

## CANCER

# Single-cell and spatial transcriptome analysis reveals the cellular heterogeneity of liver metastatic colorectal cancer

Fei Wang<sup>1,2†</sup>, Jie Long<sup>1†</sup>, Liang Li<sup>2</sup>, Zi-Xin Wu<sup>3</sup>, Tian-Tian Da<sup>4</sup>, Xiao-Qing Wang<sup>4</sup>, Chuan Huang<sup>4</sup>, Yi-Hua Jiang<sup>1</sup>, Xue-Qing Yao<sup>5</sup>, Hai-Qing Ma<sup>2\*</sup>, Zhe-Xiong Lian<sup>1,6\*</sup>, Zhi-Bin Zhao<sup>2\*</sup>, Jie Cao<sup>3\*</sup>

In this study, we comprehensively charted the cellular landscape of colorectal cancer (CRC) and well-matched liver metastatic CRC using single-cell and spatial transcriptome RNA sequencing. We generated 41,892 CD45<sup>-</sup> nonimmune cells and 196,473 CD45<sup>+</sup> immune cells from 27 samples of six CRC patients, and found that CD8<sup>+</sup> CXCL13 and CD4<sup>+</sup> CXCL13 subsets increased significantly in liver metastatic samples that exhibited high proliferation ability and tumor-activating characterization, contributing to better prognosis of patients. Distinct fibroblast profiles were observed in primary and liver metastatic tumors. F3<sup>+</sup> fibroblasts enriched in primary tumors contributed to worse overall survival by expressing protumor factors. However, MCAM<sup>+</sup> fibroblasts enriched in liver metastatic tumors might promote generation of CD8<sup>+</sup> CXCL13 cells through Notch signaling. In summary, we extensively analyzed the transcriptional differences of cell atlas between primary and liver metastatic tumors of CRC by single-cell and spatial transcriptome RNA sequencing, providing different dimensions of the development of liver metastasis in CRC.

## INTRODUCTION

Colorectal cancer (CRC) is the third most common malignant tumor worldwide (1). About 21 to 26% of patients present with synchronous metastatic diseases (2), in which 14.5 to 17.5% of patients develop liver metastases (3, 4). Metastatic CRCs are associated with poor survival rates, and CRC-derived liver metastasis (CRCLM) leads to a great clinical challenge (3).

The metastatic process is a multistep event that entails cancer cells to escape from the primary site, survive in circulation, seed at distant sites, and grow (5, 6). It is well established that metastatic spread is promoted by communication between cancer cells and stromal cells via secretion of cytokines, growth factors, and proteases that remodel the tumor microenvironment (TME) (7). The cellular compositions of TME is highly complex, including diverse populations of fibroblasts and immune cells that play important roles in cancer evasion, metastasis, and responses to treatment (8, 9). Cancer-associated fibroblast (CAF) is a predominant stromal cellular component in many types of malignancies including CRCs, breast cancers, and pancreatic cancers (10–12). Several subsets of CAFs with different functions have been identified in a variety of tumors. CD10<sup>+</sup>GPR77<sup>+</sup> CAFs can promote breast and

lung cancer chemoresistance and tumor formation by providing a survival niche for cancer stem cells (13). There are also antigen-presenting CAFs, which can activate CD4<sup>+</sup> T cells in an antigen-specific fashion (14). The tumor immune microenvironment (TIME) is composed of diverse populations of lymphocytes and myeloid cells. CD8<sup>+</sup> T cells have been proven as formidable “soldiers” in antitumor immune responses. The infiltration of CD8<sup>+</sup> T cells in solid tumors predicts a favorable prognosis (15). The immune milieu of primary CRCs has been described by single-cell RNA sequencing (scRNA-seq) (16, 17), but the global cellular landscape of CRCLM and the interaction network differences between primary and metastatic CRC are rarely reported.

To comprehensively chart the cellular landscape of CRC and matched liver metastasis, in this study, we generated high-resolution cell maps using scRNA-seq in conjunction with spatial transcriptome (ST) and immunohistochemistry (IHC) staining as well as flow cytometry. Different phenotypic profiles and highly variable frequencies of major cell types, such as CAF, CD8<sup>+</sup> T cells, and tumor-associated macrophages, were observed between CRC primary tumor and liver metastasis, which would indicate the molecular heterogeneity within TME and provide potential targets for future cancer therapy.

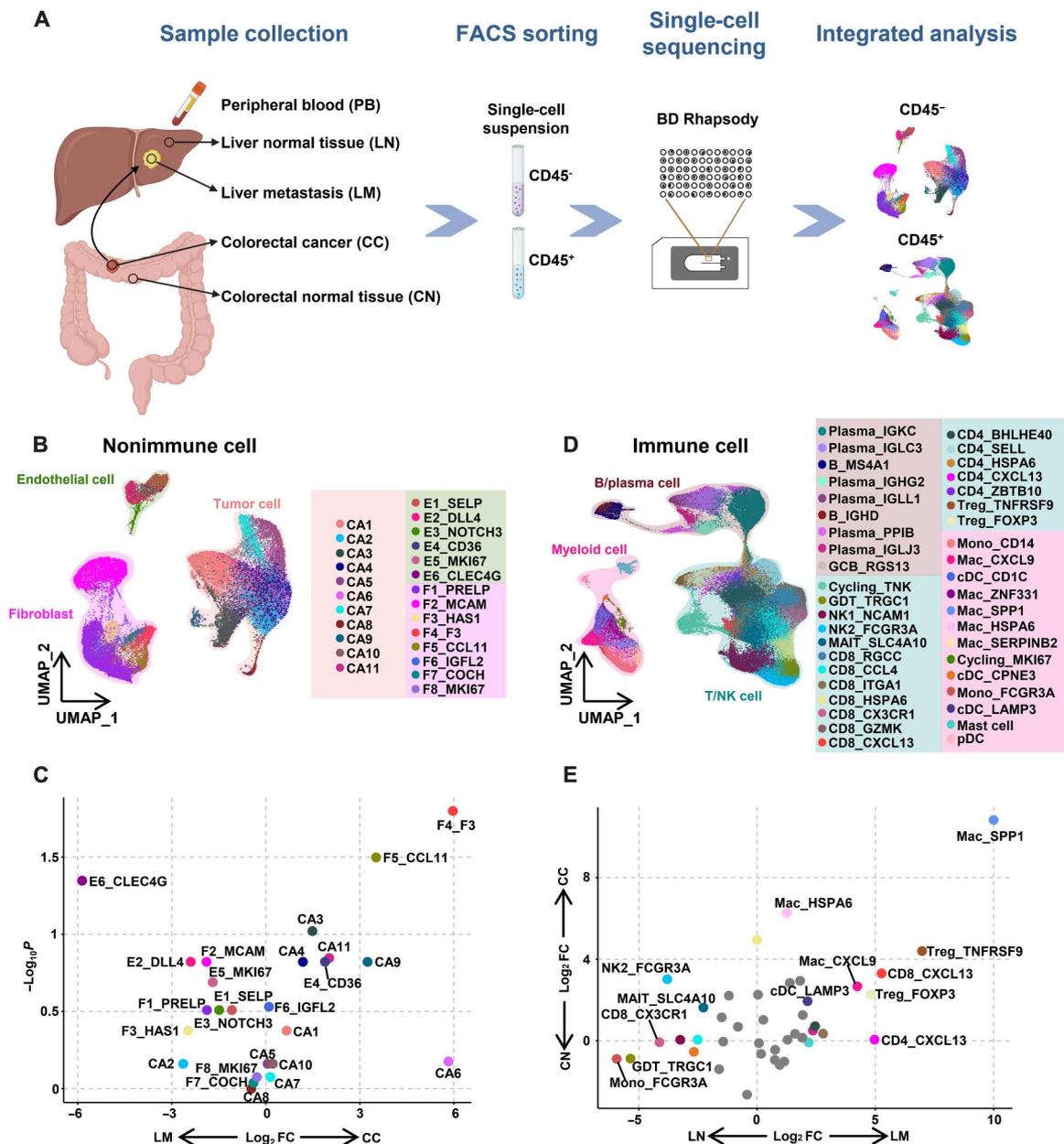
## RESULTS

### Global single-cell transcriptome map of CRC primary tumors and CRC-derived liver metastatic tumors

To elucidate the cellular heterogeneity of CRC primary tumors and CRC-derived liver metastatic tumors, CD45<sup>-</sup> nonimmune cells and CD45<sup>+</sup> immune cells were collected from primary colorectal cancer (CC), adjacent normal colorectal mucosa (CN), liver metastasis (LM), adjacent normal liver tissue (LN), and peripheral blood (PB) from six patients with CRC for single-cell transcriptome analysis (Fig. 1A). After quality control and filtration, we obtained

Copyright © 2023 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution NonCommercial License 4.0 (CC BY-NC).

<sup>1</sup>Guangdong Cardiovascular Institute, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, Guangzhou, Guangdong 510080, China. <sup>2</sup>Medical Research Institute, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, Guangzhou, Guangdong 510080, China. <sup>3</sup>Department of General Surgery, Guangzhou Digestive Disease Center, Guangzhou First People's Hospital, School of Medicine, South China University of Technology, Guangzhou, Guangdong 510080, China. <sup>4</sup>School of Medicine, South China University of Technology, Guangzhou, Guangdong 510006, China. <sup>5</sup>Department of General Surgery, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, Guangzhou, Guangdong 510080, China. <sup>6</sup>Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, Guangzhou, Guangdong 510080, China. \*Corresponding author. Email: zxljan1@ustc.edu.cn (Z.-X.L.); zzb@mail.ustc.edu.cn (Z.-B.Z.); eycaojie@scut.edu.cn (J.C.); mahaiqing@gdph.org.cn (H.-Q.M.) †These authors contributed equally to this work.



**Fig. 1. Global cellular landscape in CRC liver metastasis.** (A) Schematic overview of the experimental design and analytical workflow. (B) UMAP visualization of non-immune cell clusters. (C) Volcano plot comparing cell type relative abundance of nonimmune cell clusters in CC versus LM ( $n = 5$  patients). The x axis represents the  $\log_2$  fold change, and the y axis represents the  $-\log_{10} P$  value according to Mann-Whitney  $U$  test. Each dot represents a cell type. (D) UMAP visualization of immune cell clusters. (E) Volcano plot comparing cell type relative abundance of immune cell clusters in the tumors versus the paratumors ( $n = 6$  patients). x and y axes represent the  $\log_2$  fold change according to Mann-Whitney  $U$  test. Each dot represents a cell type.

41,892 CD45<sup>-</sup> nonimmune cells and 196,473 CD45<sup>+</sup> immune cells from 27 samples and identified 23 clusters of nonimmune cells and 41 clusters of immune cells. With marker-based annotations, tumor cells (*EPCAM* and *SOX9*), fibroblasts (*COL1A1* and *COL1A2*), and endothelial cells (*PECAM1* and *CD34*) were identified from nonimmune cells (fig. S1A), while T cells (CD3. pAbO), natural killer (NK) cells (CD56. pAbO), B cells/plasma cells (CD19. pAbO), monocytes/macrophages (CD14. pAbO), dendritic cells (DCs) (*HLA.DRA*), and mast cells (*TPSAB1*) were identified from

immune cells (fig. S1B). Each subset had a distinct gene expression pattern (fig. S1, C and D).

Further unsupervised clustering in the major population of non-immune cells gave rise to 11 tumor cell clusters, 8 fibroblast clusters, and 6 endothelial cell clusters (Fig. 1B). CA2 and F2\_MCAM were enriched in LM. F4\_F3 and F5\_CCL11 were significantly enriched in CC (Fig. 1C and fig. S1E). On the basis of the expression of known markers, the immune cells were divided into 41 populations (Fig. 1D). In the primary TME, the percentages of regulatory T cells

( $T_{\text{regs}}$ ) and macrophages were higher compared with those in the CN control. Although monocytes, NK cells, and mucosal-associated invariant T (MAIT) cells were predominant in LN, TIME was remodeled by metastatic tumor cells with significant up-regulation of  $T_{\text{regs}}$  and macrophages (Fig. 1E and fig. S1F). In summary, we identified multiple cell populations with distinct distribution patterns in different tissues of CRCLM.

### Spatial distribution profile of CRC primary tumors and CRC-derived liver metastatic tumors

To comprehensively analyze the spatial distribution profile of CRC primary tumors and CRC-derived liver metastatic tumors, we collected six tissue specimens from six patients including four cases of CRC primary tumors (C1 to C4) and two cases of liver metastatic tumors (L1 and L2). The tumor area (T) and paratumor area (PT) were identified by hematoxylin and eosin (H&E) staining and gene expression features of each sample (Fig. 2A and fig. S2). Compared with paratumor tissues, tumor tissues were enriched with cell cycle-related pathways, such as G<sub>2</sub>-M checkpoint, E2F targets, and mitotic spindle (Fig. 2B). According to the gene expression profiles, tumor tissues and paratumor tissues were divided into different regions (Fig. 2A). Considering that each spot contained multiple cells, we adopted a signature-based strategy, which integrated the ST and scRNA-seq data to estimate the proportions of different cell types for each captured spot (18). B cells, T cells, NK cells, plasma cells, myeloid cells, tumor cells, fibroblasts, and endothelial cells were identified in the ST tissues. The proportions and scores of these cell clusters varied in each region. The ratios of plasma cells decreased in tumor tissues compared with those in paratumor tissues of C1, C3, and C4 (Fig. 2, C and D, and fig. S3).

### Transcriptomic heterogeneity of tumor cells between primary and liver metastatic tumors

Studies have revealed that human colorectal tumor cells contain multiple cell types whose transcriptional characteristics mirror that of the normal colorectal epithelium (19). Stem-like/transit amplifying cells, colonocytes, goblet cells, and tuft cells have been identified in the normal epithelium (16). Thus, subclustering analysis was performed and revealed divergent differentiation lineages of tumor cells, namely, CA1 to CA11 (fig. S4A). The differentiation marker genes of tuft cells, such as *LRMP* and *TRPM5*, were identified in CA8. CA9 showed higher expression of *MUC2*, a marker for goblet cells (fig. S4B). The heterogeneity of tumor cell populations suggests the differentiation potential of tumor cells. It has been reported that tumor tissues generated from a single tumor cell can recapitulate the cellular diversity of the parental tumors (19). The population diversity of tumor cells in LM was similar to that of CC (Fig. 1C and fig. S4C). To further investigate the functional pathways and transcriptional programs of different tumor cell clusters, we performed gene set variation analysis (GSVA) and transcription factor (TF) analysis. The results showed that CA2 was enriched with Wnt- $\beta$ -catenin signaling, and the TF *LEF1*, which plays an important role in the Wnt- $\beta$ -catenin signaling pathway, was up-regulated in CA2. Furthermore, CA2 exhibited the up-regulation of the somatic stem cell division pathway (fig. S4D). Consistently, CA2 had the highest expression level of *LGR5*, a marker of intestinal stem cells. These results indicated that CA2 might be the stem cell subset in TME. In addition, CA2 had high expression levels of *EPCAM*, *CDH1*, and *PRSS2*, facilitating the adhesion and

colonization of tumor cells (fig. S4, B and E). Collectively, cluster identification and characterization analysis demonstrated the functional diversification of tumor cells, and metastatic tumors can recapitulate the cellular diversity of the parental tumors.

