Original research

Single-cell transcriptomic analysis deciphers heterogenous cancer stem-like cells in colorectal cancer and their organ-specific metastasis

Rui Li,[1] Xuefei Liu,[1,2,3] Xudong Huang,[1] Di Zhang,[4] Ziming Chen,[1] Jialiang Zhang,[1] Ruihong Bai,[1] Shaoping Zhang,[1] Hongzhe Zhao,[1] Zilan Xu,[1] Lingxing Zeng,[1] Lisha Zhuang,[1] Shujuan Wen,[1] Shaojia Wu,[1] Mei Li,[5] Zhixiang Zuo,[1] Junzhong Lin,[6] Dongxin Lin,[1,7,8] Jian Zheng ![ORCID] [1,8,9]

Correspondence to
Prof. Jian Zheng and Prof. Dongxin Lin, State Key Laboratory of Oncology in South China and Guangdong Provincial Clinical Research Center for Cancer, Sun Yat-sen University Cancer Center, Guangzhou, China; zhengjian@sysucc.org.cn, lindx@cicams.ac.cn and Dr. Junzhong Lin, Department of Colorectal Surgery, Sun Yat-sen University Cancer Center, Guangzhou, China; linjzh@sysucc.org.cn

RL, XL, XH and DZ contributed equally.

ABSTRACT

Objective Metastasis is the major cause of cancer death. However, what types of heterogenous cancer cells in primary tumour and how they metastasise to the target organs remain largely undiscovered.

Design We performed single-cell RNA sequencing and spatial transcriptomic analysis in primary colorectal cancer (CRC) and metastases in the liver (lCRC) or ovary (oCRC). We also conducted immunofluorescence staining and functional experiments to examine the mechanism.

Results Integrative analyses of epithelial cells reveal a stem-like cell cluster with high *protein tyrosine phosphatase receptor type O* (*PTPRO*) and *achaete scute-like 2* (*ASCL2*) expression as the metastatic culprit. This cell cluster comprising distinct subpopulations shows distinct liver or ovary metastatic preference. Population 1 (P1) cells with high *delta-like ligand 4 (DLL4)* and *MAF bZIP transcription factor A (MAFA)* expression are enriched in primary CRC and oCRC, thus may be associated with ovarian metastasis. P3 cells having a similar expression pattern as cholangiocytes are found mainly in primary CRC and lCRC, presuming to be likely the culprits that specifically metastasise to the liver. Stem-like cells interacted with cancer-associated fibroblasts and endothelial cells via the DLL4-NOTCH signalling pathway to metastasise from primary CRC to the ovary. In the oCRC microenvironment, myofibroblasts provide cancer cells with glutamine and perform a metabolic reprogramming, which may be essential for cancer cells to localise and develop in the ovary.

Conclusion We uncover a mechanism for organ-specific CRC metastasis.

INTRODUCTION

Colorectal cancer (CRC) is the third most common and lethal cancer in the world.[1] Tumour metastasis accounts for 90% of CRC death,[2] with reported 5-year survival rates being <20%.[3] Although the most common sites CRC metastasises to are the lymph node, liver and lung, a proportion of CRCs can metastasise to the bone and ovary.[4] The ovarian metastasis of CRC is believed to be directly colonised by malignant cells spreading from the serosal surface into the peritoneal cavity through ascites.[5] Over the last five decades, studies on various types of cancer have proposed that despite the presence of many heterogeneous cells in tumours, only certain type of malignant cells can drive the relapse and distant metastasis.[6–8] However to date, there appear to be no systematic studies on what subtype of heterogeneous CRC cells prefers distant metastasis.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Colorectal cancer (CRC) is a highly heterogeneous malignant disease that can metastasise to lymph node, liver, lung, bone and ovary.

⇒ Heterogeneity in CRC primary tumour leads to organ-specific metastasis that remains largely undiscovered.

WHAT THIS STUDY ADDS

⇒ We decipher subtypes of heterogenous cancer cells in primary CRC tissue, CRC liver and ovarian metastasis tissue, and profile characteristic for organ-specific CRC metastasis at single cell resolution.

⇒ We identify that a stem-like cell cluster with high expression of achaete scute-like 2 and protein tyrosine phosphatase receptor type O is the cause of CRC liver and ovarian metastasis.

⇒ Malignant cells with high delta-like ligand 4 (DLL4) expression interact with cancer-associated fibroblasts and endothelial cells via the DLL4-NOTCH signalling pathway to metastasise from primary CRC to the ovary.

⇒ In ovarian metastasis, myofibroblasts provide cancer cells with glutamine and induce metabolic remodelling in malignant cells.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ The comprehensive transcriptome atlas is a valuable resource for primary CRC and distant metastasis, especially for CRC ovarian metastasis.

⇒ Our findings provide an important theoretical basis for organ-specific CRC metastasis.

Single-cell RNA sequencing (scRNA-seq) can dissect heterogeneous tumour and decipher the cell compositions. Recent studies have compared cell compositions in the microenvironments of primary CRC and their liver metastatic CRC (lCRC) and found that certain subtypes of immune cells such as tumour-associated macrophages and pericytes were significantly enriched in lCRC compared with primary CRC.[9–11] Another study has reported increased immature plasma cells but decreased activated B cells in lCRC than primary CRC.[12] Although these studies have indicated the importance of different immune features in the microenvironments of CRC and lCRC, it remains unrevealed what subtype of malignant cells are the metastatic culprit. Furthermore, to the best of our knowledge, the tumour microenvironment and cell composition of ovarian metastatic CRC (oCRC) has never been explored yet.

In this study, we have performed scRNA-seq and spatial transcriptomic analysis of primary CRC and lCRC or oCRC, focusing on deciphering which subtype of malignant cells in primary CRC may metastasise to the ovary and liver and how these metastases grow in the target organs. We have discovered that a stem-like cell cluster, which expresses protein tyrosine phosphatase receptor type O (PTPRO), is the metastatic initiator in CRC. These stem-like cells also exhibit heterogeneity in terms of the transcription pattern that determines the cells to metastasise to the liver or ovary. The development of oCRC relies on the interaction network of NOTCH signalling pathway between the metastatic cells and cancer-associated fibroblasts (CAFs) or endothelial cells, and myofibroblasts also create a metabolically active state in the microenvironment to help the oCRC growth. Overall, our results have provided an important insight into CRC metastasise to the ovary and liver and identified a subtype of CRC cells for the organ-specific metastasis. These results may provide an important theoretical basis for organ-specific CRC metastasis.

METHODS
Cell culture
Human microsatellite stable CRC cell lines HT-29 and SW-480, human embryonic kidney cell line 293T and human umbilical vein endothelial cell (HUVEC) were purchased from the Cell Bank of Type Culture Collection of the Chinese Academy of Sciences Shanghai Institute of Biochemistry and Cell Biology. Human ovarian fibroblast cell line was purchased from Biotechnology Company. CRC cells and HUVEC were cultured in RPMI-1640 with 10% fetal bovine serum (FBS) and 293T cells were cultured in Dulbecco's Modified Eagle Medium with 10% FBS in an atmosphere of 5% CO_2 and 99% relative humidity at 37°C. Human ovarian fibroblast cell line was cultured in ICell Primary Fibroblast Culture System.

A detailed description of all methods used in this study can be found in online supplemental file 1.

RESULTS
Single-cell transcriptomic landscape in CRC
We performed scRNA-seq on 29 samples from 7 patients with CRC with distant metastases, including 8 primary CRC, 5 oCRC, 2 lCRC, 7 adjacent normal tissues and 7 peripheral blood mononuclear cells (PBMC). We also downloaded the publicly available scRNA-seq data of 41 samples from 24 patients, including 29 CRC and 12 adjacent normal tissues, to increase the statistical power. Among these 24 patients, 3 had distant metastases but the detailed metastatic information is unavailable.[13] Thus, for the final analysis, we had 31 patients with CRC including 10

distant metastases. All CRC were diagnosed as microsatellite stable (MSS) CRC. The working flow and the detailed information of patients and tissue samples are shown in figure 1A and online supplemental tables 1 and 2.
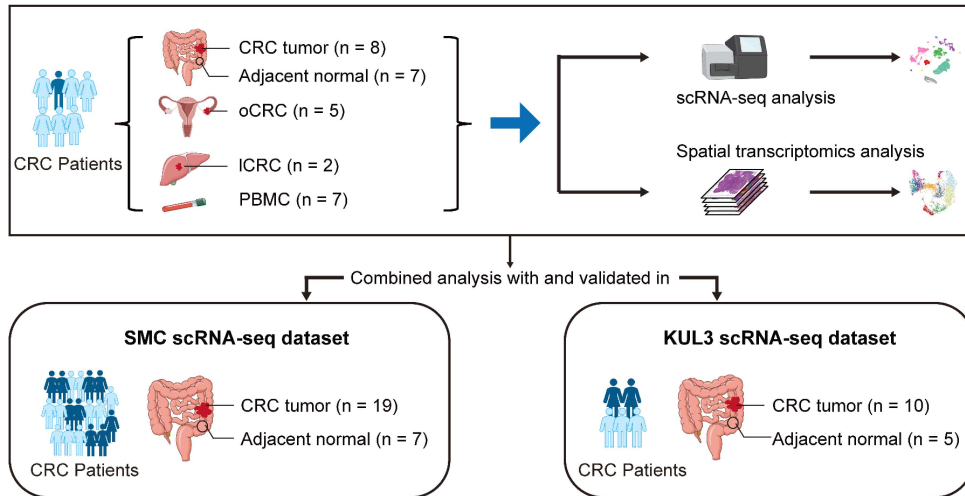
After quality control and batch correction, we analysed scRNA-seq data of 230 818 cells, in which 167 205 (72.44%) were from our own samples (online supplemental figure 1A). Using the classical markers of different cell clusters, we identified 10 cell types including 4 non-immune cell types such as epithelial cells, fibroblasts, endothelial cells and glial cells and 6 immune cell types such as B cells, T cells, myeloid cells, mast cells, megakaryocytes and plasma B cells (figure 1B–C and online supplemental figure 1B–D). We then calculated the proportions of these 10 cell types in different tissue types, which are shown in figure 1D–E and online supplemental table 3.

