CANCER

Comprehensive analysis of spatial architecture in primary liver cancer

Rui Wu[1,2]†, Wenbo Guo[3]†, Xinyao Qiu[4]†, Shicheng Wang[3], Chengjun Sui[2], Qiuyu Lian[3], Jianmin Wu[5], Yiran Shan[3], Zhao Yang[2], Shuai Yang[4], Tong Wu[1], Kaiting Wang[5], Yanjing Zhu[1], Shan Wang[4], Changyi Liu[3], Yangqianwen Zhang[1], Bo Zheng[1], Zhixuan Li[1], Yani Zhang[5], Siyun Shen[1], Yan Zhao[5], Wenwen Wang[4], Jinxia Bao[1], Ji Hu[1], Xuan Wu[6], Xiaoqing Jiang[2], Hongyang Wang[1,5,7,8]*, Jin Gu[3]*, Lei Chen[1,4,7,9]‡*

Heterogeneity is the major challenge for cancer prevention and therapy. Here, we first constructed high-resolution spatial transcriptomes of primary liver cancers (PLCs) containing 84,823 spots within 21 tissues from seven patients. The progressive comparison of spatial tumor microenvironment (TME) characteristics from nontumor to leading-edge to tumor regions revealed that the tumor capsule potentially affects intratumor spatial cluster continuity, transcriptome diversity, and immune cell infiltration. Locally, we found that the bidirectional ligand-receptor interactions at the 100-μm-wide cluster-cluster boundary contribute to maintaining intratumor architecture and the *PROM1*$^+$ and *CD47*$^+$ cancer stem cell niches are related to TME remodeling and tumor metastasis. Last, we proposed a TLS-50 signature to accurately locate tertiary lymphoid structures (TLSs) spatially and unveiled that the distinct composition of TLSs is shaped by their distance to tumor cells. Our study provides previous unknown insights into the diverse tumor ecosystem of PLCs and has potential benefits for cancer intervention.

INTRODUCTION

Large-scale cancer genome projects have already revealed extensive intertumor and intratumor heterogeneities (*1*). Recent single-cell omics studies, especially by single-cell RNA sequencing (scRNA-seq) technology, have greatly advanced our understandings of the tumor cell heterogeneities (*2*), tumor-infiltrated immune cell subpopulations (*3*), and the features of tumor-associated stromal cells (*4*) at single-cell level. These studies provided many novel insights into tumor subtyping, tumor initiation and evolution, drug resistance, and therapeutic targets. However, the scRNA-seq technology still has limitations. The most critical point is that the spatial and morphologic information is lost after the tissue dissociation into single-cell suspension, making it hard to investigate the tumor spatial architecture. Although some in situ hybridization (ISH)–based methods, such as multiplexed error-robust fluorescence in situ hybridization (MERFISH) (*5*) and sequential fluorescence in situ hybridization (seqFISH) (*6*), can obtain the spatial information, they can only detect a few known target genes simultaneously.

The recently developed spatial transcriptomics (ST) technology could overcome the above limitations. By positioning histological cryosections on arrayed reverse transcription primers with unique positional barcodes, ST provides high-quality genome-wide transcriptome data with intact two-dimensional positional information (*7*). It has been applied to analyze the spatial heterogeneity of human primary breast cancer (*8*), prostate cancer (*9*), pancreatic ductal adenocarcinomas (*10*), etc. However, because of the relatively lower resolution of the former ST method (about 1007 spots of 100-μm diameter and 200-μm interval) (*10*) and the lack of progressive comparison from adjacent normal to tumor inside region, the spatial architecture and heterogeneous tumor microenvironment (TME) have not been fully addressed.

Primary liver cancer (PLC) is the second most mortality tumor, of which hepatocellular carcinoma (HCC) and intrahepatic cholangiocarcinoma (ICC) are the two major histologic subtypes (*11*). Etiological and biological diversities, comprising chronic hepatitis virus infection, excess stress, drug-induced liver injury, aflatoxin B exposure, unresolving inflammation, and complicated TME, contribute to the high degree of intratumor heterogeneity of PLC (*12*). Until now, there are few effective nonsurgical strategies for PLC and lack of specific drug targets for effective therapeutic intervention (*13*). Only a limited proportion of patients with HCC could benefit from existing tyrosine kinase inhibitor drugs (such as sorafenib) and immune checkpoint inhibitors (*14*, *15*), largely resulting from both intertumor and intratumor heterogeneities (*16*). Several efforts have been made to explore the tumor heterogeneity and its clinical significance for liver cancer (*17*). For instance, scRNA-seq revealed that specific T cell subsets such as exhausted CD8$^+$ T cells and regulatory T cells (T$_{\text{regs}}$) are preferentially enriched and clonally expanded in HCC (*3*); similarly, a vascular endothelial growth factor (VEGF)/NOTCH-involved immunosuppressive oncofetal TME was identified in HCC tumorigenesis (*18*). Moreover, our recent study also reported that the enrichment of CD4/CD8/PD1 triple-positive T cells in tumor leading-edge region notably indicates better prognosis (*19*), which reinforces that it is indispensable for comprehensive and accurate assessment of spatial heterogeneity to understand the tumor cell community.

Here, we determined the spatial transcriptome architecture of seven PLCs including a total of 84,823 tissue spots within 21 sections

---

[1] The International Cooperation Laboratory on Signal Transduction, Eastern Hepatobiliary Surgery Hospital, Shanghai 200438, China. [2] Eastern Hepatobiliary Surgery Hospital, Shanghai 200438, China. [3] MOE Key Laboratory of Bioinformatics, BNRIST Bioinformatics Division, Department of Automation, Tsinghua University, Beijing 100084, China. [4] Fudan University Shanghai Cancer Center, Department of Oncology, Shanghai Medical College, Fudan University, Shanghai 200032, China. [5] Institute of Metabolism and Integrative Biology, School of Life Sciences, Fudan University, Shanghai 200438, China. [6] Department of Laboratory Medicine, The Tenth People's Hospital of Shanghai, Tongji University, Shanghai 200072, China. [7] National Center for Liver Cancer, Shanghai 200438, China. [8] Key Laboratory of Signaling Regulation and Targeting Therapy of Liver Cancer (SMMU), Ministry of Education, Shanghai, China. [9] Shanghai Key Laboratory of Hepatobiliary Tumor Biology (EHBH), Shanghai, China.
*Corresponding author. Email: hywangk@vip.sina.com (H.W); jgu@tsinghua.edu.cn (J.G.); chenlei@smmu.edu.cn (L.C.)
†These authors contributed equally to this work.
‡Lead contact.

and characterized the TME features including stromal and immune cell distribution, tumor cluster interaction, cancer stem cell (CSC)–niche diversity, and tertiary lymphoid structure (TLS) composition. These findings provide novel insights into the complex ecosystem of liver cancer and have the potential to improve individualized cancer prevention and drug discovery.

RESULT

Exploration of PLC architecture with ST

To comprehensively analyze the spatial heterogeneity of PLC, we collected 21 tissue specimens from seven patients including five cases of HCC (HCC-1 to HCC-5), one case of combined hepatocellular and cholangiocarcinoma (cHC-1), and one case of ICC (ICC-1) and applied ST sequencing via 10x Genomics Visium platform (Fig. 1A). For HCC-1/HCC-2/HCC-3/HCC-4 and cHC-1, we used three spatially sequential sections (N, nontumor section; L, leading-edge section; T, tumor section). For ICC-1, only L section was collected because of massive necrosis inside the tumor. For HCC-2, an extra section from the portal vein tumor thrombus (P section) was collected (fig. S1A). For HCC-5, the intact tumor nodule (diameter of about 1 cm) was cut into four parts (designated as HCC-5A, HCC-5B, HCC-5C, and HCC-5D) to form a complete plane for ST analysis (table S1). The bulk tissues of all the sections were also used for whole-exome sequencing (WES) with peripheral blood mononuclear cell (PBMC) as control.

For the ST technology in this study, the diameter of a spot is 55 μm (containing approximately 8 to 20 cells; fig. S1B), and each section contains up to 5000 spots in the capture area (6.5 mm by 6.5 mm). Data showed the median sequencing depth of a single spot at approximately 30,000 unique molecular identifiers (UMIs) and 3000 genes in this study (fig. S1C and table S2). In general, the numbers of UMIs in tumor regions were larger than that in normal regions, in line with previous studies (Fig. 1B and fig. S1C) (*20*).

Further, considering that each spot contained multiple cells, we proposed a signature-based strategy to evaluate the enrichment degree of different cell types in each spot (see Fig. 1C and Materials and Methods). It can be seen that the normal hepatocyte scores were higher in the spots of N sections, while the malignant cell scores were higher in some spots of L/T/P sections. In addition, the immune and stromal cell scores were relatively variable, reflecting the different infiltration degrees.

To verify whether the transcriptomic features are consistent with the histological information, we compared the hematoxylin and eosin (H&E) staining images with their counterpart ST data regarding the expression of several marker genes. Results confirmed that the regions defined by cell type marker genes' expressions were highly consistent with their pathological images. Specifically, *ALB* and *CYP2E1* were highly expressed in normal regions, *GPC3* and *ARK1B10* in tumor regions, *KRT19* in cholangiocarcinoma regions, and *COL1A1* in capsule and stromal regions (figs. S2A and S3A).

Different patterns of PLC spatial heterogeneities

To characterize the spatial diversity of the PLCs, we combined the spots from different sections for each patient and performed clustering analyses (*21*–*23*). The distribution of the clusters was presented in both the UMAP (Uniform Manifold Approximation and Projection) projection space and tissue physical space. As shown in Fig. 2A, we found that the clusters in HCC-1T, HCC-3T, and HCC-4T had the characteristics of regional distribution, whereas the clusters in HCC-2T and cHC-1T were intertwined. Cluster 5 in HCC-3L was a unique cluster that was not appeared either in HCC-3N or HCC-3T.

To examine the subtypes at cluster level across different patients, we performed hierarchical clustering and diffusion map analysis (*24*) and found that the clusters from the same type of regions were more similar in general (Fig. 2B and fig. S4A, left). In the diffusion map, a three-branch structure was formed: The normal and tumor region clusters were projected into the branches, and the stromal region clusters were on the junction (fig. S4A, right).

Next, we also used the signature-based strategy to compare the scores of different cell types across different clusters (table S3) (*25*). Notably, the fibroblast and endothelial cell scores were significantly higher in stromal regions, and the immune cell scores in the different clusters of tumor regions exhibited a high degree of diversity (fig. S4B). As a validation, we performed the multimodal intersection analysis (MIA) (*10*) by integrating our ST data with a liver cancer single-cell dataset (*12*) and got similar results (fig. S4, C and D). Then, to analyze the composition changes of cell types from the outside to the inside of the tumor, we used diffusion map algorithm to project the spots into a one-dimensional pseudo-order (the first diffusion component, DC1) (*24*), which can be seen as generally from normal to tumor regions by comparing the distribution of clusters in this order. By fitting the variation curves of cell type scores [including T, B, natural killer (NK), myeloid, endothelial cells, and fibroblasts], we found that the clusters related to the stromal regions shaped the variation patterns of fibroblasts, endothelial cells, and even immune cells (Fig. 2C).

Microenvironment characteristics in leading-edge area

As seen in Fig. 2A, the tumor clusters' spatial distribution presented two distinct patterns. One was block-like with a clear boundary between clusters (e.g., HCC-1T and HCC-3T), while the other was discontinuous and mixed (e.g., HCC-2T). To measure this characteristic quantitatively, we introduced a metric named "spatial continuity degree," which was calculated by comparing the consistency of cluster identity between each spot and its neighbors. Together with another metric, "transcriptome diversity degree," which measured the global transcriptomic heterogeneity of tumor regions in each section, we quantitatively found that the tumor regions of L and T sections in patients with HCC-1, HCC-3, and HCC-4 had higher spatial continuity and lower transcriptome diversity (Fig. 3A).

To further explore the common features of those three samples with higher spatial continuity and lower transcriptome diversity (HCC-1L, HCC-3L, and HCC-4L), the pathologists investigated the macroscopic tissue appearances and histopathological features of tumor specimens (table S1) and found that all of them had complete fibrous capsules. Whereas cHC-1L and HCC-2L had discontinuous bands consisting of stromal cells, and ICC-1L had no capsule at the border of the tumor nodule. It was observed that the expression of stromal and immune cell markers displayed sharply reduction across the capsule from the normal side to tumor side (such as in HCC-4L), indicating that the complete fibrous capsule may affect the stromal and immune cell distribution (Fig. 3B). We thus defined the fibrous capsules of HCC-1L, HCC-3L, and HCC-4L and stromal cell clusters of cHC-1L and HCC-2L as the "transition area" in L sections (see Materials and Methods) and investigated the spatial characteristics of the "normal region" and "tumor region" on the two sides of the transition area. Meanwhile, we named