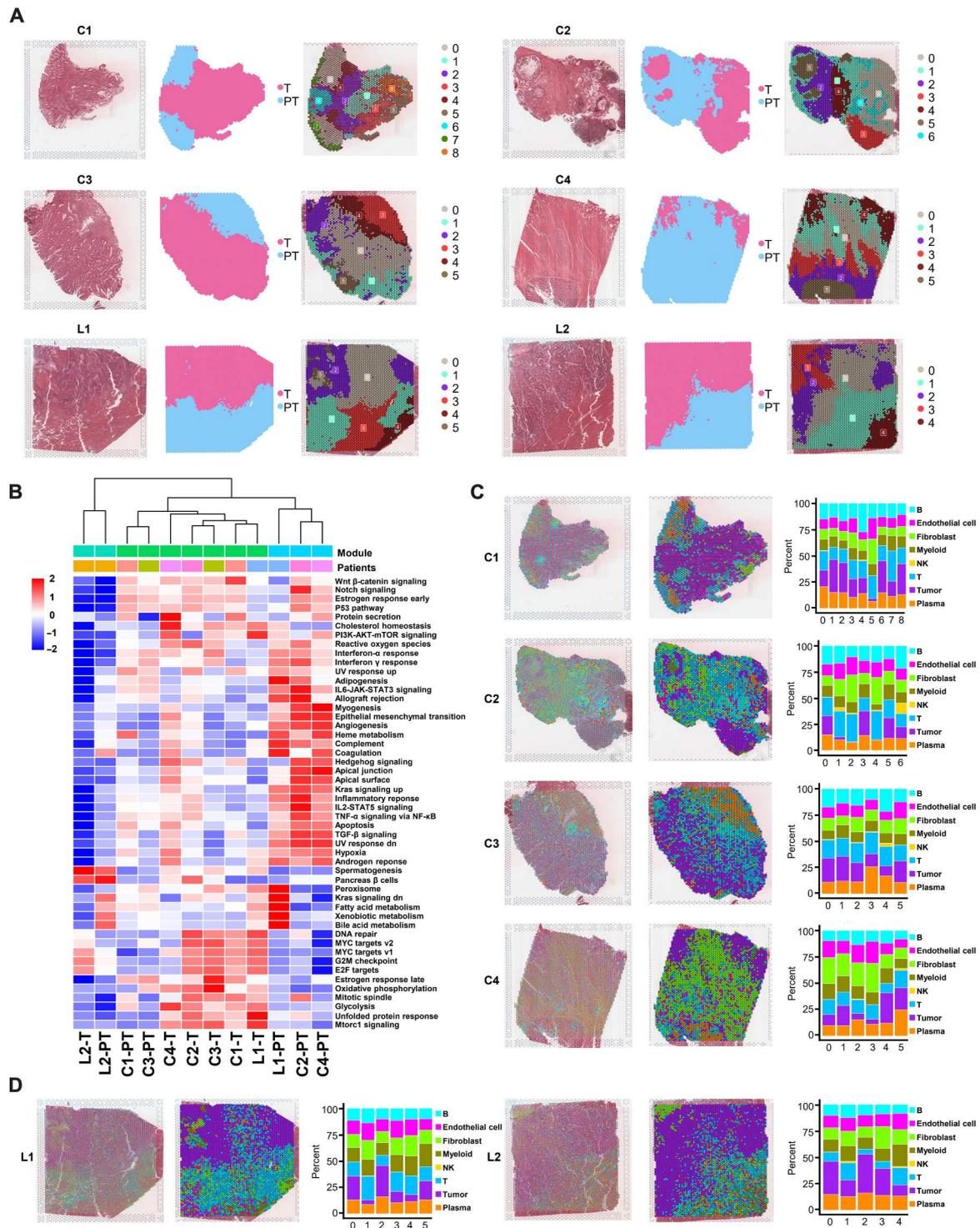
### TME reshapes the composition of myeloid cells

Myeloid cells are heterogeneous subsets in TME, and three kinds of myeloid cells were identified: monocytes/macrophages, DCs, and mast cells. Using subclustering analysis, we identified eight clusters of monocytes/macrophages and three clusters of conventional DCs (cDCs) (Fig. 3A), and each subset had a distinct gene expression pattern (fig. S5, A and B). There were distinct distribution patterns of myeloid cells in each site (Fig. 3B), demonstrating the organ-specific characterization. *CD14*<sup>+</sup> monocyte was the most enriched subset in PB (fig. S5C). Compared with those in the paratumors, the Mac\_SPP1 subset was enriched in the tumors, and most of the cells in the Mac\_SPP1 subset were from the primary tumors. In addition, the proportion of Mac\_CXCL9 was significantly increased in LM (Fig. 3C and fig. S5D), which had high expression level of *CXCL9*, suggesting that this subset could recruit CXCR3-positive effector T cells into tumors. Genes involved in recruiting myeloid cells especially granulocytes, such as *CXCL3*, were enriched in the Mac\_SPP1 subset (Fig. 3D). Myeloid-derived suppressive cells can express high level of CXCR2, which is the receptor of *CXCL3* (20) and can suppress the CD8<sup>+</sup> T cells by expressing arginase and inducible nitric oxide synthase (iNOS) (21). GSVA analysis showed that the Mac\_SPP1 subset was engaged in the pathway related to inflammatory response. However, the Mac\_CXCL9 subset was enriched in the interferon- $\gamma$  (IFN- $\gamma$ ) response and T cell activation pathways (Fig. 3E). The trajectory analysis showed that both the Mac\_CXCL9 and Mac\_SPP1 subsets were terminally differentiated, indicating that they were tumor-activated subsets (Fig. 3F).

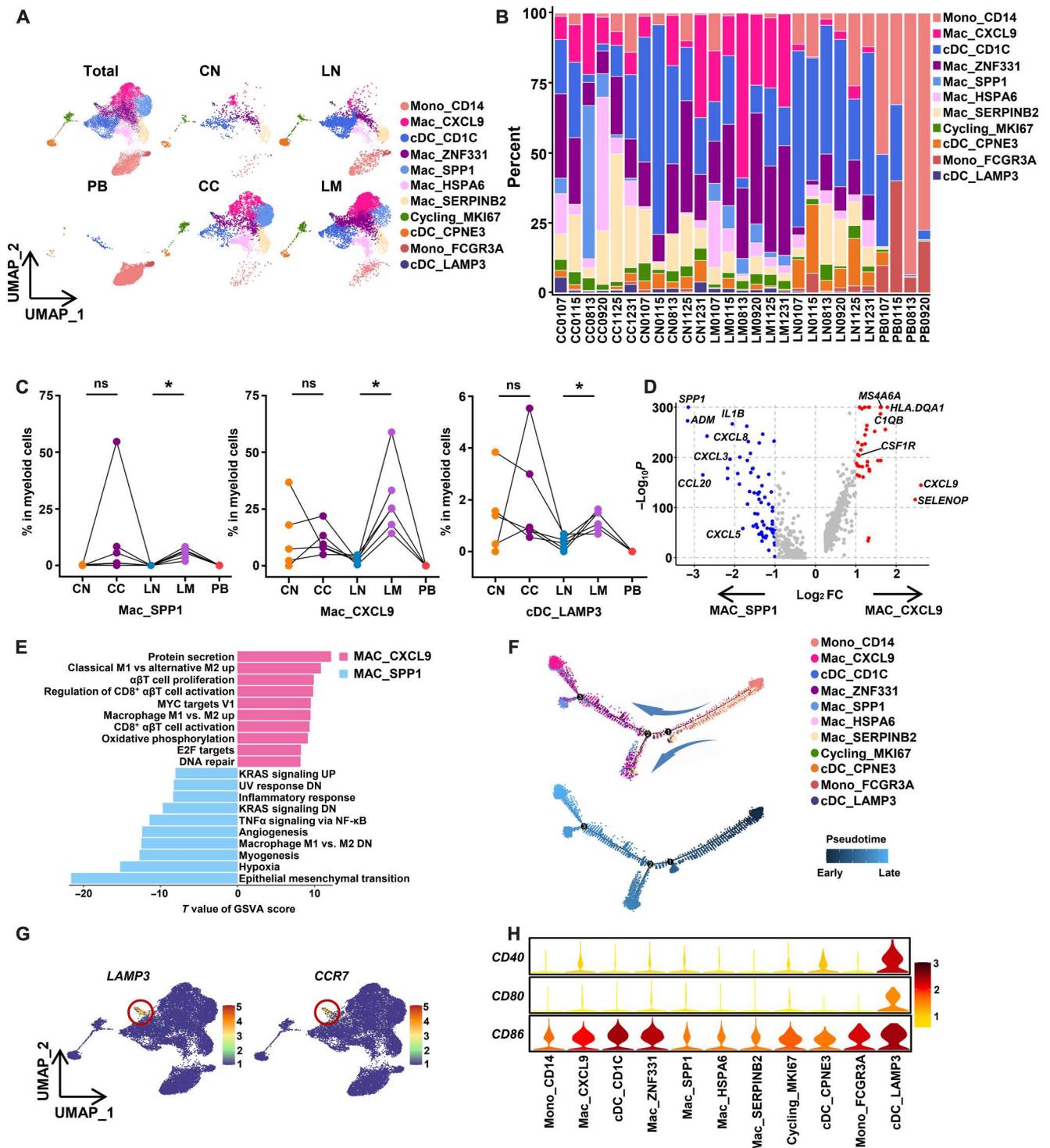
DCs can be divided into plasmacytoid DCs (pDCs) and cDCs according to their cytokine expression and function. Different subsets of cDCs, including cDC1 and cDC2, have been identified in TME (22). In our data, both cDC1 (cDC\_CPNE3) and cDC2 (cDC\_CD1c) were identified in CC and LM (Fig. 3A). Moreover, cDC\_LAMP3 was also identified and enriched in LM (Fig. 3C and fig. S5D). cDC\_LAMP3 expressed high levels of *CCR7* (Fig. 3G), which is the receptor of CCL19 and CCL21, indicating that it had migration ability to the lymph node. Furthermore, the cDC\_LAMP3 subset exhibited high expression levels of costimulatory molecules *CD40*, *CD80*, and *CD86*, markers for DC maturation (Fig. 3H). The percentage of the cDC\_LAMP3 subset was increased significantly in LM, indicating the education of tumor cells on TIME.

### CXCL13<sup>+</sup> T cells are enriched in liver metastatic tumors

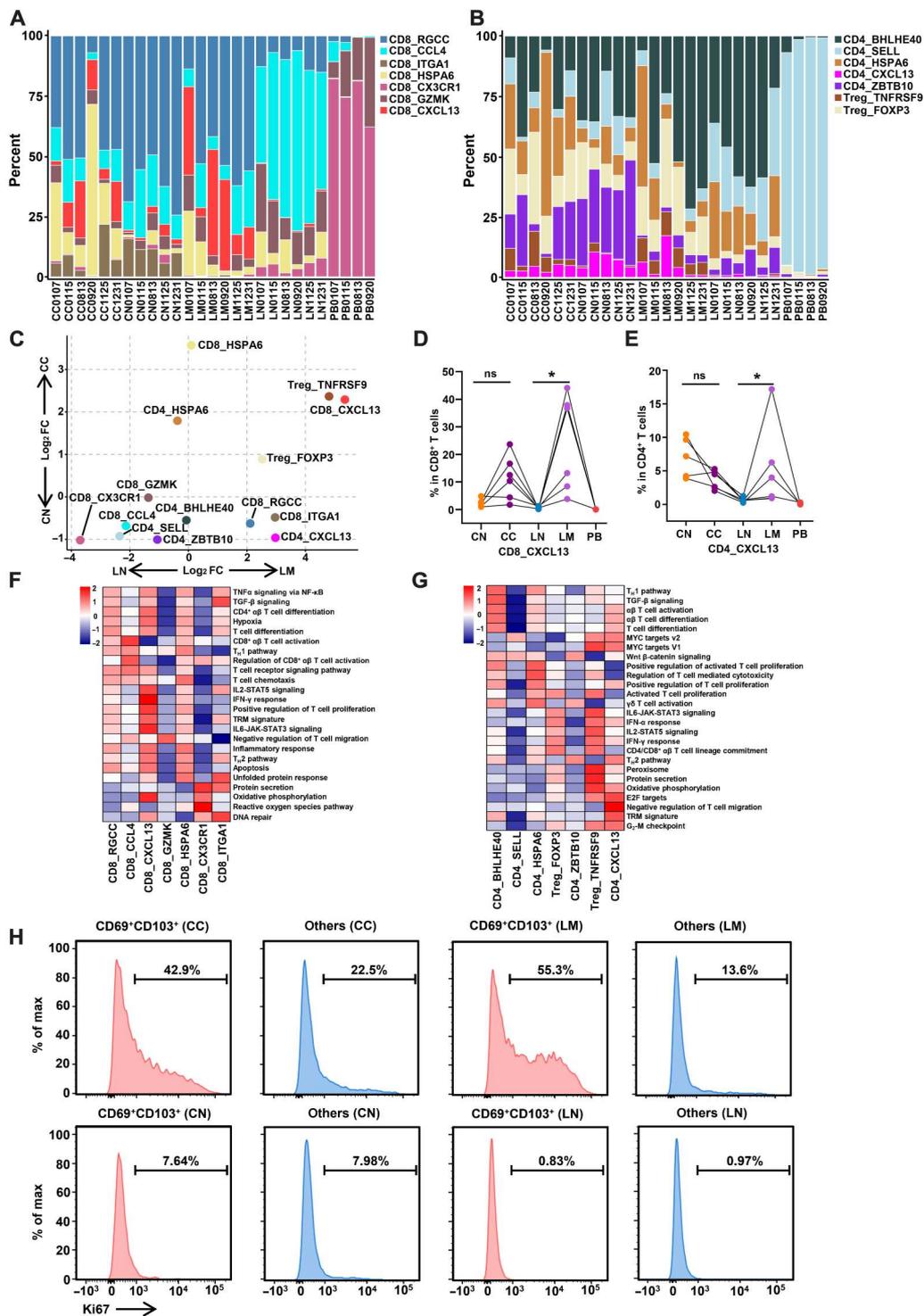
T cells, especially CD8<sup>+</sup> T cells and CD4<sup>+</sup> T cells, are predominant in the adaptive immune response. Among T cells, seven clusters of CD8<sup>+</sup> T cells, five clusters of conventional CD4<sup>+</sup> T (cCD4) cells, and two clusters of  $T_{\text{regs}}$  were identified. Both CD8<sup>+</sup> T cells and CD4<sup>+</sup> T cells included a cluster expressing *CXCL13*, a chemokine for CXCR5 (Fig. 4, A and B, and fig. S6, A to D). The percentage of the CD8\_CXCL13 cells was significantly increased in LM compared with those in LN; however, this subset was rarely detected in CN, LN, and PB (Fig. 4, C and D). In the steady state of the colon, about 3 to 12% of CD4 cells were *CXCL13* positive, which is denoted as follicular T helper ( $T_{\text{FH}}$ ). In CC, the percentage of



**Fig. 2. Cellular identification in spatial transcriptomic samples.** (A) Overview of the spatial transcriptomic sections. H&E staining of spatial transcriptomic sections (left). Tumor tissue and paratumor tissue identification of each section (middle). Spatial cluster distribution of each section (right). (B) Heatmap showing the enrichment scores of hallmark gene sets of tumor tissue and paratumor tissue in each section. (C) Cluster identification combined with single-cell RNA expression profiles in primary CRC sections. Left: SPOTlight deconvolution results, which show the cell cluster proportions in each spot. Middle: Cluster definition of each spot, which is identified as the most dominant cell cluster there. Right: Proportions of identified cell clusters in different regions, which represent the mean proportions of cell clusters in all spots of each region. (D) Cluster identification combined with single-cell RNA expression profiles in liver metastatic sections.