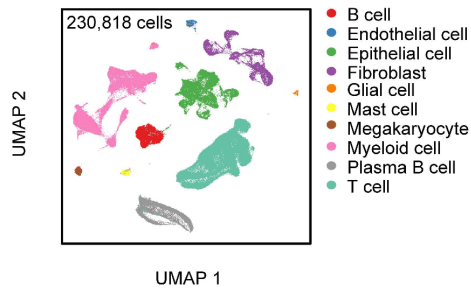### Stem-like cells are the originator of CRC metastasis
We subclassified the epithelial cells expressing *EPCAM* and *KRT18* (online supplemental figure 2A) and found 14 cell subtypes that were present in all patients (figure 2A and online supplemental figure 2B). The marker genes of these 14 cell subtypes are shown in figure 2B. Among the 14 cell subtypes, we identified 5 subtypes of malignant cells including stem-like cells highly expressing *ASCL2* and *PTPRO*, DNAJB1[+] malignant cells expressing *DNAJB1* and *RND3*, FDPS[+] malignant cells expressing *FDPS* and *SCD*, GPRC5A[+] malignant cells expressing *GPRC5A* and *SLC2A1* and IL32[+] malignant cells expressing *IL32* and *S100A11* (figure 2B and online supplemental figure 2C–E). We found that enterocytes were enriched in normal colorectal tissues, cholangiocytes were enriched in lCRC and stem-like cells and other malignant cell subtypes were enriched in both primary and metastatic CRC samples (figure 2C). The stem-like malignant cell subtypes from primary CRC and metastatic CRC had significantly higher DNA copy number variations than epithelial cells from adjacent normal tissues (figure 2D–E and online supplemental figure 2F). We then applied CytoTRACE[14] to predict the differentiation states of five malignant cell subtypes and the results showed that the stem-like cells had a significantly higher differentiation potential than other cell types in the lCRC or oCRC (online supplemental figure 2G). The partition-based graph abstraction analysis showed that the stem-like cells had a closer evolutionary relationship with the malignant cell subtypes than other epithelial cell subtypes (figure 2F). The RNA velocity analysis indicated that the flow directions were from the stem-like cells to other malignant cell subtypes (figure 2F). Moreover, the stem-like cells had the highest stemness signature score among all epithelial cell subtypes (figure 2G). These results suggested that the stem-like cells might be the originator of CRC metastasis.

We applied trajectory-inference algorithm to predict the flow directions of the malignant cell subtypes in primary tumours and metastatic tumours. The results showed that in both primary and metastatic CRC, other malignant cells seemed to come from stem-like cells (online supplemental figure 2H). We further performed differentially expressed gene analysis and found that compared with other cell subtypes, stem-like cells had significant upregulation of *PTPRO* and some stemness-associated markers including *ASCL2*, *LGR5*, *AXIN2*, *EPHB2*, *SOX9*, *PROM1*, *CD44* and *ALCAM* (figure 2H). We also revealed a significant expression elevation of genes in the MYC target V1 and WNT/β-catenin signalling pathways in the stem-like cells (figure 2I). Nevertheless, the expression program were heterogeneous among the malignant cell subtypes; for example, some
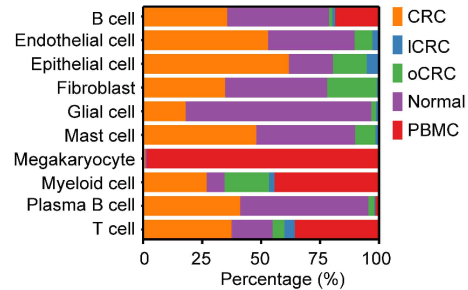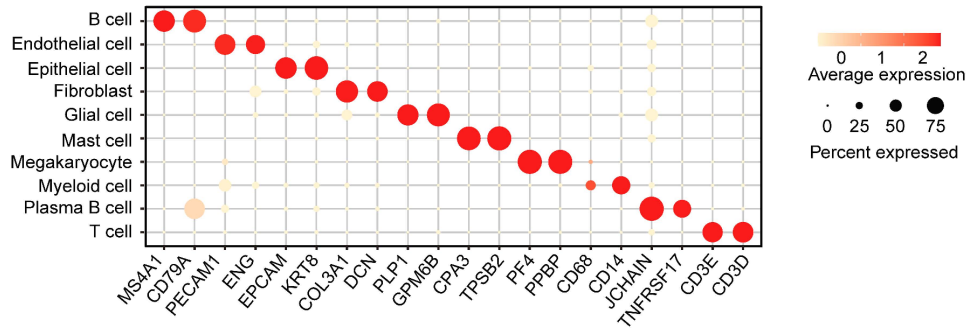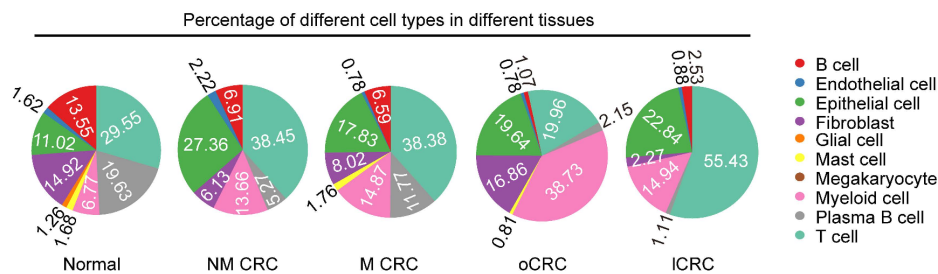
Figure 1 Single-cell transcriptomic landscape in human CRC. (A) The schematics of sample collection, scRNA-seq and spatial transcriptomic analysis of CRC and combined analysis with public datasets. CRC, colorectal cancer; lCRC, liver metastatic CRC; oCRC, ovarian metastatic CRC; PBMC, peripheral blood mononuclear cells. (B) Uniform manifold approximation and projection (UMAP) of 230 818 cells analysed by scRNA-seq across all samples. Clusters were annotated by the canonical markers. (C) The expression levels of the selected markers in different cell subtypes. Dot size indicates the fraction of expressing cells and the colours represent normalised gene expression levels. (D) Relative proportions of cell subtypes across different tissue types. (E) The proportion of different cell subtypes across normal, primary CRC from patients without distant metastases (NM-CRC), primary CRC from patients with distant metastases (M-CRC), oCRC and lCRC. Only cell subtypes with the proportion being >0.5% are shown.
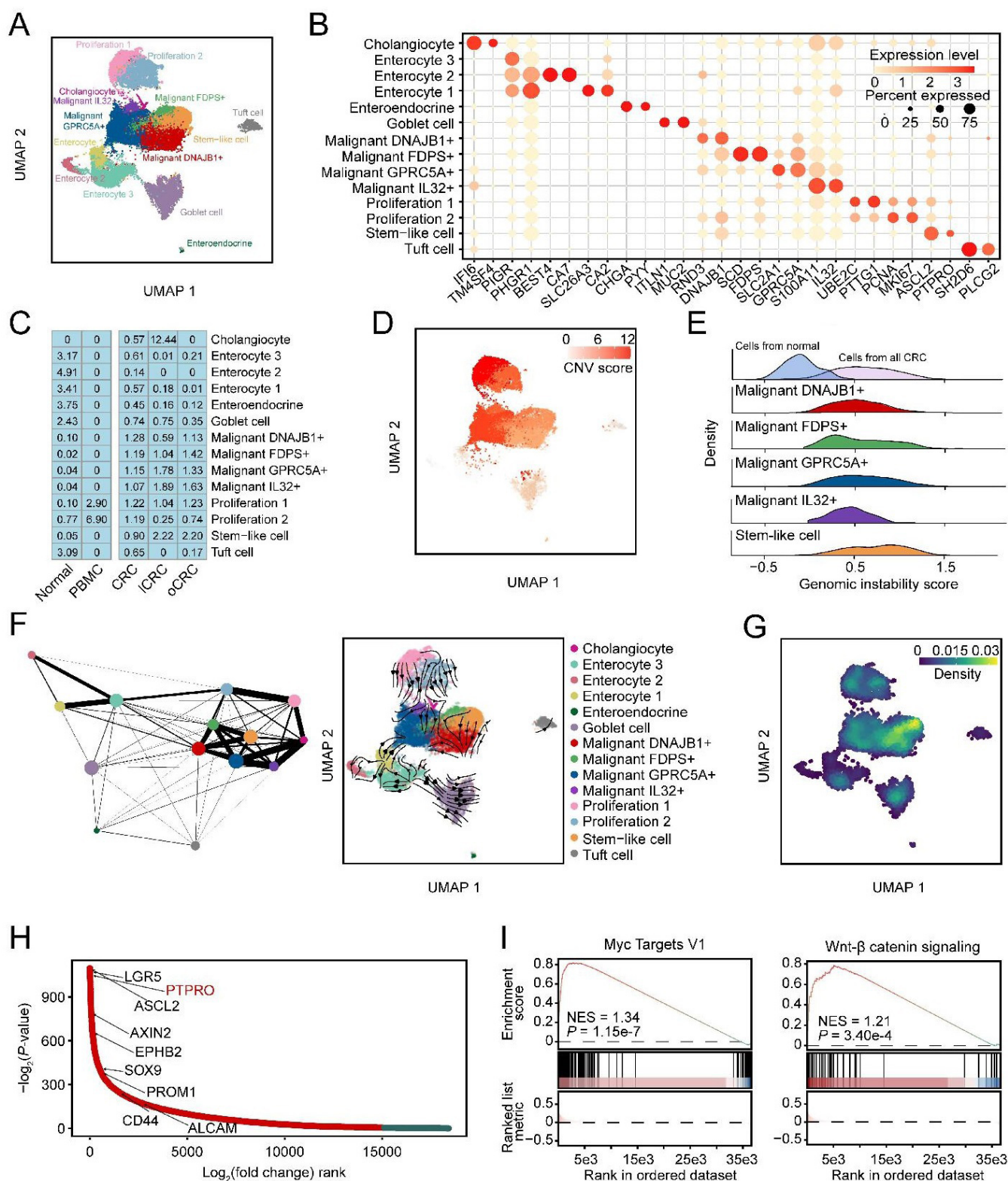
Figure 2 Stem-like cells are the originator of CRC metastasis. (A) UMAP showing 14 epithelial cell subtypes. (B) Bubble heatmap showing expression levels of marker genes (indicated as different colours) and the proportion of expressing cells (indicated as different dot sizes) across 14 epithelial cell subtypes shown in (A). (C) Tissue distribution of different epithelial cells estimated by Ro/e score. (D) UMAP plot showing the copy number variation (CNV) score of epithelial cells. (E) Ridgeline plot showing the genomic instability score of five malignant cell subtypes. (F) Partition-based graph abstraction analysis of 14 epithelial cell subtypes (left panel) and UMAP of RNA velocity of 14 epithelial cell subtypes (right panel). (G) Density of stemness scores of all epithelial cells. (H) Differential expressed genes between stem-like cells and the other cell subtypes. Y-axis indicates the -$\log_2$p value, genes ordered by $\log_2$fold change along the x-axis. (I) Gene set enrichment analysis showing MYC target V1 and WNT/β-catenin signalling pathway were upregulated in stem-like cells. The normalised enrichment score (NES) and false discovery rate (FDR) are included. CRC, colorectal cancer; lCRC, liver metastatic CRC; oCRC, ovarian metastatic CRC; PBMC, peripheral blood mononuclear cells.

genes in the fatty acid metabolism, glycolysis and gluconeogenesis pathways were highly expressed in FDPS$^+$ malignant cells while some genes in the chemokine signalling and leucocyte trans-endothelial migration pathways were highly expressed in GPRC5A$^+$ malignant cells. IL32$^+$ malignant cells had overexpression of ribosome genes (online supplemental figure 2I). We obtained similar results by using the Gene Expression Omnibus (GEO) and The Cancer Genome Atlas (TCGA)-colon adenocarcinoma (COAD) RNA-seq datasets, showing that the stem-like gene signature was significantly higher in the primary CRC with distant metastasis or in the metastatic CRC (online supplemental figure 3A,B). We analysed TCGA-COAD RNA-seq dataset and identified three CRC subtypes based on the stem-like signature score. Patients with higher stem-like signature score had significantly shorter survival time (online supplemental figure 3C–F). Additionally, we found that the *ASCL2* and *PTPRO* expression levels were higher in sigmoid colon cancers with high stem-like signature score than those with low stem-like signature score (online supplemental figure 3G–I).

