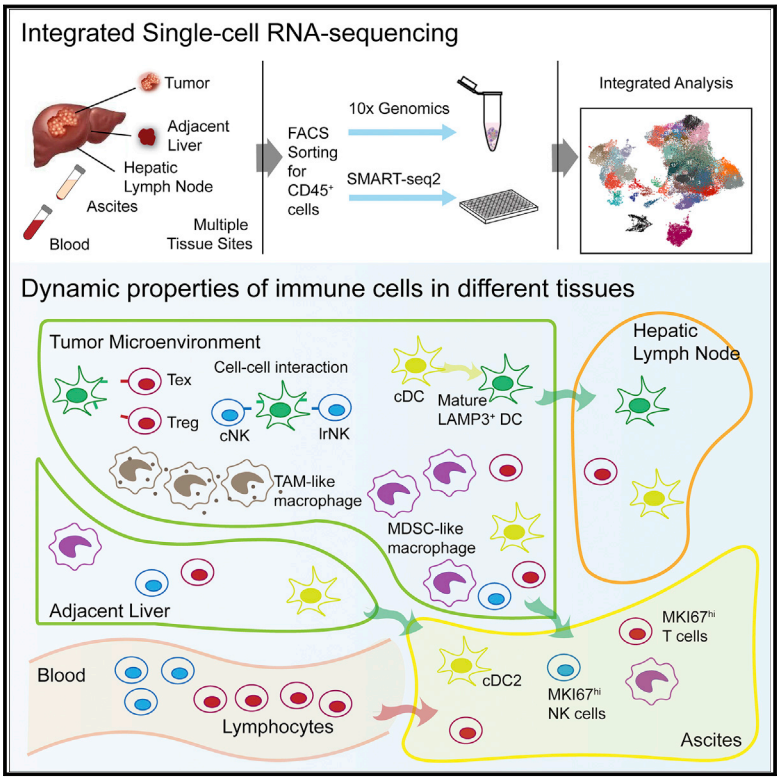
Article

Cell

Landscape and Dynamics of Single Immune Cells in Hepatocellular Carcinoma

Graphical Abstract



Highlights

- Two scRNA-seq platforms reveal a high-resolution dynamic immune landscape in HCC
- *LAMP3*$^{+}$ DCs, arising from cDCs, can migrate from tumors to hepatic lymph nodes
- Paired ligand-receptor analyses implicate the regulation of lymphocytes by *LAMP3*$^{+}$ DCs
- Macrophage subsets in tumors show distinct states and potentials to egress to ascites

Authors

Qiming Zhang, Yao He, Nan Luo, ..., Jirun Peng, Xianwen Ren, Zemin Zhang

Correspondence

kang_3.liu@boehringer-ingelheim.com (K.L.), pengjr@medmail.com.cn (J.P.), renxwise@pku.edu.cn (X.R.), zemin@pku.edu.cn (Z.Z.)

In Brief

Integrated single-cell RNA sequencing technologies and bioinformatics approaches reveal a high-resolution immune landscape of hepatocellular carcinoma, identifying inflammatory signatures and functional states of myeloid cells as well as predictions of complex cell-cell interactions.

Zhang et al., 2019, Cell *179*, 829–845
October 31, 2019
https://doi.org/10.1016/j.cell.2019.10.003

CellPress

Cell

Landscape and Dynamics of Single Immune Cells in Hepatocellular Carcinoma

Qiming Zhang,[1,14] Yao He,[2,14] Nan Luo,[3,4,14] Shashank J. Patel,[5] Yanjie Han,[1] Ranran Gao,[1] Madhura Modak,[6] Sebastian Carotta,[7] Christian Haslinger,[8] David Kind,[8] Gregory W. Peet,[5] Guojie Zhong,[1] Shuangjia Lu,[1] Weihua Zhu,[9] Yilei Mao,[10] Mengmeng Xiao,[11] Michael Bergmann,[12] Xueda Hu,[1] Sid P. Kerkar,[5] Anne B. Vogt,[13] Stefan Pflanz,[6] Kang Liu,[5,*] Jirun Peng,[3,4,*] Xianwen Ren,[1,*] and Zemin Zhang[1,2,15,*]