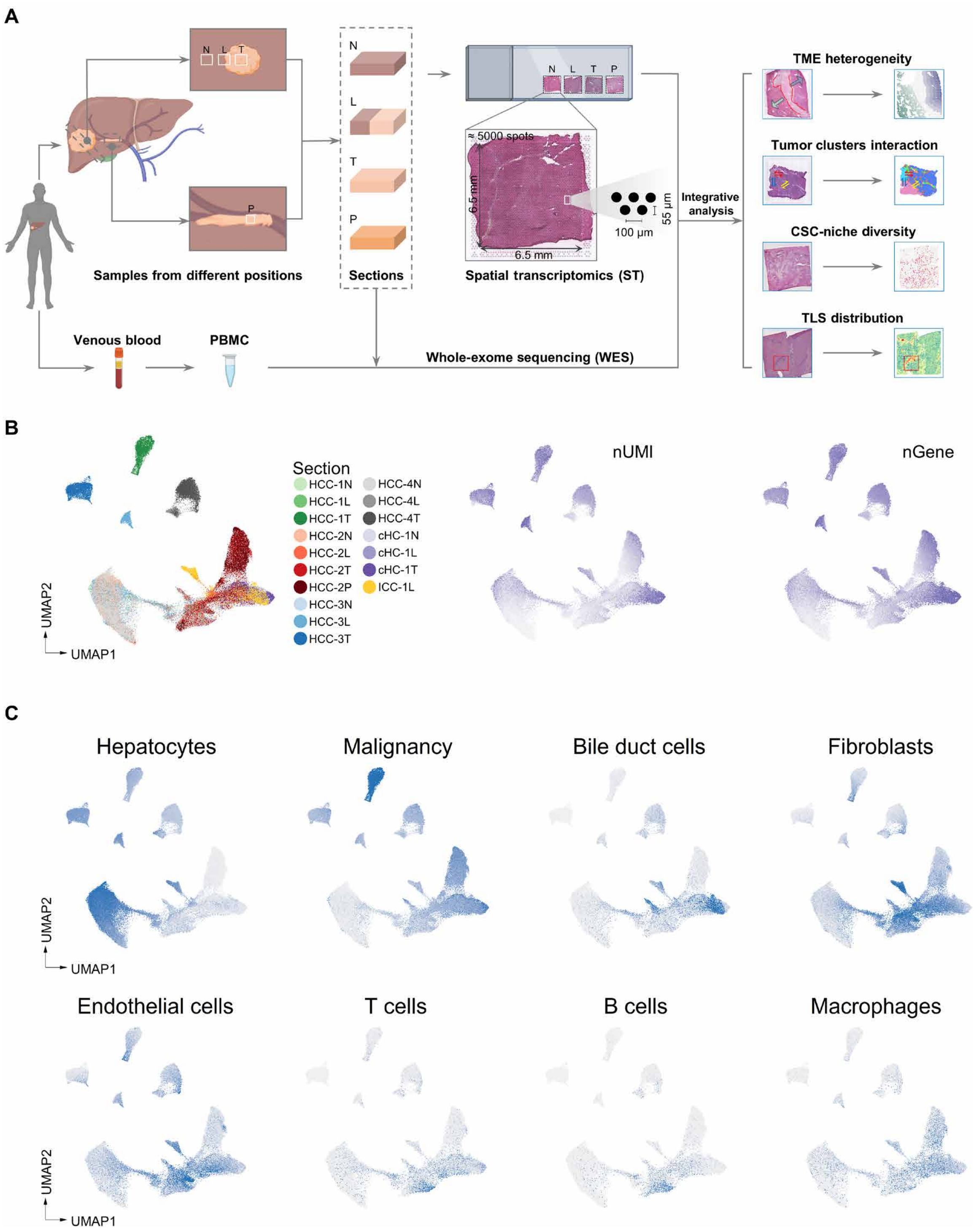

Fig. 1. Exploration of PLC architecture with ST. (**A**) Workflow of PLC samples collection, processing for ST and WES, and data analysis. (**B**) UMAP (Uniform Manifold Approximation and Projection) plot of spots from all sections, colored by their sample source, the number of expressed unique molecular identifiers (nUMI) and genes (nGene), respectively. HCC-1N represented the N section of HCC-1. (**C**) UMAP plot of cell type enrichment scores from all sections. Each dot represents a spot of ST sections, and the shade of color represents the scores of the corresponding cell type.

HCC-1L/HCC-3L/HCC-4L as "complete capsule (CC)" group and HCC-2L/cHC-1L/ICC-1L as "non- or discontinuous capsule (NC)" group (Fig. 3C). As shown in Fig. 3D and fig. S5A, the scores of T, B, and myeloid cells were much higher in the normal area than in the tumor area in CC group but not in NC group. The fibroblast and endothelial cell subtype scores were lower in both tumor and normal regions of CC group compared with NC group. CD8$^+$ T$_{em}$ cells (effective memory T cells), T$_{regs}$, CD4$^+$ memory T cells, CD4$^+$ T$_{em}$

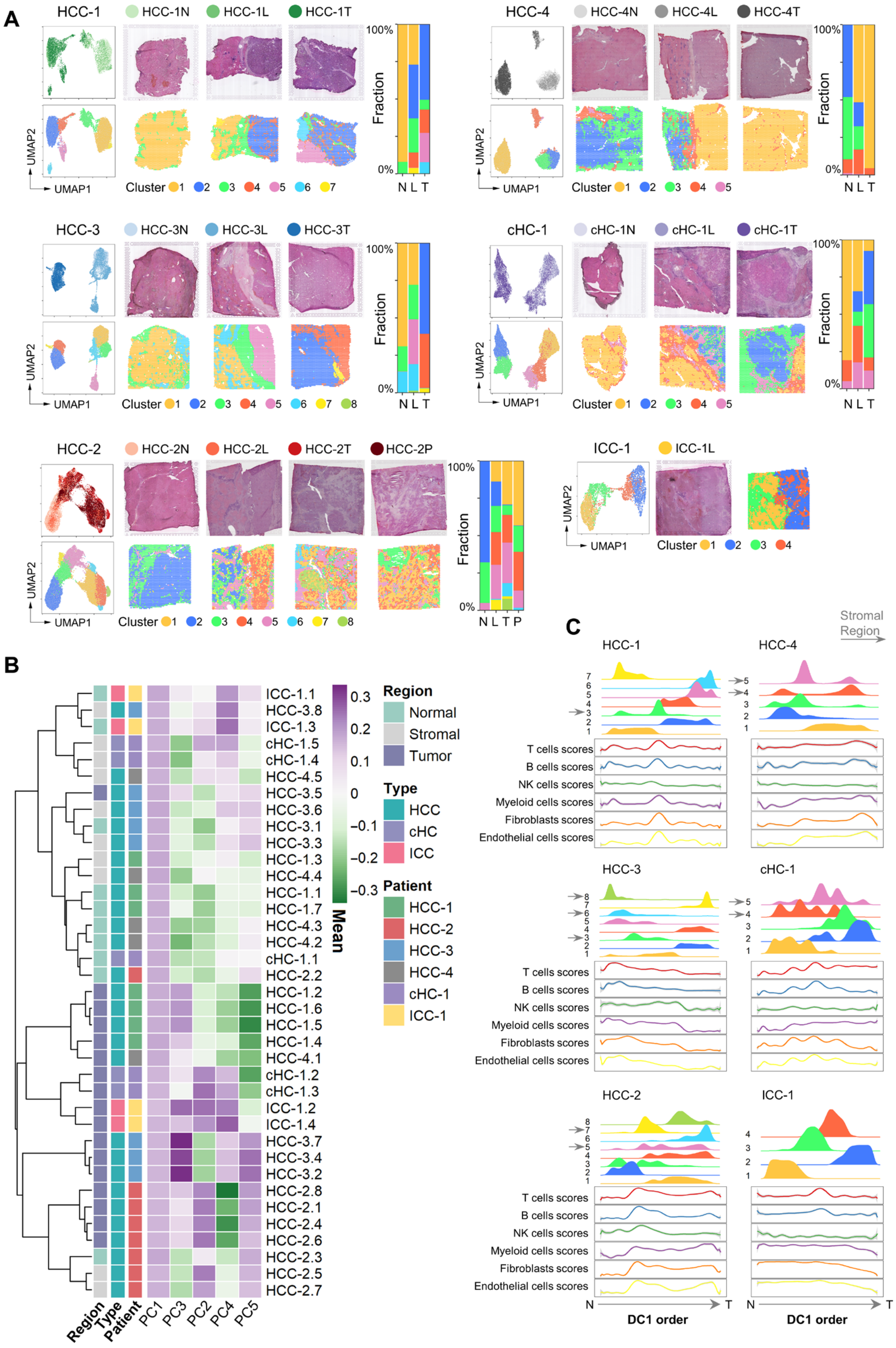

Fig. 2. Different patterns of PLC spatial heterogeneities. (**A**) Left: For each patient: UMAP of the spots colored by their section sources (top) and cluster identities (bottom), respectively. Middle: H&E staining (top) and the spatial cluster distribution (bottom) of each section in order of N/L/T/P. Right: Fraction of clusters in each section. (**B**) Similarity comparison of the clusters across different patients. The clusters' tissue regions, histopathological types, and patient information were annotated on the left. HCC-1.1 represented the cluster 1 of HCC-1. The clusters' normal, stromal, and tumor region labels were annotated according to the expression profile of marker genes and the corresponding tissue types in H&E image. (**C**) Distribution of the clusters (top half plot of each patient; *y* axis, the cluster's density) and the main stromal and immune cell type scores (bottom half plot; *y* axis, the fitted cell type enrichment score) along the direction of the first diffusion map component (DC1 order, the shared *x* axis).

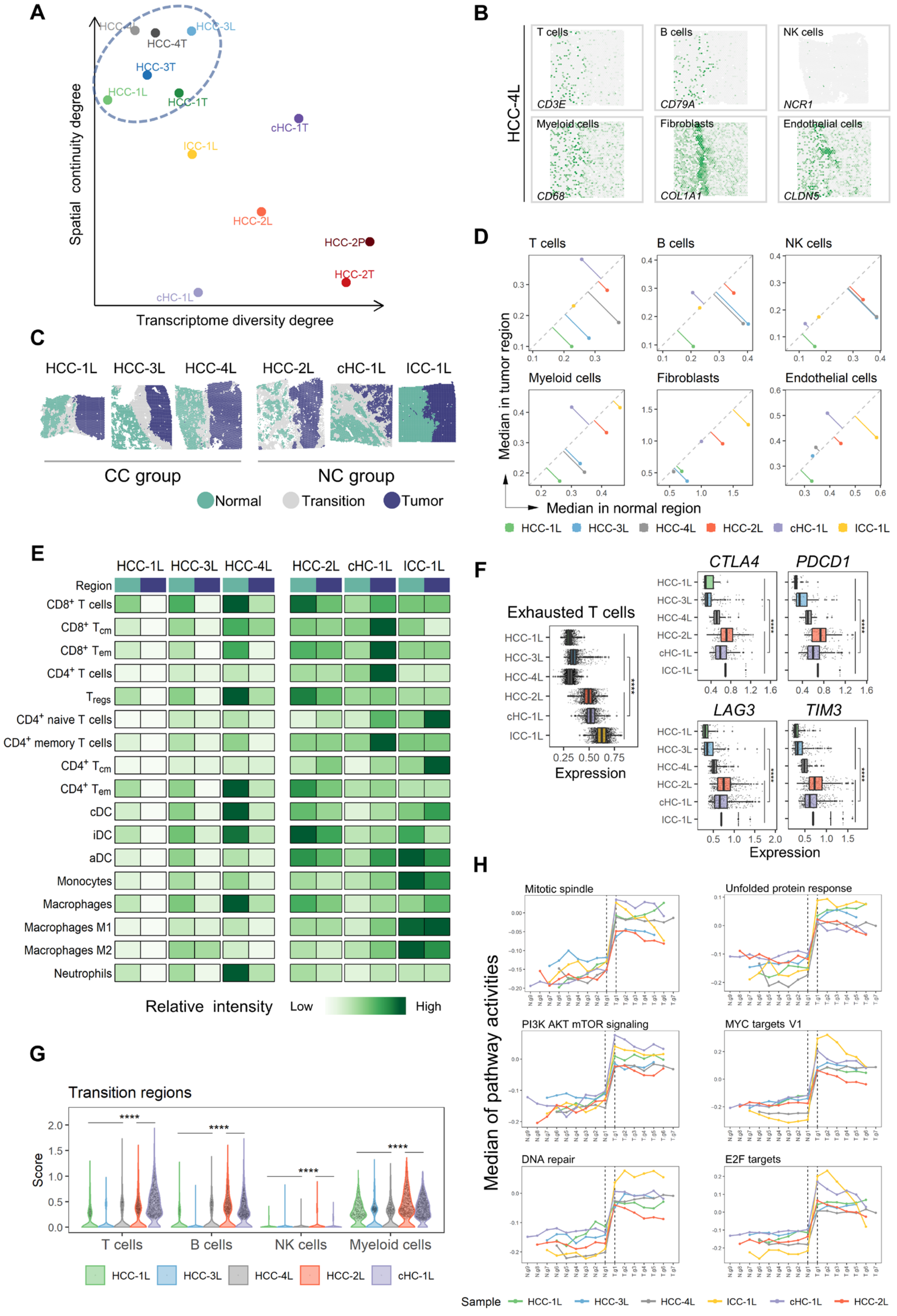

Fig. 3. Microenvironment characteristics in leading-edge area. (**A**) Transcriptome diversity degree and spatial continuity degree of tumor regions in L/T/P sections. (**B**) Spatial feature plots of six marker genes of stromal and immune cell types in HCC-4L. (**C**) Distribution of normal, tumor, and transition regions in L section and the grouping results of the CC (complete capsule) group and NC (non- or discontinuous capsule) group. (**D**) Comparison of the median of stromal and immune cell type scores between the normal (*x* axis) and the tumor (*y* axis) regions in each L section. (**E**) Comparison of the relative intensity (each row shared a color scale, while different rows did not) of stromal and immune cell subtype scores between the normal and tumor regions in each L section. iDC, interdigitating dendritic cell; aDC, active dendritic cell. (**F**) Comparison of the expression levels of exhausted T cell signature, *CTLA4*, *PDCD1*, *LAG3*, and *TIM3*, between the CC and NC groups. Two-sided Wilcoxon rank sum tests on the CC and NC groups were used to analyze the significance of their differences. ****$P < 0.0001$. (**G**) Comparison of the immune cell scores in the "transition regions" between the CC and NC groups. One-sided Wilcoxon rank sum tests (the CC group was less than NC group) were used to calculate the statistical significance. ****$P < 0.0001$. (**H**) Changes of hallmark pathways' activities along with the gradient divisions on both sides of the transition region. Each dot indicated the median of the pathway activity in the corresponding area. PI3K, phosphoinositide 3-kinase; mTOR, mammalian target of rapamycin.

cells, cDCs (conventional dendritic cells), monocytes, macrophages M1, and neutrophils were found significantly enriched in normal regions in CC group (Fig. 3E and fig. S5B). By comparing the function of T cells in the tumor region between two groups, we found that exhausted T cells were markedly increased in NC group in coupled with the increased expression of *PDCD1*, *CTLA4*, *LAG3*, and *TIM3* (Fig. 3F). In addition, the scores of T, B, NK, and myeloid cells in the transition regions were significantly lower in the CC group (Fig. 3G).

Since swelling and invasive growth are the biological characteristics of tumor cells, we wondered whether the signaling pathway activities had gradient-like changes in the direction perpendicular to the boundary on both sides of the transition area. To investigate it, we divided the normal and tumor regions of L sections into continuous zones parallel to the shape of the dividing line at intervals of five spots (fig. S6A) and analyzed the activities of hallmark pathways by gene set variation analysis (GSVA) (*26*). In general, the majority of hallmark pathways did not show a gradient in neither normal nor tumor region, and few consistent patterns were observed in different patients (fig. S6B). Individually, the ICC-1L showed several distinct changes: Hypoxia-associated signals showed a sudden decrease from N.g1 to T.g1 zone, whereas increased markedly from T.g1 to T.g6 zone, and inflammatory response and interferon response (α and γ) pathways were observed gradually reduced from T.g1 to T.g6 zone (fig. S6B). Note that six classical tumor–associated pathways (including phosphoinositide 3-kinase, MYC, mitotic spindle, unfolded protein response, E2F, and DNA repair) exhibited a sudden increase from N.g1 to T.g1 zone regardless of the width of the transition area or the presence of capsule (Fig. 3H). Together, our study here indicates that the integrity of the capsule is closely associated with the spatial heterogeneity of tumor cells and the distribution of their surrounding stromal and immune cells but has few effects on the activities of hallmark pathways in neither normal nor tumor regions.