Overexpression of *ASCL2* and *PTPRO* confers CRC stem-like cells metastatic phenotype

ASCL2, a basic helix-loop-helix transcription factor (TF), is a key regulator of intestinal stem cells and is involved in CRC proliferation and motility[15 16] and since stem-like cells in our CRC samples had high *PTPRO* RNA levels (figure 3A), we therefore analysed the relationship between the *ASCL2* and *PTPRO* expression levels and found a significant and positive correlation (online supplemental figure 4A). This result encouraged us to perform the spatial transcriptomic and immunofluorescence analyses on our primary CRC samples. As a result, we found that a subtype of epithelial cells highly expressed ASCL2 and PTPRO (figure 3B–C), suggesting that ASCL2 might be a TF of *PTPRO*. We then conducted gene operation experiments in two CRC cell lines to test this notion and the results showed that in cells with *ASCL2* depletion, the PTPRO mRNA and protein levels were significantly decreased; but in cells with *PTPRO* depletion, the ASCL2 expression levels were not significantly changed (figure 3D and online supplemental figure 4B,C). These results further indicated that ASCL2 is a TF regulating *PTPRO* expression. We also looked at the *PTPRO* RNA levels in primary CRC tumour centres, invasive fronts and peritoneal metastases in the GSE75117 dataset and revealed that the levels were significantly higher in metastases than in primary tumour centres or invasive fronts (figure 3E). Kaplan-Meier estimation and multivariate analysis of survival time based on the *PTPRO* RNA level showed that patients with MSS CRC with high *PTPRO* levels had significantly shorter survival time than those with low *PTPRO* levels (figure 3F and online supplemental figure 4D). Finally, we examined the effects of *PTPRO* expression changes on CRC cell phenotypes and found that *PTPRO* knockdown significantly reduced CRC cell migration, invasion (figure 3G and online supplemental figure 4E) and sphere-propagating abilities (online supplemental figure 4F). Together, these results indicate that primary CRC contains a stem-like cell subtype featured by high expression of *ASCL2* and *PTPRO*, which confers CRC metastatic phenotype.

Heterogeneity of CRC stem-like cells in organ-specific CRC metastasis

We next wanted to explore whether the stem-like cell subtype was of heterogeneity that might link to organ-specific metastasis. Based on the transcriptomic profiles, we divided the

stem-like cells into eight distinct subpopulations (P1 to P8) as shown in figure 4A. All these cells exhibited the stem cell property and one of the populations displayed high proliferative capacity (online supplemental figure 5A). We found that among the eight subpopulations without significant individual difference, P1 and P2 were enriched in the oCRC while P3 and P4 were enriched in the lCRC. In primary CRC and oCRC, the percentage of P1 cells was 48.57% and 45.85%, but in primary CRC and lCRC, the percentage of P3 cells was 65.35% and 24.64% (figure 4B–C, online supplemental figure 5B,C and table 4). The high proportion of P1 and P3 cells in oCRC and lCRC suggested that these stem-like cells were the originators for CRC to metastasise to these two organs, respectively. We analysed the localisation of cell subpopulations in each of oCRC from five patients and the results supported the evolutionary trend from P1 to P2; similarly, the results in each of lCRC from two patients also showed the potential evolutionary trend from P3 to P4 (online supplemental figure 5D,E). Spatial transcriptomic analysis showed a spatial distribution pattern of P1 and P3 cells that are consistent with these results (figure 4D). Gene expression analysis showed that P1 and P2 cells had uniquely high expression of *MAFA* and *DLL4* while P3 and P4 cells had high expression of *TOMM6*, *CXCL14*, *ATP6V0C*, *PSMA6*, *CALML4*, *DBNDD2*, *RNASE4* and *DEFB1* (figure 4E). Cholangiocytes also had expression pattern like P3 and P4 cells (online supplemental figure 5F). P6 cells were not linked to metastasis but expressed a set of genes, including *CXCL1*, *CXCL2*, *CXCL3*, *VEGFA*, *ASPSCR1* and *IGF2* (figure 4E). The difference in the gene expression feature between P1 and P3 cells was validated in primary CRC (figure 4F) and the different expression patterns formed in primary CRC were retained in the oCRC or lCRC (figure 4G and online supplemental figure 5G). Immunohistochemical staining of clinical tissue samples showed that primary CRCs metastasised to the ovary had higher DLL4 expression levels than that metastasised to the liver (figure 4H and online supplemental figure 6A). The calculated area under the receiver operating characteristic curve indicated the ability of DLL4 expression level in predicting ovarian metastasis (figure 4I). We also transplanted *ASCL2*, *PTPRO* or *DLL4* knockdown CRC cells into mice through tail vein and the results showed that the ovarian metastases rate was significantly decreased as compared with that transplanted with control cells (figure 5A–C). Since TF plays a critical role in keeping cell phenotype and fate by regulating the expression of certain genes, we thus explored whether there are such TF(s) that determine the metastatic potential of stem-like cells. The master TF analysis suggested that ELF3, a gene playing leading role in the malignant phenotype of ovarian cancer,[17] is a master TF in P1 cells while ETV4, a gene linking to the development of liver cancer,[18] is a master TF in P3 cells (online supplemental figure 6B). Immunohistochemical staining of clinical tissue samples showed that primary CRCs that metastasised to the ovary had higher ELF3 levels than that metastasised to the liver while primary CRCs that metastasised to the liver had higher ETV4 levels than that metastasised to the ovary (online supplemental figure 6C). These results suggest that the heterogeneity of stem-like cells caused by different master TFs may determine organ-specific metastasis of primary CRC.

Interactions of stem-like cells with CAFs and differentiated endothelial cells in CRC metastasis

