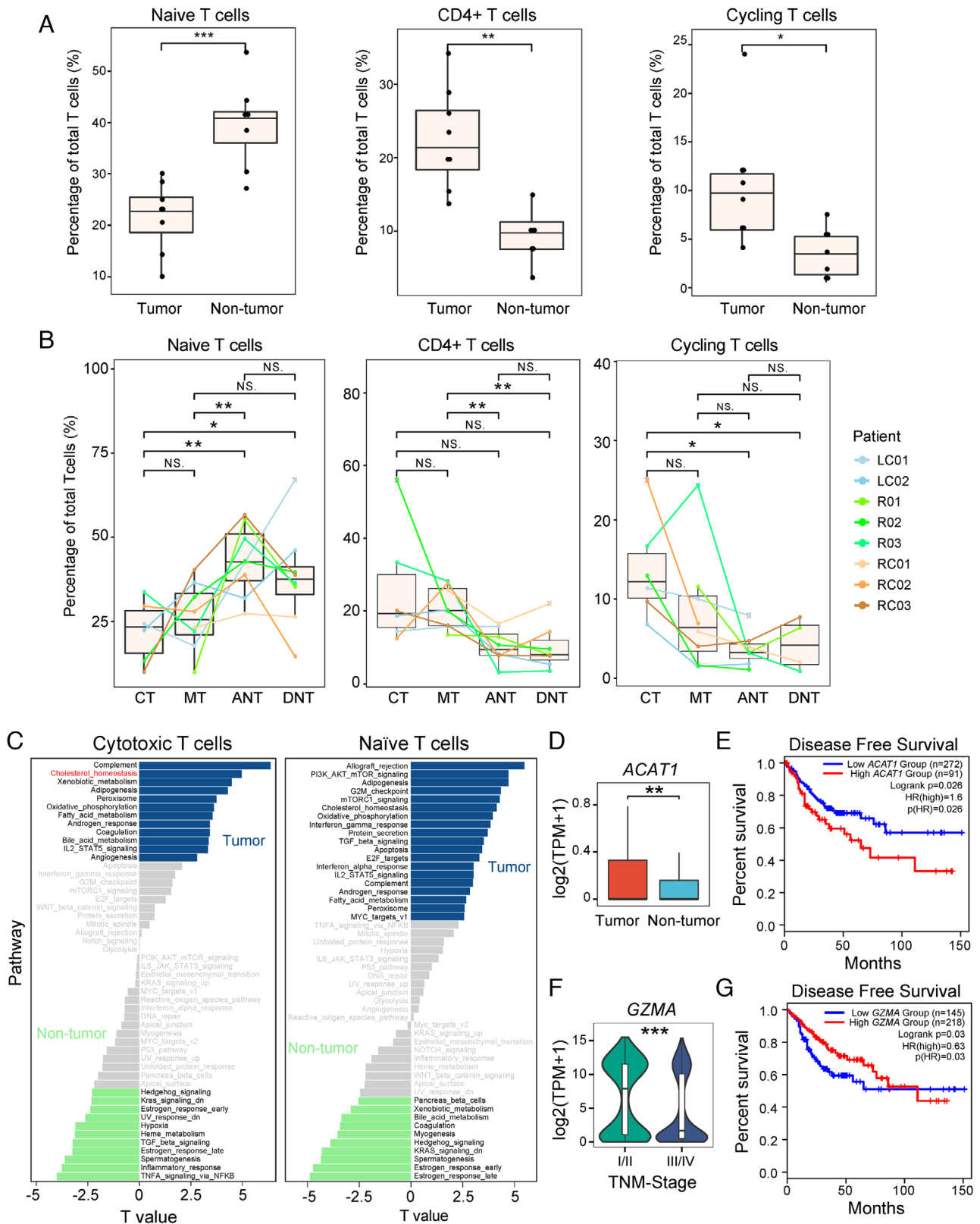
Taken together, we observed the distinct T-cell immune profiles of CRC tumor and normal tissues. Additionally, the same T-cell subpopulations possessed different expression patterns in several pathways depending on their origins of either tumor or normal tissues.

### 2.3 | Landscape of the heterogeneity and diversity of B cells in tumor microenvironment

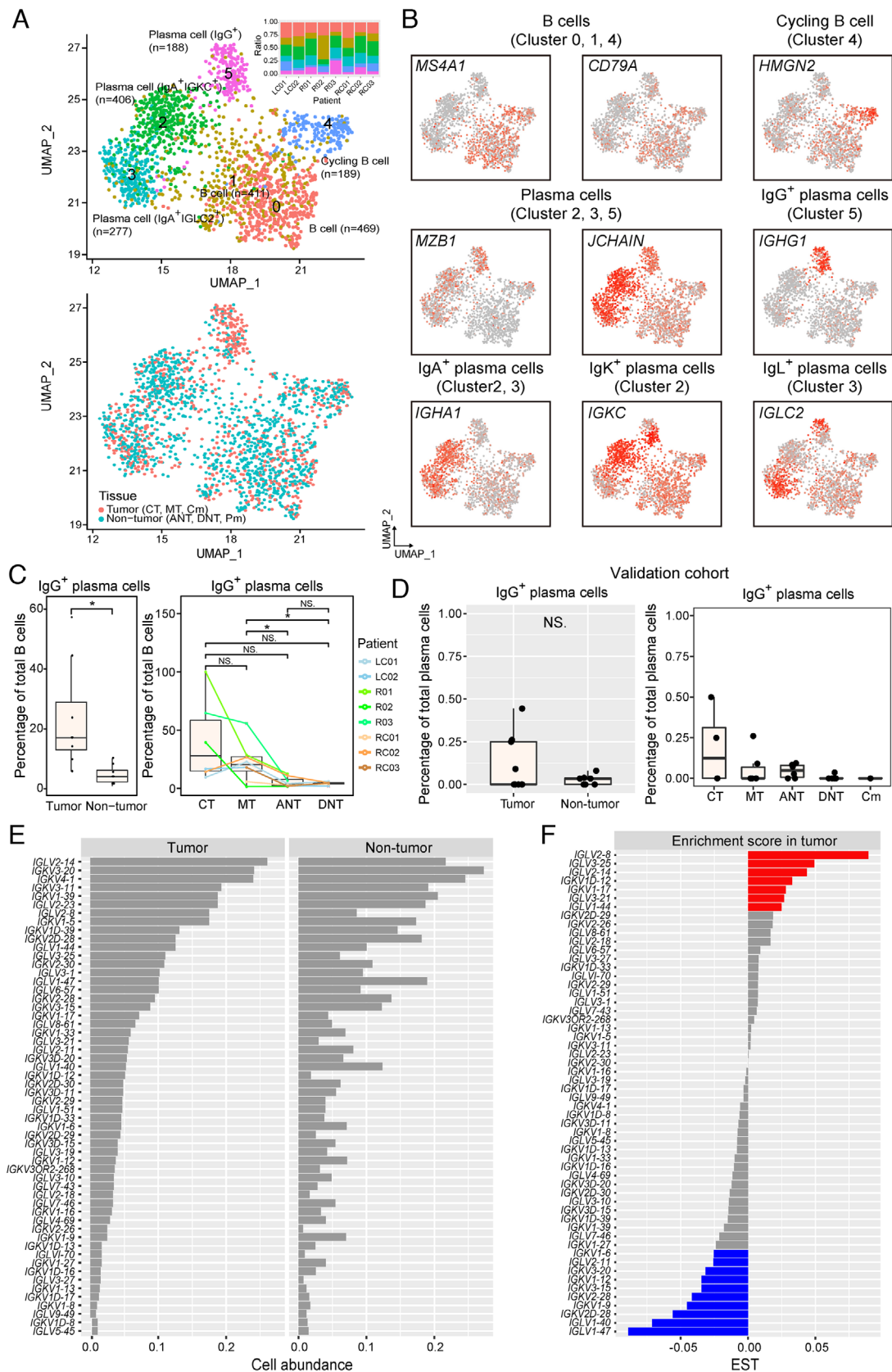
In this study, a large number of B lymphocytes were found in tumor tissues, less than T lymphocytes, and accounted for 30.6% of all immune cells in primary CRC (Figure 1C). To further understand the heterogeneity of B cells and their immunological properties, we performed refined clustering of B cells, and identified three subpopulations of *MS4A1*<sup>+</sup> B cells and three plasma cell subsets (Figure 4A and Figure S3A, Table S4). Two of the B-cell subsets were designated as tissue-resident memory B cells as they showed the hallmark *CD69* and *CD44* expression (Figure S3B).<sup>27</sup>