Intratumor heterogeneity in PLCs

To investigate the interior heterogeneities of tumor regions in both L and T sections, we calculated hallmark pathways' activities for the spots in tumor regions. By performing hierarchical clustering on the averaged pathway activities in each tumor cluster, two modules were identified (Fig. 4A). Module 1 showed high activities of cell cycle and metabolism-related pathways (e.g., MYC targets v1, G_2-M checkpoint, E2F targets, cholesterol homeostasis, and bile acid and fatty acid metabolism), while module 2 had much higher activities in inflammation, angiogenesis, and epithelial-mesenchymal transition (EMT) pathways.

The tumor clusters of HCC-1 belonged to two different modules, T.2 (represented cluster 2 in HCC-1T) to module 2 and T.5 and T.6 to module 1, showing distinct gene expression patterns (Fig. 4B and table S4). In comparison with other metabolism-related clusters, T.5 and T.6 in HCC-1 had a relatively higher level of vitamin D, carbonic acid, fatty acid, and propanoate metabolism signals (fig. S7A) (*27*). By comparing these three tumor clusters with the samples of two HCC bulk transcriptome datasets [The Cancer Genome Atlas (TCGA) cohort and the Liver Cancer Institute (LCI) cohort] (*28*, *29*), we found that bulk samples more similar to T.2 had a better prognosis than those similar to T.5 and T.6 (Fig. 4C and fig. S7B). Moreover, bulk samples more similar to T.5 showed even worse outcomes than that similar to T.6 in TCGA cohort (fig. S7B), implying that the up-regulated creatine, tyrosine, ethanol, and retinol metabolism pathways in T.5 may enhance cell malignant behaviors and new strategy targeting those metabolism pathways could be tested for HCC intervention.

To investigate the communication and interaction between tumor clusters in HCC-1, we selected the interface regions of clusters with the range of four spots wide (two spots wide for each cluster; the spots in stromal regions were excluded) (Fig. 4D) (*30*, *31*). It was found that the enriched gene pairs including NRP1 (neuropilin 1)–VEGFB (*32*, *33*), FLT1 (fms related receptor tyrosine kinase 1)-VEGFB (*34*), EFNB2 (ephrin B2)–EPHB4 (EPH receptor B4) (*35*), MDK (midkine)–SORL1 (sortilin-related receptor 1) (*36*), and EFNA1 (ephrin A1) - EPHA1 (EPH receptor A1) (*37*) may contribute to T.5-induced cell angiogenesis/proliferation/migration in T.2. In turn, T.2 may help to maintain the metabolic activity in T.5 through LGALS9 (galectin 9) - LRP1 (low-density lipoprotein receptor related protein 1) and PDGFB (platelet-derived growth factor subunit B)–LRP1 (*38*). SORL1 has been reported involving endosomal trafficking and oncogenic fitness, which might help to induce drug resistance and tumor cell growth (*36*) in T.2. Meanwhile, LRP1 is a multifunctional receptor involved in endocytosis and metabolism homeostasis (*39*), the potential role of which should be explored in the future. Similar patterns also existed in HCC-1T.2 and HCC-1T.6 (HCC-1T.2 represented cluster 2 in HCC-1T) as well as in HCC-1T.5 and HCC-1T.6 interactions (Fig. 4E). Together, our data here suggest that those high-expressed molecules in each cluster could mediate reciprocal communications and might be used as potential targets to disrupt tumor cell communities for clinical treatment.

To investigate possible genomic drivers of different tumor clusters spatially, we inferred the copy number variations (CNVs) from ST data and WES data of matched bulk tissues (*23*, *40*, *41*). It can be seen that most CNVs inferred from ST data were consistent with bulk data, which suggests that the inferred CNVs from ST data were reliable. Further, ST data generated more subtle CNV heterogeneities across different tumor clusters. For example, in HCC-1, (i) HCC-1L.2 had the same CNV pattern as in HCC-1T.2, verifying that the same clusters across different sections had identical CNV characteristics; (ii) the observation that most CNV regions were shared across T.2/T.5/T.6 in HCC-1 suggested that those three clusters might be derived from the same clone; and (iii) the gain of chromosome 11q13 was found in HCC-1T.5 and HCC-1T.6, but not in HCC-1L.2/HCC-1T.2, whereas chromosome 8q13 amplifications were in HCC-1L.2/HCC-1T.2, but not in HCC-1T.5 and HCC-1T.6 (Fig. 4F and fig. S8A). In contrast, for HCC-3, CNV regions in L.5 were obviously different from T.2/T.4/T.7, indicating the distinct origin of L.5 (fig. S8, B to D).

Together, these results suggest that spatial intratumor heterogeneities exist widely. The different clusters within certain tumor nodule have diversified pathway activities and distinct origins. The reciprocal communications across different clusters might be useful for investigating tumor ecosystem and evolution.

Functional analysis of heterogeneous CSC niches in PLCs

CSCs are regarded as an important factor for intratumor heterogeneity in liver cancer (*42*). Different CSC populations defined by various stemness markers may contain different carcinogenic driving factors. Here, we focused on five commonly used liver cancer stemness–related markers (*CD47*, *EPCAM*, *KRT19*, *PROM1*, and *SOX9*) (*43*): For each marker, the top 5% spots with the highest

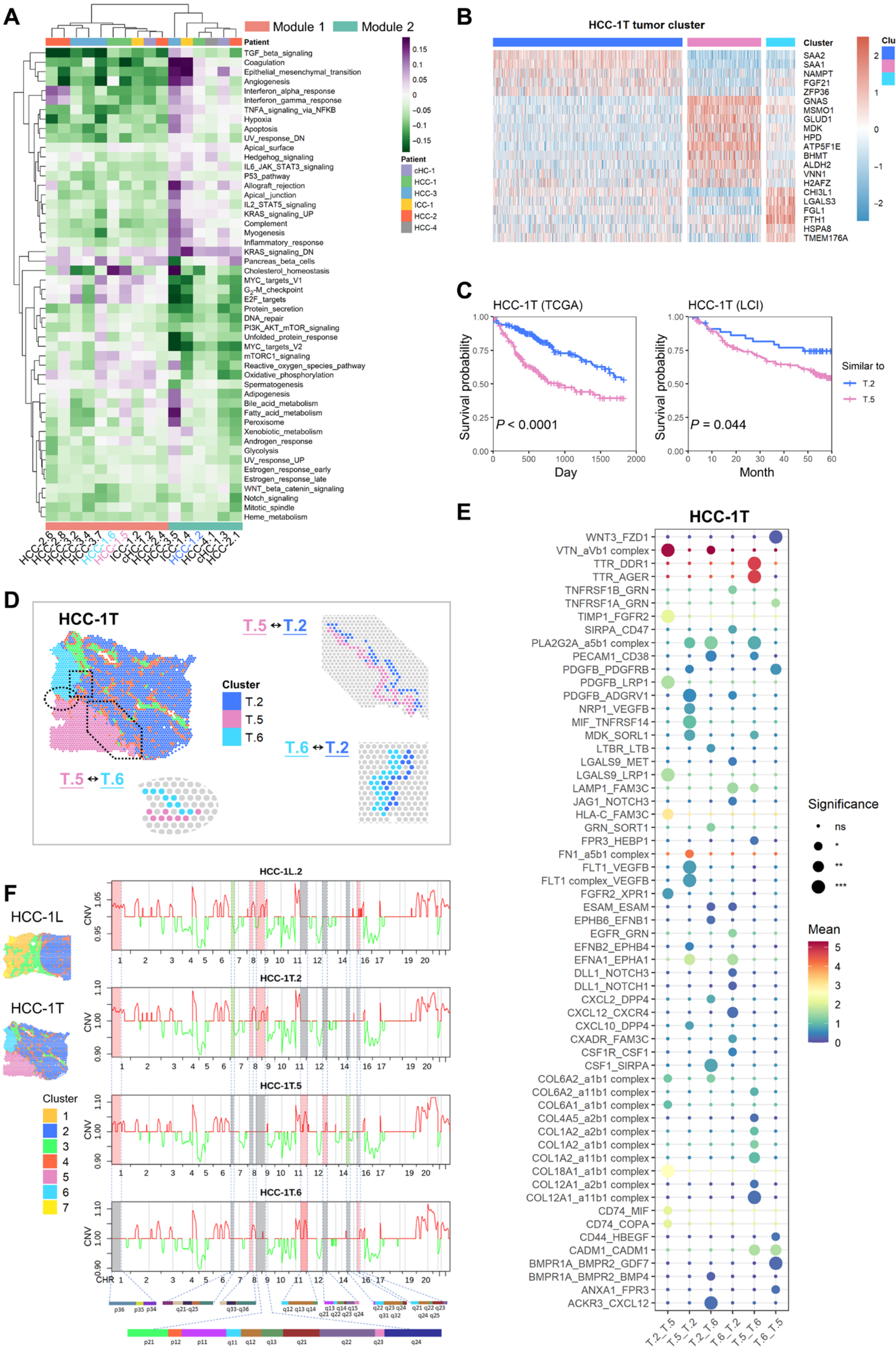

Fig. 4. Intratumor heterogeneity in PLCs. (**A**) Clustered heatmap of tumor clusters' hallmark pathway average activities. The tumor clusters were grouped into two functional modules. HCC-1.2 represented cluster 2 of HCC-1. TGF, transforming growth factor; TNFA, tumor necrosis factor a; NFKB, nuclear factor kappa B subunit; UV, ultraviolet; IL6, interleukin 6; STAT3, signal transducer and activator of transcription 3; PI3K, phosphatidylinositol 3-kinase; mTOR, mammalian target of rapamycin. (**B**) Expression profiles of some differential expression genes of the clusters 2/5/6 in HCC-1T. T.2 represented cluster 2 in HCC-1T. (**C**) Survival curves of two groups of patients in The Cancer Genome Atlas (TCGA) and Liver Cancer Institute (LCI) cohorts to compare the relative malignancy of ST tumor cluster pairs (cluster 2 versus cluster 5 in HCC-1T). These two groups were divided according to which ST tumor cluster the bulk samples were more similar to an expression level. Log-rank test was used to measure the statistical significance of their relative malignancy degrees. (**D**) Definition of the boundary areas to study the interaction between two neighbor tumor clusters in HCC-1T. The regions with four spots wide along the boundary lines in each cluster were selected, and the spots of stromal clusters were excluded. (**E**) Bubble heatmap showing the mean interaction strength between the neighbor clusters at the boundaries for ligand-receptor pairs. Dot size indicated the statistical significances by permutation test. $^*P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$. Dot color indicated the mean interaction strength levels. HCC-1T.2 represented cluster 2 in HCC-1T. ns, not significant. (**F**) Averaged copy number variation (CNV) profiles for each tumor cluster in HCC-1, inferred from spatial transcriptomes. The color of the lines indicated the amplification (red) and deletion (green). The differences between clusters were highlighted by background colors (red, green, and gray), and their detailed chromosome band labels were also presented. HCC-1L.2 represented cluster 2 in HCC-1L.

expression in each patient's tumor spots were defined as the marker positive CSC niches.

By comparing the fractions of CSC niches in L/T/P sections, we found the gradually increased fraction of *CD47*-, *SOX9*-, and *PROM1*-positive CSC niches but not *EPCAM* or *KRT19* from L to T to P sections in HCC-2 (Fig. 5A). It was observed that these CSCs were scattered in tumor regions and did not gather in a certain area (Fig. 5B). We found that *PROM1*$^+$ CSC niches not only achieved the largest fraction in portal vein tumor thrombosis (Fig. 5B) but also had higher number of double-positive spots than other positive CSC niches (Fig. 5C). By analyzing the signaling pathway activities in the five types of CSC niches of HCC-2P, we found that EMT, hypoxia, tumor necrosis factor–α, and apoptosis pathways were up-regulated in *PROM1*$^+$ CSC niches (Fig. 5D). Similar patterns were also observed in *CD47*$^+$ CSC niches, which may result from the abundance of *PROM1/CD47* double-positive CSC niches in this section. Convincingly, the positive correlation between the level of *PROM1/CD47* and the activities of the above four pathways was confirmed in our independent cohort (bulk sequencing data of 239 HCC cases) (Fig. 5E) and also in TCGA and LCI cohorts (fig. S9).

Next, we compared the microenvironment heterogeneity within different CSC niches of HCC-2P and found that CD4$^+$ T$_{em}$ cells and cDCs were significantly higher in *PROM1*$^+$ and *CD47*$^+$ CSC niches (Fig. 5F). Similar correlations were also found in our independent cohort (Fig. 5G). Briefly, these data first decipher the heterogeneity across different CSC niches and indicate that *PROM1*$^+$CSC niches may play a pivotal role in tumor progression through enhancing certain signals in tumor cells and remodeling their surrounding TME.

The global view heterogeneity within one intact HCC nodule

To better investigate the intratumor heterogeneity on a global scale, we collected a case of HCC with a diameter of about 1 cm (table S1) and performed Visium ST analysis with four sections to cover its central plane (Fig. 6A). We integrated the spots from these four sections and got six clusters by unsupervised clustering (*21*–*23*). The six clusters had different fractions in each section (Fig. 6, B and C) and can be annotated as normal region (cluster 4), stromal cell region (cluster 3/6), and tumor regions (cluster 1/2/5) according to H&E staining and different gene expression patterns (Fig. 6, A and D, and fig. S10A). In addition, the activities of oncogenic pathways in tumor clusters were different, among which clusters 2 and 5 were more similar but different from cluster 1 (fig. S10B).

To investigate the spatial variations of signaling pathway activities from tumor core to the periphery, we divided the intact plane (taking A to D sections as a whole) into 16 identical sectors around the circle center (each section was divided into four sectors named A-1, A-2, A-3, and A-4 in a clockwise longitude direction) (Fig. 6E, middle), and into g1 to g18 annuluses with five spots wide for each (indicated by circular dashed lines) (Fig. 6E, bottom). In this case, cluster 1 was spanning D-3 to B-2 sectors, and cluster 5 was regionally distributed in C-1 and C-2 sectors, whereas cluster 2 showed more complex pattern (regional distribution in C-3 to D-2 and mixed distribution in B-1 to B-4 with cluster 1) (Fig. 6E). To analyze the signaling pathway activity of different tumor clusters from g1 to g18 in each sector, we calculated the Spearman correlation coefficients between the median of signaling pathway activities and the distance to tumor center (Fig. 6F). In general, the heterogeneity of pathway activities was frequently observed. Even in the same cluster, most pathways presented diversified activities across different sectors and annulus (A-1.1 represented cluster 1 in A-1 sector) (Fig. 6F). In particular, the gradients of hallmark pathway activities mainly occurred in B and D sections. For example, the activities of most pathways in B-2.1 showed gradient descent, while gradient ascent was observed in D-3.1 and D-4.1. The level of oxidative phosphorylation signal showed a continuous gradient ascent in D-3.1, D-4.1, B-4.2, and D-1.2 (Fig. 6, G and H).