It is known that cancer metastasis is determined by cancer cells themselves and by the orchestration of other cell types including CAFs and endothelial cells in the tumour microenvironment.[19 20]
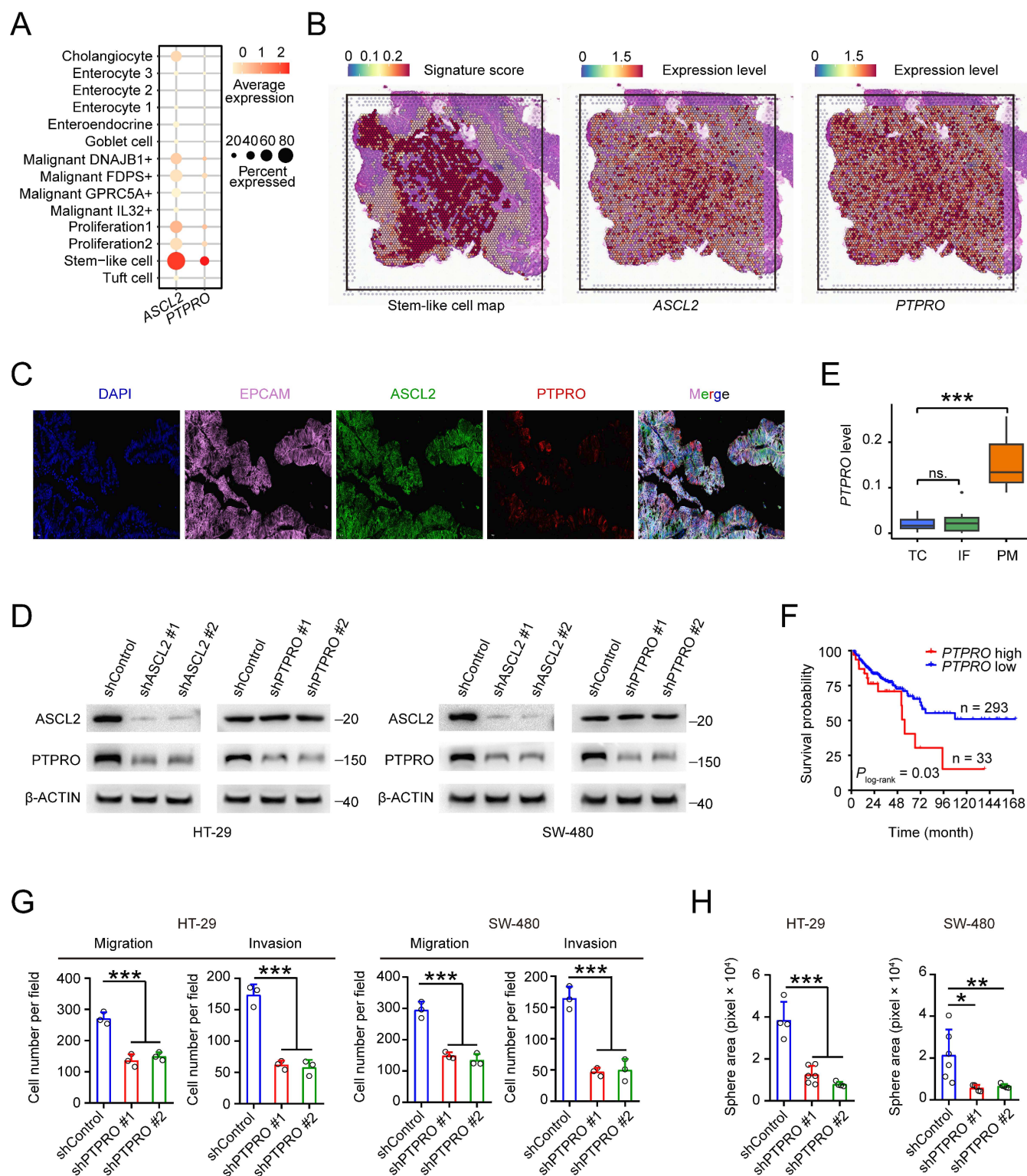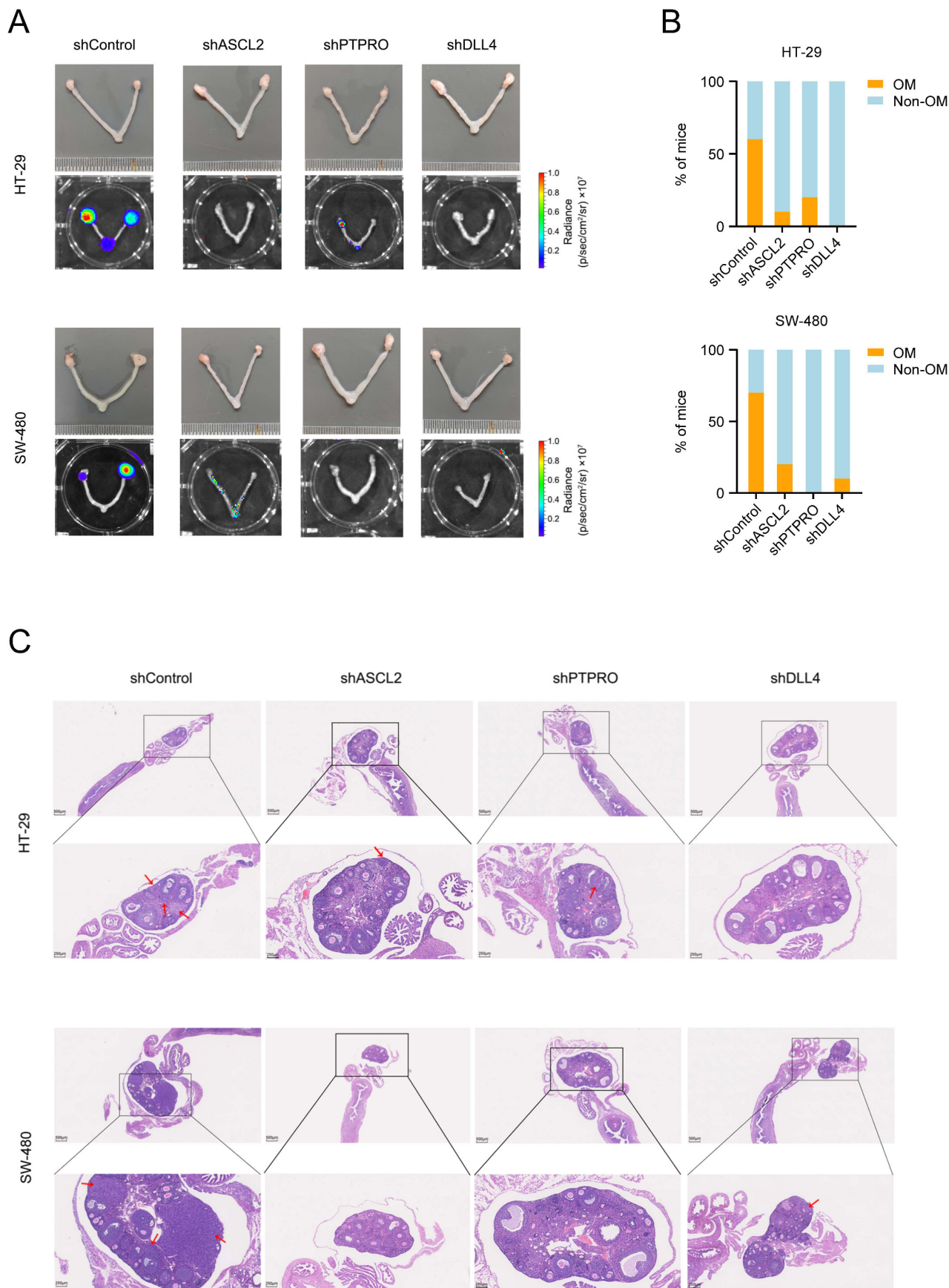
Figure 3 Overexpression of *achaete scute-like 2* (*ASCL2*) and *protein tyrosine phosphatase receptor type O* (*PTPRO*) confers colorectal cancer (CRC) stem-like cells metastatic phenotype. (A) The expression levels of *ASCL2* and *PTPRO* across 14 epithelial cell subtypes. Dot size indicates the fraction of expressing cells and the colour represents normalised expression levels. (B) Stem-like signature score (left panel) and the expression levels of *ASCL2* (middle panel) and *PTPRO* (right panel) based on spatial transcriptomic analysis of primary CRC sample from patient 2. (C) Immunofluorescence showing stem-like cells aggregate in a primary CRC sample. Scale bars, 50 µm. (D) The effect of *ASCL2* depletion (shASCL2) on the expression level of PTPRO. (E) Comparison of *PTPRO* expression levels in peritoneal CRC metastasis and primary CRC tumour centre or invasive front (from GSE75117 dataset). P values were determined by Wilcoxon rank-sum tests. (F) Kaplan-Meier estimation of overall survival time in patients with microsatellite stable CRC by the expression level of *PTPRO*. (G) The effect of *PTPRO* depletion (shPTPRO) on CRC cell migration and invasion abilities. (H) The effect of *PTPRO* expression change (shPTPRO) on CRC cell sphere-propagating capacity. Data in (G) and (H) are mean±SD from three independent experiments, *p<0.05, **p<0.01, ***p<0.001 were determined by one-way analysis of variance (ANOVA) test with Dunnett's T3 multiple-comparison. IF, invasive front; ns, not significant; PM, peritoneal metastasis; TC, tumour centre.
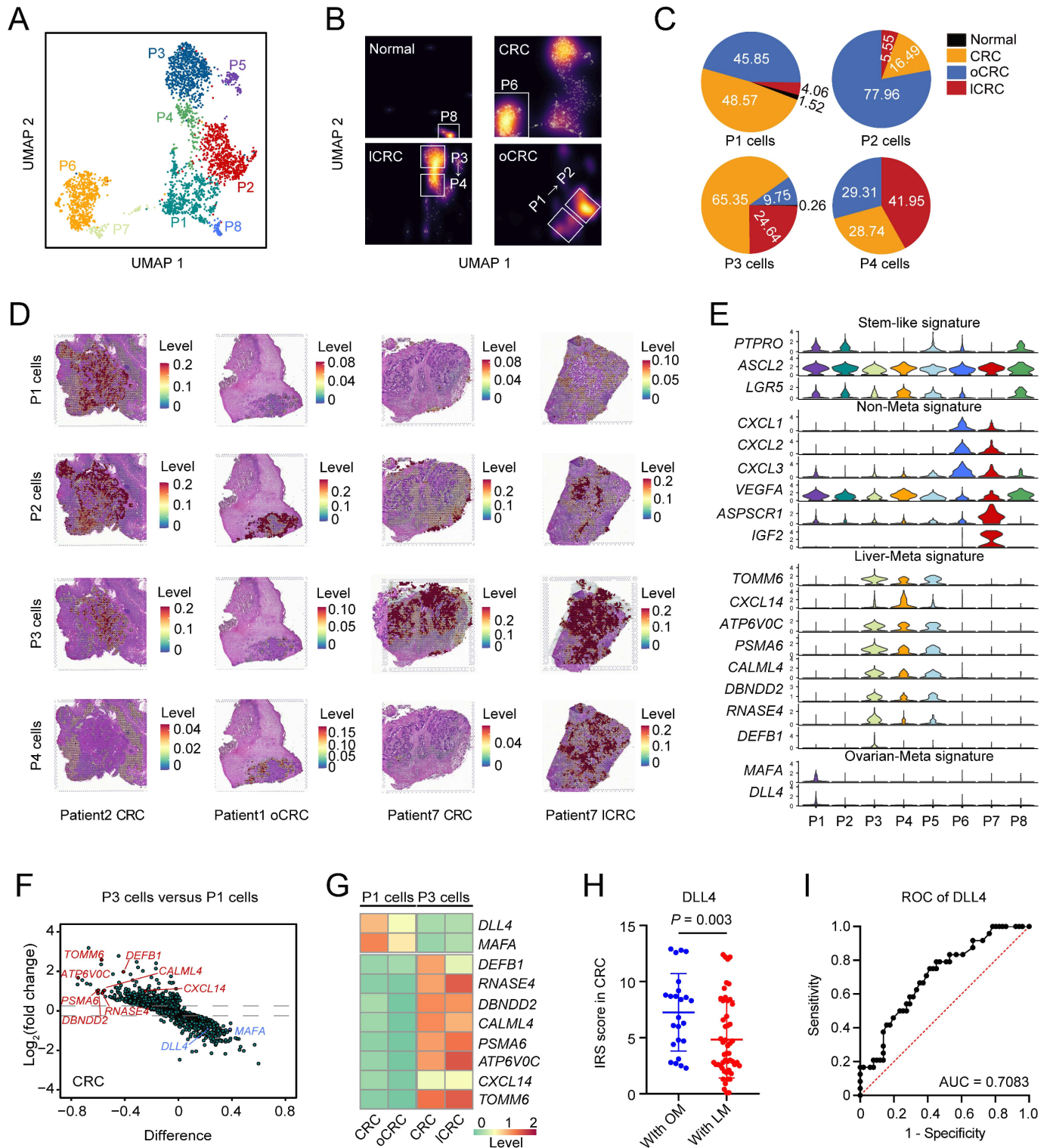
Figure 4 Heterogeneity of CRC stem-like cells in organ-specific CRC metastasis. (A) UMAP plot showing eight stem-like cell subtypes. (B) Density plot showing the distribution of stem-like cell subtypes in different organs. White box highlights high proportion of stem-like cell subtypes in each organ. (C) The proportions of P1, P2, P3 and P4 cells across normal, CRC, lCRC and oCRC. (D) Assignment of cell subtypes and their spatial distributions in CRC, lCRC and oCRC. (E) The signature expression of each stem-like cell subtype. (F) Differentially expressed genes (DEGs) between P1 cells and P3 cells in primary CRC. The x-axis represents the difference in the fraction of genes between the two groups. (G) Heatmap showing the expression level of some DEGs in P1 cells and P3 cells from primary CRC, oCRC and lCRC. (H) Comparison of DLL4 levels in primary CRC from patient with liver metastasis (LM) or ovarian metastasis (OM). P values were calculated by two-sided Wilcoxon rank-sum test. (I) Receiver operating characteristic (ROC) curves show predictive efficiency of DLL4 protein level in primary CRC for ovarian metastasis (area under the curve (AUC)=0.7083). CRC, colorectal cancer; lCRC, liver metastatic CRC; oCRC, ovarian metastatic CRC.