Two of the plasma cell subpopulations (Cluster\_2 and Cluster\_3) highly expressed IgA-related genes (*IGHA1/2*) (Figure 4A,B). Immunoglobulin  $\kappa$  constant (*IGKC*) was highly expressed in Cluster\_2, also designated as IgA<sup>+</sup>*IGKC*<sup>+</sup> cells. Cluster\_3, also classified as IgA<sup>+</sup>*IGLC2*<sup>+</sup> cells, were distinguished in the high expression of Ig lambda (*IGL*) genes (*IGLC2* and *IGLC3*), but low expression of *IGKC*. IgAs are a class of antibodies that are mainly distributed in mucosal areas, such as the gastrointestinal tract, respiratory tract, and genitourinary tract, and play an important role in preventing pathogen colonization.<sup>28</sup> *IGKC* has been found to be associated with favorable prognosis in breast cancer, CRC, and non-small cell lung cancer.<sup>29,30</sup> The third one, Cluster\_5 plasma cells (IgG<sup>+</sup>), expressed immunoglobulin heavy constant gamma genes (*IGHG1-4*), indicating that they specifically





**FIGURE 3** Interregion comparisons of T-cell subpopulations. **A**, The percentage of cells in different T-cell subpopulations in tumor and nontumor tissues. **B**, The percentage of cells in different T-cell subpopulations in different dissection regions. **C**, Differences in pathway activities between tumor and nontumor T subtypes as scored per cell by GSEA. The  $t$ -values from a linear model corrected for patient of origin are shown. **D**, The expression of *ACAT1* in cytotoxic T cells from tumor and nontumor tissues. **E**, Survival analysis based on the expression status of *ACAT1* (normalized by *CD3G*) in TCGA COAD and READ. **F**, *GZMA* was highly expressed in T cells from early CRC tumor. **G**, Survival analysis based on the expression status of *GZMA*. NS,  $P > 0.05$ ; \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ , as determined by Student's  $t$ -test.



**FIGURE 4** A distinctive landscape of B-cell subpopulations in colorectal adenocarcinoma (CRC). A, Uniform Manifold Approximation and Projection algorithm (UMAP) plot of 1940 B cells, color-coded by their associated cluster and bar plot shows distribution of cell clusters among patients (top) or sample tissue origin (bottom). B, UMAP plot, color-coded for the expression level (gray to red) of marker genes in each cell type. C, The percentage of IgG plasma cells in tumor and nontumor tissues. D, The percentage of IgG plasma cells in tumor and nontumor

secreted IgG antibodies (Figure 4B). IgG<sup>+</sup> plasma cells were significantly enriched in tumor tissues and their infiltration level gradually decreased along the center to the periphery of the tumor (Figure 4C). In the validation cohort, we also identified four plasma-cell subtypes including IgA<sup>+</sup>IGKC<sup>+</sup>, IgA<sup>+</sup>IGLC2<sup>+</sup>, IgG<sup>+</sup>, and IgM<sup>+</sup> plasma cells, and a group of plasma cells with unique profile of immunoglobulin variable region genes, which resided only in normal tissue from an individual patient (Figure S4A,B). The infiltration of IgG<sup>+</sup> plasma cells was also found to be higher in tumor in validation cohort (Figure 4D), but the difference was insignificant. This may be due to the biases associated with the two platforms (Smart-seq2 and DNBelab C4), which led to loss in certain cell populations in validation cohort. Consistent with the results in the discovery and validation cohorts, FACS analysis showed the abundance of IgG<sup>+</sup> plasma cells in the CRC tumor tissues (Figure S4C).

B-cell receptor (BCR) diversity of infiltrating B cells is important for TME. It has been previously reported that in some cancers the BCR light chain diversity of infiltrating B cells in the tumor region was significantly higher than the B cells in the nontumor region.<sup>31</sup> Interestingly in our study, the BCR light chain showed high diversity and abundance in both tumor and nontumor tissues; however, these two tissues did not differ significantly from each other (Figure 4E). This can be explained by the fact that the immune system of the gastrointestinal mucosa is influenced or complicated by antigens from the gastrointestinal microbiota and food intake.<sup>15,16</sup> We next asked whether there were specific types of BCR genes expressed only in the tumor tissue of CRC. Although most of the light chain variable region genes (IGLV genes and IGKV genes) were not uniquely expressed in B cells in the tumor region, the abundance of each variable region genes in tumor and nontumor regions was indeed different (Figure 4F). The variable region genes that showed significant differences in abundance between tumor and normal regions, such as *IGLV2-8*, *IGLV3-25*, and *IGLV2-14* (Figure 4F), might be associated with antitumor antibodies or with disordered gastrointestinal microbiota in CRC.<sup>32</sup>

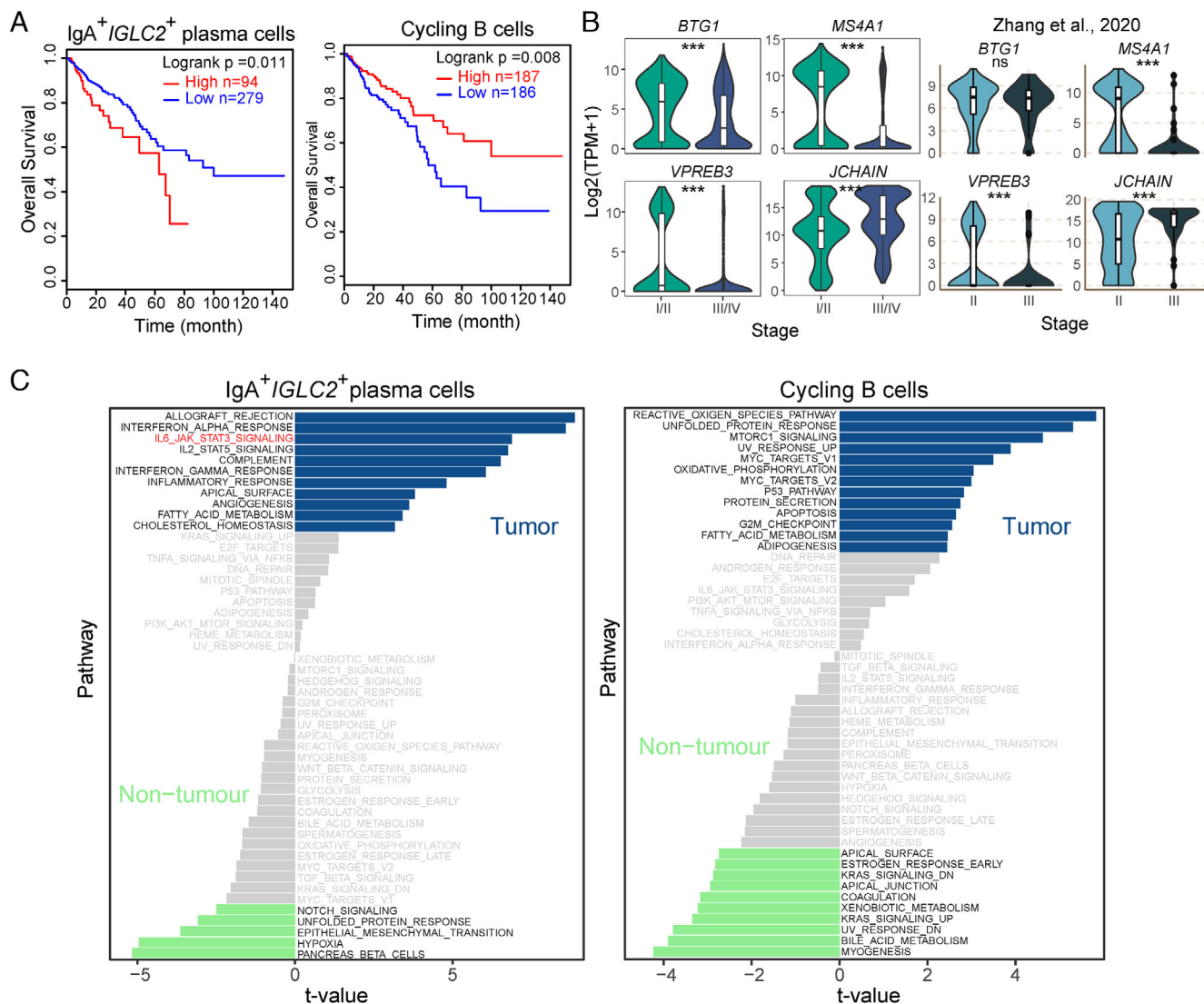
In summary, we provided a landscape of B lymphocytes in CRC, and found the different cell components in different regions. Furthermore, we found a set of BCR light chain variable region genes expressed differentially in the tumor region, which should help to gain deeper insight into the humoral immunity and the relationship between gastrointestinal microbiota and tumor progress for CRC.