Together, our data prove that the spatial heterogeneity has emerged in a very early stage of HCC (Barcelona Clinic Liver Cancer stage 0) from a global scale. Even in this tiny tumor nodule with a 1-cm diameter, the distribution and biological behaviors of tumor cell subpopulations were not symmetrical and uniform.

Spatial distribution and clinical significance of TLS

TLS is an ectopic lymphoid structure that provides an area for DC maturation, antigen presentation, and the activation and differentiation of T and B cells, which are related to antitumor immune reactions and potentially immunotherapeutic responses (*44*). Here, we noticed that cluster 6 of HCC-5 was characterized by the high expression levels of *CXCL13*, *CCL19*, *CCL21*, *LTF*, *LTB*, *CD79A/B*, etc., which were essential molecules for TLS formation (fig. S10A) (*45*). By integrating with H&E staining, this cluster 6 was identified as TLS regions. Currently, the gene signature for TLS has not been well determined. Therefore, we selected the top 50 highly and specifically expressed genes of cluster 6 as a signature (TLS-50) (table S5) and tested its performance to identify TLSs. The pathologists confirmed that TLS-50 could identify the TLS regions with the higher signal-to-noise ratio than other signatures, including breast cancer, 12-chemokine, T helper 1, and B cell, T follicular helper cell, *CXCL13*, and plasma cell signatures (Fig. 7A and fig. S11, A and B) (*45*, *46*). In addition, we found high expression of CD52 (ranked in the top five of the TLS-50), CD53 (ranked sixth from the bottom of the TLS-50), and HLA-DR (the major histocompatibility complex, class II, DR) (*HLA-DQB1* and *HLA-DQA1* are ranked the 42nd and 50th in the TLS-50 signature, respectively) in TLS regions by immunohistochemistry (*47*) on HCC tissue microarrays (fig. S12, A and B), which also confirmed that the TLS-50 signature was accurate.

By using TLS-50 to locate the TLS spots in different sections, we found that they mainly existed in L sections but not T and N sections, especially in the nontumor region of L sections (Fig. 7B and fig. S13A). Consistent with previous reports, TLS spots were enriched with T cells (mainly CD4$^+$/CD8$^+$ T cells), B cells, and DCs in contrast to the stromal background (*48*) (Fig. 7C).

Higher scores of CD8$^+$ T$_{cm}$ cells (central memory T cells) and CD8$^+$ T$_{em}$ cells in stromal region than in TLS region were observed in HCC-5 (fig. S13B), indicating the possible influence of tumor cell on TLS composition. We then analyzed possible correlations between the TLS compositions and their distances to tumor region (fig. S14A). The scores of CD8$^+$ T$_{cm}$ cells (cHC-1 and HCC-3), CD8$^+$ T$_{em}$ cells (cHC-1, HCC-1, HCC-4, and ICC-1), and CD4$^+$ T cells (cHC-1 and ICC-1) were found negatively correlated with the distance to tumor (Fig. 7D and fig. S14B). Meanwhile, the expression levels of several functional genes *APOA1*, *EGR1*, *SELENOP*, *DUSP1*, *TIMP1*, and *HLA*-related genes (at least three of six cases) were also correlated with the distance to tumor (Fig. 7D and fig. S14C). These results suggest that the composition and function of TLSs may be affected by tumor cells.

We further checked whether the gene expression gradients exist in the surrounding region of TLSs. Besides the canonical chemokines (*CCL19*, *CCL21*, and *LTB*) that have been reported for promoting TLS formation (*49*), the expressions of *CXCR4* (a marker of the

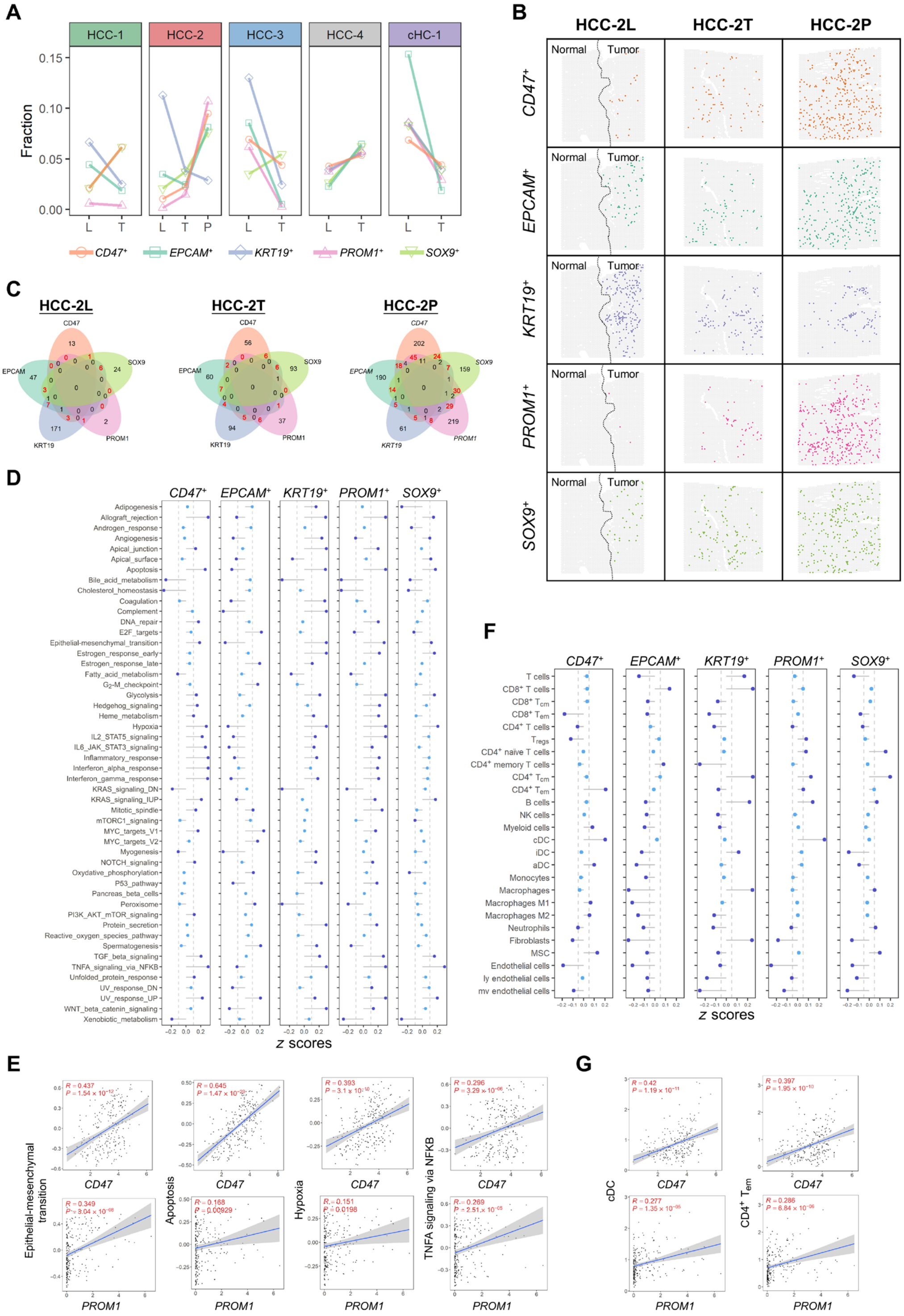

Fig. 5. Functional analysis of heterogeneous CSC niches in PLCs. (**A**) Fractions of the positive spots of the five CSC markers in indicated sections. (**B**) Distribution of the putative CSC niches (spots) of five CSC markers in the L/T/P sections of HCC-2. (**C**) Venn diagrams of the positive niches of the five CSC markers in L/T/P sections of HCC-2. The numbers of double-positive spots were highlighted in red. (**D**) Hallmark pathway activities (average *z* scores of HCC-2P) in the five types of CSC niches. TGF, transforming growth factor; TNFA, tumor necrosis factor a; NFKB, nuclear factor kappa B subunit; UV, ultraviolet; IL6, interleukin 6; STAT3, signal transducer and activator of transcription 3; PI3K, phosphatidylinositol 3-kinase; mTOR, mammalian target of rapamycin. (**E**) Scatter plots showing the correlation between hallmark pathway activities and the expression levels of the CSC markers (*CD47* and *PROM1*) on bulk sequencing data of 239 HCC cases. Linear regression and Pearson correlation were used to measure their relation. (**F**) Cell type enrichment scores (average *z* scores of HCC-2P) in the five types of CSC niches. MSC, mesenchymal stem cells. (**G**) Scatter plots showing the correlation between cell type enrichment scores and the expression levels of the CSC markers (*CD47* and *PROM1*) on bulk sequencing data of 239 HCC cases. Linear regression and Pearson correlation were used to measure their relation.

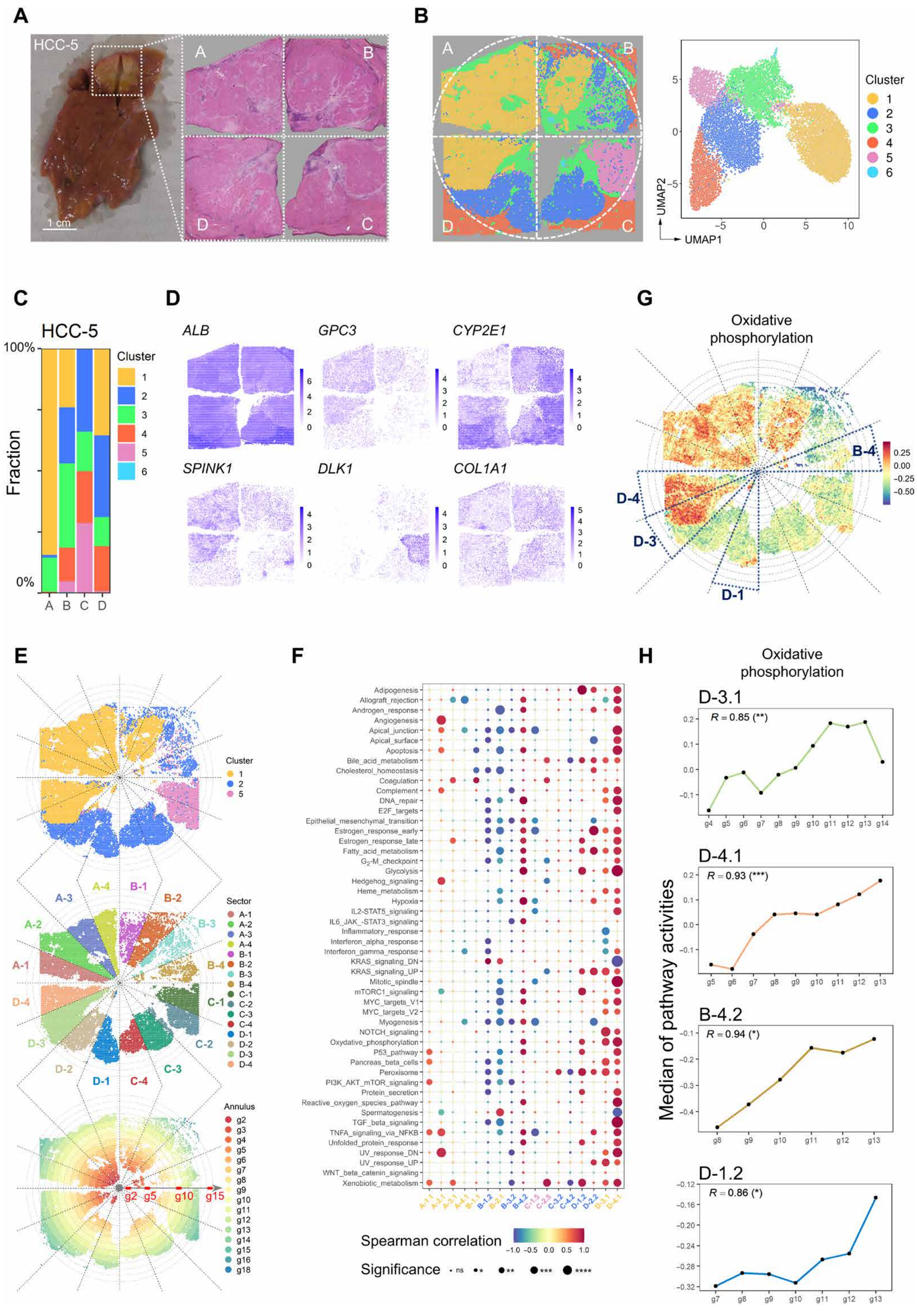

Fig. 6. The global view tumor heterogeneity within one intact HCC nodule. (**A**) A case of HCC (HCC-5) with a diameter of about 1 cm was divided into four sections (A, B, C, and D sections) to analyze the global spatial heterogeneity. Photo credit: Rui Wu, International Cooperation Laboratory on Signal Transduction, Eastern Hepatobiliary Surgery Hospital, Shanghai 200438, China. (**B**) Distribution of clusters in the tissue space and UMAP space. (**C**) Fractions of the clusters in each section. (**D**) Spatial feature plots of six differentially expressed genes on HCC-5. (**E**) Distribution of the divided sectors and annuluses on the tumor region (top). Each section was equally divided into four sectors (middle). Each annulus area was five spots wide (bottom). (**F**) Bubble heatmap showing the Spearman correlation between the median of the hallmark pathway activities and the annular orders in the tumor clusters of different sectors. A-1.1 represented cluster 1 in sector A-1. Dot color indicated the correlation values. Dot size indicated the statistical significance. $^*P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$; $^{****}P < 0.0001$. (**G**) Spatial distribution of the oxidative phosphorylation pathway activities. (**H**) Change of oxidative phosphorylation pathway median activities along with the annulus orders from inside to the outside in the D-3.1, D-4.1, B-4.2, and D-1.2 (bottom). The Spearman correlations and statistical significances were labeled. $^*P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$.