Figure 5 Delta-like ligand 4 (DLL4), achaete scute-like 2 (ASCL2) and protein tyrosine phosphatase receptor type O (PTPRO) promotes colorectal cancer (CRC) ovarian metastasis. (A) Effect of *ASCL2*, *PTPRO* or *DLL4* depletion on the CRC ovarian metastasis. (B) The proportions of mice with ovarian metastasis (OM) or without ovarian metastasis (non-OM) as function of *ASCL2*, *PTPRO* or *DLL4* depletion. (C) Representative H&E images of mouse OM. Red arrows indicate CRC metastases. Scale bars, 500 µm and 250 µm.

Based on the specific markers, we identified 13 fibroblast subtypes in all samples, including fibroblasts, myofibroblasts, pericytes and smooth muscle cells (figure 6A and online supplemental figure 7A–C), and among these subtypes, 9 were enriched in primary CRC samples (online supplemental figure 7D–E). We then calculated the attraction strength of ligand-receptor pairs to seek the molecules that may mediate the interactions of CAFs with stem-like cells and promote CRC metastasis. We found that among the ligand-receptor pairs between stem-like cells and CAFs, PDGFA-PDGFRA, DLL4-NOTCH2 and DLL4-NOTCH3 pairs were significantly enriched in P1 cells despite that many other ligand-receptor pairs were presented in P6 cells that were not linked to metastasis (figure 6B). Analysis using another tool iTALK yielded similar results showing that P1 cells had stronger crosstalk with CAFs through the ligand-receptor interaction of DLL4-NOTCH2, DLL4-NOTCH3 and PDGFA-PDGFRA than P6 cells (online supplemental figure 7F). We then calculated the interaction scores of DLL4-NOTCH2, DLL4-NOTCH3 and PDGFA-PDGFRA pairs, which were corrected by the corresponding cell fractions, and analysed their correlations with patient survival time in TCGA-COAD cohort. The results showed that patients with MSS CRC with high scores had significantly shorter survival time than those with low scores (figure 6C and online supplemental figure 7G).

We also subclassified endothelial cells and identified 10 subpopulations (figure 6D–E, online supplemental figure 8A,B). Among these, Art_NOTCH4 (another receptor of DLL4), Tip_COL4A1, Veins_ACKR1 and Veins_SERPINE1 subtypes were enriched in primary CRC (figure 6F). Enrichment score analysis of gene signatures for these four cell subtypes in TCGA-COAD dataset revealed that patients with metastases had significantly higher scores of Art_NOTCH4 and Tip_COL4A1 than patients without metastases. Conversely, patients without metastases had higher scores of Veins_ACKR1 and Veins_SERPINE1 than those with metastases (figure 6G). Analysing the Cancer Cell Line Encyclopaedia (CCLE) database obtained a similar result showing that the signature scores of Art_NOTCH4 and Tip_COL4A1 were positively correlated but the signature scores of Veins_ACKR1 and Veins_SERPINE1 were negatively correlated with the metastatic potential of CRC cell lines (online supplemental figure 8C). Differential expression analysis revealed that Art_NOTCH4 and Tip_COL4A1 cells had some similar signalling pathway activation, for example, the NOTCH signalling pathway; however, Veins_ACKR1 and Veins_SERPINE1 cells had reduced expression levels of genes in the NOTCH signalling pathway (figure 6H). Single sample gene set enrichment analysis indicated that Art_NOTCH4 and Tip_COL4A1 cells had a significantly higher NOTCH signalling score than two veins endothelial cell types (online supplemental figure 8D). We examined TCGA-COAD dataset and the results also showed that the signature score of the NOTCH signalling pathway was significantly higher in patients with MSS CRC with metastases than those without metastases, and a high score was associated with poor survival (online supplemental figure 8E). In addition, Art_NOTCH4 and Tip_COL4A1 cells expressed significantly higher levels of the key components in the NOTCH pathway, that is, *NOTCH1*, *NOTCH4*, *HEY1*, *JAG2* and *RBPJ* than two veins endothelial cell types (figure 6I). We found that P1 cells had strong crosstalk with endothelial cells via the ligand-receptor interaction of DLL4 and JAG1 with NOTCH1 and NOTCH4, and in TCGA-COAD cohort, high score of P1 cells and endothelial cell interaction was associated with significantly shorter survival time (online supplemental figure 8F). Since arterial and venous capillaries in tumour tissues are formed by differentiated

(D) endothelium derived from endovascular progenitor (EVP) and transit-amplifying cells,[21] we thus identified the differential expressed genes between EVP and D endothelial cells in two bulk RNA-seq datasets and examined the expression pattern of overlapped genes in four endothelial cell subtypes. The results showed that 34 genes were upregulated in EVP while 65 genes were upregulated in D (online supplemental figure 8G–I). We found that the signature score consisting of the 65 upregulated genes was significantly higher in Art_NOTCH4 and Tip_COL4A1 cells than in veins_ACKR1 and veins_SERPINE1 cells (figure 6J), suggesting that Tip_COL4A1 and Art_NOTCH4 cells are likely the D endothelial cells. Analysis of transendothelial migration in vitro showed that compared with control cells, CRC cells with *ASCL2*, *PTPRO* or *DLL4* knockdown had significant decreased ability to migrate through endothelial cells (online supplemental figure 9A–C). Together, these results suggest that P1 cells can interact with CAFs and D endothelial cells in the primary CRC microenvironment via their overexpressed ligands DLL4 with the NOTCH receptors on CAFs and D endothelial cells, which may promote them to metastasise to the ovary (figure 6K).
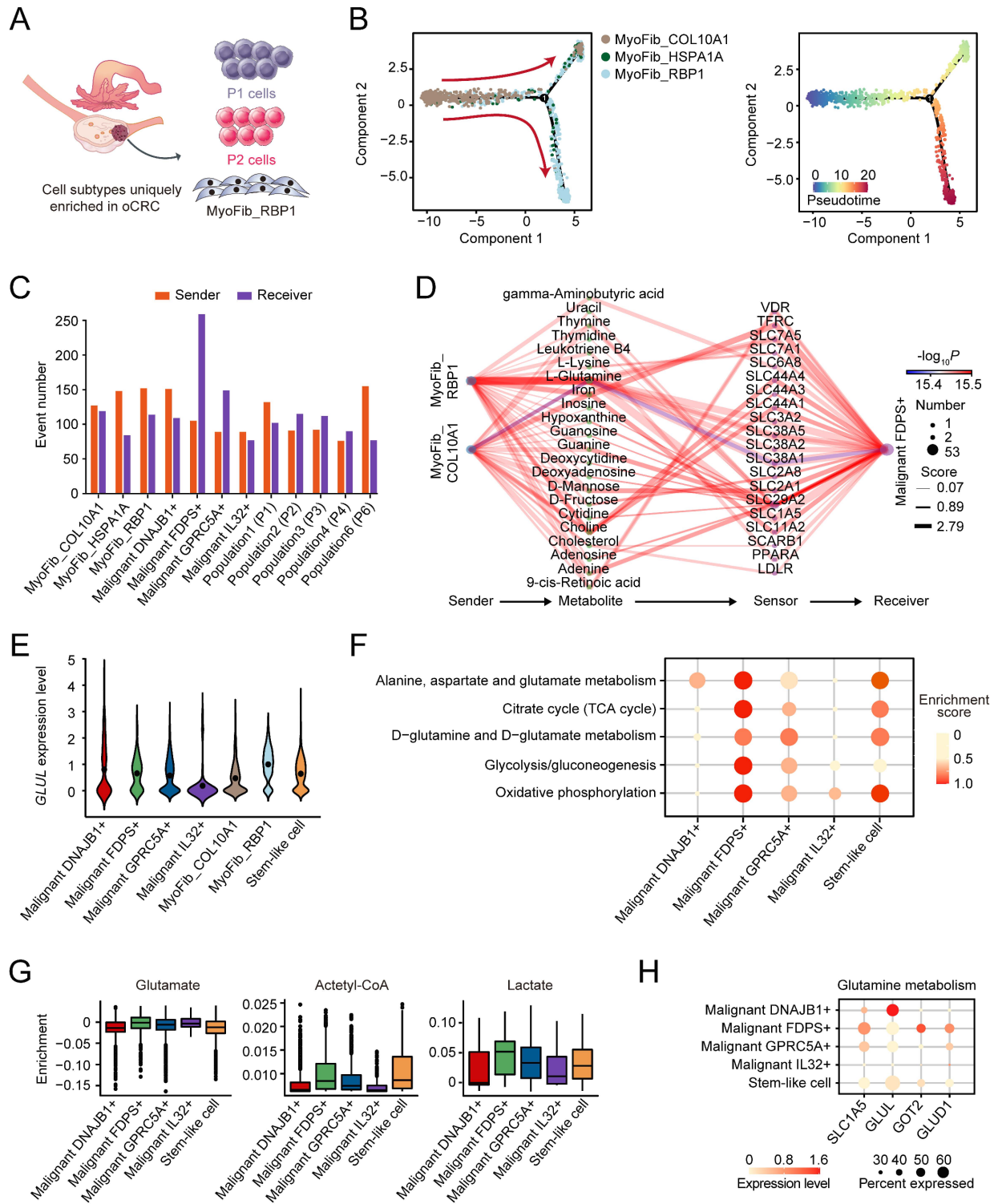
Characterisation of the metabolic features in ovarian CRC metastasis