## 2.4 | Various B-cell subtypes execute different antitumor responses in CRC

Infiltrating B cells are commonly observed in a variety of tumor tissues,<sup>33–35</sup> yet their reported correlation with patient outcomes has been inconsistent.<sup>34,36,37</sup> For example, the infiltration of *MS4A1*<sup>+</sup> B cells and plasma cells in tumor lesion was related to favorable prognosis in breast cancer, non-small cell lung carcinoma, melanoma, and CRC.<sup>30,38–40</sup> However, other studies have shown that increased infiltration of *MS4A1*<sup>+</sup> B and plasma cells is associated with poor prognosis in epithelial ovarian cancer and invasive ductal breast cancer<sup>41,42</sup> We hypothesized that this inconsistency may be due to the presence of distinct B-cell subpopulations performing distinct biological functions masked by bulk analysis. To reveal the association of different B-cell subtypes with CRC prognosis, the features of B-cell subtypes were quantified, and the relationship between these features and the CRC prognosis was analyzed in a CRC cohort from TCGA. The results showed that CRC patients with higher levels of IgA<sup>+</sup>IGLC2<sup>+</sup> plasma cells tended to have poor prognoses (Figure 5A and Figure S3D). Conversely, high cycling B-cell (Cluster\_4) scores were associated with prolonged survival. These results indicate that B-cell subpopulations may indeed have distinct roles in the antitumor response in CRC.

We next compared the expression profiles of tumor-derived and normal tissue-derived B-cell subsets IgA<sup>+</sup>IGLC2<sup>+</sup> plasma and cycling B cells. For IgA<sup>+</sup>IGLC2<sup>+</sup> plasma cells, allograft rejection, interferon-alpha response, interferon-gamma response, and inflammatory response were enhanced in the tumor-derived cells, indicating the involvement in promoting inflammatory responses in tumors. However, angiogenesis and activation of IL-6-mediated JAK/STAT3 signaling was also observed in these cells. According to a previous study, activation of angiogenesis in tumor-derived cells has been shown to be associated with poor CRC prognosis.<sup>43</sup> In addition, The IL-6-mediated JAK/STAT3 signaling pathway is closely related to the formation of inflammatory bowel disease and diverse human solid tumors, including CRC.<sup>47,48</sup> For cycling B cells, stress-related responses including reactive oxygen species pathways, unfolded protein responses, and ultraviolet responses were activated in the tumor-derived cells, and so was the enhancement of cell proliferation-related pathways (Figure 5C).

tissues from validation cohort. E, Frequency of B cells expressing immunoglobulin (Ig) light chain variable genes in B-cell population. Only 56 genes having expression across 20 B cells (1%) were selected to compare the frequency between tumor and nontumor region. F, The calculated enrichment score in tumor of 56 light chain variable genes in (E). Enrichment score in tumor = [frequency in tumor region] – [frequency in nontumor region]. NS,  $P > 0.05$ ; \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ , as determined by Student's *t*-test.



**FIGURE 5** Characterization of B-cell subpopulations in colorectal adenocarcinoma (CRC). **A**, The Kaplan-Meier overall survival curves of TCGA CRC patients grouped by the mean expression of feature of  $IgA^+ IGLC2^+$  plasma cells and cycling B cells. **B**, Violin plot comparing the gene expression of B cells from early and advanced CRC tumors from discovery cohort (left panel) and Zhang et al. (right panel). NS,  $P > 0.05$ ; \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ , as determined by Student's  $t$ -test. **C**, Differences in pathway activities between tumor and nontumor B-cell subtypes as scored per cell by GSEA. The  $t$ -values from a linear model corrected for patient of origin are shown; dn, down; UV, ultraviolet; v1, version 1; v2, version 2.

To explore the association between TNM-stage and B-cell activity, we conducted differential expression analysis for B cells from early and advanced CRC tumors (Figure S3E). An antiproliferative protein gene *BTG1* was highly expressed in infiltrating B cells of the early CRC tumor (Figure 5B). Pre-B lymphocyte protein 3 gene (*VPREB3*), which is thought to be involved in B-cell maturation and development,<sup>41</sup> was also upregulated in early CRC tumor (Figure 5B). Moreover, the B-cell marker *MS4A1* was upregulated, and the plasma cell marker *JCHAIN* was downregulated in early CRC tumor (Figure 5B). *TCL1A* and *BTG1*, which are highly expressed in B cells from early CRC tumor (Figure 5B and Figure S3E),

are reported as tumor suppressors in cervical cancer and CRC, respectively<sup>44–46</sup>. In addition, plasma cells were more abundant in advanced CRC and vice versa (Figure S3C). These results indicate that B cells from early CRC tumors are pre-B like cells and may have antitumor capabilities, while B cells from advanced CRC tumors fully develop into plasma cells.

Taken together, these analyses of the B-cell subpopulations in CRC expand our understanding of B-cell activities during the progression of CRC. Prognostic correlation analysis revealed a link between two B-cell subsets and prognosis, which provides a useful clue for the clinical treatment of CRC.

## 2.5 | Transitions in cell-cell interactions in the CRC microenvironment reshape immune capability

The immune microenvironment formed by the interaction between immune cells affects the occurrence and progression of tumor cells.<sup>43</sup> We used CellPhoneDB<sup>47</sup> to analyze the interaction between immune cells in CRC (See Section 4). By comparing the interaction network between tumor tissues and nontumor tissues, we found that the communication of myeloid cells with other immune cells was significantly increased in tumor tissues (Figure 6A,C). This suggests that myeloid cells may act as a hub for regulating the immune responses in CRC tumors. Interactions of IgA<sup>+</sup>IGLC2<sup>+</sup> plasma cells,  $\gamma\delta$  T cells, and cycling T with other immune cells also increased, while interactions of the CD4<sup>+</sup> T cells with the myeloid cells and most T cells decreased (Figure 6A,C). It can, therefore, be hypothesized that immunoregulatory functions in CRC primarily depend on the myeloid cells and B cells rather than on the CD4<sup>+</sup> regulatory T cells. Comparing cell-cell interactions for early and advanced CRC tumors indicates that cell-cell interactions were more complex, and the B lymphocytes interaction increased in advanced CRC tumors (Figure 6B). Notably, IgA<sup>+</sup>IGLC2<sup>+</sup> plasma cells, which were associated with poor prognosis in CRC, showed a significant interaction with the myeloid cells and cytotoxic T cells in advanced CRC tumors.

TILs sense tumor antigens and are activated by antigen-presenting cells to become the primary final effectors of tumor cell death.<sup>48</sup> The ability of T cells to fight tumors depends on their invasive site and the balance between costimulation and immune suppression.<sup>49</sup> To reveal how the state of T cells is regulated by other immune cells in CRC, we evaluated the expression of immune response, immune stimulation, inhibition, and chemokine-related genes of T cells in tumor and adjacent normal tissues, as well as the interactions between T cells and other immune cells (Figure S5A). In CRC tumor infiltration sites, CD28-CD80 costimulation signaling was observed in the myeloid cell-T cell interactions (Figure 6D). Interestingly, CD48-CD244 interactions were also enriched in the myeloid cell-T cell and B cell (B0 and B1)-T cell interactions. Previous studies have shown that CD244 is associated with T-cell exhaustion in tumors.<sup>50</sup> CD52 is a regulatory molecule that binds to SIGLEC10 and inhibits T-cell activation.<sup>51</sup> In our study, B cells (B0, B1, and B4) and myeloid cells expressing CD52 tended to interact with and inhibit various SIGLEC10<sup>+</sup> T cells (Figure 6D). The CCL8/CCR5 signaling axis is known to recruit CCR5<sup>+</sup> T cells.<sup>52</sup> In our study, IGLC2<sup>+</sup> plasma cells and cycling B cells (B3 and B4) expressing CCL8 interacted with CCR5<sup>+</sup> T cells in CRC (Figure 6D). These

results indicate that the myeloid cells exhibit both proimmune and immunosuppressive activities in CRC, and the IGLC2<sup>+</sup> plasma cells and cycling B cells in CRC can recruit CCR5<sup>+</sup> T cells to tumor lesions.