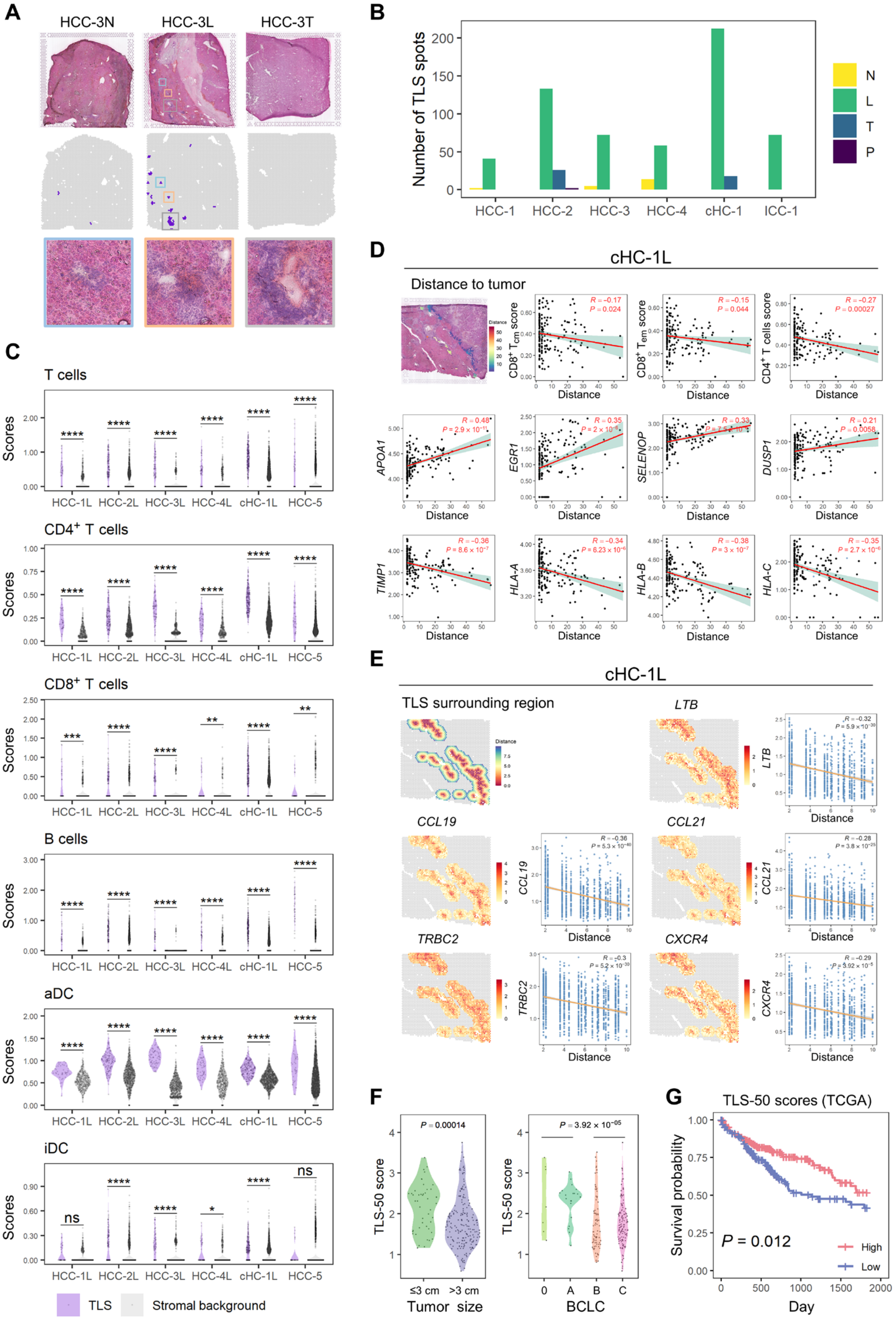

Fig. 7. Spatial distribution and clinical significance of TLS. (**A**) TLSs identified by TLS-50 signature and their corresponding H&E staining images in HCC-3. (**B**) Bar plot showing the distribution of the identified TLS spots in the different sections of each patient. (**C**) Violin plots showing the signature difference of immune cell type scores in the TLSs and their surrounding stromal backgrounds. Statistical significances were determined by two-sided Wilcoxon rank sum tests. $*P < 0.05$; $**P < 0.01$; $***P < 0.001$; $****P < 0.0001$. (**D**) Scatter plots showing the correlation between the cell type enrichment degrees or gene expression levels of TLS spots and their distances to tumor regions in cHC-1 L. These distances were presented at the H&E staining (the first subplot). Linear regression and Pearson correlation were used to measure the relation. (**E**) Spatial plots showing the surrounding regions (five spots wide) of TLSs in cHC-1 L, which were colored by the distances to TLS spots or some genes' expression profiles. Scatter plots showing the correlation between these genes' expression levels and the distances to TLS spots. Linear regression and Pearson correlation were used to measure the relation. (**F**) Violin plots showing the distribution of TLS-50 scores of the bulk sequencing data of 239 HCC cases, which were grouped by tumor sizes (left) and BCLC stages (right), respectively. (**G**) Survival curves of cases with high (top 50%) and low (bottom 50%) TLS-50 signature scores in TCGA cohorts. Log-rank test was used to measure the statistical significance.

mature germinal center) (*50*) and *TRBC2* (a gene-encoding T cell receptor constant region) (*51*) also showed gradient descent in all the seven PLC cases, indicating their potential roles in TLS formation and development (Fig. 7E and fig. S15, A and B).

At last, in our bulk sequencing data of 239 HCC samples, we confirmed that TLS-50 signature was comparable to 12-chemokine signature in evaluating TLS abundance (fig. S11C). TLS-50 scores were found to be lower in HCCs with a tumor size of >3 cm and BCLC stage B/C (Fig. 7F). The higher score of TLS-50 was found significantly associated with a better prognosis in TCGA patients with HCC (Fig. 7G).

DISCUSSION

In this study, genome-wide transcriptome heterogeneity of seven PLCs was analyzed comprehensively by Visium ST technology at 55-μm resolution. By conducting progressive comparison from normal to leading-edge to tumor regions, we found that the tumor capsule potentially affects intratumor spatial cluster continuity, transcriptome diversity, immune cell infiltration, and cancer hallmark pathway activities. Meanwhile, several local spatial features of liver cancers, including the 100-μm-wide tumor cluster-cluster boundary, CSC-niches, and TLS composition, were comprehensively analyzed.

The TME comprises tumor, stromal and immune cells, extracellular matrix, and signaling molecules (*52*). The spatially and temporally dynamic variations in TME are considered as the key factors for tumor heterogeneity (*53*). To our knowledge, this is the first study to analyze the genome-wide TME characteristics from normal to leading-edge to tumor regions, which provides us the opportunity to investigate both the global and local variation tendency of different cell populations. These results show that the complete fibrous capsules, mainly consisting of fibroblasts and endothelial cells, can act as a barrier preventing the infiltration of immune cells, which is supported by a few previous studies (*54*). We found that the absence of capsules in L sections could lead to both lower spatial continuity and higher transcriptome diversity in tumor. Together with the observation that the relatively small TLS spots in L and T sections in HCC-1/3/4 (fig. S11A), our data here reinforce the key role of complete fibrous capsule for both TME architecture and intratumor heterogeneity.

The tumor heterogeneity includes interpatient heterogeneity, intertumor heterogeneity (different tumor nodules within the same patient), and intratumor heterogeneity (different regions in the identical tumor nodule) (*55*). Globally, the spatial distribution of clusters within tumors has two distinct patterns: regional distribution and intertwined distribution. The regionally distributed clusters tend to have higher spatial continuity and lower transcriptome diversity. By comparing the inferred CNVs from ST data with the CNVs from matched bulk WES data, we found most of them were consistent, reinforcing that it is practical to infer CNV by ST data. Note that in comparison with cluster 2/4/7 in HCC-3T, cluster 5 displayed a distinct CNV pattern and a significant lower level of genomic disorder. Together with the pathological stage that cluster 5 was focal nodular hyperplasia (the premalignant nodule) instead of neoplasia, it suggested that the evolution trajectory of cluster 5 was different from cluster 2/4/7 in HCC-3 (fig. S8, C and D). By building a spatial transcriptome landscape of an intact tumor nodule (diameter around 1 cm), we found that the heterogeneous tumor clusters have already existed in such a small tumor nodule, implying the formation of diverse populations during tumorigenesis. Regarding the changes of signal activities, we uncovered widely different trends of hallmark pathway activities in all directions from inside to outside of tumor, which might result from the interplay with other cluster cells and multiple stresses locally. This complex heterogeneity of HCC explains the reason why the current drug treatment for HCC is usually ineffective and also suggests that a comprehensive detection of the genetic characteristics of different parts of the tumor may make a breakthrough in immunotherapy.

It is well known that CSCs play multiple roles in tumor development, progression, and drug resistance (*56*). Several CSC populations with different surface markers have been identified in HCCs (*43*). However, because of the lack of understanding of CSC niche composition, the function and underlying mechanism of different CSC populations remain unclear. Our study here firstly investigated the characteristics of diverse CSC niches defined by five different stemness markers spatially. We found that the CSC niches are scattered distributed in tumors, their abundances vary from leading-edge to tumor regions, and their cell compositions and pathway activities are diversified. *PROM1*$^+$ and *CD47*$^+$ CSCs were found to be gradually increased from leading-edge to tumor to portal vein tumor thrombus and closely related to EMT activation and suppressive immune cells infiltration, suggesting that the *PROM1* and *CD47* are potential targets for preventing tumor vascular metastasis.

On the basis of the ST data, we defined a novel TLS-50 signature to identify TLS with high confidence and to assess its abundance from bulk sequencing data. The TLS-50 high-scoring regions were confirmed as the TLSs by pathologists based on histopathological features (see Materials and Methods for details). Immunohistochemistry and CODEX (*47*) confirmed that the selected genes in TLS-50 (*CD52*, *CD53*, *HLA-DQB1*, and *HLA-DQA1*) were highly expressed in TLSs in additional patients with HCC. Because of the existence of TLSs in both adjacent normal and tumor regions, the TLS-50 signature could be applied with TCGA data, and we found that HCC with a tumor size of >3 cm or BCLC stage B/C has lower TLS-50 scores, which is in accordance with previous reports that the existence of TLSs in HCC tumor area favors the prognosis (*45*). Taking advantage of ST technology, the heterogeneity of TLS function in different spatial locations was first revealed. CD8$^+$ T$_{cm}$ and CD8$^+$ T$_{em}$ cell scores of TLS in the tumor core region (HCC-5) were distinct from that in the tumor-invasive margin or peritumor area. Consistently, the CD8$^+$ T$_{cm}$ and CD8$^+$ T$_{em}$ cell scores and the expression level of indicated genes (*APOA1*, *EGR1*, *TIMP1*, *SELENOP*, *DUSP1*, *HLA-A*, etc.) were significantly correlated with the distance from TLS to the tumor cells, suggesting that TLS function was largely affected by tumor cells, whereas the underlying mechanisms and functional molecules need to be further clarified. Transcription factor early growth response 1 (encoding by EGR1) has the ability to stimulate T cell activity and promotes B cell proliferation (*57*). Dual-specificity phosphatases (DUSPs), the major regulator of mitogen-activated protein kinases signaling pathway, widely express in lymphoid tissues and manipulate T cell function and immune response (*58*). Tissue inhibitor of matrix metalloproteinase 1 (TIMP-1) is activated primarily in B cells and serves as a survival factor for B cells growth (*59*). It is rational to speculate that EGR1/DUSP1/TIMP-1 might involve in the proliferation and homeostasis of TLS near the tumor. The combination of single-cell and ST technologies with a larger sample size and higher precision is expected to further elucidate the function of TLS and provide more theoretical basis for tumor immunotherapy.

The emerging ST technology largely solves the shortcomings of missing cell spatial location information by single-cell sequencing and plays an important role in many research fields. However, the major limitations of ST are still its resolution: (i) Although the diameter of each spot in ST section in our present study has reached 55 μm instead of 100 μm in the former version (8 to 20 cells versus 100 to 200 cells each spot) (*10*), it is still unable to provide the comparable accuracy at single-cell scale; (ii) ST method can only provide the transcript information of cells within spots, whereas the information of interval spaces between two spots is missed. To deal with the first issue, we applied the cell type signature–based strategy to calculate the enrichment of TME cell types in each spot, which achieved a similar result as the previous developed MIA method did (fig. S4, B to D). For the second issue, the improved technology with shorter interval distance and advanced analytical methods needs to be developed in the future.

Tumor heterogeneity is the major obstacle for liver cancer diagnosis and therapy. Our study presents the first genome-wide spatial transcriptome map of three major liver cancer subtypes. Extensive global and local intratumor heterogeneities of tumors and TMEs have been found. In addition, the tumor clusters from different patients show distinct spatial patterns and transcriptomic diversities. These findings provide meaningful insights to find new drug targets and develop novel therapeutic strategies.

MATERIALS AND METHODS

Human PLC samples and blood

The adjacent normal and tumor (PLC) tissues were collected under a protocol approved by the Ethics Committee of Eastern Hepatobiliary Surgery Hospital (EHBH) (EHBHKY2018-1-001) following Institutional Review Board guidelines. Individuals donating fresh surgical tissue provided informed consent. All diagnoses were verified by histological review by a board-certified pathologist. The samples were delivered within MACS Tissue Storage Solution (Miltenyi, catalog no. 130-100-008). For tumors of adequate size, small fragments of each tumor were snap-frozen in optimum cutting temperature (OCT) compound (Sakura, catalog no. 4583) and stored at −80°C until use. Peripheral venous blood (5 ml) was collected in an anticoagulation tube (BD Vacutainer, catalog no. 367525).

Collection and preparation of PLC tissue

The surgically resected PLC tissue was immediately submerged in MACS Tissue Storage Solution and sent to the laboratory for processing as soon as possible. Then, PLC tissue was gently washed with cold phosphate-buffered saline (PBS) (Gibco, catalog no. 20012-043) and cut into about 6.5-mm^3 pieces (bulks) according to the experimental design. In addition, the same pieces immediately adjacent to them were directly frozen at −80°C for WES. The tissue bulks were placed in OCT-filled mold and snap-frozen in isopentane and liquid nitrogen. Cryosections were stored at −80°C until use.

PBMC preparation

Fresh anticoagulated blood was centrifuged (500*g* for 5 min), and then the upper layer of plasma was removed. After adding 10 ml of PBS and mixing thoroughly, the mixed solution was slowly added to the surface of the Ficoll solution (GE Healthcare, catalog no. 17544203). After centrifugation (800*g* for 20 min at room temperature, brake off), the white mononuclear cell layer in the middle of the solution was taken out and mixed with 10 ml of PBS. After centrifugation again (500*g* for 5 min), the solution was discarded. Last, the sediment at the bottom was resuspended in 200 μl of cell-freezing medium (10% dimethyl sulfoxide + fetal bovine serum) and stored at −80°C.

Spatial transcriptome sequencing

This experiment is based on the Visium Technology Platform of 10x Genomics company. The reagents and consumables in the experiment are provided by this platform, and the specific product numbers can be found at www.10xgenomics.com/products/spatial-gene-expression.