**Fig. 3. Immune landscape of myeloid cells.** (A) UMAP visualization of myeloid cell clusters. (B) Proportions of the myeloid cell clusters in CN, CC, LN, LM, and PB. (C) Proportions of MAC\_SPP1, MAC\_CXCL9, and cDC\_LAMP3 subsets in myeloid cells. *P* values were determined by the paired nonparameter test. (D) Volcano plot showing differentially expressed genes between MAC\_CXCL9 and MAC\_SPP1 subsets. (E) GSEA analysis showing pathways enriched in MAC\_CXCL9 and MAC\_SPP1 subsets. (F) Monocle analysis showing the developmental trajectory of myeloid cells. (G) Feature plots showing expressions of *LAMP3* and *CCR7* in myeloid cells. (H) Violin plots showing expressions of *CD40*, *CD80*, and *CD86* in myeloid cells.



**Fig. 4. Transcriptional reprogramming of tumor-infiltrated T cells.** (A) Proportions of the CD8<sup>+</sup> T cell clusters in CN, CC, LN, LM, and PB. (B) Proportions of the CD4<sup>+</sup> T cell clusters in CN, CC, LN, LM, and PB. (C) Volcano plot comparing cell type relative abundance of T cell clusters in the tumors versus the paratumors (*n* = 6 patients). *X* and *y* axes represent the log<sub>2</sub> fold change according to Mann-Whitney *U* test. Each dot represents a cell type. (D) Proportions of CD8\_CXCL13 subsets in CD8<sup>+</sup> T cells. (E) Proportions of CD4\_CXCL13 subsets in CD4<sup>+</sup> T cells. (F) GSEA pathway analysis of CD8<sup>+</sup> T cell clusters. (G) GSEA pathway analysis of CD4<sup>+</sup> T cell clusters. (H) Ki67 expression levels in different cell subsets in CN, CC, LN, and LM by flow cytometry. Gated on tumor-infiltrated CD8<sup>+</sup>CD45RO<sup>+</sup> T cells.

CD4\_CXCL13 was decreased, but the percentage of this subset was increased in LM compared with that in LN (Fig. 4E).

To further investigate the characteristics of the CXCL13<sup>+</sup> T subset, we analyzed the pathways up-regulated in each subset. GSVA analysis revealed that the CD8\_CXCL13 subset was enriched with T cell proliferation pathway, and the CD4\_CXCL13 subset was enriched with G<sub>2</sub>-M checkpoint pathway, showing the proliferative properties of these two subsets (Fig. 4, F and G). The gene analysis of each cluster showed that CD8\_CXCL13 cells expressed high level of *ITGAE*, which is a marker for tissue-resident memory T (T<sub>RM</sub>) cells. Considering that the T<sub>RM</sub> cells were also CD69 positive (fig. S6E), we identified CD69<sup>+</sup>CD103<sup>+</sup>CD8<sup>+</sup> T cells by flow cytometry. Most CD103<sup>+</sup> cells were CD69-positive cells, and the CD69<sup>+</sup>CD103<sup>+</sup>CD8<sup>+</sup> T cells were rarely presented in LN (fig. S6F). Consistently, the expression levels of Ki67 in CD69<sup>+</sup>CD103<sup>+</sup>CD8<sup>+</sup> T cells from both CC and LM were higher than those in other CD8<sup>+</sup> T cells. However, the proliferative properties of different CD8<sup>+</sup> T cell subsets from CN and LN were almost identical (Fig. 4H). Together, these results indicated that both CD8\_CXCL13 and CD4\_CXCL13 cells were up-regulated in LM of CRC because of their high proliferation abilities.

### The CXCL13<sup>+</sup> T cells are associated with good prognosis of CRC patients

In aforementioned results, we showed that CD8\_CXCL13 was enriched in CC and LM compared with the CD4\_CXCL13 subset (Fig. 4, D and E), indicating that the CD8\_CXCL13 cells might be a tumor-activating subset. Thus, we further investigated the characteristics of the CD8\_CXCL13 subset. High levels of exhausted markers, such as *PDCD1*, *HAVCR2*, *LAG3*, *CTLA4*, and *TIGIT*, were observed in the CD8\_CXCL13 subset compared with other subsets (Fig. 5, A and B). Previous studies showed that tumor-reactive T cells exhibited an exhaustion phenotype due to the persistent tumor antigen stimulation (23–25), and the exhaustion phenotype indicated their tumor reactivity. The trajectory analysis showed that the CD8\_CXCL13 cells were terminally differentiated (Fig. 5C). Furthermore, CXCL13<sup>+</sup> T cell subsets showed remarkable clonal expansion ability in TME, further indicating their antigen experience properties (26). In addition, the CD8\_CXCL13 subset also expressed high levels of effector molecules, such as *GZMB* (Fig. 5, A and B), which indicated that this subset might preserve partial antitumor function. Because it was difficult to detect CXCL13 in the tissues, we attempted to use CD69 and CD103 to label CXCL13<sup>+</sup> cells, and the IHC results showed that CD69<sup>+</sup>CD103<sup>+</sup>CD8<sup>+</sup> T cells were more adjacent to the CK19<sup>+</sup> tumor cells than other CD8<sup>+</sup> T cells in the tumor tissues (Fig. 5, D and E), facilitating their tumor-killing function. Notably, CD103<sup>+</sup>CD8<sup>+</sup> T cells were present in the tertiary lymphoid structures (TLSs) of CRC tissue, and patients with a higher TLS score had a higher CD8\_CXCL13 score, which indicated that this subset might participate in the formation of TLS (Fig. 5, F and G).

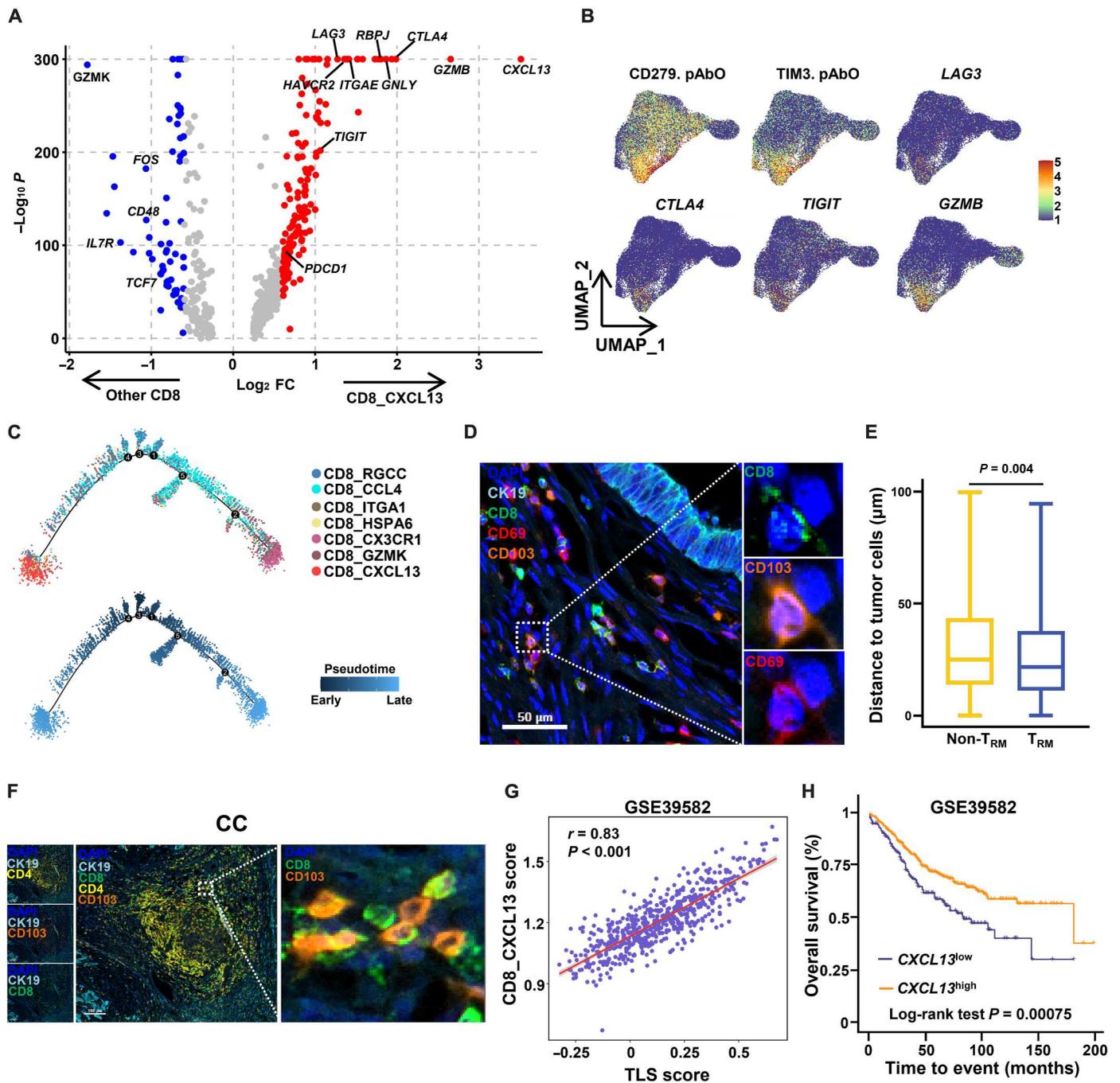
To explore the prognostic value of CXCL13<sup>+</sup> T cells in CRC, we divided CRC patients from the Gene Expression Omnibus (GEO) cohort into CXCL13<sup>high</sup> and CXCL13<sup>low</sup> groups and found that the higher expression of CXCL13 in CC predicted better overall survival (Fig. 5H). To summarize, CXCL13<sup>+</sup> T cells enriched in TME were a tumor-reactive subset and contributed to a better prognosis of CRC patients.

### Distinct subsets of fibroblasts exist in the primary and liver metastatic tumors of CRC

Fibroblasts are the major types of stromal nonimmune cells in TME (27). Various fibroblast subsets have been identified in a wide range of tumors (10–12). Thus, we further characterized the fibroblasts to explore their heterogeneity in CRC primary and liver metastatic tumors. Eight clusters of fibroblasts with distinct gene expression patterns were identified in TME of CRC, which were designated as F1\_PRELP, F2\_MCAM, F3\_HAS1, F4\_F3, F5\_CCL11, F6\_IGFL2, F7\_COCH, and F8\_MKI67 (Fig. 6A and fig. S7, A and B). All the fibroblast highly expressed *ACTA2* (fig. S7B), which was widely reported as an important marker of CAF. The percentage of the F2\_MCAM cluster was higher in LM compared with that in CC. The F4\_F3 and F5\_CCL11 clusters mostly contained cells derived from CC (Fig. 6B and fig. S7C). The infiltration of F4\_F3 was increased in CC compared with that in CN according to the bulk RNA-sequencing data. However, the percentage of the F5\_CCL11 cluster was decreased in CC, which might turn into the F4\_F3 cluster based on the trajectory analysis (Fig. 6C and fig. S7, D and E). Trajectory analysis also predicted that F2\_MCAM and F4\_F3 were two distinct terminally differentiated clusters (Fig. 6C). IHC analysis verified that F2\_MCAM existed in CC and LM (Fig. 6D). According to the IHC result, the MCAM<sup>+</sup> CAF is very rare in LN. Most of the MCAM<sup>+</sup> cells were endothelial cells (fig. S7F). However, consistent with the single-cell analysis, the F4\_F3 subset only existed in CC, but not in LM, and this phenomenon might be caused by the absence of F4\_F3 in LN and the presence in CN (Fig. 6, E and F). ST analysis also confirmed that there was more infiltration of F4\_F3 subset in CC than in LM (Fig. 6, G and H). The F4\_F3 subset closely laid around the epithelial cells in both CN and CC, facilitating its interaction with the epithelial cells (Fig. 6I). Furthermore, the increased infiltration of F4\_F3 in CC might lead to worse prognosis (Fig. 6J). In summary, different phenotypic profiles and highly variable frequencies of fibroblasts were observed between primary and liver metastatic tumors of CRC, which indicated the cellular heterogeneity within TMEs in different cancer settings.