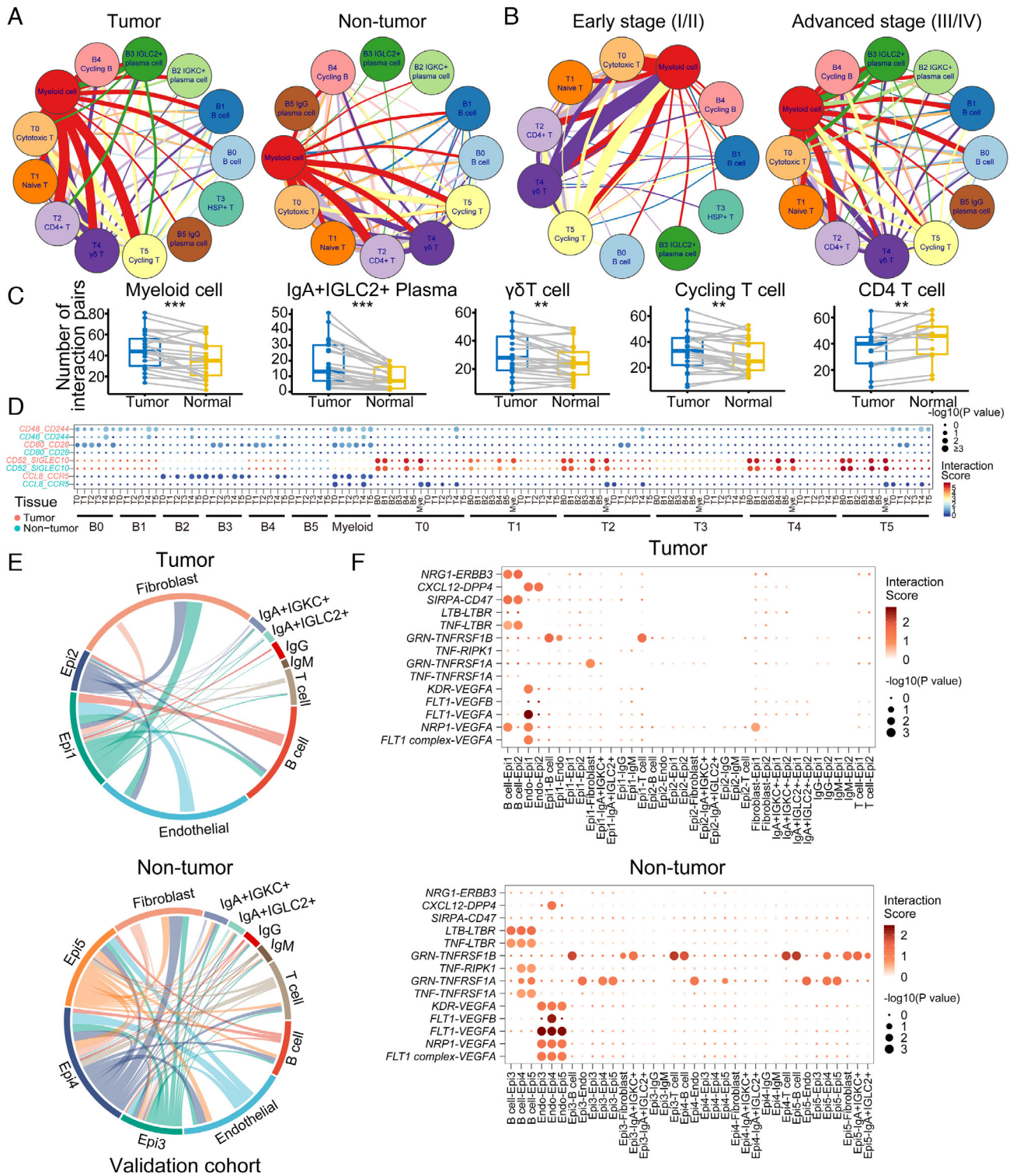
Moreover, nonimmune cells were obtained in validation cohort. To further investigate cell-cell interactions between immune/stroma cell and epithelial cells, we inferred the interactions between epithelial cells and immune cells as well as stroma cells. First of all, we grouped epithelial cells into five subtypes according to cell distribution in sampling site and their transcriptome profile (Figure S5B-D). Epi1 and Epi2, mainly from tumor tissue, show lower expression of metallothionein genes (Figure S5C). Previous study reported that metallothionein genes were downregulated in CRC tumor tissue.<sup>53</sup> Epi1 consisting of cluster 2 and 4 of epithelia also showed higher cell cycle, epithelial-mesenchymal transition (EMT), invasion, and metastasis score (Figure S5D), which indicated that Epi1 may be malignant cell. In tumor microenvironment, epithelial cells show fewer potential interactions with immune cells compared with normal tissue (Figure 6D). We found that endothelial cells from normal tissue would receive stronger angiogenic stimulatory signals from epithelia cells through VEGF and its receptor FLT1/NRP1/KDR. Epithelia cells in tumor show a lower activation of tumor necrosis factor signaling from B cells. Several interaction pairs were found only between epithelial cells and B cells in tumor tissue, including SIRPA-CD47 and NRG1-ERBB3 (Figure 6E), which were related to immune escape and EMT-related metastasis. In total, epithelial cells may regulate endothelial cells to form the microvessel in nontumor tissue, and B cells in CRC may contribute to progress of tumor.

Taken together, these results suggest that myeloid cells and B cells play a critical role in CRC immune regulation. The antitumor response of T cells in CRC tumors was activated and conversely attenuated by myeloid cells in the TME. This reflects “accelerator” and “brake” mechanisms of the immune system in CRC.

## 3 | DISCUSSION

It has been widely accepted that TME, and in particular its immune response, is crucial for the modulation of tumor progression and responsiveness to therapy.<sup>4</sup> Here, we present a comprehensive single-cell transcriptomic atlas of 5,345 immune cells isolated from multiple sites of eight CRC patients, providing a rich resource for understanding the multidimensional characterization of immune cells in CRC.

T and B lymphocytes are the predominant infiltrating cell types in the CRC immune ecosystem. In this study, we



**FIGURE 6** Cell-cell communication in the colorectal adenocarcinoma (CRC) tumor microenvironment. **A** and **B**, Network of potential cell interactions in tumor and normal tissues (**A**) as well as in early and advanced CRC tumors (**B**). Width of lines indicates the number of ligand-receptor pairs between the indicated cell types (interactions with fewer than 20 ligand-receptor pairs are not shown). **C**, The difference of number of interaction pairs in different cell types between tumor and nontumor. NS,  $P > 0.05$ ; \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ , as determined by Student's *t*-test. **D**, Overview of selected ligand-receptor interaction pairs in tumor and nontumor tissues. *P*-value is indicated by the size of the circle; color indicates the interaction score, which refers to the mean total of all individual ligand-receptor partner average expression values. **E**, Network of potential interactions between epithelial cells and immune cells or stroma cells in validation cohort. Width of lines indicates the number of ligand-receptor pairs between the indicated cell types. **F**, Overview of selected ligand-receptor interaction pairs in tumor and nontumor tissues in validation cohort.

focused on the heterogeneity between different sampling sites. The distribution of different types of TILs varied between tumor and normal tissues. In particular, the proportion of the naïve T cells in tumor tissues was significantly lower than in normal tissues, whereas the cycling T cells exhibited the opposite distribution pattern, indicating that T cells in tumor tissues were more likely to be activated and be in the proliferation stage. Hepatic metastasis is the dominant metastatic site for patients with CRC.<sup>54</sup> We documented that the T cells in hepatic metastatic sites were highly abundant, considering the limitation of sample size, we validated the existence of HSP<sup>+</sup> T cells by immunofluorescence (Figure S2C). In addition, these HSP<sup>+</sup> T cells were characterized by the naïve and inactive features, as well as HSP genes, which were reported to be related to T-cell suppression in many studies.<sup>21,55</sup> Stimulating these hepatic metastasis-derived T cells from their naïve state to the activated state may be a potential strategy to treat CRC-associated liver metastatic disease. Also, in the present study, the tumor-derived cytotoxic T cells showed significantly enhanced expression of cholesterol homeostasis-related genes and higher expression of *ACAT1*. Therefore, blocking cholesterol synthesis enzymes (such as *ACAT1*) in the T cells in tumor tissues may facilitate treatment of CRC.

B cells are a dominant immune cell type in CRC that up to now have only been studied to a very limited extent, and with incompatible opinions on their role in tumor progression. In this study, we for the first time identified six subtypes of B cells in CRC and provided a comprehensive transcriptomic profile, and we suggested that the previous conflicting results on B-cell effects on tumors may be due to the diversity in the roles for the B-cell subtypes. Importantly, we identified *MS4A1*<sup>+</sup> B cells and plasma cells that expressed different Ig isotypes (IgA and IgG). IgG-secreting plasma cells were dominant in the tumor center and decreased progressively from the tumor center to normal adjacent tissue. Previous studies have shown that higher levels of ex vivo IgG responses to tumor-associated antigens were related to shorter recurrence-free survival (RFS) in breast cancer, whereas IgA levels were not significantly associated with RFS.<sup>56</sup> Beyond this, several studies have demonstrated that IgG was expressed in various cancer cell types and was involved in the development and growth of tumor cells including in CRC, LNCaP prostate cancer cells, breast, liver, and lung cancers.<sup>57–59</sup> These results imply that the blockage of IgG might be a targeted therapy of CRC. Two B-cell subpopulations (IgA<sup>+</sup>IGLC2<sup>+</sup> plasma cells and cycling B cells) were found to have opposite effects on CRC prognosis, as the former correlated with poor patient

survival, whereas the latter correlated with favorable patient survival. According to cell-cell communication analysis, the highly proliferative *IGLC2*<sup>+</sup> plasma cells and cycling B cells were associated with a better prognosis and may recruit *CCR5*<sup>+</sup> T cells via *CCL8* in CRC tumors.