We found that some unique cell subtypes including P1 and P2 cells and RBP1+ myofibroblasts (MyoFib_RBP1) were enriched in the oCRC microenvironment with MyoFib_RBP1 cells occurred only in oCRC (figure 7A and online supplemental figure 7D). We investigated the lineage relationship among three myofibroblast subtypes using Monocle2[22] and the results showed that the developmental routes began from MyoFib_COL10A1 and ended in MyoFib_RBP1 (figure 7B). We then applied MEBOCOST to examine the metabolite-sensor communications between myofibroblasts and malignant cells and found that MyoFib_RBP1 cells had the most communications (152 events) among myofibroblasts as sender and FDPS+ malignant cells had the most communications (259 events) as receiver among all malignant epithelial cells (figure 7C). We also found that among all metabolite-sensor partners between MyoFib_RBP1 cells and FDPS+ malignant cells, the communication scores of L-glutamine and its transporters were substantially higher than others (figure 7D). In line with this, MyoFib_RBP1 cells had significantly elevated expression of *GLUL* encoding the enzyme that converts glutamate to glutamine compared with MyoFib_COL10A1 cells. The level of *GLUL* RNA was lower in FDPS+ malignant cells than in MyoFib_RBP1 cells, suggesting that FDPS+ malignant cells had a low ability to produce glutamine and the external supply was necessary (figure 7E). Analysis of TCGA-COAD dataset also revealed a significant and negative correlation between *GLUL* RNA levels and FDPS+ malignant signature score (online supplemental figure 10A). In addition, FDPS+ malignant cells in oCRC showed abnormally activated oxidative phosphorylation, glycolysis, D-glutamine/D-glutamate metabolism and TCA cycle (figure 7F) as indicated by the increase in the expression levels of some genes and their catalysed metabolites (figure 7G–H and online supplemental figure 10B–D). We found that the *GLUL* mRNA level in human CRC cells and glutamine content in their culture medium were lower than that in human ovarian fibroblast (online supplemental figure 10E–F). We then cultured CRC cells with the medium collected from the cultivation of human ovarian fibroblasts without glutamine. Analysis of intracellular glutamine of cultured CRC cells showed that the glutamine level was significantly higher than that of CRC cells cultured
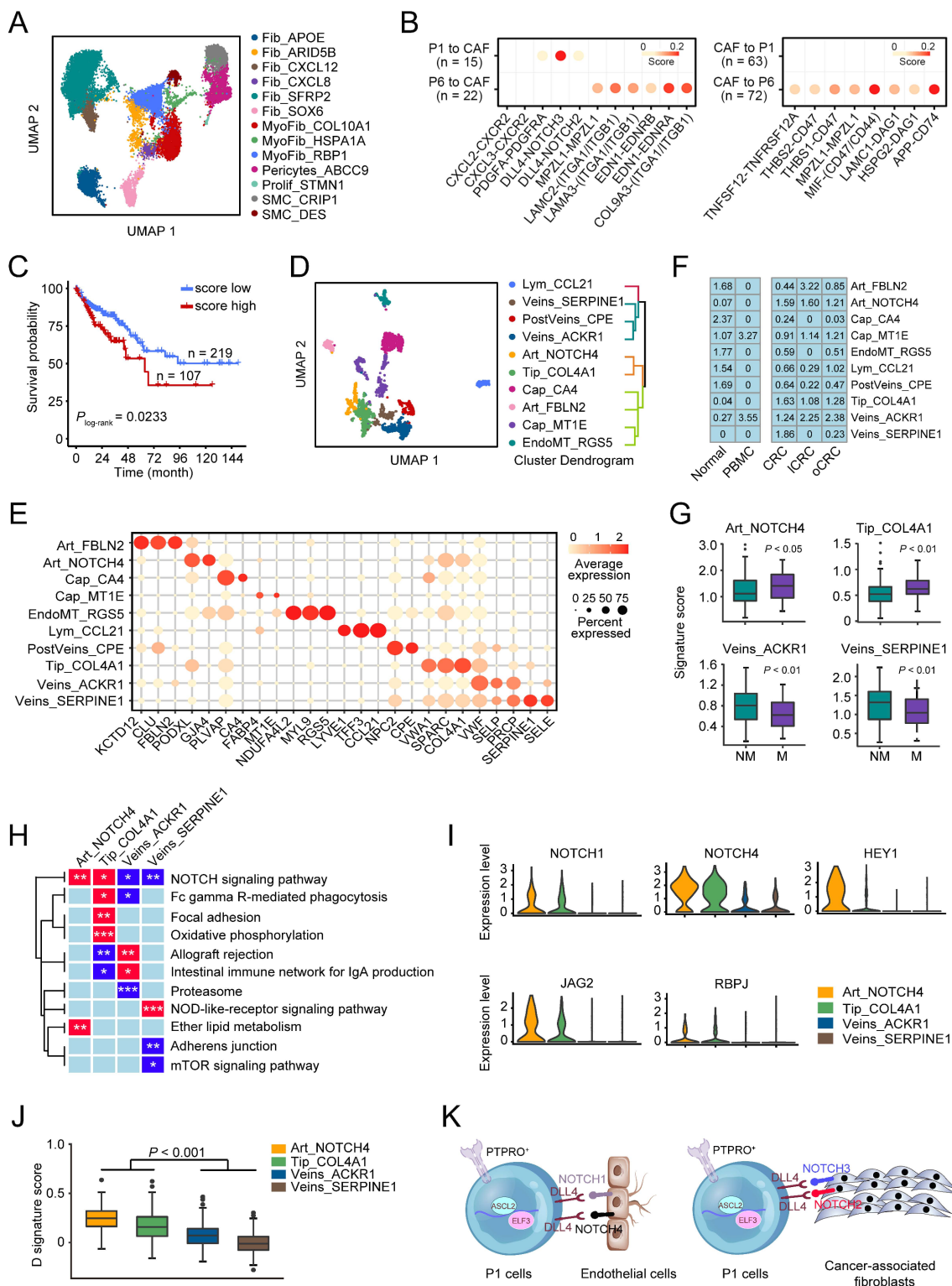
Figure 6 Interactions of stem-like cells with CAFs and differentiated endothelial cells in colorectal cancer (CRC) metastasis. (A) UMAP plot of 13 fibroblast cell subtypes. (B) Selected ligand-receptor pairs between CAFs and P1/P6 cells in primary CRC. (C) Kaplan-Meier estimation of survival time of patients with microsatellite stable CRC in The Cancer Genome Atlas (TCGA)-COAD dataset based on the score of P1-CAFs interactions. (D) UMAP plot of 10 endothelial cell subtypes (left panel) and hierarchical clustering of these endothelial cell subtypes (right panel). (E) Bubble heatmap showing expression levels of selected markers in different endothelial cell subtypes. Dot size indicates the fraction of expressing cells and the colours indicate the normalised expression level. (F) Tissue prevalence of endothelial cell subtypes estimated by Ro/e score. (G) The signature scores of four endothelial cell subtypes in patients with or without CRC metastases. Data were from TCGA-COAD dataset. P values by Wilcoxon rank-sum tests. (H) Heatmap of the gene expression levels in some pathways in four endothelial cell subtypes. (I) The differential expression levels of some genes in the NOTCH signalling pathway in four endothelial cell subtypes. (J) The comparison of the D signature score of differentiated endothelial cells in four endothelial cell subtypes. P values by Wilcoxon rank-sum tests. (K) The network involved in DLL4-NOTCH ligand-receptor pairs between P1 and CAFs or Art_NOTCH4/Tip_COL4A1.

Figure 7 Characterisation of the metabolic features in ovarian metastatic colorectal cancer (oCRC). (A) The schematics of cell subtypes present in the oCRC microenvironment. (B) Pseudotime-ordered analysis of three myofibroblast subtypes inferred by Monocle2. Cell subtypes are labelled by different colours (left panel) and pseudotime (right panel). (C) Bar plot showing the number of communications for the senders and receivers. The x-axis is the myofibroblast and cancer cell subtypes while the y-axis is the number of communications. The orange bars and purple bars indicate the number of communications for sender and receiver cells, respectively. (D) The diagram showing the information flow of metabolite-sensor communications from myofibroblast to FDPS$^+$ malignant cells through metabolites and sensors. The size of dots represents the number of connections. The lines connect the sender, metabolite, sensor and receiver. The colour of the line indicates the $-\log_{10}$(p value) and the width of line represents the communication score. (E) Violin plot showing the *GLUL* mRNA level of two myofibroblasts and five malignant cell subtypes. (F) The metabolic scores of five malignant cell subtypes in oCRC, showing that FDPS$^+$ malignant cells had the highest metabolic activity. The circle size and colour darkness both represent the scaled metabolic score. (G) Box plot showing the enrichment level of select metabolites in five malignant cell subtypes. (H) The expression levels of genes in the glutamine metabolism pathway in five malignant cell subtypes. Dot size indicates the fraction of expressing cells while the colour indicates the normalised expression levels.

with control medium. Furthermore, the cultured medium of ovarian fibroblast significantly promoted CRC cell proliferation compared with control medium. However, treatment of CRC cells with V-9302, a glutamine transporter ASCT2 (SLC1A5) inhibitor, significantly abolished the pro-proliferation effect of the cultured medium of ovarian fibroblasts and intracellular glutamine in CRC cells (online supplemental figure 10G–H). Gene regulatory network analysis of five malignant cell subtypes indicated that HOXA13 was highly expressed only in FDPS⁺ malignant cells colonised in the ovary (online supplemental figure 10I,J). These results indicated that FDPS⁺ malignant cells in oCRC had very active metabolisms and myofibroblasts can provide these cancer cells with glutamine to perform glutamine metabolism.

TIGIT is a potential target for immunotherapy of oCRC

We performed unsupervised clustering of CD4 cells and identified nine subtypes expressing CD4 but not CD8A and these cells expressed high levels of CD40LG or FOXP3 (online supplemental figure 11A,B). We also clustered CD8 cells and identified 12 subtypes expressing high CD8A but not CD4, CD40LG or FOXP3 (online supplemental figure 11C–D). We divided CD4 cells into nine subtypes based on their marker gene expression, that is, naïve-like (*SELL*, *TCF7*, *LEF1*), memory (*FOS*, *JUN*, *IL7R*, *CXCR4*), effector (*GZMA*, *GZMB*, *IFNG*, *NKG7*), exhausted (*PDCD1*, *CXCL13*, *CTLA4*) and regulated (*FOXP3*, *IL2RA*, *IKZF2*). We found that the CD4_ALOX5AP subtype with high levels of *GZMA*, *GZMB*, *IFNG*, *PDCD1* and *CXCL13*, presented mainly in primary CRC and metastatic CRC while FOXP3⁺ regulatory T cells were predominantly in primary CRC and oCRC but not lCRC (online supplemental figure 11E). We identified 12 subtypes of CD8 cells, including naïve-like (*SELL*, *TCF7*, *LEF1*), central memory (*FOS*, *JUN*, *IL7R*, *CXCR4*), effector memory (*GZMK*, *EOMES*), stress response (*HSPA1A*, *HSPA1B*), effector (*GZMA*, *GZMB*, *IFNG*, *NKG7*), exhausted (*PDCD1*, *CXCL13*, *CTLA4*), MAIT (*SLC4A10* and *ZBTB16*), dg T cells (*TRGC2*, *TRDC*) and proliferated (*MKI67*, *STMN1*). CD8 cells with high expression of heat shock protein were specifically enriched in oCRC (online supplemental figure 11F). Previous studies have shown that this subtype of CD8 cells was associated with poor prognosis of immune checkpoint inhibitor therapy.[23] We then calculated the attraction strength of ligand-receptor pairs to seek the molecules that may mediate the interactions of CAFs related to ovarian metastasis and T cells enriched in the oCRC and found that the THBS1-CD47, NECTIN2-TIGIT, MDK-NCL, LAMA4-CD44 and CXCL12-CXCR4 pairs were significantly enriched (online supplemental figure 11G,H). We also found that *TIGIT* expression level in T cells was significantly higher than that of *PDCD1* (programmed cell death protein 1 (PD-1)), the corresponding immune checkpoint ligand CD274 (programmed death-ligand 1 (PD-L1)) was barely expressed in either cancer cells or fibroblasts in the tumour microenvironment and TIGIT immune checkpoint ligand, *NECTIN2*, was highly expressed in cancer cells and fibroblasts (online supplemental figure 11I). These results indicate that the PD-1/PD-L1 pair may not be the dominant immune checkpoint signalling molecules in oCRC. Instead, the TIGIT-NECTIN2 pair might be the important immune checkpoint molecules.