### The F3-expressing fibroblast subset enriched in the primary tumors secretes protumor factors and is associated with poor prognosis of CRC patients

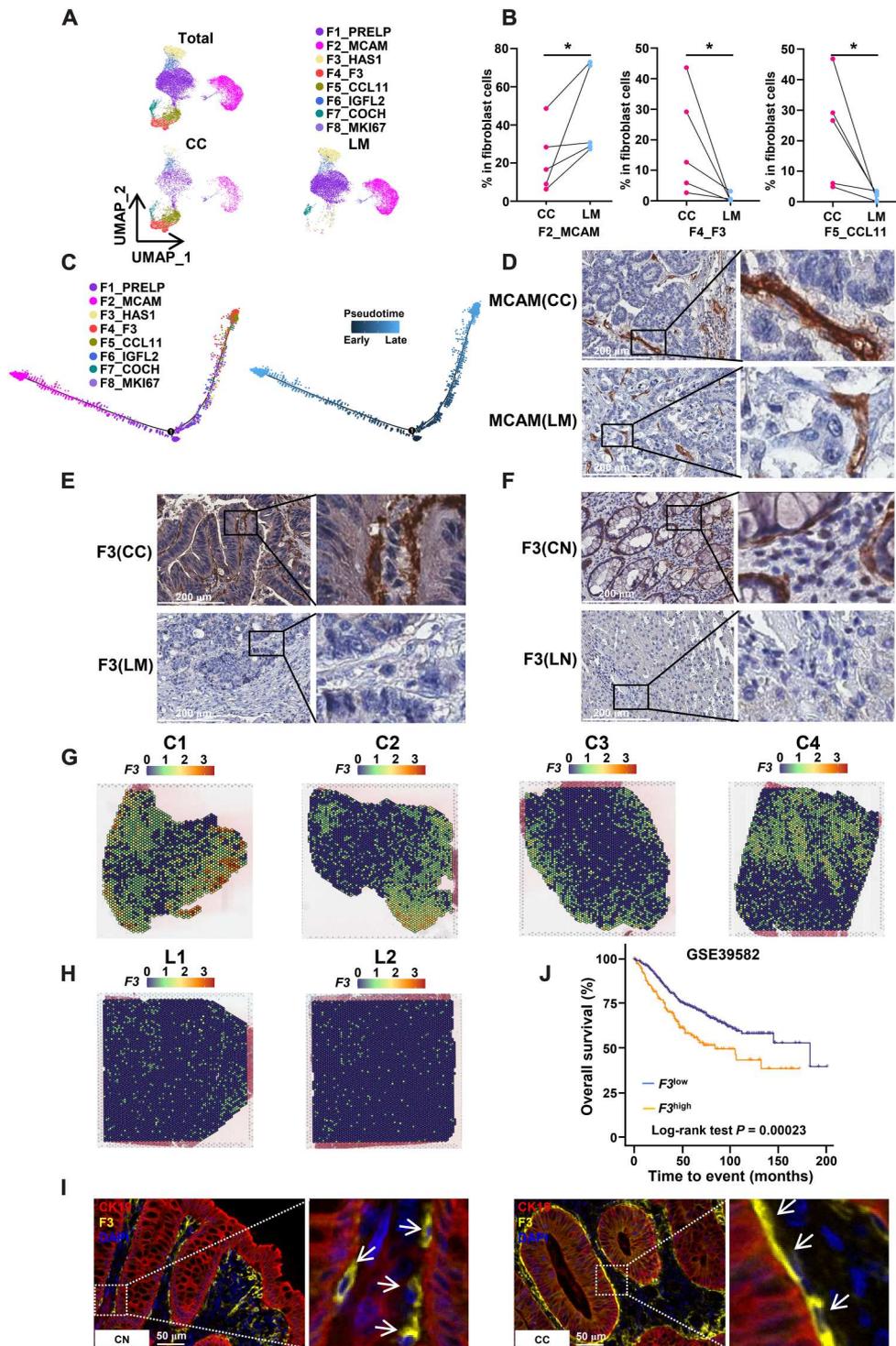
To investigate the remodeling of TME by different fibroblasts enriched in CC and LM, we further analyzed the characteristics of F2\_MCAM and F4\_F3. F2\_MCAM was enriched with *JAG1* and *NOTCH3*, which participated in the NOTCH signaling pathway. F4\_F3 was enriched with *C3* and *CXCL1*, indicating that this subset was involved in the complement and inflammatory response pathway (Fig. 7, A and B). F4\_F3 also expressed high levels of *MMP2* and *MMP3*, which might contribute to the extracellular matrix organization (Fig. 7A and fig. S7G). In addition, protumor factors participating in angiogenesis and tumor invasiveness, such as *VEGFA*, *NRG1*, *HGF*, *GDF15*, *AREG*, and *BMP2*, were enriched in F4\_F3 (Fig. 7C and fig. S7G). The interaction analysis between F4\_F3 and tumor cells further revealed their cross-talk through the *NRG1* and Erb-B2 receptor tyrosine kinase 3 (*ERBB3*) pathway, and *ERBB3* was almost enriched in the tumor cells and could form heterodimer with *ERBB2*, promoting tumor cell proliferation and resistance to cetuximab of CRC patients (Fig. 7D and fig. S7, H and I) (28, 29). ST analysis also indicated the



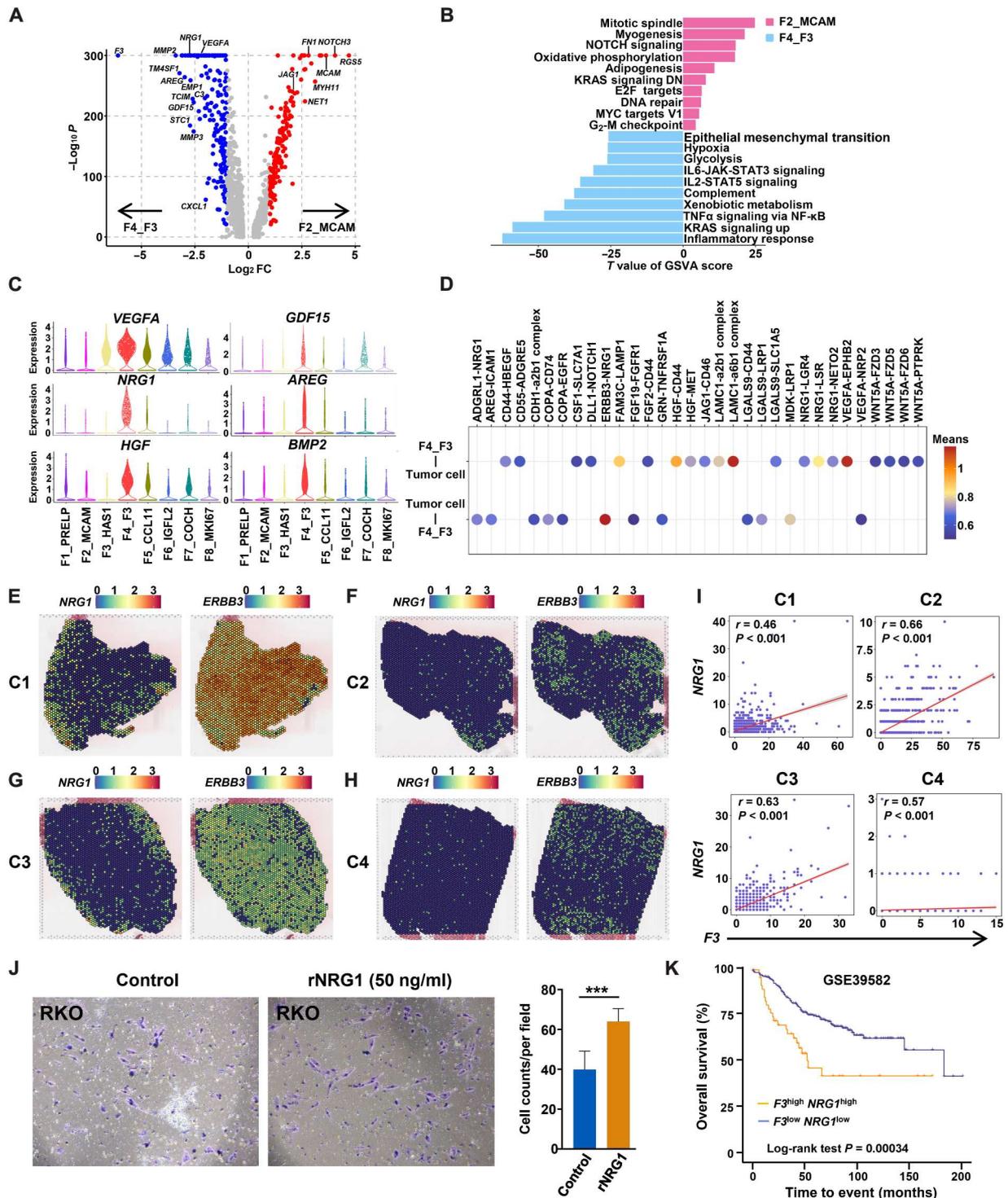
**Fig. 5. Characterization and prognostic effect of CXCL13<sup>+</sup> T cells.** (A) Volcano plot showing differentially expressed genes between CD8\_CXCL13 and other CD8<sup>+</sup> T cells. (B) Feature plots showing the checkpoint and effector molecules in CD8<sup>+</sup> T cells. (C) Monocle analysis showing the developmental trajectory of CD8<sup>+</sup> T cells. (D) Multicolor IHC staining of CD69<sup>+</sup>CD103<sup>+</sup>CD8<sup>+</sup> T cells in CRC tertiary lymphoid structure. (E) Distances of CD69<sup>+</sup>CD103<sup>+</sup>CD8<sup>+</sup> T cells and other CD8<sup>+</sup> T cells to cancer cells. (F) Multicolor IHC staining of CD103<sup>+</sup>CD8<sup>+</sup> T cells in CRC tertiary lymphoid structure. (G) Relationship of tertiary lymphoid structure score and CD8\_CXCL13 infiltration score in the GSE39582 staining dataset. (H) Overall survival of CXCL13<sup>high</sup> and CXCL13<sup>low</sup> patients from the GSE39582 dataset.

colocalization of *F3* with *NRG1*, which surrounded the *ERBB3*<sup>+</sup> tumor cells (Figs. 6G and 7, E to I). Transwell migration assay showed that recombinant human *NRG1* (rNRG1) can promote the migration of RKO and SW620 cells (Fig. 7J and fig. S7J). CRC patients with high expression levels of *F3* and *NRG1* had worse overall survival (Fig. 7K). These results demonstrated that the

CC-enriched fibroblast F4\_F3 could lead to poor prognosis of CRC patients by secreting protumor factors.



**Fig. 6. Transcriptome signatures and heterogeneity of fibroblasts in primary and metastatic tumors.** (A) UMAP visualization of fibroblast clusters. (B) Proportions of the F2\_MCAM, F4\_F3, and F5\_CCL11 clusters in CC and LM ( $P$  values were determined by the paired nonparameter test). (C) Monocle analysis showing the developmental trajectory of fibroblasts. (D) IHC staining of F2\_MCAM in CC and LM. (E) IHC staining of F4\_F3 in CC and LM. (F) IHC staining of F4\_F3 in CN and LN. (G) F4\_F3 fibroblast distributions in CC (C1 to C4) by ST. (H) F4\_F3 fibroblast distributions in LM (L1 and L2) by ST. (I) Multicolor IHC staining of F4\_F3 fibroblasts and tumor cells (CK19). (J) Overall survival of  $F3^{high}$  and  $F3^{low}$  patients from the GSE39582 dataset.



**Fig. 7. Characterization and prognostic effect of F3<sup>+</sup> fibroblasts.** (A) Volcano plot showing differentially expressed genes between F2\_MCAM and F4\_F3. (B) GSEA analysis showing pathways enriched in F2\_MCAM and F4\_F3 subsets. (C) Violin plots showing expressions of *VEGFA*, *NRG1*, *HGF*, *GDF15*, *AREG*, and *BMP2* in fibroblasts. (D) Communication network between tumor cells and F4\_F3 fibroblasts by CellPhone DB. (E) *NRG1* and *ERBB3* distributions in C1 sample. (F) *NRG1* and *ERBB3* distributions in C2 sample. (G) *NRG1* and *ERBB3* distributions in C3 sample. (H) *NRG1* and *ERBB3* distributions in C4 sample. (I) Pearson correlation of *F3* (x axis) and *NRG1* (y axis). (J) Effects of rNRG1 on migration of the RKO cells. The migrated cell number in six different fields was counted in the experiment, and the values were averaged. Three independent experiments were performed. The bars represent means  $\pm$  SD (\*\**P* < 0.001). (K) Overall survival of F3<sup>high</sup> NRG1<sup>high</sup> and F3<sup>low</sup> NRG1<sup>low</sup> CRC patients from the GSE39582 dataset.

### The *MCAM*-expressing fibroblast in TME of LM modulates the generation of CD8\_CXCL13 cells through the Notch signaling pathway

The Notch signaling TF *RBPJ* was preferentially expressed in both the CD8\_CXCL13 and CD4\_CXCL13 subsets (Fig. 5A and fig. S6, C and D). It has been reported that the Notch signaling pathway can be activated by the ligand and receptor interaction (30). Following the interaction with its ligands, the intracellular domain of Notch can be cleaved and translocate into the nucleus, where it modulates the transcription of its target genes. The *RBPJ* expression in LM was positively correlated with *CXCL13* and *ITGAE*, which was not observed in CC (Fig. 8A and fig. S8A). The CD8\_CXCL13 and CD4\_CXCL13 subsets mainly expressed the *NOTCH1* receptor (Fig. 8B). To further gain insights into what types of cells modulate Notch signaling in the CD8\_CXCL13 and CD4\_CXCL13 subsets, we analyzed the expression levels of Notch ligands, including *DLL1*, *DLL3*, *DLL4*, *JAG1*, and *JAG2*, and we found that Notch ligands were predominantly expressed in the fibroblasts and endothelial cells (fig. S8B). The F2\_MCAM subset was enriched with *JAG1*, and the F5\_COCH subset with *DLL1*, while the E2\_DLL4 subset was enriched with *DLL4*, *JAG1*, and *JAG2* (Fig. 8C). The interaction analysis of Notch and its ligands using the CellPhone DB showed that among the endothelial cells, the E2\_DLL4 subset exhibited the strongest interaction with *CXCL13*<sup>+</sup> T cells, and among the fibroblasts, the F2\_MCAM cluster interacted with the CD8\_CXCL13 and CD4\_CXCL13 subsets by *JAG1-NOTCH1* (Fig. 8D and fig. S8C). Because of the scatter distribution pattern of the fibroblasts in TME, we speculated that the F2\_MCAM subset contributed to the activation of Notch signaling in *CXCL13*<sup>+</sup> T cells.

Next, we divided the patients into two groups according to the ratio of F2\_MCAM from LM and found that patients in the F2\_MCAM-high group had higher proportion of CD8\_CXCL13 subset in LM (Fig. 8E). GEO dataset analysis showed that patients with a higher F2\_MCAM infiltration score had a higher CD8\_CXCL13 subset infiltration score in LM, but not the CD4\_CXCL13 subset, and not in CC (Fig. 8F and fig. S8, D and E). Furthermore, we detected the locations of F2\_MCAM and CD8\_CXCL13 in LM through ST. Consistent with the single-cell results, regions with a higher F2\_MCAM infiltration score had a higher CD8\_CXCL13 infiltration score in the ST samples (Fig. 8, G and H). Moreover, the strength of the Notch signaling interaction in LM was much stronger than that in CC (fig. S8F), which might be due to the higher proportion of F2\_MCAM subset in LM (Fig. 6B).