As for the myeloid cells, these cells were shown to be involved in T-cell exhaustion and activation through different ligand-receptor interactions, such as CD48-CD244 and CD28-CD80. This may also be attributed to the heterogeneity of myeloid cells. As for the mast cells, we found that they accumulated more in advanced CRC tumors, which supports the previous finding that mast cells promote CRC growth.<sup>18</sup> However, the role of mast cells and their interaction with colon cancer cells is still poorly understood. Further study on mast cells will contribute to our understanding of its function and immune mechanism in tumors.

Using the validation cohort, we elucidated the intercellular interactions between epithelial cells and immune and other stroma cells in the TME of CRC. Angiogenic stimulatory signals were found to be higher in endothelial cells from normal tissue, which indicates that vascular formed around tumor may facilitate metastasis of tumor cells. The *SIRPA*-CD47, *CXCL12*-DPP4, and *NRG1-ERBB3* interactions were found between *MS4A1*<sup>+</sup> B cells and epithelial cells in tumor. It has been suggested that blockade of the *SIRPα*-CD47 checkpoint may provide a potential new way to treat cancer.<sup>60</sup> In addition, *CXCL12* inhibits DPP4 and accelerates EMT and metastasis in breast cancer,<sup>41</sup> and the inhibition of *ERBB3* signaling suppresses EMT of hepatocellular carcinoma.<sup>42</sup> The B cells in CRC may also facilitate immune escape and metastasis of tumor, and this provides a potential target for the treatment of CRC.

Our study suggests that the antibody-producing cells in TME of CRC were mainly B cells and plasma cells. This is consistent with increasing evidences that link the gastrointestinal microbiota and the CRC progression.<sup>61–63</sup> Our single-cell analyses provide a new perspective for the relationship, as we profiled the BCR in both tumor and non-tumor regions for CRC patients. While we are unable to assign specific BCRs to the antigens from gastrointestinal microbiota, it is reasonable to speculate that the TME of CRC is coconstructed by gastrointestinal microbiota and tumor. The relationship between the immune repertoire of CRC and the gastrointestinal microbiota needs to be further explored, which is essential for the development of novel immune therapies for CRC.

In summary, our study highlights the diverse phenotypes of immune cells, in particular the B cells, in the CRC microenvironment. This immune-cell atlas adds to the resources for identifying clinically relevant predictive markers for immunotherapy.

## 4 | METHODS

### 4.1 | CRC patients and tissue samples

This study was approved by the local ethics committee of Guangdong Academy of Medical Sciences and General Hospital (License No. 2017233H [R2]) and complied with all relevant ethical regulations. All patients signed informed consent forms before recruitment.

Fourteen male and four female patients who were pathologically diagnosed with CRC at the Guangdong Academy of Medical Sciences and General Hospital were enrolled in this study. Their ages ranged from 30 to 83 years, with a median age of 61. Among these patients, two were diagnosed at stage I, four at stage II, nine at stage III, and four at stage IV. One patient (NR01) with rectal cancer had received neoadjuvant chemoradiation therapy before tumor resection, all other patients had not received radiotherapy or chemotherapy prior to tumor resection. The available clinical characteristics of these patients are summarized in Table S1.

Human tissue specimens were collected at the time of surgery resection under the supervision of a qualified pathologist. For each patient, three types of fresh tissues were collected during the operation, including primary tumor tissue, adjacent noncancerous tissue (to the brim of matched tumor 3-5 cm), and DNT (to the brim of matched tumor  $\geq 10$  cm). Then, two representative areas of the tumor center (CT) and the invasive margin (MT) were selected and cut from the primary tumor tissue. For patient LC02 and NLC08, liver metastases (Cm) and adjacent noncancerous tissues (Pm, to the brim of matched tumor 3 cm) of the liver were also collected. All collected samples were kept on ice-cold RPMI-1640 medium (Invitrogen) before processing.

### 4.2 | Preparation of single-cell suspensions

Two methods were used to prepare single cell suspension.

For patients RC01, RC02, RC03, R01, R02, R03, LC01, LC02, NR01, NLC03, and NLC04, cancerous tissues and noncancerous tissues, collected from the primary surgical or liver metastatic specimens, were mechanically and enzymatically disaggregated into single-cell suspensions, following previously published methods.<sup>64,65</sup> Briefly, after the tissue dissection, each tissue was washed with HPBS twice and cut into small pieces (0.2-0.5 mm<sup>3</sup>) with scissors, resuspended in RPMI-1640 medium (Invitrogen) containing 10% fetal bovine serum (FBS; GIBCO) and digested into single-cell suspensions with 1 mg/mL type I collagenase (GIBCO, #17100017) and 0.5 mg/mL type

IV collagenase (GIBCO, #17104019) at 37°C for 2 hours, flick every 30 minutes. After dissociation, cell suspensions were serially filtered with a 40  $\mu$ m cell strainer (BD) and centrifuged at 500  $\times$  g for 5 minutes, the supernatant was discarded and the cell pellet was resuspended with 1 mL of freezing buffer (90% FBS with 10% DMSO). Cells were then transferred into a Nalgene Mr Frosty Cryo 1°C freezing container (Thermo Fisher Scientific, #5100-0001) and placed at -80°C for 12 hours before being transferred to liquid nitrogen for storage.

For the other patients, cancerous tissues and noncancerous tissues were separately cut into approximately 1-2 mm<sup>3</sup> pieces in the DMEM medium (Biosharp) with 10% FBS (GIBCO), and enzymatically digested with MACS tumor dissociation kit (Miltenyi Biotec) for 30 minutes on a rotor at 37°C, according to manufacturer's instruction. After dissociation, cell suspensions were serially filtered with a 40  $\mu$ m cell strainer (BD) and centrifuged at 400  $\times$  g for 8 minutes, then the supernatant was discarded and the cell pellet was resuspended in red blood cell lysis buffer (Solarbio) and incubated on ice for 2 minutes to lyse red blood cells. After washing twice with 1 $\times$  PBS (Invitrogen), the cell pellets were resuspended in 2 mL 1 $\times$  PBS containing 0.04% bovine serum albumin (Sangon Biotech, A600903). Quantification of cell yields was performed by both Trypan blue dye exclusion staining counted with a hemocytometer and a handheld automated cell counter based on the Coulter principle (Scepter 2.0, Millipore). The final cell concentration was adjusted to 1,000 cells/ $\mu$ L for single-cell library preparation.

### 4.3 | Single-cell RNA-seq process for CRC patients (Smart-seq2)

For Patient RC01, RC02, RC03, R01, R02, R03, LC01, and LC02, single-cell RNA-seq were performed by the Smart-seq2 platform. The cryopreserved single-cell suspensions from the above eight patients were placed in a 37°C water bath for rapid recovery and washed twice with 1 $\times$  PBS, then single-cell suspensions were stained with antibodies against CD45 (antihuman CD45, BD Biosciences) for FACS sorting. Single cells of different subtypes of immune cells were sorted into 72  $\times$  72-microwell chips (WaferGen Biosystems) with lysis buffer (10% Triton X-100 0.5 nL, 40 U/ $\mu$ L RNase inhibitor 1.25 nL, 10  $\mu$ M Oligo-dT primer 12.5 nL, 10 mM dNTP mix 12.5 nL, and spike-in RNAs 10 nL, nuclease-free water 13.25 nL) dispensed into every microwell on the chip.

We prepared single-cell transcriptome amplifications by the MIRALCS method following a modified Smart-seq2 protocol.<sup>66</sup> The External RNA Controls Consortium (ERCC) spike-in mRNAs (1:50,000, Ambion, Life Tech-



nologies) were added into the lysis buffer as the exogenous spike-in control before the reverse transcription. We quantified the amplified cDNA products of each single cell by SmartChip real-time PCR with the values of cycle threshold, melting temperature, and Agilent Bioanalyzer 2100. The single-cell samples with high quality after amplification were extracted by an automatic extractor from the chip to a 96-well plate and then used for library construction. We prepared the libraries according to the MIRALCS method,<sup>66</sup> but cyclized the libraries into ssDNA libraries at the last step. Each single-cell sample was labeled with a barcode. All single-cell samples were sequenced by a BGISEQ500 sequencer with 100-bp single-end reads.