DISCUSSION

Cancer cells within a tumour are heterogeneous in terms of their biological behaviours. Although several published studies using scRNA-seq have examined the cell compositions and

transcriptomic alterations in primary CRC and lCRC and described an intricate atlas,[9 10] it remains unrevealed what types of CRC cells are the metastasis determinants. In the present study, we have used the integrative analysis of scRNA-seq and spatial transcriptomic assays to explore the primary CRC and their oCRC or lCRC and deciphered the cell compositions including immune and non-immune cells, which are consistent with the previous studies.[24 25] Differently, our study has more focused on malignant cells and identified a stem-like cell cluster marked with *PTPRO* and *ASCL2* as the metastatic initiator. This stem-like cell population shows further different organotropism, with P1 and P2 cells being enriched in oCRC while P3 and P4 cells being enriched in lCRC. Furthermore, P1 cells interact with CAFs and D endothelial cells in the primary CRC microenvironment through the NOTCH signalling pathway to metastasise.

It has been proposed that cancer stem cells (CSCs) may be the culprit to regenerate cancer metastasis. For example, it has been shown that despite numerous malignant cells in a tumour, only CSCs can drive the persistence and relapse of disseminated cancer and regenerate tumours at distant sites.[6 7] In orthotopically xenografted CRC organoids, ablation of LGR5⁺ CSCs halts tumour growth and re-emergence leads to tumour regrowth.[26] We have revealed that stem-like cells in CRC tumours overexpress ASCL2 and PTPRO. ASCL2, a helix-loop-helix TF, is overexpressed and associated with CRC malignant phenotype,[27] inhibited immune infiltration[28] and drug resistance[29] and has been recognised as CRC stem cell marker.[15 27 30] PTPRO has been reported to play oncogenic[31 32] or antitumour[33–35] roles. Our results have clearly demonstrated that *PTPRO* knockdown significantly inhibits CRC cell migration, invasion and sphere-propagating capacity, indicating that PTPRO plays an oncogenic role in CRC. Interestingly, we have found that knockdown of *ASCL2* significantly suppresses PTPRO expression, and a previous study has shown that ASCL2 can bind to the enhancer of PTPRO in human CRC cells,[27] suggesting a direct regulation of *PTPRO* expression by ASCL2. Results of scRNA-seq, spatial transcriptomic and immunofluorescent analyses all have demonstrated the occurrence of ASCL2⁺ and PTPRO⁺ stem-like cells in primary CRC.

Cancer metastasis always show organotropism or organ-specific metastasis, which is probably regulated by multiple factors, including tumour-intrinsic factors, the circulation pattern, organ-specific niches and the interaction between tumour cells and the host microenvironment.[36 37] Accumulating evidence indicated that cancer cells in primary tumours exhibit tropism towards the particular organ by evolving fitness or adaptability, which is regulated by epigenomic programmes and manifested by the activation of specific genes.[38 39] This process is called 'seed preselection', which may explain why metastatic tropism often associated with specific gene expression patterns in primary tumours. These studies and the resultant hypothesis are corroborated by our findings. We have found that the stem-like cells have eight distinct subpopulations based on their transcriptomic features and these populations show organ-specific metastasis. P1 cells with high *DLL4* are enriched in primary CRC and oCRC, thus are presumed to be likely the culprits that specifically metastasise from primary CRC to the ovary. P3 cells having a similar expression pattern to cholangiocytes are mainly found in primary CRC and lCRC and thus may be associated with liver metastasis. However, P6 cells with high *CXCL1*, *CXCL2* and *CXCL3* are uniquely enriched in adjacent normal tissues and are not associated with metastasis, although the CXCL3-CXCR2 signalling axis has been considered to drive pancreatic ductal adenocarcinoma (PDAC) metastasis.[37]

Although some CSCs are recognised as the culprit of CRC metastasis, our study is the first to show that CRC CSCs are heterogenous in the transcription pattern and organ-metastatic specificity. The activation of specific genes may make CSCs to acquire organ tropism adaptability, known as 'seed preselection', and might be the underlying mechanism for organotropism of metastasis. Further studies are warranted to examine the relationship between unique expression program, especially the master TFs in CSCs and organ-metastatic specificity.

Endothelial cells and fibroblasts play important roles in both metastases of primary cancer to distant organs and in situ in target organs the localisation and growth of metastasised cancer. We have found that P1 cells can uniquely interact with CAFs and D endothelial cells through the NOTCH signalling pathway in the primary CRC microenvironment, suggesting that NOTCH signalling activation in primary tumour is necessary for ovarian metastasis. The activation is mediated by DLL4, a ligand to the NOTCH receptor family in P1 cells, contributing to angiogenesis,[40] which can enhance cancer metastasis. Previous study has shown that silencing DLL4 in ovarian cancer cells or endothelial cells suppressed both tumour growth and angiogenesis and induced hypoxia in the tumour microenvironment.[41] We have found that the interaction of DLL4 in P1 cells and NOTCH1/NOTCH4 in D endothelial cells resulted in upregulation of many genes such as *JAG2*, *RBPJ* and *HEY1*, which play important roles in either tumour angiogenesis[42] or cancer metastasis.[43] These results indicate that the activation of DLL4-NOTCH signalling pathway may mediate primary CRC metastasis to other organs such as the ovary. For the interaction between oCRC cells and fibroblasts in the ovarian microenvironment, we have also demonstrated crosstalk between FDPS⁺ malignant cells and myofibroblasts. We have found that myofibroblasts provided oCRC cells with glutamine, indicating that oCRC cells performed a metabolic reprogramming. This metabolic reprogramming may be critical for oCRC cells to localise and develop in the ovary since it is well-known that glutamine pathway is an important carbon source required for cellular bioenergy and biosynthesis.[44] It is worth noting that FDPS⁺ malignant cells uniquely enriched in oCRC displayed the activation of other metabolic pathways including TCA cycle, glycolysis and oxidative phosphorylation, all of which can facilitate oCRC growth. The roles of adipocytes and lipid metabolic reprogramming in supporting cancer growth and metastasis are well studied. Cancer cells transmit signals to cancer-associated adipocytes (CAAs), the latter produce various molecules or metabolic products that affect tumour cell growth and invasion. It has been suggested that CAAs could reside in distal tissues and organs and not directly physically contact with cancer cells yet influencing tumour cells by growth factors or cytokines.[45] Iwamoto *et al*[46] reported that lipid-dependent metabolic reprogramming is associated with antiangiogenic drug (AAD) resistance in patients with cancer. Mechanistically, AAD-induced tumour hypoxia activates fatty acid oxidation and promotes uptake of free fatty acid, eventually promoting cancer cell proliferation. Inhibition of fatty acid oxidation increases the therapeutic efficacy of AAD, which uncovers the regulatory role of lipid metabolism in the cancer cell behaviour and clinical therapy. However, the present study did not focus on the relationship between lipid metabolism and CSCs, an interesting issue that is warranted to investigate in the future.

We acknowledge some limitations in the present study. It would be more convincing if the spatial transcriptomics data were obtained from more than one section of the tumours. Thus, further validations using multiple tissue sections are needed in the future. In addition, a limitation might also exist in the comparison analysis between in-house and publicly available data due to patient heterogeneity and disbalance in cell numbers. Therefore, caution should be taken in interpreting any conclusions drawn from such analysis.

In conclusion, by using integrative scRNA-seq, spatial transcriptomic analysis and functional assays, we have uncovered a stem-like cell cluster marked with PTPRO and ASCL2 as the metastatic culprit, whose subpopulations show further different liver or ovary organotropism. In the primary CRC, metastasised CRC cells interact with CAFs and D endothelial cells through the NOTCH signalling pathway to obtain the nutrients and energy for their metastases. Our findings reveal the metastatic culprit of organ-specific CRC metastasis. Moreover, previous studies have reported that Wnt target genes, including LGR5 and AXIN2, are functional stem cell markers in epithelial tissues[47–49] while ASCL2 is restricted to gastric and intestinal stem cells.[27 50] Wang *et al* showed that high level of ASCL2 expression is critical for maintaining self-renewal properties and tumourigenicity of gastric CSCs.[51] Therefore, our findings might also be implicated in other gastrointestinal (GI) cancers, especially gastric cancer, which warrants investigations.

Author affiliations
[1]State Key Laboratory of Oncology in South China and Guangdong Provincial Clinical Research Center for Cancer, Sun Yat-sen University Cancer Center, Guangzhou, China
[2]Department of Biochemistry, School of Medicine, Southern University of Science and Technology, Shenzhen, China
[3]Shenzhen Institute of Pediatrics, Shenzhen Children's Hospital, Shenzhen, China
[4]Department of General Surgery (Colorectal Surgery), Guangdong Provincial Key Laboratory of Colorectal and Pelvic Floor Diseases, Guangdong Institute of Gastroenterology, Biomedical Innovation Center, The Sixth Affiliated Hospital, Sun Yat-sen University, Guangzhou, China
[5]Department of Pathology, Sun Yat-sen University Cancer Center, Guangzhou, China
[6]Department of Colorectal Surgery, Sun Yat-sen University Cancer Center, Guangzhou, China
[7]Department of Etiology and Carcinogenesis, National Cancer Center/National Clinical Research Center/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China
[8]Collaborative Innovation Center for Cancer Personalized Medicine, Nanjing Medical University, Nanjing, China
[9]Affiliated Cancer Hospital and Institute of Guangzhou Medical University, Guangzhou, China

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval The study was approved by the institutional review board of the SYSUCC (B2022-511-01) and informed consent was obtained from each participant.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request. Single-cell RNA sequencing (scRNA-seq) and spatial transcriptome raw data generated in this study have been deposited and accessed in the Genome Sequence Archive in BIG Data Center, Beijing Institute of Genomics, Chinese Academy of Sciences (https://ngdc.cncb.ac.cn/gsa-human, accession number HRA002863). We acquired the publicly available scRNA-seq data from previous study (GSE132465 and GSE144735, termed as KUL3 dataset and SMC dataset).[13] Public bulk RNA-seq datasets (GSE50760, GSE75117) from the Gene Expression Omnibus database (GEO, https://www.ncbi.nlm.nih.gov/geo/).[52 53] Bulk RNA-seq data of EVP and D endothelial cells were obtained from the GEO accession number GSE114528[21] and the ArrayExpress (https://www.ebi.ac.uk/arrayexpress) accession number E-MTAB-7148.[54] Transcriptomic data and clinical information of The Cancer Genome Atlas (TCGA)-COAD cohort were downloaded from the UCSC Xena data portal (https://xenabrowser.net).[55]