The Notch signaling pathway has been reported to modulate the antitumor immunity of CD8<sup>+</sup> T cells (31); however, little is known about whether the Notch signaling pathway can modulate the expression of *CXCL13*. To explore such modulation, we predicted the binding sites of *RBPJ* on the *CXCL13* promoter using JASPAR (<http://jaspar.genereg.net/>). Several potential binding sites were detected (table S1), indicating that *RBPJ* might act as a TF influencing the expression of *CXCL13*.

### Intercellular interaction network in ST tissues

We also investigated the cellular interactions between different clusters in ST of the primary tumors and liver metastatic tumors. It showed that the *VEGFA-NRP1* and *VEGFA-NRP2* ligand-receptor pairs were enriched in both the primary and liver metastatic tumors. We also found there were various different enriched ligand-receptor

pairs between CC and LM. The *ERBB3-NRG1* pair was enriched in both C1 and C3, but absent in both L1 and L2 (Fig. 9A), indicating the potential role of the *ERBB3-NRG1* interaction in the development and metastasis in the primary tumors.

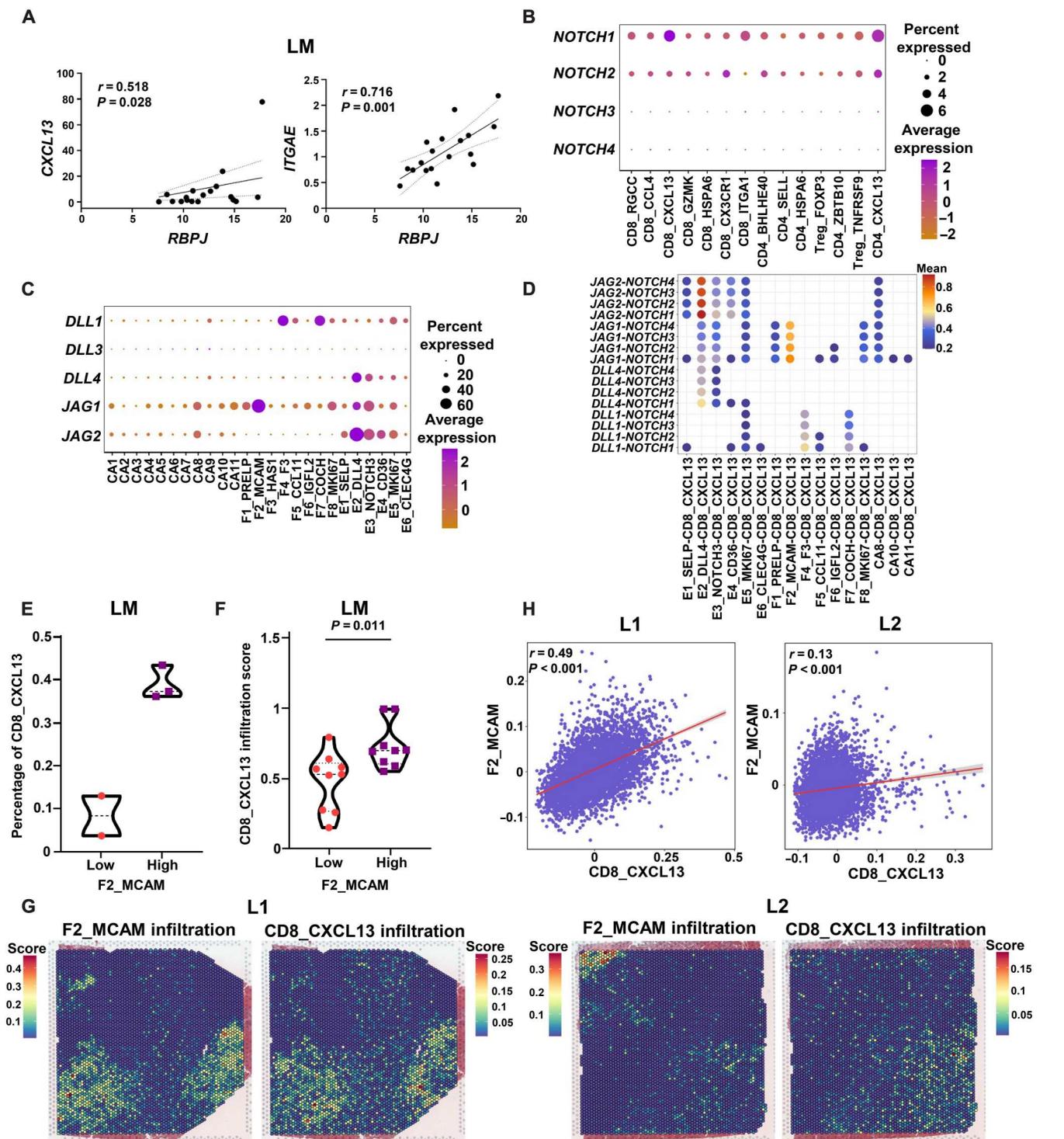
Together, we found that *F3*<sup>+</sup> fibroblast enriched in the CRC primary tumor could modulate the development and/or migration by producing various protumor factors, including *NRG1*, which interacted with *ERBB3* expressed on tumor cells to exert their protumor functions. However, the *MCAM*<sup>+</sup> fibroblast enriched in LM could modulate the generation of CD8\_CXCL13 cells through the Notch signaling pathway, different from the protumor subset *F3*<sup>+</sup> fibroblast in CC, indicating the cellular heterogeneity of stromal cells in different cancer settings (Fig. 9B).

## DISCUSSION

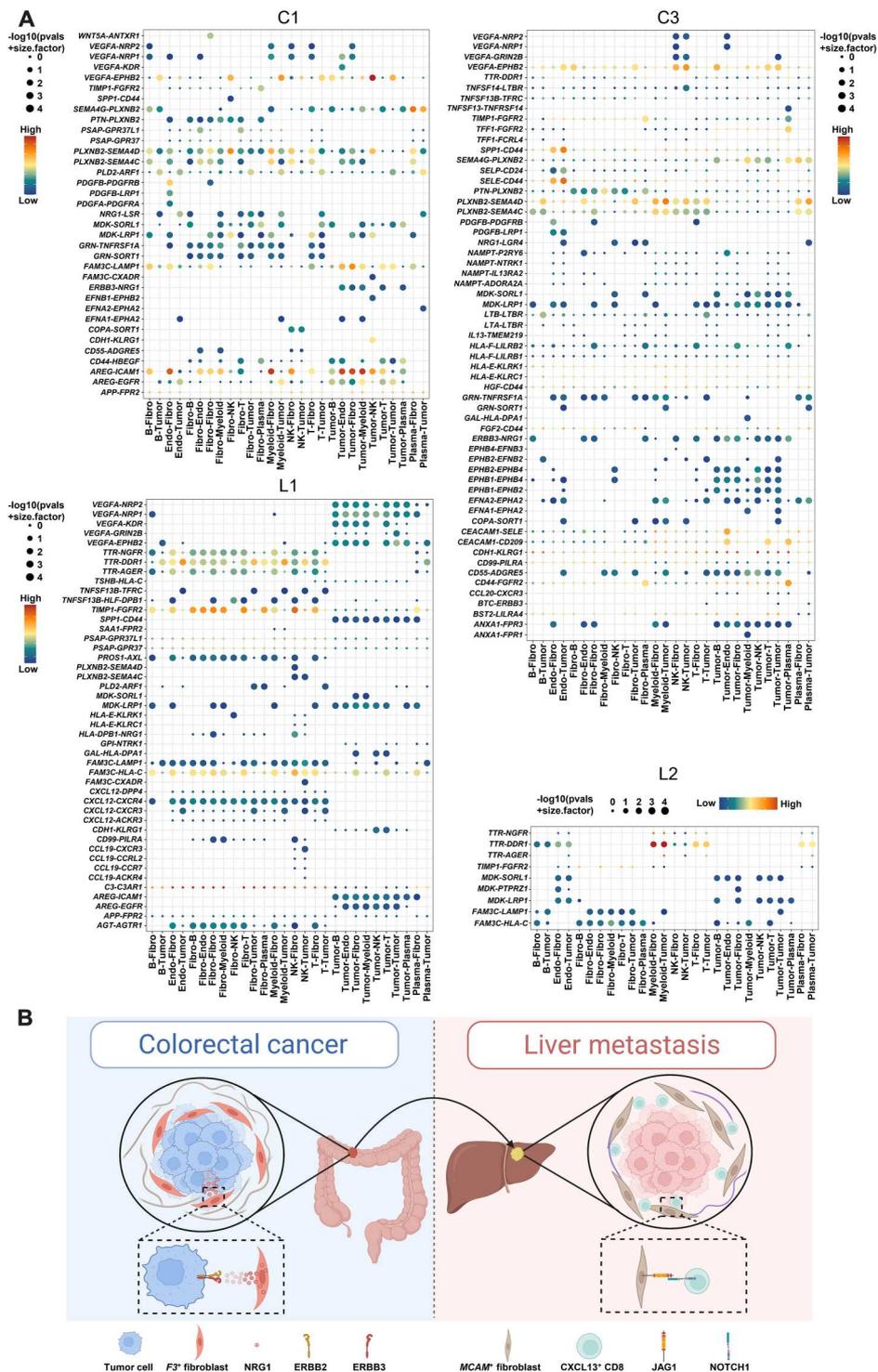
CRCLM is a multistep process and fatal in most of the cases (3). In this study, we presented a comprehensive single-cell transcriptomic characterization of the primary and liver metastatic tumors of CRC, covering 41,892 nonimmune cells and 196,473 immune cells from six patients. We observed that colorectal tumor cells could recapitulate the multilineage differentiation processes of normal colon epithelium. Distinct cellular profiles of CAF were identified from CC and LM. The CC-enriched CAF subset F4\_F3 was involved in promoting tumor invasiveness, while the LM-enriched CAF subset could promote the generation of tumor antigen-specific *CXCL13*<sup>+</sup> CD8<sup>+</sup> T cells through the Notch signaling pathway.

It has been reported that tumor tissues generated from a single tumor cell can recapitulate the cellular diversity of the parental tumors (19). In this study, we found that tumor cells in LM contained multiple cell types whose transcriptome profiles mirror those in the primary tumors. The intratumoral heterogeneity of tumor cells with different biological processes may contribute to the chemotherapy failure of CRC, thus investigating the cellular components of each patient may facilitate personal precision treatment.

Fibroblasts are critical to normal tissue homeostasis. However, they are functionally subverted in TME (32). Different types of fibroblasts have been identified from a wide range of malignant tumors, including breast cancer (9), pancreatic cancer (33), CRC (34), lung cancer (35), and gastric cancer (8). CAFs have been ascribed pleiotropic protumorigenic functions, such as extracellular matrix remodeling, tissue stiffening, escape from immune surveillance, and promotion of therapeutic resistance (36, 37). Tumor-specific *FAP*<sup>+</sup> fibroblast was identified in primary CRC and proved to be positively correlated with *SPP1*<sup>+</sup> macrophage, which would stimulate the formation of immune-excluded desmoplastic structure and limit the T cell infiltration (38). In this study, we identified eight subpopulations of fibroblasts with distinct gene expression profiles in CC and LM. The F4\_F3 subset expressing protumor factors was enriched in the primary tumors and probably contributed to tumor invasiveness through the *NRG1-ERBB3* pathway. Considering the characterized genes, the F4\_F3 subset was distinct from *FAP*<sup>+</sup> fibroblast (38). It has been reported that *NRG1* fusion can be detected in CRC, and *NRG1* amplification was found in anti-epidermal growth factor receptor (EGFR) therapy intrinsic-resistant CRC patients. Bone marrow-derived mesenchymal stem cells stimulate invasion, survival, and tumorigenesis of CRC through the release of soluble *NRG1*, activating the *HER2/HER3*-dependent



**Fig. 8. NOTCH signaling in TME modulates the generation of CD8\_CXCL13 cells.** (A) Scatterplots showing correlations of *CXCL13*, *ITGAE*, and *RBPJ* in LM using the GSE50760 dataset. (B) Dot plot showing the expressions of *NOTCH1*, *NOTCH2*, *NOTCH3*, and *NOTCH4* in T cell clusters. (C) Dot plot showing the expressions of *DLL1*, *DLL3*, *DLL4*, *JAG1*, and *JAG2* in nonimmune cell clusters. (D) Communication network of the Notch signaling pathway between CD8\_CXCL13 cluster and nonimmune cells by CellPhone DB. (E) Percentages of CD8\_CXCL13 in F2\_MCAM infiltration high and low groups in LM of scRNA-seq data. (F) Infiltration scores of CD8\_CXCL13 in F2\_MCAM infiltration high and low groups in LM of GSE50760 dataset. (G) Infiltration of F2\_MCAM and CD8\_CXCL13 subsets in L1 and L2 by ST. (H) Pearson correlation of signature score of CD8\_CXCL13 (x axis) and F2\_MCAM (y axis) in L1 and L2.



**Fig. 9. Cellular ligand and receptor interactions in ST tissues. (A)** Bubble heatmap showing the mean interaction strength between two clusters for ligand-receptor pairs in C1, C3, L1, and L2. Dot size indicated the statistical significances by permutation test. Dot color indicated the mean interaction strength levels. The size factor used here is  $10^{-4}$ . **(B)** Cross-talk model between the fibroblasts and other cells in CC and LM.

phosphatidylinositol 3-kinase (PI3K)/AKT signaling cascade in CRC cells (39–41). In our study, rNRG1 can promote the migration of CRC cell lines in vitro. These results hint that NRG1-ERBB3 signaling can promote the progression of CRC. However, the F2\_MCAM fibroblasts enriched in LM expressed Notch ligands, revealing that different organs had distinct stromal microenvironment. It is reported that the NOTCH-RBPJ regulatory network is vital for persistence of T<sub>RM</sub> state in non-small cell lung cancer (42), which expressed high level of CXCL13, indicating that Notch signaling is vital for the generation of CXCL13<sup>+</sup>CD8<sup>+</sup> T cells. By analyzing the ligand-receptor pair communication, we found that the LM-enriched F2\_MCAM subset could interact with tumor-specific CD8\_CXCL13 cells through JAG1-NOTCH1, which was confirmed by the colocalization of F2\_MCAM cells with CD8\_CXCL13 cells using the ST.