#### 4.4 | Single-cell RNA-seq process for CRC patients (DNBelab C4)

For Patients NRC01, NRC02, NR01, NLC03, NLC04, NLC05, NLC06, NLC07, NLC08, NLC09, and NLC10, single-cell RNA-seq were performed by the DNBelab C4 platform. The DNBelab C Series Single-Cell Library Prep Set (MGI, #1000021082) was utilized as previously described.<sup>17</sup> In brief, single-cell suspensions were used for droplet generation, emulsion breakage, beads collection, reverse transcription, and cDNA amplification to generate barcoded libraries. Indexed single-cell RNA-seq libraries were constructed according to the manufacturer's protocol. The sequencing libraries were quantified by Qubit™ ssDNA Assay Kit (Thermo Fisher Scientific, #Q10212). DNA nanoballs (DNBs) were loaded into the patterned nano arrays and sequenced on the ultra-high-throughput DIPSEQ T1 sequencer using the following read length: 30 bp for read 1, inclusive of 10 bp cell barcode 1, 10 bp cell barcode 2, and 10 bp unique molecular identifier (UMI), 100 bp of transcript sequence for read 2, and 10 bp for sample index.

#### 4.5 | Comparing IgG<sup>+</sup> plasma cells from different tissue sites by flow cytometry

Single cells from fresh primary tumors, adjacent non-cancerous tissue, and distal noncancerous tissue were obtained. After blocking with Fc Receptor Blocking Solution (Biolegend Cat #422301), cell surface staining was performed in FACS buffer containing antibody cocktails (anti-CD45, anti-CD38, anti-CD19) on ice for 1 hour. For detection of total IgG, cells were stained for both surface IgG and intracellular IgG. The intracellular IgG was strained by using an Intracellular Fixation and Permeabilization Buffer Set (BD Biosciences Cat #558126) according to the manufacturer's protocols.

#### 4.6 | HE and IF staining of hepatic metastatic tumor in CRC patients

Human tissue specimens were provided by Guangdong Academy of Medical Sciences and General Hospital. The specimens were collected within 30 minutes after the tumor resection and fixed in formalin for 48 hours. Dehydration and embedding in paraffin was performed following routine methods. Specimen slices were cut from paraffin-embedded tissue blocks with a microtome, deparaffinized and rehydrated. For HE staining, the slices were sequentially immersed in HE. For IF staining, the slices were further processed, including antigen retrieval, blocking of endogenous peroxidase, primary antibody incubation (included CD3 epsilon antibody [1:200, Affinit #DF6594] and HSP70 antibody [1:200, Affinit #AF5466]), secondary antibody incubation (included Alexa Fluor 647-labeled goat anti-rabbit IgG [H+L], [1:200, Beyotime #A0468], and FITC-labeled goat anti-rabbit IgG [H+L] [1:200, Beyotime #A0562]), and nuclear staining (DAPI). Finally, the slices were sealed with mounting medium for imaging. Images were taken on a LSM800 (Carl Zeiss) confocal microscope with 100× oil immersion lens.

#### 4.7 | Acquisition of single-cell gene expression matrices and cell clustering

Raw reads were cleaned using Cutadapt (Version 1.15) and mapped to hg38 using STAR (Version 020201). Gene expression levels of each cell were quantified using RSEM (Version 1.3.0) and combined in R (Version 3.5.1). Cell clustering was accomplished by the Seurat R package (Version 3.0.1).<sup>67</sup> Cells with less than 500 genes (TPM > 1) or over 20% TPM derived from the mitochondrial genome were removed, after which 5345 cells remained. Only protein-coding, TCR, and IG genes were used for clustering, and genes expressed (TPM > 1) in less than three cells were discarded. TPM matrices from each patient were log-normalized, and 2,000 variably expressed genes were selected using Seurat's FindVariableFeatures function. To remove batch effects, TPM matrices of different patients were integrated by Anchors, using the FindIntegrationAnchors and the IntegrateData functions from the Seurat R package. To reduce dimensionality, variably expressed genes were summarized by principle component analysis. The *t*-distributed Stochastic Neighbor Embedding algorithm loses global structure due to a focus on local information,<sup>68</sup> and as such we chose the Uniform Manifold Approximation and Projection (UMAP) algorithm to visualize the data.<sup>69</sup> The top 10 principle components (PCs) were used as input to the RunUMAP function with default settings. The top 10 PCs were used as inputs in the

FindClusters function with the “resolution” parameter set to 1.0 to find clusters. Cell clusters in the resulting two-dimensional representation were annotated to known biological cell types using canonical marker genes (Figure S1A). Of note, very few lymphocytes (~4%) were positive for cell proliferation markers. We therefore opted not to correct our gene expression matrices for effects of cell cycle. To better identify T-cell and B-cell subtypes, we extracted T cells and B cells and ran Seurat, respectively. Parameters were the same as introduced previously, except resolution was adjusted to 0.6 for T cells.

For DNBelab C4 data, raw sequencing reads were filtered and demultiplexed using PISA (Version 439 0.2; <https://github.com/shiquan/PISA>). The filtered reads were aligned to hg38 genome using STAR and sorted by sambamba (Version 441 0.7.0). Cell versus gene UMI count matrix was generated using PISA. Cells with less than 500 genes (UMI > 1) or over 25% UMI derived from the mitochondrial genome were removed, after which 9,770 cells remained. The clustering parameter was similar to Smart-seq data, with PC usage set to 20.

We used harmony (<https://github.com/immunogenomics/harmony>) to integrate the data from two platforms with 20 PCs.

#### 4.8 | Identifying marker genes

We used the Seurat’s FindAllMarkers function to identify marker genes for each cluster. This function contrasted cells from each cluster to all other cells of that cluster. The “RNA” assay and the “Wilcox” test were used for the function. Marker genes found by the FindAllMarkers function were required to have an average expression in that cluster that was >1.5-fold higher than the average expression in the other clusters, and a detectable expression in >15% of cells from that cluster.

Marker genes of main cell types (Table S2), T-cell subtypes (Table S3), and B-cell subtypes (Table S4) were identified separately as described above.

#### 4.9 | Cell developmental trajectory

Cell lineage trajectory of T cells was inferred by using Monocle2.<sup>22</sup> We first used the “relative2abs” function in Monocle2 to convert TPM into normalized mRNA counts and created an object with parameter “expressionFamily = negbinomial.size” following the Monocle2 tutorial. We used the “differentialGeneTest” function to derive differentially expressed genes from each cluster and genes with  $q$ -value <  $1e-5$  were used to order the cells in a pseudotime analysis. After the cell trajectories were constructed,

differentially expressed genes along the pseudotime were detected using the “differentialGeneTest” function.

#### 4.10 | TCR analysis

We used the TraCeR<sup>70</sup> method to assemble TCR sequences for T cells. TraCeR can identify rearranged TCR chains and calculate their TPM values. After assembly, we arranged the productive TCR chains of every T cell by their TPM values. For example, if two TCR $\alpha$  chains were assembled in one single cell and were both productive, the chain with the higher TPM was defined as TCR $\alpha$ 1, while the chain with the lower TPM being defined as TCR $\alpha$ 2. Nonproductive TCR chains were excluded. The same arrangement was deployed on TCR $\beta$ . We kept cells with at least one pair of productive TCR $\alpha$  and TCR $\beta$  chain for subsequent analysis. We used a strict standard to define TCR clones: cells with the same TCR $\alpha$ 1 and TCR $\beta$ 1 were considered to be one TCR clone, while the expanded clones were defined as clones with at least two cells sharing the same TCR $\alpha$ 1 and TCR $\beta$ 1 in a given cell population.

#### 4.11 | GSVA and differential activities of pathways

Pathway analyses were predominantly performed on the 50 hallmark pathways described in the molecular signature database. To assign pathway activity estimates to individual cells, we applied GSVA<sup>71</sup> using standard settings, as implemented in the GSVA package (Version 1.30.0). Differential activated pathways between cells from tumor and nontumor tissues were tested using the generalized linear model from the Limma package (Version 3.38.3), and  $P$ -values were Benjamini-Hochberg-corrected. Pathways with adjusted  $P$ -value < .05 were considered to be significantly differentially activated. Tumor features, cell cycle, EMT, invasion, and metastasis, were quantified using genes involved in these pathways by GSVA (Table S5).

#### 4.12 | Survival analysis

The TCGA provisional data were used to evaluate the prognostic effect of gene sets derived from specific cell clusters. The provisional gene expression and survival data of the TCGA were accessed using CBioPortal. We used Limma to perform differential expression gene analysis for each cluster and defined genes with more than two-fold increase in expression in one cluster compared to the other clusters as “signature genes.” For the signature gene sets of each cluster, we calculated a “gene signature score” for

each patient using fold-change values for each gene in the signature to weigh genes in calculations of the average. Then, patients were grouped into high- and low-expression groups by the median value of the “gene signature score.” To correct for clinical covariates including age and histological grade, we performed multivariable analyses using Cox proportional hazard survival models to obtain the hazard ratio and adjusted *P*-value.