[1]BIOPIC, Beijing Advanced Innovation Center for Genomics, School of Life Sciences, Peking University, Beijing 100871, China
[2]Peking-Tsinghua Center for Life Sciences, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing 100871, China
[3]Department of Surgery, Beijing Shijitan Hospital, Capital Medical University, Beijing 100038, China
[4]Ninth School of Clinical Medicine, Peking University, Beijing 100038, China
[5]Department of Cancer Immunology and Immune Modulation, Boehringer Ingelheim Pharmaceuticals, Inc., 900 Ridgebury Road, Ridgefield, CT 06877, USA
[6]Department of Cancer Immunology and Immune Modulation, Boehringer Ingelheim Pharma, Birkendorfer Str. 65, 88400 Biberach, Germany
[7]Department of Cancer Cell Signaling, Boehringer Ingelheim RCV GmBH & Co KG., Dr. Boehringer Gasse 5–11, 1120 Vienna, Austria
[8]Department of Computational Biology and Genomics, Boehringer Ingelheim Pharma, Birkendorfer Str. 65, 88400 Biberach, Germany
[9]Department of Hepatobiliary Surgery, Peking University People's Hospital, Beijing 100044, China
[10]Department of Liver Surgery, Peking Union Medical College Hospital, Beijing 100730, China
[11]Department of Retroperitoneal Tumor Surgery, Peking University International Hospital, Beijing 102206, China
[12]Department of Surgery and Comprehensive Cancer Center, Medical University of Vienna, Vienna, Austria
[13]Department of Human Cancer Immunology, Boehringer Ingelheim RCV GmBH & Co KG., Dr. Boehringer Gasse 5–11, 1120 Vienna, Austria
[14]These authors contributed equally
[15]Lead Contact
*Correspondence: kang_3.liu@boehringer-ingelheim.com (K.L.), pengjr@medmail.com.cn (J.P.), renxwise@pku.edu.cn (X.R.), zemin@pku.edu.cn (Z.Z.)
https://doi.org/10.1016/j.cell.2019.10.003