The CD8\_CXCL13 cell subset has been identified in various tumors, such as breast cancer (26) and uterine cancer (43), and proven to be associated with increased efficacy of immunotherapy (26). It has been proven that CXCL13<sup>+</sup> T cells can attract CXCR5<sup>+</sup> B cells into tissues, participating in the formation of TLS (43), which was also verified in our study. TLS can promote antitumor immunity, contribute to the efficacy of immunotherapy, and improve prognosis (44, 45). Our data also showed that high infiltration of CXCL13<sup>+</sup> T cells predicted better prognosis in CRC patients. Despite the universal existence of this subset in many kinds of tumors, the mechanism of CD8\_CXCL13 cells activation remains unknown. In our study, both CD8\_CXCL13 cells and CD4\_CXCL13 cells expressed high levels of RBPJ, the characteristic TF of the Notch signaling pathway. In addition, RBPJ was correlated with the expression of CXCL13 and ITGAE, which was highly enriched in the CD8\_CXCL13 cell subset according to the analysis of bulk RNA-sequencing dataset of CRCLM, demonstrating that the Notch signaling pathway could influence the generation of tumor-specific T cells. Furthermore, we also identified the possible binding sites of RBPJ on the promoter of CXCL13, indicating the role of Notch signaling in the antitumor immunity of T cells. Because Notch signaling mutations contribute to the antitumor immune responses in CRC (46, 47), further studies should be performed to address the interaction of Notch signaling in TME and the possibility of targeting this signaling pathway.

There were some limitations of our study. First, all the patients included have undergone neoadjuvant therapy and were unable to reflect the original state of CRC. Second, because the subjects included received different treatments, it was difficult to value the influence of each kind of treatment. Third, the characterizations of stromal cells from the paratumors were unavailable, and the cause of the divergent atlas of stromal cells in CC and LM remains to be elucidated. Last but not least, because our study used a relatively small sample size, a larger population cohort is needed to verify the conclusion.

In conclusion, we unveiled the cellular profiles of TME from tumors and paratumor tissues of CRCLM. Our analysis uncovered the different lineages and functions of fibroblasts as well as the dynamic nature of immune cells in different cancer settings, which can be used for further identification of regulatory mechanisms and for potential therapeutic targets.

## MATERIALS AND METHODS

### Patients

The study was approved by the Ethical Committees of the Guangzhou First People's Hospital and the Sixth Affiliated Hospital of Sun Yat-sen University (no. KY2020-361-01-01). For scRNA-seq, six CRC patients with liver metastasis were enrolled after written informed consent was obtained. Primary CC, CN (at least 2 cm from matched tumors), LM, LN (at least 2 cm from matched tumors), and PB of the enrolled patients were collected following resection. All the patients received preoperative chemotherapy and/or radiotherapy. For ST sequencing, formalin-fixed paraffin-embedded (FFPE) tissues of six CRC patients were prepared. The clinical and demographic information of all the patients were summarized in table S2, and the neoadjuvant regimens of the patients for scRNA-seq were listed in table S3.

### Processing of human tissues and cell sorting

Freshly resected tumors were rinsed with 1× phosphate-buffered saline (PBS) for three times, cut into small pieces, and transferred to 15 ml of digestion medium containing collagenase IV (1 mg/ml, C5138, Sigma-Aldrich) in Dulbecco's modified Eagle's medium (DMEM)/F-12 1:1 (70100300, Biosharp). Tissues were digested for 45 min at 37°C under agitation at 120 rpm; afterward, the samples were filtered using a 150-μm nylon mesh. Following the centrifugation at 450g for 5 min, the supernatants were discarded and the cell pellets were resuspended in 3 ml of red blood cell lysis buffer (C3702, Beyotime) at 4°C for 10 min. After washing with 1× PBS, the cell pellets were suspended in 1× PBS containing 0.2% bovine serum albumin (4240GR500, BioFroxx) and filtered using a 74-μm nylon mesh. Single-cell suspensions were stained with allophycocyanin (APC)/Cy7-CD45 (HI30, BioLegend) and 7-aminoactinomycin D (7-AAD; B226396, BioLegend) for fluorescence-activated cell sorting using BD Aria III (BD Biosciences). Nonimmune cells were sorted as 7-AAD<sup>-</sup>CD45<sup>-</sup>, and immune cells were sorted as 7-AAD<sup>-</sup>CD45<sup>+</sup>. The information of cells obtained from different sites of all the patients was listed in table S4.

### Single-cell RNA library preparation and sequencing

After sorting, cells from different sites were labeled with sample tags (633781, BD Single Cell Multiplexing Kit, BD Biosciences) at 4°C for 45 min and washed by cell staining buffer (B302741, BioLegend) for three times. For immune cells, the expressions of seven proteins on cell surface were detected simultaneously using BD AbSeq Oligo-Conjugated Antibodies (BD Biosciences). The information of AbSeq Ab-Oligos was listed in table S5. After washing and centrifugation at 400g, 4°C for 5 min, the cell pellet was resuspended using BD stain buffer (554656, BD Biosciences) to a density of 700 to 1200 cells/μl. Single-cell library preparation was carried out according to the BD Rhapsody Single-Cell Whole Transcriptome Analysis alpha protocol (<https://bdbiosciences.com/>). Before loading cells to BD Rhapsody Cartridge, nonimmune cell suspensions from primary and metastatic sites were pooled together, so were immune cells from primary and metastatic sites. About 20,000 cells from the cell suspension mixture were loaded onto a microwell array, followed by adding barcoded magnetic BD Rhapsody Cell Capture Beads and cell lysis buffer. Upon cell lysis, the mRNAs and sample tag barcodes of each cell were captured by probes via polyA/polyT hybridization. Beads were subsequently

retrieved from the microwells by magnets and pooled into a single tube for reverse transcription to synthesize complementary DNAs (cDNAs). Then, cDNAs were amplified following multiple amplification schemes (<https://bdbiosciences.com/>). Amplified cDNAs were purified using SPRIselect beads (B23318, Beckman Coulter Life Sciences). Last, the whole transcriptome libraries, AbSeq libraries, and sample tag libraries were obtained for sequencing. Libraries were sequenced on NovaSeq 6000 (Illumina) after quality examination by Agilent 2100 Bioanalyzer (G2940CA, Agilent Technologies).

### scRNA-seq data quality control and preprocessing

The sequencing data were mapped to the human reference sequence (GRCh38). The raw gene expression matrix from each sample was aggregated and converted into a Seurat object via Seurat R package (V4.0). Nonimmune cells with >6000 or <200 genes or >40% mitochondrial genes were discarded. Immune cells with >4000 or <200 genes or >25% mitochondrial genes were filtered out. To further eliminate the data of doublets, we performed the scrublet pipeline (48) for each batch of our scRNA-seq data, which were expected objectively to exclude doublets, and the expected\_doublet\_rate was set at 0.05. We got 238,365 cells for further analysis, including 41,892 nonimmune cells and 196,473 immune cells. The genes expressed in more than 50 cells were selected. Last, 17,515 genes from nonimmune cells and 17,066 genes from immune cells met the criteria. The gene expression matrices were normalized to the total unique molecular identifiers (UMI) counts per cell and transformed to the natural log scale. To correct the technical and biological variations and increase the accuracy of cell type designation, we applied canonical correlation analysis implemented in Seurat to all the samples before cell type identification. After selecting 2500 highly variable genes by the FindVariableFeatures function in Seurat, principal components analysis (PCA) was performed using these genes. The numbers of principal components (PCs) of different kinds of cells were differentially selected according to the knee point of the scree plot for each cell type to accommodate different population complexities based on the ElbowPlot function in Seurat, and we chose the top 20 PCs to run the FindNeighbors and RunUMAP functions. We used the FindClusters function to cluster the cells, and resolutions from 0.1 to 0.8 were explored for better cell clustering, and we set this argument at 0.5 for clustering immune cells and 0.1 for clustering tumor cells. Two-dimensional Uniform Manifold Approximation and Projection (UMAP) was used to represent cell clusters.

### Unsupervised cell clustering and annotation

According to the differentially expressed genes of each lineage, non-immune cells were divided into tumor cells, fibroblasts, and endothelial cells. Immune cells were divided into T cells, NK cells, B cells, plasma cells, monocytes/macrophages, DCs, and mast cells. To identify subpopulations within these major cell types, second-round UMAP reduction was performed. Differentially expressed genes of each subset were identified using the FindAllMarkers function implemented in Seurat. Clusters were named by the highly expressed genes of each cluster.

### Single sample gene set enrichment analysis

To calculate the subset infiltration score in the bulk RNA-sequencing dataset, we adopted single sample gene set enrichment analysis (ssGSEA) according to Senbabaoglu *et al.* (49). Marker genes for

F2\_MCAM, CD8\_CXCL13, CD4\_CXCL13, and TLS were listed in table S6.

### Gene set variation analysis

To compare signaling pathway enrichment of different cell clusters, GSVA was performed to calculate gene set enrichment scores (50). Briefly, we run the gsva function for each single cell and used the "ssGSEA" method to calculate the enrichment score of indicated gene sets, and then we used limma packages (version 3.52.3) to analyze the difference of enrichment score of every gene set between each cell cluster and all other clusters, respectively. We chose the top five up-regulated gene sets for each cluster, scaled the score among the clusters, and exhibited in heatmap.

### SCENIC analysis

The single-cell regulatory network inference and clustering (SCENIC) package (v1.1.2.1) (51) was used to analyze the TF activity. We generated the super cells, which combined the data of every 20 single cells in each cluster to reduce the computing resource consumption. The mean values of normalized counts of 20 single cells were calculated as the raw input data of SCENIC. The Wilcoxon rank sum test was used to identify the differentially activated TFs for each subcluster, with  $\log_2$  fold change > 0.1 and adjusted  $P < 0.05$  as significantly changed.

### Cell differentiation trajectory inference

To infer the differentiation trajectory of intratumor T cells and fibroblasts, we used Monocle2 (52) to infer the pseudotime of each cell. The differentially expressed genes across the clusters were identified by FindMarkers in Seurat for T cells and fibroblasts, respectively. The dimensionality reduction was performed with the DDRTree algorithm, using the most highly variable genes (top 2000) to arrange the cells in order.

### Flow cytometry

The single-cell suspensions were stained to detect the tissue-resident CD8<sup>+</sup> T cells (CD8<sup>+</sup> T<sub>RM</sub>) with the following antibodies, including mouse anti-human BV510-CD45 (304036, BioLegend), mouse anti-human BV711-CD3 (317328, BioLegend), mouse anti-human BV785-CD56 (362550, BioLegend), mouse anti-human BUV563-CD4 (566000, BD Bioscience), mouse anti-human Alexa Fluor 700-CD8A (300920, BioLegend), mouse anti-human BV421-CD45RO (304224, BioLegend), mouse anti-human BV650-CD45RA (304136, BioLegend), mouse anti-human phycoerythrin (PE)-CD69 (310906, BioLegend), mouse anti-human APC-CD103 (350216, BioLegend), and mouse anti-human BV421-Ki67 (350505, BioLegend). After staining with surface markers, cells were fixed and permeabilized using staining buffer set (00-5523-00, eBioscience) to detect the Ki67 expression. Flow cytometry data were obtained using BD Fortessa (BD Biosciences) and analyzed with FlowJo software (V10, BD Biosciences).

### Molecular interaction network analysis

We used CellPhone DB (<http://CellPhoneDB.org>) to analyze the interactions between the tumor cells and the fibroblasts as well as the communications between the CXCL13<sup>+</sup> T cells and nonimmune cells through the Notch signaling pathway (53). CellChat (<http://cellchat.org/>) was applied to compare the differences of the ligand-receptor interactions between CC and LM (54).

### IHC staining

IHC was performed on 5- $\mu$ m FFPE tissue sections prepared from primary sites. The tissues were dewaxed, hydrated, and subjected to antigen retrieval. After blocking with goat serum at room temperature, tissues were stained with primary antibodies, followed by staining with horseradish peroxidase (HRP)-conjugated secondary antibodies. For single-color IHC, 3,3'-diaminobenzidine tetrahydrochloride (DAB) was used as substrate for staining. For multiplex IHC, a PANO 7-plex IHC kit (TSA-RM) (0004100100, Panovue) was used. Tyramide signal amplification (TSA) dye was applied to amplify the signal following the secondary antibodies. The mouse anti-human CD8 antibody (70306s, Cell Signaling Technology), rabbit anti-human CD69 antibody (ab233396, Abcam), rabbit anti-human CD103 antibody (ab224202, Abcam), rabbit anti-human CK19 antibody (ab52625, Abcam), rabbit anti-human CD4 antibody (ab133616, Abcam), rabbit anti-human F3 antibody (ab228968, Abcam), and rabbit anti-human MCAM antibody (ab75769, Abcam) were used as primary antibodies. The nuclei were stained with 4',6-diamidino-2-phenylindole (DAPI) (C1006, Beyotime). Multiplex IHC slides were viewed, and images were recorded with a Vectra Polaris multispectral imaging system (Akoya Biosciences). Phenochart software (v1.0.12, Akoya Biosciences) and HALO software (V3.2, Indica Labs) were used to view images and calculate the distance of different subsets to tumor cells.

### CRC patient survival analysis

The expression information and survival information of CRC patients were downloaded from GEO (<https://ncbi.nlm.nih.gov/geo/>), and the accession number is GSE39582. The Kaplan-Meier method was used to generate survival curves, and the log-rank test was used to determine the statistical significance of differences in survival. Survival curves were constructed with R packages survival (version 3.4.0) and survminer (version 0.4.9). The "maxstat.test" function of R package maxstat was run to find the cutoff value of gene expressions for maximum rank statistic.

### Gene expression correlation analysis

Gene expression data of CRCLM patients generated by bulk RNA sequencing were downloaded from the GEO public database (<https://ncbi.nlm.nih.gov/geo/>) under the data series accession number GSE50760. The processed fragments per kilobase per million (FPKM) expression matrices of *CXCL13*, *ITGAE*, and *RBPJ* of liver metastasis were extracted. The Pearson correlation analysis was applied to calculate the relationship of different genes.