#### 4.13 | Cell-cell interaction analysis

To analyze cell-cell interactions between different cell types, we used CellPhoneDB<sup>47</sup> to identify significant ligand-receptor pairs within early and advanced CRC. The cell type-specific receptor-ligand interactions between cell types were identified based on specific expression of a receptor by one cell type and a ligand by another cell type. The interaction score refers to the mean total of all individual ligand-receptor partner average expression values in the corresponding interacting pairs of cell types. The expression of any complex's output by CellPhoneDB was calculated as the sum of the expression of the component genes.

#### 4.14 | Statistics and reproducibility

Box plots were generated using the ggplot2 package (version 3.2.0) and default parameters. Hence, the boxes span the interquartile range (IQR; from the 25th to the 75th percentiles), with the centerline corresponding to the median. The upper whisker extends from the hinge to the largest value no further than  $1.5 \times$  IQR from the hinge (where IQR is the distance between the first and third quartiles). The lower whisker extends from the hinge to the smallest value at most  $1.5 \times$  IQR of the hinge. Data beyond the end of the whiskers are called “outlying” points and are plotted individually. Each data point is displayed in the box plots.

Comparisons of cell ratios between two groups were performed using unpaired two-tailed *t*-tests.

#### ACKNOWLEDGMENTS

We thank the eighteen colorectal cancer patients who participated in this study at Guangdong Provincial People's Hospital. We thank Xiaoyun Huang for assistance with the single-cell RNA-seq. We thank Shishang Qin and Miaomiao Jiang for their assistance with data analysis. This work was supported by the Fundamental Research Funds for the Central Universities of China (No. X2SWD2172910). Xueqing Yao was supported by grants from the Science and Technology Planning Project of Guangdong Province, China (No. 2017A0300223006), Science and Technology

Project of Guangzhou, China (No. 201704020077), Science and Technology Planning Project of Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences, No. DFJH201913), and CSCO-Roche Cancer Research Foundation (No. Y-2019Roche-190). Fei Ling was supported by grants from the Science and Technology Planning Project of Guangdong Province, China (No. 2019A030317013) and Science and Technology Program of Guangzhou, China (No. 201804010318). Shiping Liu and Yong Hou were supported by Natural Science Foundation of Guangdong Province, China (No. 2018A030313379), Shenzhen Key Laboratory of Single-Cell Omics (No. ZDSYS201909020936138 31), and Science, Technology and Innovation Commission of Shenzhen Municipality under grant No. GJHZ20180419190827179. Guangdong-Hong Kong-Macao-Joint Labs Program from Guangdong Science and Technology (No. 2019B121205005).

#### AUTHOR CONTRIBUTIONS

Xueqing Yao, Zhanglin Lin, Fei Ling, and Shiping Liu jointly supervised the research. Fei Ling, Xueqing Yao, Shiping Liu, Liang Wu, Jianhua Yin, and Wei Wang designed the experiments. Xueqing Yao recruited patients and provided clinical samples from CRC patients. Wei Wang, Yueer Lu, Jiarui Xie, Hang Yu, Shanshan Wang, Chunqing Wang, Yuanhang Zhang, Zhiqiang Jiang, Yilin Huang, Chongyin Han, Zhenggang Zhong, Jialin Hu, Ying Ouyang, Huisheng Liu, Chengzhi Huang, and Mengya Yu prepared the samples under the supervision of Lizhen Huang and Fei Ling. Wei Wang, Yueer Lu, Jiarui Xie, Xiaochan Wei, and Dandan Chen performed all scRNA-seq data acquisition experiments. Wei Wang, Yu Zhong, Zhenkun Zhuang, Yan Sun, and Jianhua Yin analyzed and interpreted the data. Wei Wang, Yu Zhong, Zhenkun Zhuang, and Yan Sun wrote the manuscript. Zhanglin Lin, Shiping Liu, Wei Wang, Yu Zhong, Zhenkun Zhuang, Chengzhi Huang, and Fei Ling revised the manuscript. All the authors discussed the results and commented on the manuscript.

#### CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

#### DATA AVAILABILITY STATEMENT

The data reported in this study are available in the CNGB Nucleotide Sequence Archive (CNSA: <https://db.cngb.org/cnsa>; accession number: CNP0000916).

#### ORCID

Yu Zhong  <https://orcid.org/0000-0001-6902-2732>

Yan Sun  <https://orcid.org/0000-0002-5080-0781>

Xueqing Yao  <https://orcid.org/0000-0003-0570-7585>

## REFERENCES

- Galon J, Costes A, Sanchez-Cabo F, et al. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science*. 2006;313:1960-1964.
- Schulz M, Salamero-Boix A, Niesel K, Alekseeva T, Sevenich L. Microenvironmental regulation of tumor progression and therapeutic response in brain metastasis. *Front Immunol*. 2019;10:1713.
- Vesely MD, Kershaw MH, Schreiber RD, Smyth MJ. Natural innate and adaptive immunity to cancer. *Annu Rev Immunol*. 2011;29:235-271.
- Binnewies M, Roberts EW, Kersten K, et al. Understanding the tumor immune microenvironment (TIME) for effective therapy. *Nat Med*. 2018;24:541-550.
- Zhang L, Li Z, Skrzypczynska KM, et al. Single-cell analyses inform mechanisms of myeloid-targeted therapies in colon cancer. *Cell*. 2020;181:442-459.e429.
- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68:394-424.
- Arnold M, Sierra MS, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global patterns and trends in colorectal cancer incidence and mortality. *Gut*. 2017;66:683-691.
- Zhang L, Yu X, Zheng L, et al. Lineage tracking reveals dynamic relationships of T cells in colorectal cancer. *Nature*. 2018;564:268-272.
- Pages F, Mlecnik B, Marliot F, et al. International validation of the consensus immunoscore for the classification of colon cancer: a prognostic and accuracy study. *Lancet*. 2018;391:2128-2139.
- Bindea G, Mlecnik B, Tosolini M, et al. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity*. 2013;39:782-795.
- Li H, Courtois ET, Sengupta D, et al. Reference component analysis of single-cell transcriptomes elucidates cellular heterogeneity in human colorectal tumors. *Nat Genet*. 2017;49:708-718.
- de Vries NL, van Unen V, Ijsselstein ME, et al. High-dimensional cytometric analysis of colorectal cancer reveals novel mediators of antitumour immunity. *Gut*. 2020;69(4):691-703.
- Zhang L, Zhao Y, Dai Y, et al. Immune landscape of colorectal cancer tumor microenvironment from different primary tumor location. *Front Immunol*. 2018;9:1578.
- Spencer J, Sollid LM. The human intestinal B-cell response. *Mucosal Immunol*. 2016;9:1113-1124.
- Palm NW, de Zoete MR, Cullen TW, et al. Immunoglobulin A coating identifies colitogenic bacteria in inflammatory bowel disease. *Cell*. 2014;158:1000-1010.
- Zarour LR, Anand S, Billingsley KG, et al. Colorectal cancer liver metastasis: evolving paradigms and future directions. *Cell Mol Gastroenterol Hepatol*. 2017;3:163-173.
- Liu C, Wu T, Fan F, et al. A portable and cost-effective microfluidic system for massively parallel single-cell transcriptome profiling. *bioRxiv*. 2019. <https://doi.org/10.1101/818450>
- Yu Y, Blokhuis B, Derks Y, Kumari S, Garssen J, Redegeld F. Human mast cells promote colon cancer growth via bidirectional crosstalk: studies in 2D and 3D coculture models. *Oncoimmunology*. 2018;7:e1504729.
- Crawford A, Angelosanto JM, Kao C, et al. Molecular and transcriptional basis of CD4<sup>+</sup> T cell dysfunction during chronic infection. *Immunity*. 2014;40:289-302.
- Plitas G, Konopacki C, Wu K, et al. Regulatory T cells exhibit distinct features in human breast cancer. *Immunity*. 2016;45:1122-1134.
- de Zoeten EF, Wang L, Sai H, Dillmann WH, Hancock WW. Inhibition of HDAC9 increases T regulatory cell function and prevents colitis in mice. *Gastroenterology*. 2010;138:583-594.
- Trapnell C, Cacchiarelli D, Grimsby J, et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol*. 2014;32:381-386.
- Jin YB, Luo W, Zhang GY, et al. TCR repertoire profiling of tumors, adjacent normal tissues, and peripheral blood predicts survival in nasopharyngeal carcinoma. *Cancer Immunol Immunother*. 2018;67:1719-1730.
- Stubbington MJT, Lönnberg T, Proserpio V, et al. T cell fate and clonality inference from single-cell transcriptomes. *Nat Methods*. 2016;13:329-332.
- Yang W, Bai Y, Xiong Y, et al. Potentiating the antitumour response of CD8(+) T cells by modulating cholesterol metabolism. *Nature*. 2016;531:651-655.
- Saraon P, Trudel D, Kron K, et al. Evaluation and prognostic significance of ACAT1 as a marker of prostate cancer progression. *Prostate*. 2014;74:372-380.
- Allie SR, Bradley JE, Mudunuru U, et al. The establishment of resident memory B cells in the lung requires local antigen encounter. *Nat Immunol*. 2019;20:97-108.
- Cerutti A, Chen K, Chorny A. Immunoglobulin responses at the mucosal interface. *Annu Rev Immunol*. 2011;29:273-293.
- Chen Z, Gerhold-Ay A, Gebhard S, et al. Immunoglobulin kappa C predicts overall survival in node-negative breast cancer. *PLoS One*. 2012;7:e44741.
- Lohr M, Edlund K, Botling J, et al. The prognostic relevance of tumour-infiltrating plasma cells and immunoglobulin kappa C indicates an important role of the humoral immune response in non-small cell lung cancer. *Cancer Lett*. 2013;333:222-228.
- Kim N, Kim HK, Lee K, et al. Single-cell RNA sequencing demonstrates the molecular and cellular reprogramming of metastatic lung adenocarcinoma. *Nat Commun*. 2020;11:2285.
- Chen W, Liu F, Ling Z, Tong X, Xiang C. Human intestinal lumen and mucosa-associated microbiota in patients with colorectal cancer. *PLoS One*. 2012;7:e39743.
- Hu X, Zhang J, Wang J, et al. Landscape of B cell immunity and related immune evasion in human cancers. *Nat Genet*. 2019;51:560-567.
- Nelson BH. CD20+ B cells: the other tumor-infiltrating lymphocytes. *J Immunol*. 2010;185:4977-4982.
- Linnebacher M, Maletzki C. Tumor-infiltrating B cells: the ignored players in tumor immunology. *Oncoimmunology*. 2012;1:1186-1188.
- Donnem T, Al-Saad S, Al-Shibli K, Andersen S, Busund LT, Bremnes RM. Prognostic impact of platelet-derived growth factors in non-small cell lung cancer tumor and stromal cells. *J Thorac Oncol*. 2008;3:963-970.
- Wouters MCA, Nelson BH. Prognostic significance of tumor-infiltrating B cells and plasma cells in human cancer. *Clin Cancer Res*. 2018;24:6125-6135.