Slide preparation

The Visium Spatial Gene Expression Slide (from Visium Spatial Gene Expression Slide Kit, 10x Genomics, PN-1000185) includes four capture areas (6.5 mm by 6.5 mm), each defined by a fiducial frame (fiducial frame + capture area is 8 mm by 8 mm). The capture area has ~5000 gene expression spots, each spot with primers that include Illumina TruSeq Read 1 (partial read 1 sequencing primer), 16–nucleotide (nt) spatial barcode (all primers in a specific spot share the same spatial barcode), 12-nt UMI; 30-nt poly(dT) sequence (captures polyadenylated mRNA for cDNA synthesis).

RNA integrity number

We use RNeasy Mini Kit (QIAGEN, catalog no. 74104) to test the integrity of RNA. After taking 10 slices of 10-mm-thick cryosections, RNA was extracted and analyzed by RNeasy Mini Kit immediately. An RNA integrity number of ≥7 is qualified.

Optimization of the permeabilization time

Before using a new tissue for generating Visium Spatial Gene Expression libraries, the permeabilization time was optimized. Briefly, the Visium Spatial Tissue Optimization workflow included placing tissue sections on seven-capture areas on a Visium Tissue Optimization slide (from Visium Spatial Gene Expression Reagent Kit, 10x Genomics, PN-1000186). The sections were fixed, stained, and then permeabilized for different times. The mRNA released during permeabilization binds to oligonucleotides on the capture areas. Fluorescent cDNA was synthesized on the slide and imaged. The permeabilization time that results in maximum fluorescence signal with the lowest signal diffusion was optimal. If the signal was the same at two time points, then the longer permeabilization time was considered optimal. Once optimal conditions had been established, each cryosection was cut at 10 mm in thickness onto Visium Slide (from Visium Slide Kit) and processed immediately. In this study, the permeabilization time ranges from 6 to 24 min depending on the samples.

Tissue fixation, staining, and imaging

Tissue sections on the Visium Slide (from Visium Slide Kit) were fixed using methanol (MilliporeSigma) by incubating 30 min at −20°C. For tissue staining, sections were incubated in isopropanol (MilliporeSigma) for 1 min, in hematoxylin (Agilent) for 7 min, in Bluing Buffer (Agilent) for 2 min, and in Eosin Mix (MilliporeSigma) for 1 min at room temperature. Last, slides were incubated for 5 min at 37°C in the Thermocycler Adaptor (10x Genomics, PN-3000380).

The slides were washed in ultrapure water after each staining step. Then, the stained tissue sections are imaged.

Tissue permeabilization and reverse transcription

For tissue permeabilization, the slides were first placed in the Slide Cassette (from the Visium Slide kit) for the optimal permeabilization time. A permeabilization enzyme (from the Visium Reagent kit) was used for permeabilizing the tissue sections on the slide for incubating for the predetermined permeabilization time. The polyadenylated mRNA released from the overlying cells was captured by the primers on the spots. After washing by 0.0.1×SSC (saline sodium citrate buffer, MilliporeSigma), RT Master Mix (provided in Visium Reagent kit) containing reverse transcription reagents was added to the permeabilized tissue sections in the Thermocycler Adaptor. Incubation with the reagents produces spatially barcoded full-length cDNA from polyadenylated mRNA on the slide.

Second strand synthesis and denaturation

After removing RT Master Mix (provided in Visium Reagent kit) from the wells, sections were incubated in 0.08 M KOH for 5 min and washed by Buffer EB (QIAGEN). Then, Second Strand Mix (provided in Visium Reagent kit) was added to the tissue sections on the slide to initiate second strand synthesis on the Thermocycler Adaptor. This is followed by denaturation and transfer of the cDNA from each capture area to a corresponding tube for amplification and library construction. The slides were washed by Buffer EB and incubated in 0.08 M KOH for 5 min. Then, samples from each well were transferred to a corresponding tube containing tris-HCl (1 M; pH 7.0) in eight-tube strip for amplification and library construction.

cDNA amplification and quality control

Denaturation sample of 1 μl was transferred to the quantitative polymerase chain reaction (qPCR) plate well containing the qPCR Mix [nuclease-free water + KAPA SYBR FAST qPCR Master Mix (KAPA Biosystems) + cDNA Primers (from Visium Reagent kit)]. The C_q value for each sample was recorded after qPCR. For cDNA amplification, cDNA Amplification Mix (from Visium Reagent kit) was added to the remaining sample from denaturation. Then, the product was incubated in Thermocycler Adaptor for a cycle. For cDNA Cleanup–SPRIselect, 60 μl of SPRIselect reagent (Beckman Coulter) was added to each sample and incubated for 5 min at room temperature. The sample was repeatedly adsorbed by the magnet·High, washed with ethanol (MilliporeSigma) and Buffer EB and transferred to a new tube strip. Then, we ran 1 μl of sample on an Agilent Bioanalyzer High Sensitivity chip (Agilent, catalog no. 50674626) for cDNA quality control (QC) and quantification.

Visium spatial gene expression library construction

Enzymatic fragmentation and size selection were used to optimize the cDNA amplicon size. P5, P7, i7, and i5 sample indexes and TruSeq Read 2 (read 2 primer sequence) were added via End Repair, A-tailing, Adaptor Ligation, and PCR. The final libraries contain the P5 and P7 primers used in Illumina amplification. Library construction was performed with Library Construction Kit (10x Genomics, catalog no. PN-1000190).

Fragmentation, End Repair, and A-tailing

Only 10 μl of purified cDNA sample from cDNA Cleanup was transferred to a tube strip. Buffer BE and Fragmentation Mix (from Library Construction kit) were added to each sample, and fragmentation was performed in a thermal cycler. After fragmentation, 30 μl of SPRIselect reagent (0.6×) was added to each sample and incubated for 5 min at room temperature. The sample was repeatedly adsorbed by the magnet·High, washed with ethanol and Buffer EB, and transferred to a new tube strip.

Adaptor ligation

Adaptor Ligation Mix (from Library Construction kit; 50 μl) was added to each 50 μl of sample and incubated in a thermal cycler.

Postligation cleanup–SPRIselect

SPRIselect reagent (0.8×) was added to each sample and incubated for 5 min at room temperature. The sample was repeatedly adsorbed by the magnet·High, washed with ethanol and Buffer EB, and transferred to a new tube strip.

Sample index PCR

Amp Mix (from Library Construction kit; 50 μl) and an individual Dual Index TT Set A (10x Genomics, catalog no. PN-1000215, 20 μl) were added to each 30 μl of sample and incubated in a thermal cycler.

Post–sample index PCR double-sided size selection – SPRIselect

SPRIselect reagent (0.6×) was added to each sample and incubated for 5 min at room temperature. After adsorbing by the magnet·High, 150 μl of supernatant was transferred to a new tube strip. Then, 20 μl of SPRIselect reagent (0.8×) was added to each sample and incubated for 5 min at room temperature. After adsorbing by the magnet·High and supernatant removed, samples were washed with ethanol and Buffer EB and transferred to a new tube strip. We ran 1 μl of the sample (1:10 dilution) on an Agilent Bioanalyzer High Sensitivity chip.

Sequencing

A Visium Spatial Gene Expression library comprises standard Illumina paired-end constructs that begin and end with P5 and P7. The 16–base pair (bp) spatial barcode and 12-bp UMI are encoded in Read 1, while Read 2 is used to sequence the cDNA fragment. i7 and i5 sample index sequences are incorporated. TruSeq Read 1 and TruSeq Read 2 are standard Illumina sequencing primer sites used in paired-end sequencing.

Raw sequencing data processing

The Visium Spatial RNA-seq output and bright-field and fluorescence microscope images were analyzed by Space Ranger (version 1.1.0) to detect tissue, align reads, generate feature-spot matrices, perform clustering and gene expression analysis, and place spots in spatial context on the slide image. These pipelines combined Visium-specific algorithms with the widely used RNA-seq aligner STAR.

Whole-exome sequencing

WES was performed per standard protocols using the Next-Generation Sequencing platform of TWIST bioscience company whose details can be found at www.twistbioscience.com/products/ngs. Briefly, for DNA extraction, snap-frozen fresh biopsy and matched whole blood samples were processed using QIAamp DNA Mini Kit (QIAGEN, catalog no. 51304) and QIAamp DNA Blood Maxi Kit (QIAGEN, catalog no. 51194) according to the manufacturer's instructions and quantified using the Qubit dsDNA BR Assay Kit (Thermo Fisher Scientific, catalog no. Q32853). Libraries were generated with Twist Library Preparation EF Kit 1 (TWIST, catalog no. 100572), Twist CD Index Adapter Set (TWIST, catalog no. 100577), and Twist Library Preparation Kit 2 (TWIST, catalog no. 100573). Subsequently, hybridization and capture were performed using the Twist Fast Hybridization and Wash Kit (TWIST, catalog no. 101175) and Twist Binding and Purification Bead (TWIST, catalog no. 100983), respectively. After capture, libraries were amplified by PCR. Then, purified libraries were validated and quantified using an Agilent Bioanalyzer High Sensitivity DNA Kit (Agilent, catalog no. 50674626) and a Qubit dsDNA High Sensitivity Quantitation Assay (Thermo Fisher Scientific, catalog no. Q32854). Last, the enriched libraries were sequenced on the Nova6000 instrument of Illumina platform (Illumina).

Pathological annotation

In this study, the macroscopic tissue appearance and histopathological and clinicopathological findings of the tumor specimen were annotated by pathologists. Tumors with complete fibrous capsule (without evidence of capsular and/or vascular invasion) were defined as the CC group, and tumors lacking fibrous capsules or with a discontinuous band consisting of stromal cells were defined as the NC group.

The TLSs identified by TLS-50 signature were all double-checked by two independent pathologists according to two criteria: (i) presenting aggregated immune cells with germinal centers in H&E image and (ii) highly expressing TLS-related marker genes. Pathologists confirmed that the candidate TLSs with >5 spots were highly reliable. For the candidates with smaller size (two to five spots), the pathologists also considered them as TLSs by carefully examining H&E staining and clinical features. All those patients had chronic hepatitis B virus infection (see table S1), which is a favorable factor for TLS formation.

ST data basic analysis

For the gene-spot matrixes generated by Space Ranger, some routine statistical analyses were performed first, including calculating the number of the detected UMIs (nUMI) and genes (nGene) in each spot. On the basis of them, the basic QCs were applied on the data. In detail, the spots with extremely low nUMI or nGene (outliers), and the spots isolated from the main tissue sections were removed. The genes were expressed in less than three spots, and mitochondrial and ribosomal genes were filtered.

After QC, we used the R package harmony (v1.0) (*21*) to integrate the expression data from different sections of each patient and used the Seurat package (v3.1.5) (*22*) and scCancer package (*23*) to perform the basic downstream analysis and visualization. In detail, we first combined the expression matrixes of each patient's all sections and performed normalization, log-transformation, centering, and scaling on them. Next, we identified 2000 highly variable genes according to their expression means and variances. On the basis of them, principal components analysis (PCA) was performed to project the spots into a low-dimensional space, which was defined by the first 20 principal components (PCs). Then, by setting the section source as the batch factor and using the "RunHarmony" function, we iteratively corrected the spots' low-dimensional PC representation to reduce the impact of batch effect. After this step, the corrected PC matrixes were used to perform unsupervised shared-nearest-neighbor (SNN)–based clustering and UMAP visualization analysis further. In addition, to compare the clusters at gene level, we identified differentially expressed genes of all or selected clusters using fold change analysis and Wilcoxon rank sum test with Bonferroni correction.

Cluster similarity analysis

For the clusters from different patients, we represented them by their spots' average expression profiles (the log-transformed normalization values). To reduce the impact of extreme values, we excluded some outlier spots in advance, whose first three PC values were beyond the range of the means ±3SD of the cluster they belonged to. Moreover, only the genes with the mean above 0.1 and the variance above 0.05 across all the cluster expression vectors were retained for the downstream comparison analyses.

To measure the clusters' similarities across patients, we performed two types of analyses, hierarchical clustering and low-dimensional projection. In detail, we first applied PCA on the centered and scaled clusters' average expression profiles and used the first five PCs to perform hierarchical clustering (Fig. 2B). Besides, the diffusion map was used to project clusters of different patients into a two-dimensional space (the first two DCs) based on the package destiny (*24*) with default parameter setting (fig. S4A). For the convenience of comparison, we annotated each cluster with a region label (normal, stromal, or tumor), which was decided by integrating the information of the cluster's marker genes and H&E staining images.

Cell type scoring by a signature-based strategy

At the current Visium ST resolution, each spot may contain approximately 8 to 20 cells, so that we could not assign a certain cell type for each spot. Considering this, to compare the distribution of cell types across the tissue sections, we proposed a signature-based strategy to score the cell type enrichments in each spot. First, we curated a set of gene signatures of common cell types in liver cancer based on the xCell signatures (*25*) and biology prior knowledge (table S3). Then, we defined the average log-transformed normalization expression values of the genes in the signature as the corresponding cell type scores. Taking advantage of these scores, the cell type relative enrichment degree across different tissue regions can be compared. By testing on some scRNA-seq datasets of liver cancer, we proved that our curated gene signatures had high sensitivity and specificity.

Furthermore, we also verified the performance of our method by comparing it with the MIA (*10*), which determined the cell type enrichment degrees by performing hypergeometric test on the overlap between the tissue region–specific genes of ST data and the cell type–specific genes of single-cell data. Here, we took advantage of cell type annotation and differential expression gene results of a liver cancer single-cell dataset (*12*) and performed MIA on the clusters of our ST data, so that we can use the *P* values of hypergeometric test to measure the enrichment of different cell types in each cluster (fig. S4C). By comparing these enrichment degrees and the mean values of our signature-based cell type scores of all ST clusters, we observed a generally high correlation (fig. S4D), which proved the reliability of our signature-based cell type scoring method. At the same time, it had the advantage of not requiring single-cell data, which was more flexible.

Intratumor spatial heterogeneity measurement

To measure the degree of intratumor heterogeneity from two aspects of transcriptome and tissue space, we proposed two metrics, transcriptome diversity degree and spatial continuity degree. For the transcriptome diversity degree, we first calculated the Pearson correlation coefficients between each pair of tumor region spots based on the highly variable genes. Then, we defined the sample's transcriptome diversity degree as the 1.4826 times median absolute deviation (MAD) of these correlations, which was an approximation of the SD but can avoid the impact of outliers. The larger this metric meant that the similarities among the sample's tumors spots had larger variance so that the sample had higher intratumor heterogeneity. Formulaically, it can be calculated as

$$\text{Transcriptome diversity degree} = 1.4826 * \text{MAD}\left(\text{cor}_{i \neq j}(\mathbf{e}_i, \mathbf{e}_j)\right)$$

where $\mathbf{e}_i$ indicated the expression vector of the tumor region spot i and the MAD was defined as

$$\text{MAD}(\mathbf{x}) = \text{median}(\,|\,x_i - \text{median}(\mathbf{x})\,|\,)$$

For the spatial continuity degree, we first compared the cluster identities of each tumor region spot with its six neighbor spots. Then, the total fraction of the neighbor spots with the same cluster identity was defined as the spatial continuity degree. This metric measured the tumor region's spatial heterogeneity. The larger this metric meant the sample's tumor region more tended to be block-like (higher spatial continuity degree and lower spatial mixed degree). Formulaically, it can be calculated as

$$\text{Spatial continuity degree} = \frac{\sum_i \sum_{j \in \text{Neigh}(i)} I(\text{cluster}_i = \text{cluster}_j)}{\sum_i \sum_{j \in \text{Neigh}(i)} 1}$$

where i indicated a tumor region spot and $I()$ was the indicative function.