### ST tissue handling, data processing, and cell type infiltration score calculation

The tissues were obtained from patients after resection, fixed by formalin, and embedded in paraffin. A 10 $\times$  FFPE gene expression slide (PN-1000185, 10X Genomics) was used for ST. A slide with 5- $\mu$ m FFPE section was dewaxed with xylene (214736, Sigma-Aldrich) and stained with hematoxylin (51275, Sigma-Aldrich) and eosin (HT110116, Sigma-Aldrich). After visualizing and scanning the whole slide, decrosslinking was performed using the tris-ethylene-diaminetetraacetic acid (TE) buffer (10-0046, GeneMed) to release the RNA. Forward and reverse human transcriptome probes (PN-1000364, 10X Genomics) were used for probe hybridization overnight. After hybridization, the cDNA libraries were constructed according to the 10X protocol

(CG000407\_VisiumSpatialGeneExpressionforFFPE\_UserGuide\_RevA). Sequencing was performed on NovaSeq 6000 (Illumina). Raw sequencing reads were processed with Cell Ranger V3 and aligned to human genome (GRCh38, ENSEMBL). Seurat (v4.0) was used to process the Space Ranger output files. We used SCTransform to normalize the data, ScaleData to scale the data, RunPCA to perform dimension reduction, FindNeighbors and FindClusters to cluster the ST spots, and RunUMAP to visualize the data. Then, we chose the scaled matrix to calculate the mean values of the F2\_MCAM and CD8\_CXCL13 gene signatures in every spot. Last, we ran the log1p function for the mean values of the F2\_MCAM and CD8\_CXCL13 gene signatures as the F2\_MCAM and CD8\_CXCL13 scores, respectively. The reference genes used were listed in table S7. SpatialFeaturePlot was performed to visualize the infiltration score of each subset. Pearson correlation analysis was applied to calculate the infiltration relationship of different subsets in each spot. The functional pathway enrichment analysis was performed to explore the function differences between the tumor and paratumor region. Briefly, we calculated the mean expression levels of each gene in the tumor and paratumor regions and then run the gsva function to calculate the gene set enrichment scores for every hallmark gene sets, which were downloaded from MSigDB database (<http://gsea-msigdb.org/gsea/msigdb/>). The results were scaled and displayed in the heatmap. To estimate the cell cluster proportions in each captured spot, we performed the SPOTlight deconvolution analysis according to the standard pipeline supplied in <https://github.com/MarcElosua/SPOTlight>. Briefly, we extracted the marker gene matrix for each cell cluster in our scRNA-seq data by running the FindAllMarkers function in Seurat package, setting logfc.threshold = 1 and min.pct = 0.8. Then, the ST data and marker gene matrix were used to run the spotlight\_deconvolution function with the default parameters. The deconvolution results were displayed in the histology image by running the spatial\_scatterpie function, and each captured spot was renamed using the most dominant cluster there. Then, to infer the cell-cell communication in ST data, the renamed spatial object was used to perform the cellphoneDB analysis pipeline (<https://github.com/Teichlab/cellphonedb>).

### Cell lines and culture

Human CRC cell line RKO was a gift from H. Zhang (Guangdong Provincial People's Hospital). Human CRC cell line SW620 was a gift from Y. Yao (Guangzhou First People's Hospital). Both RKO and SW620 were maintained in DMEM (Gibco) supplemented with 10% fetal bovine serum (FBS) (ExCell Bio) and antibiotics (Gibco).

### Transwell migration assay

Transwell migration assay was performed to determine the influence of rNRG1 on the cell migration of CRC cell lines. Cells were seeded in 24-well Transwell inserting chambers (BD Biosciences) that contained serum-low medium. FBS (10%) with (50 ng/ml) or without (0 ng/ml) rNRG1 (R&D Systems, 396-HB-050) was added to the lower chamber as a source of chemoattractant. After 24 hours, the cells having not migrated through the chamber membrane were removed, and those having passed through the membrane were subjected to fixation and staining with crystal violet. They were subsequently detected and photographed using a microscope.

## Statistical analysis

Measurement data were presented as mean  $\pm$  SD. The Wilcoxon paired nonparametric test was used to compare percentages of different cell clusters among CN, CC, LN, LM, and PB. The infiltration scores of subsets were calculated by unpaired nonparametric test. The analysis was performed by SPSS 21 (IBM). Two-sided  $P < 0.05$  was considered significant.

## Supplementary Materials

This PDF file includes:

Figs. S1 to S8

Tables S1 to S7

Legend for supplementary code

Other Supplementary Material for this manuscript includes the following:

Supplementary code

[View/request a protocol for this paper from Bio-protocol.](#)

## REFERENCES AND NOTES

- H. Sung, J. Ferlay, R. L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, F. Bray, Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **71**, 209–249 (2021).
- R. L. Siegel, K. D. Miller, A. Jemal, Cancer statistics, 2020. *CA Cancer J. Clin.* **70**, 7–30 (2020).
- S. Manfredi, C. Lepage, C. Hatem, O. Coatmeur, J. Faivre, A.-M. Bouvier, Epidemiology and management of liver metastases from colorectal cancer. *Ann. Surg.* **244**, 254–259 (2006).
- C. Hackl, P. Neumann, M. Gerken, M. Loss, M. Klinkhammer-Schalke, H. J. Schlitt, Treatment of colorectal liver metastases in Germany: A ten-year population-based analysis of 5772 cases of primary colorectal adenocarcinoma. *BMC Cancer* **14**, 810 (2014).
- D. G. DeNardo, M. Johansson, L. M. Coussens, Immune cells as mediators of solid tumor metastasis. *Cancer Metastasis Rev.* **27**, 11–18 (2008).
- L. R. Rudmik, A. M. Magliocco, Molecular mechanisms of hepatic metastasis in colorectal cancer. *J. Surg. Oncol.* **92**, 347–359 (2005).
- A. Mantovani, P. Allavena, A. Sica, F. Balkwill, Cancer-related inflammation. *Nature* **454**, 436–444 (2008).
- Y. Yang, Y. Ma, S. Yan, P. Wang, J. Hu, S. Chen, J. Zhu, J. Wang, G. Chen, Y. Liu, CAF promotes chemoresistance through NRP2 in gastric cancer. *Gastric Cancer* **25**, 503–514 (2022).
- Z. Yan, Z. Sheng, Y. Zheng, R. Feng, Q. Xiao, L. Shi, H. Li, C. Yin, H. Luo, C. Hao, W. Wang, B. Zhang, Cancer-associated fibroblast-derived exosomal miR-18b promotes breast cancer invasion and metastasis by regulating TCEAL7. *Cell Death Dis.* **12**, 1120 (2021).
- Y. Gao, Z. Sun, J. Gu, Z. Li, X. Xu, C. Xue, X. Li, L. Zhao, J. Zhou, C. Bai, Q. Han, R. C. Zhao, Cancer-associated fibroblasts promote the up-regulation of PD-L1 expression through Akt phosphorylation in colorectal cancer. *Front. Oncol.* **11**, 748465 (2021).
- B. Y. Lee, E. K. J. Hogg, C. R. Below, A. Kononov, A. Blanco-Gomez, F. Heider, J. Xu, C. Hutton, X. Zhang, T. Scheidt, K. Beattie, A. Lamarca, M. McNamara, J. W. Valle, C. Jorgensen, Heterocellular OSM-OSMR signalling reprograms fibroblasts to promote pancreatic cancer growth and metastasis. *Nat. Commun.* **12**, 7336 (2021).
- H. Kobayashi, K. A. Gieniec, T. R. M. Lannagan, T. Wang, N. Asai, Y. Mizutani, T. Iida, R. Ando, E. M. Thomas, A. Sakai, N. Suzuki, M. Ichinose, J. A. Wright, L. Vrbancac, J. Q. Ng, J. Goyne, G. Radford, M. J. Lawrence, T. Sammour, Y. Hayakawa, S. Klebe, A. E. Shin, S. Asfaha, M. L. Bettington, F. Rieder, N. Arpaia, T. Danino, L. M. Butler, A. D. Burt, S. J. Leedham, A. K. Rustgi, S. Mukherjee, M. Takahashi, T. C. Wang, A. Enomoto, S. L. Woods, D. L. Worthley, The origin and contribution of cancer-associated fibroblasts in colorectal carcinogenesis. *Gastroenterology* **162**, 890–906 (2022).
- S. Su, J. Chen, H. Yao, J. Liu, S. Yu, L. Lao, M. Wang, M. Luo, Y. Xing, F. Chen, D. Huang, J. Zhao, L. Yang, D. Liao, F. Su, M. Li, Q. Liu, E. Song, CD10<sup>+</sup>GPR77<sup>+</sup> cancer-associated fibroblasts promote cancer formation and chemoresistance by sustaining cancer stemness. *Cell* **172**, 841–856.e16 (2018).
- E. Elyada, M. Bolisetty, P. Laise, W. F. Flynn, E. T. Courtois, R. A. Burkhart, J. A. Teinor, P. Belleau, G. Biffi, M. S. Lucito, S. Sivajothi, T. D. Armstrong, D. D. Engle, K. H. Yu, Y. Hao, C. L. Wolfgang, Y. Park, J. Preall, E. M. Jaffe, A. Califano, P. Robson, D. A. Tuveson, Cross-species single-cell analysis of pancreatic ductal adenocarcinoma reveals antigen-presenting cancer-associated fibroblasts. *Cancer Discov.* **9**, 1102–1123 (2019).
- W. H. Fridman, F. Pages, C. Sautes-Fridman, J. Galon, The immune contexture in human tumours: Impact on clinical outcome. *Nat. Rev. Cancer* **12**, 298–306 (2012).
- H.-O. Lee, Y. Hong, H. E. Etlioglu, Y. B. Cho, V. Pomella, B. Van den Bosch, J. Vanhecke, S. Verbandt, H. Hong, J.-W. Min, N. Kim, H. H. Eum, J. Qian, B. Boeckx, D. Lambrechts, P. Tsantoulis, G. De Hertogh, W. Chung, T. Lee, M. An, H.-T. Shin, J.-G. Joong, M.-H. Jung, G. Ko, P. Wirapati, S. H. Kim, H. C. Kim, S. H. Yun, I. B. H. Tan, B. Ranjan, W. Y. Lee, T.-Y. Kim, J. K. Choi, Y.-J. Kim, S. Prabhakar, S. Tejpar, W.-Y. Park, Lineage-dependent gene expression programs influence the immune landscape of colorectal cancer. *Nat. Genet.* **52**, 594–603 (2020).
- W. Wang, Y. Zhong, Z. Zhuang, J. Xie, Y. Lu, C. Huang, Y. Sun, L. Wu, J. Yin, H. Yu, Z. Jiang, S. Wang, C. Wang, Y. Zhang, Y. Huang, C. Han, Z. Zhong, J. Hu, Y. Ouyang, H. Liu, M. Yu, X. Wei, D. Chen, L. Huang, Y. Hou, Z. Lin, S. Liu, F. Ling, X. Yao, Multiregion single-cell sequencing reveals the transcriptional landscape of the immune microenvironment of colorectal cancer. *Clin. Transl. Med.* **11**, e253 (2021).
- M. Elosua-Bayes, P. Nieto, E. Mereu, I. Gut, H. Heyn, SPOTlight: Seeded NMF regression to deconvolute spatial transcriptomics spots with single-cell transcriptomes. *Nucleic Acids Res.* **49**, e50 (2021).
- P. Dalerba, T. Kalisky, D. Sahoo, P. S. Rajendran, M. E. Rothenberg, A. A. Leyrat, S. Sim, J. Okamoto, D. M. Johnston, D. Qian, M. Zabala, J. Bueno, N. F. Neff, J. Wang, A. A. Shelton, B. Visser, S. Hisamori, Y. Shimono, M. van de Wetering, H. Clevers, M. F. Clarke, S. R. Quake, Single-cell dissection of transcriptional heterogeneity in human colon tumors. *Nat. Biotechnol.* **29**, 1120–1127 (2011).
- L. A. Sokolsky, K. Garcia-Netto, T. H. Nguyen, J. L. N. Girkin, A. Collison, J. Mattes, G. Kaiko, C. Liu, N. W. Bartlett, M. Yang, P. S. Foster, A critical role for the CXCL3/CXCL5/CXCR2 neutrophilic chemotactic axis in the regulation of type 2 responses in a model of rhinoviral-induced asthma exacerbation. *J. Immunol.* **205**, 2468–2478 (2020).
- D. I. Gabrilovich, S. Nagaraj, Myeloid-derived suppressor cells as regulators of the immune system. *Nat. Rev. Immunol.* **9**, 162–174 (2009).
- S. Cheng, Z. Li, R. Gao, B. Xing, Y. Gao, Y. Yang, S. Qin, L. Zhang, H. Ouyang, P. Du, L. Jiang, B. Zhang, Y. Yang, X. Wang, X. Ren, J. X. Bei, X. Hu, Z. Bu, J. Ji, Z. Zhang, A pan-cancer single-cell transcriptional atlas of tumor infiltrating myeloid cells. *Cell* **184**, 792–809.e23 (2021).
- T. Duhen, R. Duhen, R. Montler, J. Moses, T. Moudgil, N. F. de Miranda, C. P. Goodall, T. C. Blair, B. A. Fox, J. E. Mcdermott, S. C. Chang, G. Grunkeimer, R. Leidner, R. B. Bell, A. D. Weinberg, Co-expression of CD39 and CD103 identifies tumor-reactive CD8 T cells in human solid tumors. *Nat. Commun.* **9**, 2724 (2018).
- K. Litchfield, J. L. Reading, C. Puttick, K. Thakkar, C. Abbosh, R. Bentham, T. B. K. Watkins, R. Rosenthal, D. Biswas, A. Rowan, E. Lim, M. A. Bakir, V. Turati, J. A. Guerra-Assuncao, L. Conde, A. J. S. Furness, S. K. Saini, S. R. Hadrup, J. Herrero, S.-H. Lee, P. Van Loo, T. Enver, J. Larkin, M. D. Hellmann, S. Turajlic, S. A. Quezada, N. Mcgranahan, C. Swanton, Meta-analysis of tumor- and T cell-intrinsic mechanisms of sensitization to checkpoint inhibition. *Cell* **184**, 596–614.e14 (2021).
- Y. Simoni, E. Becht, M. Fehlings, C. Y. Loh, S.-L. Koo, K. W. W. Teng, J. P. S. Yeong, R. Nahar, T. Zhang, H. Kared, K. Duan, N. Ang, M. Poidinger, Y. Y. Lee, A. Larbi, A. J. Khng, E. Tan, C. Fu, R. Mathew, M. Teo, W. T. Lim, C. K. Toh, B.-H. Ong, T. Koh, A. M. Hillmer, A. Takano, T. K. H. Lim, E. H. Tan, W. Zhai, D. S. W. Tan, I. B. Tan, E. W. Newell, Bystander CD8<sup>+</sup> T cells are abundant and phenotypically distinct in human tumour infiltrates. *Nature* **557**, 575–579 (2018).
- Y. Zhang, H. Chen, H. Mo, X. Hu, R. Gao, Y. Zhao, B. Liu, L. Niu, X. Sun, X. Yu, Y. Wang, Q. Chang, T. Gong, X. Guan, T. Hu, T. Qian, B. Xu, F. Ma, Z. Zhang, Z. Liu, Single-cell analyses reveal key immune cell subsets associated with response to PD-L1 blockade in triple-negative breast cancer. *Cancer Cell* **39**, 1578–1593.e8 (2021).
- S. V. Puram, I. Tirosh, A. S. Parikh, A. P. Patel, K. Yizhak, S. Gillespie, C. Rodman, C. L. Luo, E. A. Mroz, K. S. Emerick, D. G. Deschler, M. A. Varvares, R. Mylvaganam, O. Rozenblatt-Rosen, J. W. Rocco, W. C. Faquin, D. T. Lin, A. Regev, B. E. Bernstein, Single-cell transcriptomic analysis of primary and metastatic tumor ecosystems in head and neck cancer. *Cell* **171**, 1611–1624.e24 (2017).
- T. Holbro, R. R. Beerli, F. Maurer, M. Koziczak, C. F. Barbas III, N. E. Hynes, The ErbB2/ErbB3 heterodimer functions as an oncogenic unit: ErbB2 requires ErbB3 to drive breast tumor cell proliferation. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 8933–8938 (2003).
- K. Yonesaka, K. Zejnullahu, I. Okamoto, T. Satoh, F. Cappuzzo, J. Souglakos, D. Ercan, A. Rogers, M. Roncalli, M. Takeda, Y. Fujisaka, J. Phillips, T. Shimizu, O. Maenishi, Y. Cho, J. Sun, A. Destro, K. Taira, K. Takeda, T. Okabe, J. Swanson, H. Itoh, M. Takada, E. Lifshits, K. Okuno, J. A. Engelman, R. A. Shivdasani, K. Nishio, M. Fukuoka, M. Varella-Garcia, K. Nakagawa, P. A. Janne, Activation of ERBB2 signaling causes resistance to the EGFR-directed therapeutic antibody cetuximab. *Sci. Transl. Med.* **3**, 99ra86 (2011).
- R. Kopan, M. X. G. Ilagan, The canonical Notch signaling pathway: Unfolding the activation mechanism. *Cell* **137**, 216–233 (2009).
- S. I. Tsukumo, K. Yasutomo, Regulation of CD8<sup>+</sup> T cells and antitumor immunity by Notch signaling. *Front. Immunol.* **9**, 101 (2018).
- H. Hu, Z. Piotrowska, P. J. Hare, H. Chen, H. E. Mulvey, A. Mayfield, S. Noeen, K. Kattermann, M. Greenberg, A. Williams, A. K. Riley, J. J. Wilson, Y. Q. Mao, R. P. Huang, M. K. Banwait, J. Ho,