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iD
Jian Zheng http://orcid.org/0000-0001-9831-7038

REFERENCES

1 Torre LA, Islami F, Siegel RL, *et al*. Global cancer in women: burden and trends. *Cancer Epidemiol Biomarkers Prev* 2017;26:444–57.
2 Gupta GP, Massagué J. Cancer metastasis: building a framework. *Cell* 2006;127:679–95.
3 Biller LH, Schrag D. Diagnosis and treatment of metastatic colorectal cancer: a review. *JAMA* 2021;325:2405.
4 Riihimäki M, Hemminki A, Sundquist J, *et al*. Patterns of metastasis in colon and rectal cancer. *Sci Rep* 2016;6:29765.
5 Bouzid H, Soualmia F, Oikonomopoulou K, *et al*. Kallikrein-related peptidase 6 (KLK6) as a contributor toward an aggressive cancer cell phenotype: a potential role in colon cancer peritoneal metastasis. *Biomolecules* 2022;12:1003.
6 Batlle E, Clevers H. Cancer stem cells revisited. *Nat Med* 2017;23:1124–34.
7 Lambert AW, Pattabiraman DR, Weinberg RA. Emerging biological principles of metastasis. *Cell* 2017;168:670–91.
8 Ganesh K, Basnet H, Kaygusuz Y, *et al*. L1Cam defines the regenerative origin of metastasis-initiating cells in colorectal cancer. *Nat Cancer* 2020;1:1128.
9 Wu Y, Yang S, Ma J, *et al*. Spatiotemporal immune landscape of colorectal cancer liver metastasis at single-cell level. *Cancer Discov* 2022;12:134–53.
10 Liu Y, Zhang Q, Xing B, *et al*. Immune phenotypic linkage between colorectal cancer and liver metastasis. *Cancer Cell* 2022;40:424–37.
11 Li X, Pan J, Liu T, *et al*. Novel TCF21 (high) pericyte subpopulation promotes colorectal cancer metastasis by remodelling perivascular matrix. *Gut* 2023;72:710–21.
12 Xu Y, Wei Z, Feng M, *et al*. Tumor-infiltrated activated B cells suppress liver metastasis of colorectal cancers. *Cell Reports* 2022;40:111295.
13 Lee H-O, Hong Y, Etlioglu HE, *et al*. Lineage-dependent gene expression programs influence the immune landscape of colorectal cancer. *Nat Genet* 2020;52:594–603.
14 Gulati GS, Sikandar SS, Wesche DJ, *et al*. Single-cell transcriptional diversity is a hallmark of developmental potential. *Science* 2020;367:405–11.
15 Basu S, Gavert N, Brabletz T, *et al*. The intestinal stem cell regulating gene ASCL2 is required for L1-mediated colon cancer progression. *Cancer Lett* 2018;424:9–18.
16 Ziskin JL, Dunlap D, Yaylaoglu M, *et al*. In situ validation of an intestinal stem cell signature in colorectal cancer. *Gut* 2013;62:1012–23.
17 Seo SH, Hwang S-Y, Hwang S, *et al*. Hypoxia-induced ELF3 promotes tumor angiogenesis through IGF1/IGF1R. *EMBO Rep* 2022;23:e52977.
18 Zheng C, Liu M, Ge Y, *et al*. HBx increases chromatin accessibility and ETV4 expression to regulate dishevelled-2 and promote HCC progression. *Cell Death Dis* 2022;13:116.

19 Olumi AF, Grossfeld GD, Hayward SW, *et al*. Carcinoma-associated fibroblasts direct tumor progression of initiated human prostatic epithelium. *Cancer Res* 1999;59:5002–11.
20 Han Y, Zhang Y, Jia T, *et al*. Molecular mechanism underlying the tumor-promoting functions of carcinoma-associated fibroblasts. *Tumour Biol* 2015;36:1385–94.
21 Donovan P, Patel J, Dight J, *et al*. Endovascular progenitors infiltrate melanomas and differentiate towards a variety of vascular beds promoting tumor metastasis. *Nat Commun* 2019;10:18.
22 Qiu X, Mao Q, Tang Y, *et al*. Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods* 2017;14:979–82.
23 Chu Y, Dai E, Li Y, *et al*. Pan-cancer T cell Atlas links a cellular stress response state to immunotherapy resistance. *Nat Med* 2023;29:1550–62.
24 Becker WR, Nevins SA, Chen DC, *et al*. Single-cell analyses define a continuum of cell state and composition changes in the malignant transformation of polyps to colorectal cancer. *Nat Genet* 2022;54:985–95.
25 Joanito I, Wirapati P, Zhao N, *et al*. Single-cell and bulk transcriptome sequencing identifies two epithelial tumor cell States and refines the consensus molecular classification of colorectal cancer. *Nat Genet* 2022;54:963–75.
26 Shimokawa M, Ohta Y, Nishikori S, *et al*. Visualization and targeting of LGR5$^+$ human colon cancer stem cells. *Nature* 2017;545:187–92.
27 van der Flier LG, van Gijn ME, Hatzis P, *et al*. Transcription factor achaete scute-like 2 controls intestinal stem cell fate. *Cell* 2009;136:903–12.
28 Yang Q, Huang G, Li L, *et al*. Potential mechanism of immune evasion associated with the master regulator ASCL2 in microsatellite stability in colorectal cancer. *J Immunol Res* 2021;2021:5964752.
29 Tanaka T, Yamashita K. ASO author reflections: ASCL2, a marker of dynamic colon stem cell, involved in WNT pathway activation reflects anti-cancer drug resistance in primary colorectal cancer. *Ann Surg Oncol* 2019;26:634.
30 Schuijers J, Junker JP, Mokry M, *et al*. ASCL2 acts as an R-Spondin/WNT-responsive switch to control stemness in intestinal crypts. *Cell Stem Cell* 2015;16:158–70.
31 Bollu LR, Mazumdar A, Savage MI, *et al*. Molecular pathways: targeting protein tyrosine phosphatases in cancer. *Clin Cancer Res* 2017;23:2136–42.
32 Elson A. Stepping out of the shadows: oncogenic and tumor-promoting protein tyrosine phosphatases. *Int J Biochem Cell Biol* 2018;96:135–47.
33 Dai W, Xiang W, Han L, *et al*. PTPRO represses colorectal cancer tumorigenesis and progression by reprogramming fatty acid metabolism. *Cancer Commun (Lond)* 2022;42:848–67.
34 Dong H, Ma J, Gan J, *et al*. PTPRO represses ERBB2-driven breast oncogenesis by dephosphorylation and endosomal internalization of ERBB2. *Oncogene* 2017;36:410–22.
35 Hou J, Xu J, Jiang R, *et al*. Estrogen-sensitive PTPRO expression represses hepatocellular carcinoma progression by control of STAT3. *Hepatology* 2013;57:678–88.
36 Gao Y, Bado I, Wang H, *et al*. Metastasis organotropism: redefining the congenial soil. *Dev Cell* 2019;49:375–91.
37 Sun X, He X, Zhang Y, *et al*. Inflammatory cell-derived CXCL3 promotes pancreatic cancer metastasis through a novel myofibroblast-hijacked cancer escape mechanism. *Gut* 2022;71:129–47.
38 Chan EC, De Ciccio A, McLaughlin R, *et al*. An inexpensive solid medium for obtaining colony-forming units of oral spirochetes. *Oral Microbiol Immunol* 1997;12:372–6.
39 Zhang X-F, Jin X, Malladi S, *et al*. Selection of bone metastasis seeds by mesenchymal signals in the primary tumor stroma. *Cell* 2013;154:1060–73.
40 Meurette O, Mehlen P. Notch signaling in the tumor microenvironment. *Cancer Cell* 2018;34:536–48.
41 Hu W, Lu C, Dong HH, *et al*. Biological roles of the delta family notch ligand Dll4 in tumor and endothelial cells in ovarian cancer. *Cancer Res* 2011;71:6030–9.
42 Zhao D, Xue C, Lin S, *et al*. Notch signaling pathway regulates angiogenesis via endothelial cell in 3D co-culture model. *J Cell Physiol* 2017;232:1548–58.
43 Shen Z, Yao C, Wang Z, *et al*. Vastatin, an endogenous antiangiogenesis polypeptide that is lost in hepatocellular carcinoma, effectively inhibits tumor metastasis. *Mol Ther* 2016;24:1358–68.
44 Zhang J, Pavlova NN, Thompson CB. Cancer cell metabolism: the essential role of the nonessential amino acid, glutamine. *EMBO J* 2017;36:1302–15.
45 Cao Y. Adipocyte and lipid metabolism in cancer drug resistance. *J Clin Invest* 2019;129:3006–17.
46 Iwamoto H, Abe M, Yang Y, *et al*. Cancer lipid metabolism confers antiangiogenic drug resistance. *Cell Metab* 2018;28:104–17.
47 Jaks V, Barker N, Kasper M, *et al*. LGR5 marks cycling, yet long-lived, hair follicle stem cells. *Nat Genet* 2008;40:1291–9.
48 Lim X, Tan SH, Yu KL, *et al*. Axin2 marks quiescent hair follicle bulge stem cells that are maintained by autocrine WNT/Β-Catenin signaling. *Proc Natl Acad Sci U S A* 2016;113:E1498–505.
49 Lee J-H, Tammela T, Hofree M, *et al*. Anatomically and functionally distinct lung mesenchymal populations marked by LGR5 and LGR6. *Cell* 2017;170:1149–63.
50 Leushacke M, Tan SH, Wong A, *et al*. LGR5-expressing chief cells drive epithelial regeneration and cancer in the oxyntic stomach. *Nat Cell Biol* 2017;19:774–86.
51 Wang T, Wu H, Liu S, *et al*. SMYD3 controls a WNT-responsive epigenetic switch for ASCL2 activation and cancer stem cell maintenance. *Cancer Lett* 2018;430:11–24.

52 Kim S-K, Kim S-Y, Kim J-H, *et al*. Data from: a nineteen gene-based risk score classifier predicts prognosis of colorectal cancer patients. *Mol Oncol* 2014;8:1653–66.

53 Muto Y, Suzuki K, Kato T, *et al*. Data from: heterogeneous expression of zinc-finger E-box-binding Homeobox 1 plays a pivotal role in metastasis via regulation of miR-200C in epithelial-mesenchymal transition. *Int J Oncol* 2016;49:1057–67.

54 Lukowski SW, Patel J, Andersen SB, *et al*. Single-cell transcriptional profiling of aortic endothelium identifies a hierarchy from endovascular progenitors to differentiated cells. *Cell Rep* 2019;27:2748–58.

55 Goldman MJ, Craft B, Hastie M, *et al*. Visualizing and interpreting cancer genomics data via the Xena platform. *Nat Biotechnol* 2020;38:675–8.