Gene set variation analysis

The pathway activities of tumor cluster spots were quantified by applying GSVA, implemented in the GSVA package (*26*). In detail, the log-transformed normalization expression matrix of tumor spots was inputted into the "gsva" function with the default parameters setting. The set of 50 cancer hallmark signatures (MSigDB, H sets), 189 oncogenic signatures (MSigDB, C6 sets) (*60*), and 96 metabolic pathways (*27*) were used to analyze. Besides, to compare the tumor clusters across patients at the pathway level, we averaged the resulting GSVA score matrixes over each cluster and performed hierarchical clustering on them with Ward's minimum variance method (Fig. 4A).

Spatial gradient change analysis

The spatial gradient distributions of hallmark pathway activities were analyzed on our leading-edge samples (L sections) and the intact HCC nodule (HCC-5). For the leading-edge samples, we focused on analyzing the gradient changes from capsules or tumor-normal boundary lines to both tumor and normal sides. The capsule (HCC-1L, HCC-3L, and HCC-4L) boundaries were determined on the basis of the clustering results and fine-adjusted manually using the software Loupe Browser. When the capsules did not exist (ICC-1L) or were incomplete (HCC-2L and cHC-1L), the tumor-normal boundaries were decided manually according to the interface of clusters and the H&E staining images in Loupe Browser. Then, we divided the normal and tumor regions into continuous zones parallel to the shape of the boundary lines at intervals of five spots (fig. S6A). In addition, the gradient changes along these zones were analyzed.

For the intact HCC nodule, to explore its global spatial gradient changes from tumor core to the periphery at different directions, we defined the junction of the four sections as the tumor center and divided the tumor region into 16 identical sectors (Fig. 6E, middle). Simultaneously, we divided the entire nodule into some concentric annuluses with five spots wide (named from g1 to g18) (Fig. 6E, bottom). Then, the gradient changes along the annuluses in each sector's main tumor clusters (across at least five gradient divisions) were analyzed.

In detail, the significance of the gradient change trends was measured by calculating the Spearman correlation coefficients between the pathway activity medians of different gradient divisions and their distance ranks to the boundaries (for L sections) or the tumor core (for HCC-5). To avoid the impact of some outliers, the gradient divisions with too few spots were discarded in the analysis.

Tumor cluster malignancy comparison analysis

To evaluate the relative malignancy degree of different ST tumor clusters, we used two liver cancer bulk datasets [TCGA-LIHC (liver hepatocellular carcinoma) and LCI cohorts] (*29*) for comparison, which were downloaded from the HCCDB website (i.e., HCCDB15 and HCCDB6 datasets) (*28*). To reduce the impact of disease-irrelevant deaths, we truncated the patients' survival times to 5 years and set their statuses as "alive" when they had longer survival times. Then, for each sample's ST tumor cluster, we excluded its outliers and represented them by their average expression values of the remaining spots (the details were the same as the method in "Cluster similarity analysis" section). This step can be regarded as transforming each tumor cluster into a pseudo-bulk sample. Next, to compare the malignancy of any pair of two ST tumor clusters, we calculated the Spearman correlation coefficients between each cluster and the bulk samples (TCGA-LIHC or LCI cohorts) across a list of about 1000 survival-related genes from HCCDB. According to the correlations, we determined which ST tumor cluster the bulk samples were more similar to, so that these bulk samples can be classified into two groups. By plotting Kaplan-Meier survival curves and performing log-rank test on these two groups, we decided the relative malignancy degrees between these two ST tumor clusters (Fig. 4C and fig. S7B).

Cluster interaction analysis

We used HCC-1T to explore the interaction between two neighbor tumor clusters because the three tumor clusters (2, 5, and 6) in HCC-1T had clear interfaces between each other and were highly heterogeneous. For each pair of neighbor tumor clusters, we selected their interface regions with four spots wide (two spots wide for each cluster) and excluded the spots identified as stromal clusters (Fig. 4D). Then, we used the CellPhoneDB (*30*, *31*) to analyze the interaction strengths, which were defined as the means of the average expression level of ligand and receptor in the corresponding cluster interface spots. For each ligand-receptor pair in each interaction analysis, we performed 1000 randomized permutations for spots' cluster labels and recalculated the mean values, which can be seen as a null distribution. By calculating the proportion of these mean values that exceed the actual interaction strength, we obtained a P value to measure the statistical significance of the interaction on the interface of two tumor clusters (Fig. 4E).

CNV comparison analysis

The CNVs of each tumor spot were estimated on the basis of their transcriptome profiles using the method of infercnv (*40*). First, for each patient, we defined the normal hepatocyte spots in their N section as normal references. Then, all the analyzed genes were sorted by their location in the chromosomes and a sliding window of 100 genes was applied on them to calculate their moving average expression values so that the initial copy numbers were estimated. By subtracting the normal reference copy number profiles from that of the tumor cluster spots, we got the CNV estimation of tumor spots. To reduce the impact of dropout, we took advantage of the spots' SNN relationships and smoothed each spot's CNVs further by calculating the weighted average of it and its SNNs (*23*).

To confirm the CNV results inferred from ST data, we also performed the bulk WES on the PBMCs, normal sections, T sections, and normal/tumor regions of the L sections of the corresponding patients. Then, the CNVs of the tumor bulk samples were called from the paired tumor-normal WES data using CNVkit software (stable version) (*41*). The normal reference adopting the PBMC or the normal section data can generate similar results. Besides, for the derived $\log_2$ copy-ratio results, the outliers were detected and filtered.

CSC-niche analysis

For each of the five common stemness genes (*CD47*, *EPCAM*, *KRT19*, *PROM1*, and *SOX9*), we tested different thresholds (3, 5, 8, and 10%) to define the CSC-niches and found that the variation trend of the CSC-niche fractions across L/T/P sections was consistent under different thresholds. Here, we used the top 5% spots with the highest expression in each patient's tumor region to define the gene's positive CSC-niches. Then, to analyze the distribution characteristics of these CSC-niches, we compared the proportions of these niches in the tumor spots of different sections (Fig. 6A), and the overlap between these five types of niches in patient's different sections (Fig. 6C).

Furthermore, to measure the cell type enrichments and hallmark pathway activities in these CSC-niches, we picked the spots belonging to only one of the stemness genes' positive CSC-niches and defined them as five stemness groups. At the same time, we scored the remaining spots by the mean of the above five stemness genes and selected the 50% of them with the lowest scores as a background group. Then, we randomly sampled spots from the background group for 100 times and calculated the means (μ_k, $k = 1\ldots100$) and SDs (σ_k) of their cell type scores or GSVA pathway scores ($\mathbf{x}_{\text{bg}}^{(k)}$), respectively. Using them, the cell type or pathway scores (x_{stem}) of the spot in each stemness group can be transformed into z-score formation

$$z_{\text{stem}} = \frac{1}{100}\sum_{k=1}^{100}\frac{x_{\text{stem}} - \mu_k}{\sigma_k}$$

These z scores measured the relative cell type enrichments or pathway activities of different types of CSC-niches than the nonstemness spots (background).

TLS analysis

We defined the TLS-50 signature as the top 50 positive differential expression genes between cluster 6 and the remaining clusters in HCC-5, which were ordered by the average log fold change (fig. S10A). Using this signature, we scored each spot in the other 17 ST samples with the average expression (log-transformed normalization values) of these 50 genes and selected putative TLS spots based on the scores, where the cutoff was set as 0.6 times the maximum of the scores in the sample. To eliminate the impact of outliers, we discarded the spots that were isolated single points. At the same time, the TLS spots' identities were further checked on H&E images by pathologists.

To analyze the abundance of cell types in TLS spots, we compared the cell type scores between the identified TLS spots and the stromal background spots in all L sections and HCC-5 and applied Wilcoxon rank sum test to measure the statistical significance (Fig. 7C and fig. S13B). Besides, to explore whether there were potential relationships between the TLS spots' characteristics and their distance to tumor regions, we calculated the spatial distances from each TLS spot to the tumor regions in L sections and performed Pearson correlation analyses on these distances and the cell type scores or gene expression levels of TLS spots (Fig. 7D and fig. S14). Furthermore, to checked the gradient changes around TLS spots, we selected the TLS spots' surrounding regions with five spots wide and analyzed the relationship between their gene expression levels (only kept the values greater than 0) and the distances to TLS spots by Pearson correlation analysis (Fig. 7E and fig. S15, A and B).

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at https://science.org/doi/10.1126/sciadv.abg3750

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES

1. I. Dagogo-Jack, A. T. Shaw, Tumour heterogeneity and resistance to cancer therapies. *Nat. Rev. Clin. Oncol.* **15**, 81–94 (2018).
2. D. A. Lawson, K. Kessenbrock, R. T. Davis, N. Pervolarakis, Z. Werb, Tumour heterogeneity and metastasis at single-cell resolution. *Nat. Cell Biol.* **20**, 1349–1360 (2018).
3. C. Zheng, L. Zheng, J. K. Yoo, H. Guo, Y. Zhang, X. Guo, B. Kang, R. Hu, J. Y. Huang, Q. Zhang, Z. Liu, M. Dong, X. Hu, W. Ouyang, J. Peng, Z. Zhang, Landscape of infiltrating T cells in liver cancer revealed by single-cell sequencing. *Cell* **169**, 1342–1356.e16 (2017).
4. 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$^+$GPR77$^+$ cancer-associated fibroblasts promote cancer formation and chemoresistance by sustaining cancer stemness. *Cell* **172**, 841–856.e816 (2018).
5. K. H. Chen, A. N. Boettiger, J. R. Moffitt, S. Wang, X. Zhuang, RNA imaging. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* **348**, aaa6090 (2015).
6. C. L. Eng, M. Lawson, Q. Zhu, R. Dries, N. Koulena, Y. Takei, J. Yun, C. Cronin, C. Karp, G. C. Yuan, L. Cai, Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH. *Nature* **568**, 235–239 (2019).
7. S. Vickovic, G. Eraslan, F. Salmén, J. Klughammer, L. Stenbeck, D. Schapiro, T. Äijö, R. Bonneau, L. Bergenstråhle, J. F. Navarro, J. Gould, G. K. Griffin, Å. Borg, M. Ronaghi, J. Frisén, J. Lundeberg, A. Regev, P. L. Ståhl, High-definition spatial transcriptomics for in situ tissue profiling. *Nat. Methods* **16**, 987–990 (2019).
8. P. L. Ståhl, F. Salmén, S. Vickovic, A. Lundmark, J. F. Navarro, J. Magnusson, S. Giacomello, M. Asp, J. O. Westholm, M. Huss, A. Mollbrink, S. Linnarsson, S. Codeluppi, A. Borg, F. Ponten, P. I. Costea, P. Sahlén, J. Mulder, O. Bergmann, J. Lundeberg, J. Frisén, Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* **353**, 78–82 (2016).
9. E. Berglund, J. Maaskola, N. Schultz, S. Friedrich, M. Marklund, J. Bergenstråhle, F. Tarish, A. Tanoglidi, S. Vickovic, L. Larsson, F. Salmén, C. Ogris, K. Wallenborg, J. Lagergren, P. Ståhl, E. Sonnhammer, T. Helleday, J. Lundeberg, Spatial maps of prostate cancer transcriptomes reveal an unexplored landscape of heterogeneity. *Nat. Commun.* **9**, 2419 (2018).
10. R. Moncada, D. Barkley, F. Wagner, M. Chiodin, J. C. Devlin, M. Baron, C. H. Hajdu, D. M. Simeone, I. Yanai, Integrating microarray-based spatial transcriptomics and single-cell RNA-seq reveals tissue architecture in pancreatic ductal adenocarcinomas. *Nat. Biotechnol.* **38**, 333–342 (2020).
11. D. Sia, A. Villanueva, S. L. Friedman, J. M. Llovet, Liver cancer cell of origin, molecular class, and effects on patient prognosis. *Gastroenterology* **152**, 745–761 (2017).
12. B. Losic, A. J. Craig, C. Villacorta-Martin, S. N. Martins-Filho, N. Akers, X. Chen, M. E. Ahsen, J. von Felden, I. Labgaa, D. D'Avola, K. Allette, S. A. Lira, G. C. Furtado, T. Garcia-Lezana, P. Restrepo, A. Stueck, S. C. Ward, M. I. Fiel, S. P. Hiotis, G. Gunasekaran, D. Sia, E. E. Schadt,

R. Sebra, M. Schwartz, J. M. Llovet, S. Thung, G. Stolovitzky, A. Villanueva, Intratumoral heterogeneity and clonal evolution in liver cancer. *Nat. Commun.* **11**, 291 (2020).