- G. S. Crowther, L. P. Harii, R. S. Heist, D. P. Kodack, L. Pinello, A. T. Shaw, M. Mino-Kenudson, A. N. Hata, L. V. Sequist, C. H. Benes, M. J. Niederst, J. A. Engelman, Three subtypes of lung cancer fibroblasts define distinct therapeutic paradigms. *Cancer Cell* **39**, 1531–1547.e10 (2021).
33. E. J. Helms, M. W. Berry, R. C. Chaw, C. C. DuFort, D. Sun, M. K. Onate, C. Oon, S. Bhattacharyya, H. Sanford-Crane, W. Horton, J. M. Finan, A. Sattler, R. Makar, D. W. Dawson, Z. Xia, S. R. Hingorani, M. H. Sherman, Mesenchymal lineage heterogeneity underlies non-redundant functions of pancreatic cancer-associated fibroblasts. *Cancer Discov.* **12**, 484–501 (2021).
34. B. Zhong, B. Cheng, X. Huang, Q. Xiao, Z. Niu, Y.-F. Chen, Q. Yu, W. Wang, X.-J. Wu, Colorectal cancer-associated fibroblasts promote metastasis by up-regulating LRG1 through stromal IL-6/STAT3 signaling. *Cell Death Dis.* **13**, 16 (2021).
35. A. Kawase, G. Ishii, K. Nagai, T. Ito, T. Nagano, Y. Murata, T. Hishida, M. Nishimura, J. Yoshida, K. Suzuki, A. Ochiai, Podoplanin expression by cancer associated fibroblasts predicts poor prognosis of lung adenocarcinoma. *Int. J. Cancer* **123**, 1053–1059 (2008).
36. C. Feig, J. O. Jones, M. Kraman, R. J. B. Wells, A. Deonarine, D. S. Chan, C. M. Connell, E. W. Roberts, Q. Zhao, O. L. Caballero, S. A. Teichmann, T. Janowitz, D. I. Jodrell, D. A. Tuveson, D. T. Fearon, Targeting CXCL12 from FAP-expressing carcinoma-associated fibroblasts synergizes with anti-PD-L1 immunotherapy in pancreatic cancer. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 20212–20217 (2013).
37. E. Sahai, I. Astsaturov, E. Cukierman, D. G. Denardo, M. Egeblad, R. M. Evans, D. Fearon, F. R. Greten, S. R. Hingorani, T. Hunter, R. O. Hynes, R. K. Jain, T. Janowitz, C. Jorgensen, A. C. Kimmelman, M. G. Kolonin, R. G. Maki, R. S. Powers, E. Pure, D. C. Ramirez, R. Scherz-Shouval, M. H. Sherman, S. Stewart, T. D. Tlsty, D. A. Tuveson, F. M. Watt, V. Weaver, A. T. Weeraratna, Z. Werb, A framework for advancing our understanding of cancer-associated fibroblasts. *Nat. Rev. Cancer* **20**, 174–186 (2020).
38. J. Qi, H. Sun, Y. Zhang, Z. Wang, Z. Xun, Z. Li, X. Ding, R. Bao, L. Hong, W. Jia, F. Fang, H. Liu, L. Chen, J. Zhong, D. Zou, L. Liu, L. Han, F. Ginhoux, Y. Liu, Y. Ye, B. Su, Single-cell and spatial analysis reveal interaction of FAP<sup>+</sup> fibroblasts and SPP1<sup>+</sup> macrophages in colorectal cancer. *Nat. Commun.* **13**, 1742 (2022).
39. A. De Boeck, P. Pauwels, K. Hensen, J.-L. Rummens, W. Westbroek, A. Hendrix, D. Maynard, H. Denys, K. Lambein, G. Braems, C. Gespach, M. Bracke, O. De Wever, Bone marrow-derived mesenchymal stem cells promote colorectal cancer progression through paracrine neuregulin 1/HER3 signalling. *Gut* **62**, 550–560 (2013).
40. S. Jonna, R. A. Feldman, J. Swensen, Z. Gatalica, W. M. Korn, H. Borghaei, P. C. Ma, J. J. Nieva, A. I. Spira, A. M. Vanderwalde, A. J. Wozniak, E. S. Kim, S. V. Liu, Detection of NRG1 gene fusions in solid tumors. *Clin. Cancer Res.* **25**, 4966–4972 (2019).
41. S. M. Bray, J. Lee, S. T. Kim, J. Y. Hur, P. J. Ebert, J. N. Calley, I. H. Wulur, T. Gopalappa, S. S. Wong, H. R. Qian, J. C. Ting, J. Liu, M. D. Willard, R. D. Novosiadly, Y. S. Park, J. O. Park, H. Y. Lim, W. K. Kang, A. Aggarwal, H. C. Kim, C. Reinhard, Genomic characterization of intrinsic and acquired resistance to cetuximab in colorectal cancer patients. *Sci. Rep.* **9**, 15365 (2019).
42. J. W. Cho, S. Park, G. Kim, H. Han, H. S. Shim, S. Shin, Y. S. Bae, S. Y. Park, S. J. Ha, I. Lee, H. R. Kim, Dysregulation of T(FH)-B-T(RM) lymphocyte cooperation is associated with unfavorable anti-PD-1 responses in EGFR-mutant lung cancer. *Nat. Commun.* **12**, 6068 (2021).
43. H. H. Workel, J. M. Lubbers, R. Arnold, T. M. Prins, P. van der Vlies, K. de Lange, T. Bosse, I. C. van Gool, F. A. Eggink, M. C. A. Wouters, F. L. Komdeur, E. C. van der Slikke, C. L. Creutzberg, A. Kol, A. Plat, M. Glaire, D. N. Church, H. W. Nijman, M. de Bruyn, A transcriptionally distinct CXCL13<sup>+</sup>CD103<sup>+</sup>CD8<sup>+</sup> T-cell population is associated with B-cell recruitment and neoantigen load in human cancer. *Cancer Immunol. Res.* **7**, 784–796 (2019).
44. B. A. Helmink, S. M. Reddy, J. Gao, S. Zhang, R. Basar, R. Thakur, K. Yizhak, M. Sade-Feldman, J. Blando, G. Han, V. Gopalakrishnan, Y. Xi, H. Zhao, R. N. Amaria, H. A. Tawbi, A. P. Cogdill, W. Liu, V. S. Lebleu, F. G. Kugeratski, S. Patel, M. A. Davies, P. Hwu, J. E. Lee, J. E. Gershenwald, A. Lucci, R. Arora, S. Woodman, E. Z. Keung, P. O. Gaudreau, A. Reuben, C. N. Spencer, E. M. Burton, L. E. Haydu, A. J. Lazar, R. Zappasodi, C. W. Hudgens, D. A. Ledesma, S. Ong, M. Bailey, S. Warren, D. Rao, O. Krijgsman, E. A. Rozeman, D. Peepker, C. U. Blank, T. N. Schumacher, L. H. Butterfield, M. A. Zelazowska, K. M. McBride, R. Kalluri, J. Allison, F. Petitprez, W. H. Fridman, C. Sautes-Fridman, N. Hacohen, K. Rezvani, P. Sharma, M. T. Tetzlaff, L. Wang, J. A. Wargo, B cells and tertiary lymphoid structures promote immunotherapy response. *Nature* **577**, 549–555 (2020).
45. R. Cabrita, M. Lauss, A. Sanna, M. Donia, M. S. Larsen, S. Mitra, I. Johansson, B. Phung, K. Harbst, J. Vallon-Christersson, A. van Schoiack, K. Lovgren, S. Warren, K. Jirstrom, H. Olsson, K. Pietras, C. Ingvar, K. Isaksson, D. Schadtendorf, H. Schmidt, L. Bastholt, A. Carneiro, J. A. Wargo, I. M. Svane, G. Jonsson, Tertiary lymphoid structures improve immunotherapy and survival in melanoma. *Nature* **577**, 561–565 (2020).
46. F. Wang, J. Long, L. Li, Z. B. Zhao, F. Wei, Y. Yao, W. J. Qiu, Z. X. Wu, Q. Q. Luo, W. Liu, Y. B. Quan, Z. X. Lian, J. Cao, Mutations in the notch signalling pathway are associated with enhanced anti-tumour immunity in colorectal cancer. *J. Cell. Mol. Med.* **24**, 12176–12187 (2020).
47. F. Wang, C. Huang, J. Long, Z. B. Zhao, H. Q. Ma, X. Q. Yao, L. Li, Z. X. Lian, Notch signaling mutations increase intra-tumor chemokine expression and predict response to immunotherapy in colorectal cancer. *BMC Cancer* **22**, 933 (2022).
48. S. L. Wolock, R. Lopez, A. M. Klein, Scrublet: Computational identification of cell doublets in single-cell transcriptomic data. *Cell Syst.* **8**, 281–291.e9 (2019).
49. Y. Senbabaoglu, R. S. Gejman, A. G. Winer, M. Liu, E. M. Van Allen, G. de Velasco, D. Miao, I. Ostrovnya, E. Drill, A. Luna, N. Weinhold, W. Lee, B. J. Manley, D. N. Khalil, S. D. Kaffenberger, Y. Chen, L. Danilova, M. H. Voss, J. A. Coleman, P. Russo, V. E. Reuter, T. A. Chan, E. H. Cheng, D. A. Scheinberg, M. O. Li, T. K. Choueiri, J. J. Hsieh, C. Sander, A. A. Hakimi, Tumor immune microenvironment characterization in clear cell renal cell carcinoma identifies prognostic and immunotherapeutically relevant messenger RNA signatures. *Genome Biol.* **17**, 231 (2016).
50. S. Hanzelmann, R. Castelo, J. Guinney, GSEA: Gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics* **14**, 7 (2013).
51. S. Aibar, C. B. Gonzalez-Blas, T. Moerman, V. A. Huynh-Thu, H. Imrichova, G. Hulselmans, F. Rambow, J. C. Marine, P. Geurts, J. Aerts, J. van den Oord, Z. K. Atak, J. Wouters, S. Aerts, SCENIC: Single-cell regulatory network inference and clustering. *Nat. Methods* **14**, 1083–1086 (2017).
52. X. Qiu, Q. Mao, Y. Tang, L. Wang, R. Chawla, H. A. Pliner, C. Trapnell, Reversed graph embedding resolves complex single-cell trajectories. *Nat. Methods* **14**, 979–982 (2017).
53. R. Vento-Tormo, M. Efrémova, R. A. Botting, M. Y. Turco, M. Vento-Tormo, K. B. Meyer, J.-E. Park, E. Stephenson, K. Polanski, A. Goncalves, L. Gardner, S. Holmqvist, J. Henriksson, A. Zou, A. M. Sharkey, B. Millar, B. Innes, L. Wood, A. Wilbrey-Clark, R. P. Payne, M. A. Ivarsson, S. Liso, A. Filby, D. H. Rowitch, J. N. Bulmer, G. J. Wright, M. J. T. Stubbington, M. Haniffa, A. Moffett, S. A. Teichmann, Single-cell reconstruction of the early maternal-fetal interface in humans. *Nature* **563**, 347–353 (2018).
54. S. Jin, C. F. Guerrero-Juarez, L. Zhang, I. Chang, R. Ramos, C. H. Kuan, P. Myung, M. V. Plikus, Q. Nie, Inference and analysis of cell-cell communication using CellChat. *Nat. Commun.* **12**, 1088 (2021).

**Acknowledgments:** We gratefully acknowledge X.-Y. Zhong and Y.-S. Xu for the assistance in the scRNA-seq experiment and data analysis. We thank H. Zhang and Y. Yao for supplying CRC cell lines. **Funding:** This work was supported by the National Natural Science Foundation of China (82120108013 to Z.-X.L., 81871943 to J.C., 82173236 to J.C., and 82072719 to H.-Q.M.), the Program for Guangdong Introducing Innovative and Entrepreneurial Teams (2017ZT07S054 to Z.-X.L.), and the Guangdong Provincial Department of Science and Technology (202206080008 for J.C.). **Author contributions:** F.W. and J.L. designed the study, collected samples, performed the experiments, analyzed the data, and wrote the manuscript. L.L. and Z.-X.W. analyzed the data of single-cell datasets. T.-T.D., X.-Q.W., C.H., and X.-Q.Y. performed the experiments. Y.-H.J. analyzed spatial transcriptomic sequencing data. J.C., Z.-B.Z., Z.-X.L., and H.-Q.M. designed and supervised the study, provided their expertise in the analysis of the data, and edited the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. The data can be obtained in GEO under the accession number: GSE225857 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE225857>). The R package was obtained at <https://r-project.org>, Seurat at <http://satijalab.org/seurat/>, SCENIC at <https://rdocumentation.org/packages/SCENIC/versions/1.1.2-01>, CellChat at <http://cellchat.org/>, and CellPhone DB at <http://CellPhoneDB.org>. The code for data cleaning and analysis is hosted at [https://github.com/jalon9358/LianLab\\_CRCLM](https://github.com/jalon9358/LianLab_CRCLM) and is included as a supplementary material, which can be used to reproduce the aforementioned analysis.

Submitted 30 October 2022

Accepted 10 May 2023

Published 16 June 2023

10.1126/sciadv.adf5464