38. Ladanyi A, Kiss J, Mohos A, et al. Prognostic impact of B-cell density in cutaneous melanoma. *Cancer Immunol Immunother.* 2011;60:1729-1738.
39. Schmidt M, Bohm D, von Torne C, et al. The humoral immune system has a key prognostic impact in node-negative breast cancer. *Cancer Res.* 2008;68:5405-5413.
40. Berntsson J, Nodin B, Eberhard J, Micke P, Jirstrom K. Prognostic impact of tumour-infiltrating B cells and plasma cells in colorectal cancer. *Int J Cancer.* 2016;139:1129-1139.
41. Lundgren S, Berntsson J, Nodin B, Micke P, Jirstrom K. Prognostic impact of tumour-associated B cells and plasma cells in epithelial ovarian cancer. *J Ovarian Res.* 2016;9:21.
42. Mohammed ZM, Going JJ, Edwards J, Elsberger B, McMillan DC. The relationship between lymphocyte subsets and clinico-pathological determinants of survival in patients with primary operable invasive ductal breast cancer. *Br J Cancer.* 2013;109:1676-1684.
43. Smillie CS, Biton M, Ordovas-Montanes J, et al. Intra- and inter-cellular rewiring of the human colon during ulcerative colitis. *Cell.* 2019;178:714-730.e722.
44. Jung YY, Sung JY, Kim JY, Kim HS. Down-regulation of B-cell translocation gene 1 by promoter methylation in colorectal carcinoma. *Anticancer Res.* 2018;38:691-697.
45. Punt S, Corver WE, van der Zeeuw SA, et al. Whole-transcriptome analysis of flow-sorted cervical cancer samples reveals that B cell expressed TCL1A is correlated with improved survival. *Oncotarget.* 2015;6:38681-38694.
46. Zhang J, Dong W. Expression of B cell translocation gene 1 protein in colon carcinoma and its clinical significance. *Recent Pat Anticancer Drug Discov.* 2020;15:78-85.
47. Vento-Tormo R, Efremova M, Botting RA, et al. Single-cell reconstruction of the early maternal-fetal interface in humans. *Nature.* 2018;563:347-353.
48. Chen DS, Mellman I. Elements of cancer immunity and the cancer-immune set point. *Nature.* 2017;541:321-330.
49. Chen L, Flies DB. Molecular mechanisms of T cell co-stimulation and co-inhibition. *Nat Rev Immunol.* 2013;13:227-242.
50. Baitsch L, Baumgaertner P, Devevre E, et al. Exhaustion of tumor-specific CD8(+) T cells in metastases from melanoma patients. *J Clin Invest.* 2011;121:2350-2360.
51. Toh BH, Kyaw T, Tipping P, Bobik A. Immune regulation by CD52-expressing CD4 T cells. *Cell Mol Immunol.* 2013;10:379-382.
52. Halvorsen EC, Hamilton MJ, Young A, et al. Maraviroc decreases CCL8-mediated migration of CCR5(+) regulatory T cells and reduces metastatic tumor growth in the lungs. *Oncoimmunology.* 2016;5:e1150398.
53. Si M, Lang J. The roles of metallothioneins in carcinogenesis. *J Hematol Oncol.* 2018;11:107.
54. Sade-Feldman M, Yizhak K, Bjorgaard SL, et al. Defining T cell states associated with response to checkpoint immunotherapy in melanoma. *Cell.* 2018;175:998-1013.e1020.
55. Akyol S, Gercel-Taylor C, Reynolds LC, Taylor DD. HSP-10 in ovarian cancer: expression and suppression of T-cell signaling. *Gynecol Oncol.* 2006;101:481-486.
56. Garaud S, Zayakin P, Buisseret L, et al. Antigen specificity and clinical significance of IgG and IgA autoantibodies produced in situ by tumor-infiltrating B cells in breast cancer. *Front Immunol.* 2018;9:2660.
57. Niu N, Zhang J, Huang T, et al. IgG expression in human colorectal cancer and its relationship to cancer cell behaviors. *PLoS One.* 2012;7:e47362.
58. Xu Y, Chen B, Zheng S, et al. IgG silencing induces apoptosis and suppresses proliferation, migration and invasion in LNCaP prostate cancer cells. *Cell Mol Biol Lett.* 2016;21:27.
59. Qiu X, Zhu X, Zhang L, et al. Human epithelial cancers secrete immunoglobulin g with unidentified specificity to promote growth and survival of tumor cells. *Cancer Res.* 2003;63:6488-6495.
60. Veillette A, Chen J. SIRP $\alpha$ -CD47 immune checkpoint blockade in anticancer therapy. *Trends Immunol.* 2018;39:173-184.
61. Tilg H, Adolph TE, Gerner RR, Moschen AR. The intestinal microbiota in colorectal cancer. *Cancer Cell.* 2018;33:954-964.
62. Lu Y, Chen J, Zheng J, et al. Mucosal adherent bacterial dysbiosis in patients with colorectal adenomas. *Sci Rep.* 2016;6:26337.
63. Lucas C, Barnich N, Nguyen HTT. Microbiota, inflammation and colorectal cancer. *Int J Mol Sci.* 2017;18(6):1310.
64. Dalerba P, Kalisky T, Sahoo D, et al. Single-cell dissection of transcriptional heterogeneity in human colon tumors. *Nat Biotechnol.* 2011;29:1120-1127.
65. Dalerba P, Dylla SJ, Park IK, et al. Phenotypic characterization of human colorectal cancer stem cells. *Proc Natl Acad Sci U S A.* 2007;104:10158-10163.
66. Wu L, Zhang X, Zhao Z, et al. Full-length single-cell RNA-seq applied to a viral human cancer: applications to HPV expression and splicing analysis in HeLa S3 cells. *Gigascience.* 2015;4:51.
67. Stuart T, Butler A, Hoffman P, et al. Comprehensive integration of single-cell data. *Cell.* 2019;177:1888-1902.e1821.
68. Maaten Lvd, Hinton G. Visualizing data using t-SNE. *J Mach Learn Res.* 2008;9:2579-2605.
69. Becht E, McInnes L, Healy J, et al. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat Biotechnol.* 2019;37:38-44.
70. Stubbington MJT, Lonnberg T, Proserpio V, et al. T cell fate and clonality inference from single-cell transcriptomes. *Nat Methods.* 2016;13:329-332.
71. Hanzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics.* 2013;14:7.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**How to cite this article:** Wang W, Zhong Y, Zhuang Z, et al. Multiregion single-cell sequencing reveals the transcriptional landscape of the immune microenvironment of colorectal cancer. *Clin Transl Med.* 2020;11:e253. <https://doi.org/10.1002/ctm2.253>