13. T. F. Greten, X. W. Wang, F. Korangy, Current concepts of immune based treatments for patients with HCC: From basic science to novel treatment approaches. *Gut* **64**, 842–848 (2015).
14. J. M. Llovet, S. Ricci, V. Mazzaferro, P. Hilgard, E. Gane, J.-F. Blanc, A. C. de Oliveira, A. Santoro, J.-L. Raoul, A. Forner, M. Schwartz, C. Porta, S. Zeuzem, L. Bolondi, T. F. Greten, P. R. Galle, J.-F. Seitz, I. Borbath, D. Häussinger, T. Giannaris, M. Shan, M. Moscovici, D. Voliotis, J. Bruix; SHARP Investigators Study Group, CollectiveName, sorafenib in advanced hepatocellular carcinoma. *N. Engl. J. Med.* **359**, 378–390 (2008).
15. A. X. Zhu, R. S. Finn, J. Edeline, S. Cattan, S. Ogasawara, D. Palmer, C. Verslype, V. Zagonel, L. Fartoux, A. Vogel, D. Sarker, G. Verset, S. L. Chan, J. Knox, B. Daniele, A. L. Webber, S. W. Ebbinghaus, J. Ma, A. B. Siegel, A. L. Cheng, M. Kudo; K.-i. CollectiveName, Pembrolizumab in patients with advanced hepatocellular carcinoma previously treated with sorafenib (KEYNOTE-224): A non-randomised, open-label phase 2 trial. *Lancet Oncol.* **19**, 940–952 (2018).
16. X. Ding, M. He, A. W. H. Chan, Q. X. Song, S. C. Sze, H. Chen, M. K. H. Man, K. Man, S. L. Chan, P. B. S. Lai, X. Wang, N. Wong, Genomic and epigenomic features of primary and recurrent hepatocellular carcinomas. *Gastroenterology* **157**, 1630–1645.e1636 (2019).
17. M. D. Lynch, F. M. Watt, Fibroblast heterogeneity: Implications for human disease. *J. Clin. Invest.* **128**, 26–35 (2018).
18. A. Sharma, J. J. W. Seow, C. A. Dutertre, R. Pai, C. Blériot, A. Mishra, R. M. M. Wong, G. S. N. Singh, S. Sudhagar, S. Khalilnezhad, S. Erdal, H. M. Teo, A. Khalilnezhad, S. Chakarov, T. K. H. Lim, A. C. Y. Fui, A. K. W. Chieh, C. P. Chung, G. K. Bonney, B. K. Goh, J. K. Y. Chan, P. K. H. Chow, F. Ginhoux, R. DasGupta, Onco-fetal reprogramming of endothelial cells drives immunosuppressive macrophages in hepatocellular carcinoma. *Cell* **183**, 377–394.e21 (2020).
19. B. Zheng, D. Wang, X. Qiu, G. Luo, T. Wu, S. Yang, Z. Li, Y. Zhu, S. Wang, R. Wu, C. Sui, Z. Gu, S. Shen, S. Jeong, X. Wu, J. Gu, H. Wang, L. Chen, Trajectory and functional analysis of PD-1^{high}CD4^{+}CD8^{+}T cells in hepatocellular carcinoma by single-cell cytometry and transcriptome sequencing. *Adv. Sci. (Weinh)* **7**, 2000224 (2020).
20. C. Y. Lin, J. Lovén, P. B. Rahl, R. M. Paranal, C. B. Burge, J. E. Bradner, T. I. Lee, R. A. Young, Transcriptional amplification in tumor cells with elevated c-Myc. *Cell* **151**, 56–67 (2012).
21. I. Korsunsky, N. Millard, J. Fan, K. Slowikowski, F. Zhang, K. Wei, Y. Baglaenko, M. Brenner, P. R. Loh, S. Raychaudhuri, Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).
22. T. Stuart, A. Butler, P. Hoffman, C. Hafemeister, E. Papalexi, W. M. Mauck III, Y. Hao, M. Stoeckius, P. Smibert, R. Satija, Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902.e21 (2019).
23. W. Guo, D. Wang, S. Wang, Y. Shan, C. Liu, J. Gu, scCancer: A package for automated processing of single-cell RNA-seq data in cancer. *Brief. Bioinform.* **22**, bbaa127 (2021).
24. P. Angerer, L. Haghverdi, M. Buttner, F. J. Theis, C. Marr, F. Buettner, Destiny: Diffusion maps for large-scale single-cell data in R. *Bioinformatics* **32**, 1241–1243 (2016).
25. D. Aran, Z. Hu, A. J. Butte, xCell: Digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol.* **18**, 220 (2017).
26. S. Hänzelmann, R. Castelo, J. Guinney, GSVA: Gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics* **14**, 7 (2013).
27. E. Gaude, C. Frezza, Tissue-specific and convergent metabolic transformation of cancer correlates with metastatic potential and patient survival. *Nat. Commun.* **7**, 13041 (2016).
28. Q. Lian, S. Wang, G. Zhang, D. Wang, G. Luo, J. Tang, L. Chen, J. Gu, HCCDB: A database of hepatocellular carcinoma expression atlas. *Genom. Proteom. Bioinf.* **16**, 269–275 (2018).
29. S. Roessler, H. L. Jia, A. Budhu, M. Forgues, Q. H. Ye, J. S. Lee, S. S. Thorgeirsson, Z. Sun, Z. Y. Tang, L. X. Qin, X. W. Wang, A unique metastasis gene signature enables prediction of tumor relapse in early-stage hepatocellular carcinoma patients. *Cancer Res.* **70**, 10202–10212 (2010).
30. M. Efremova, M. Vento-Tormo, S. A. Teichmann, R. Vento-Tormo, CellPhoneDB: Inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. *Nat. Protoc.* **15**, 1484–1506 (2020).
31. R. Vento-Tormo, M. Efremova, R. A. Botting, M. Y. Turco, M. Vento-Tormo, K. B. Meyer, J. E. Park, E. Stephenson, K. Polański, 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. Lisgo, 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).
32. M. Bry, R. Kivelä, V. M. Leppänen, K. Alitalo, Vascular endothelial growth factor-B in physiology and disease. *Physiol. Rev.* **94**, 779–794 (2014).
33. S. Soker, S. Takashima, H. Q. Miao, G. Neufeld, M. Klagsbrun, Neuropilin-1 is expressed by endothelial and tumor cells as an isoform-specific receptor for vascular endothelial growth factor. *Cell* **92**, 735–745 (1998).
34. L. Seetharam, N. Gotoh, Y. Maru, G. Neufeld, S. Yamaguchi, M. Shibuya, A unique signal transduction from FLT tyrosine kinase, a receptor for vascular endothelial growth factor VEGF. *Oncogene* **10**, 135–147 (1995).
35. B. D. Ferguson, R. Liu, C. E. Rolle, Y. H. Tan, V. Krasnoperov, R. Kanteti, M. S. Tretiakova, G. M. Cervantes, R. Hasina, R. D. Hseu, A. J. Iafrate, T. Karrison, M. K. Ferguson, A. N. Husain, L. Faoro, E. E. Vokes, P. S. Gill, R. Salgia, The EphB4 receptor tyrosine kinase promotes lung cancer growth: A potential novel therapeutic target. *PLOS ONE* **8**, e67668 (2013).
36. M. Pietilä, P. Sahgal, E. Peuhu, N. Z. Jäntti, I. Paatero, E. Närvä, H. Al-Akhrass, J. Lilja, M. Georgiadou, O. M. Andersen, A. Padzik, H. Sihto, H. Joensuu, M. Blomqvist, I. Saarinen, P. J. Bostrom, P. Taimen, J. Ivaska, SORLA regulates endosomal trafficking and oncogenic fitness of HER2. *Nat. Commun.* **10**, 2340 (2019).
37. K. Ieguchi, Y. Maru, Roles of EphA1/A2 and ephrin-A1 in cancer. *Cancer Sci.* **110**, 841–848 (2019).
38. A. Kinoshita, T. Shah, M. M. Tangredi, D. K. Strickland, B. T. Hyman, The intracellular domain of the low density lipoprotein receptor-related protein modulates transactivation mediated by amyloid precursor protein and Fe65. *J. Biol. Chem.* **278**, 41182–41188 (2003).
39. H. Mao, P. Lockyer, L. Li, C. M. Ballantyne, C. Patterson, L. Xie, X. Pi, Endothelial LRP1 regulates metabolic responses by acting as a co-activator of PPARγ. *Nat. Commun.* **8**, 14960 (2017).
40. A. P. Patel, I. Tirosh, J. J. Trombetta, A. K. Shalek, S. M. Gillespie, H. Wakimoto, D. P. Cahill, B. V. Nahed, W. T. Curry, R. L. Martuza, D. N. Louis, O. Rozenblatt-Rosen, M. L. Suvà, A. Regev, B. E. Bernstein, Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* **344**, 1396–1401 (2014).
41. E. Talevich, A. H. Shain, T. Botton, B. C. Bastian, CNVkit: Genome-wide copy number detection and visualization from targeted DNA sequencing. *PLoS Comput. Biol.* **12**, e1004873 (2016).
42. T. Oikawa, Cancer stem cells and their cellular origins in primary liver and biliary tract cancers. *Hepatology* **64**, 645–651 (2016).
43. H. P. Zheng, Y. Pomyen, M. O. Hernandez, C. Y. Li, F. Livak, W. Tang, H. E. Dang, T. F. Greten, J. L. Davis, Y. M. Zhao, M. Mehta, Y. Levin, J. Shetty, B. Tran, A. Budhu, X. W. Wang, Single-cell analysis reveals cancer stem cell heterogeneity in hepatocellular carcinoma. *Hepatology* **68**, 127–140 (2018).
44. M.-C. Dieu-Nosjean, N. A. Giraldo, H. Kaplon, C. Germain, W. H. Fridman, C. Sautès-Fridman, Tertiary lymphoid structures, drivers of the anti-tumor responses in human cancers. *Immunol. Rev.* **271**, 260–275 (2016).
45. C. Sautès-Fridman, F. Petitprez, J. Calderaro, W. H. Fridman, Tertiary lymphoid structures in the era of cancer immunotherapy. *Nat. Rev. Cancer* **19**, 307–325 (2019).
46. A. Andersson, L. Larsson, L. Stenbeck, F. Salmén, A. Ehinger, S. Z. Wu, G. Al-Eryani, D. Roden, A. Swarbrick, Å. Borg, J. Frisén, C. Engblom, J. Lundeberg, Spatial deconvolution of HER2-positive breast cancer delineates tumor-associated cell type interactions. *Nat. Commun.* **12**, 6012 (2021).
47. C. M. Schürch, S. S. Bhate, G. L. Barlow, D. J. Phillips, L. Noti, I. Zlobec, P. Chu, S. Black, J. Demeter, D. R. McIlwain, S. Kinoshita, N. Samusik, Y. Goltsev, G. P. Nolan, Coordinated cellular neighborhoods orchestrate antitumoral immunity at the colorectal cancer invasive front. *Cell* **183**, 838 (2020).
48. V. H. Engelhard, A. B. Rodriguez, I. S. Mauldin, A. N. Woods, J. D. Peske, C. L. Slingluff, Immune cell infiltration and tertiary lymphoid structures as determinants of antitumor Immunity. *J. Immunol.* **200**, 432–442 (2018).
50. T. S. Weber , Cell cycle-associated CXCR4 expression in germinal center B cells and its implications on affinity maturation. *Front. Immunol.* **9**, 1313 (2018).
49. E. M. Pimenta, B. J. Barnes, Role of tertiary lymphoid structures (TLS) in anti-tumor immunity: Potential tumor-Induced cytokines/chemokines that regulate TLS formation in epithelial-derived cancers. *Cancers* **6**, 969–997 (2014).
51. P. M. Maciocia, P. A. Wawrzyniecka, B. Philip, I. Ricciardelli, A. U. Akarca, S. C. Onuoha, M. Legut, D. K. Cole, A. K. Sewell, G. Gritti, J. Somja, M. A. Piris, K. S. Peggs, D. C. Linch, T. Marafioti, M. A. Pule, Targeting the T cell receptor β-chain constant region for immunotherapy of T cell malignancies. *Nat. Med.* **23**, 1416–1423 (2017).
52. D. F. Quail, J. A. Joyce, Microenvironmental regulation of tumor progression and metastasis. *Nat. Med.* **19**, 1423–1437 (2013).
53. Z. F. Lim, P. C. Ma, Emerging insights of tumor heterogeneity and drug resistance mechanisms in lung cancer targeted therapy. *J. Hematol. Oncol.* **12**, 134 (2019).
54. R. Rahmanzade, Redefinition of tumor capsule: Rho-dependent clustering of cancer-associated fibroblasts in favor of tensional homeostasis. *Med. Hypotheses* **135**, 109425 (2020).
55. A. J. Craig, J. von Felden, T. Garcia-Lezana, S. Sarcognato, A. Villanueva, Tumour evolution in hepatocellular carcinoma. *Nat. Rev. Gastroenterol. Hepatol.* **17**, 139–152 (2020).
56. M. Najafi, B. Farhood, K. Mortezaee, Cancer stem cells (CSCs) in cancer progression and therapy. *J. Cell. Physiol.* **234**, 8381–8395 (2019).
57. M. Gururajan, A. Simmons, T. Dasu, B. T. Spear, C. Calulot, D. A. Robertson, D. L. Wiest, J. G. Monroe, S. Bondada, Early growth response genes regulate B cell development, proliferation, and immune response. *J. Immunol.* **181**, 4590–4602 (2008).
58. R. Lang, M. Hammer, J. Mages, DUSP meet immunology: Dual specificity MAPK phosphatases in control of the inflammatory response. *J. Immunol.* **177**, 7497–7504 (2006).

59. P. Gaudin, C. Trocmé, S. Berthier, S. Kieffer, J. Boutonnat, C. Lamy, A. Surla, J. Garin, F. Morel, TIMP-1/MMP-9 imbalance in an EBV-immortalized B lymphocyte cellular model: Evidence for TIMP-1 multifunctional properties. *Biochim. Biophys. Acta* **1499**, 19–33 (2000).
60. Y. Subbannayya, S. M. Pinto, K. Bösl, T. S. K. Prasad, R. K. Kandasamy, Dynamics of dual specificity phosphatases and their interplay with protein kinases in immune signaling. *Int. J. Mol. Sci.* **20**, 2086 (2019).

Acknowledgments: We acknowledge the members of the International Co-operation Laboratory on Signal Transduction and Genergy Biotechnology (Shanghai) Co. Ltd. for excellent technical assistance. We thank the operating room of EHBH for providing human tumor specimens. Funding: This work was supported by the National Key Research and Development Program of China (2017YFA0505803, 2017YFC0908102, and 2020YFA0712403), the State Key Project for Liver Cancer (2018ZX10732202, 2018ZX10302207, and 2018ZX10301202), National Natural Science Foundation of China (81790633, 61922047, 81830045, 61721003, and 81902412), and the National Natural Science Foundation of Shanghai (21XD1404600). Author contributions: R.W., W.G., and X.Q. contributed equally and developed the concept and discussed experiments. R.W. and X.Q. performed all experiments and wrote the manuscript. W.G. performed all data analysis and wrote the manuscript. Shicheng Wang performed WES data preprocessing. C.S. and Z.Y. provided human specimens and clinical information. Q.L., J.W., Y.S., S.Y., T.W., K.W., Y. Zhu, S. Wang, C.L., Yangqianwen Zhang, B.Z., Z.L., Yani Zhang, S.S., Y. Zhao, W.W., J.B., J.H., X.W., and X.J. provided technical assistance. H.W., J.G., and L.C. designed research, supervised the study, guided the discussion, and revised the manuscript. Competing interests: The authors declare that they have no competing interests. Data and materials availability: The accession number for the raw sequencing data deposited in Genome Sequence Archive (GSA) is HRA000437. Details of the ST and WES data can be found at https://ngdc.cncb.ac.cn/gsa-human/browse/HRA000437. The major steps for processing the ST have been integrated as an R package stCancer (http://lifeome.net/software/stcancer). All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials.

Submitted 1 January 2021
Accepted 3 November 2021
Published 17 December 2021
10.1126/sciadv.abg3750