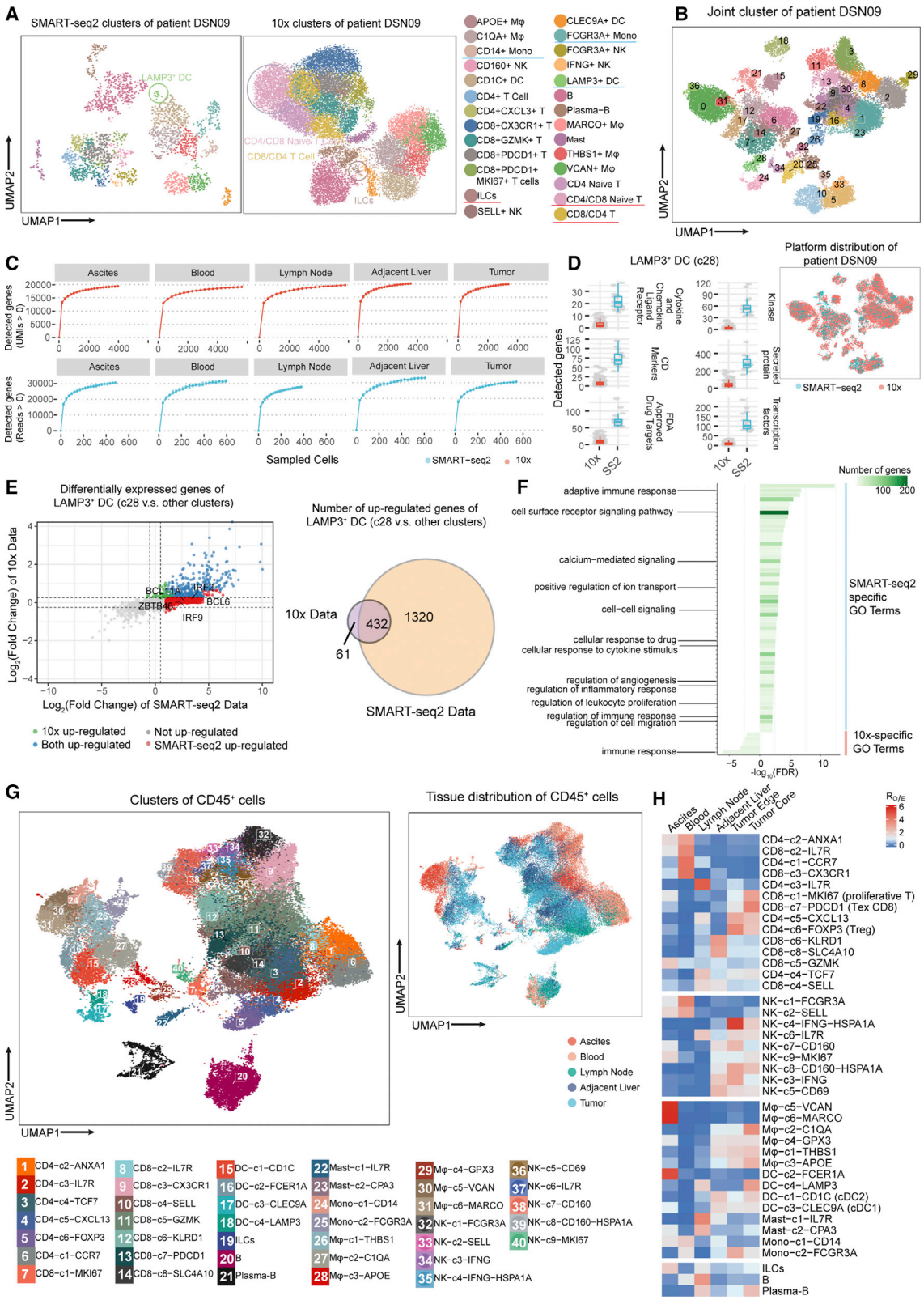
SUMMARY

The immune microenvironment of hepatocellular carcinoma (HCC) is poorly characterized. Combining two single-cell RNA sequencing technologies, we produced transcriptomes of CD45$^+$ immune cells for HCC patients from five immune-relevant sites: tumor, adjacent liver, hepatic lymph node (LN), blood, and ascites. A cluster of *LAMP3*$^+$ dendritic cells (DCs) appeared to be the mature form of conventional DCs and possessed the potential to migrate from tumors to LNs. *LAMP3*$^+$ DCs also expressed diverse immune-relevant ligands and exhibited potential to regulate multiple subtypes of lymphocytes. Of the macrophages in tumors that exhibited distinct transcriptional states, tumor-associated macrophages (TAMs) were associated with poor prognosis, and we established the inflammatory role of *SLC40A1* and *GPNMB* in these cells. Further, myeloid and lymphoid cells in ascites were predominantly linked to tumor and blood origins, respectively. The dynamic properties of diverse CD45$^+$ cell types revealed by this study add new dimensions to the immune landscape of HCC.

INTRODUCTION

Liver cancer is the third leading cause of cancer-related mortality in the world (Forner et al., 2018), and hepatocellular carcinoma (HCC) accounts for approximately 90% of the incidence of all liver cancers (Bray et al., 2018). Although treatments with sorafenib and regorafenib lead to a modest survival benefit, overall anti-tumor responses are still limited (Llovet et al., 2008; Ray, 2017). Although immunotherapies have clinical benefits for other cancer indications, the response rates in HCC are much lower (El-Khoueiry et al., 2017). Because parameters of the immune contexture have been associated with treatment efficacy (Gnjatic et al., 2017), it is important to characterize the baseline HCC immune milieu to clarify the composition and property of tumor-infiltrating immune cells in comparison with ones in other immune-relevant anatomical compartments.

The cellular components of the tumor microenvironment (TME) are highly complex, with diverse populations of myeloid cells and lymphocytes playing important roles in inflammation, cancer immune evasion, and responses to immunotherapy treatment (Hackl et al., 2016; Ringelhan et al., 2018). The presence of myeloid cells in the TME is often linked to altered patient survival (Engblom et al., 2016). Tumor-associated macrophages (TAMs) have been reported to prevent T cells from recognizing and killing cancer cells (Peranzoni et al., 2018; Engblom et al., 2016; Georgoudaki et al., 2016). Conventional dendritic cell (DC) subsets (cDC1 and cDC2) have been reported to migrate into tumor-draining LN and prime CD8$^+$ or CD4$^+$ T cells in mouse models (Binnewies et al., 2019; Salmon et al., 2016). However, the characteristics and functions of TAM and DC subsets in HCC patients are still poorly understood. We have previously characterized T cells in HCC by single-cell RNA sequencing (scRNA-seq) (Zheng et al., 2017a), but the global immune landscape is still unknown. Recently, studies characterizing immune



(legend on next page)

cell subsets of the TME have emerged, including mass cytometry studies (Chevrier et al., 2017; Lavin et al., 2017; Wagner et al., 2019) and scRNA-seq studies on melanoma (Li et al., 2019; Tirosh et al., 2016), non-small-cell lung carcinoma (Lambrechts et al., 2018), breast carcinoma (Azizi et al., 2018), head and neck cancer (Puram et al., 2017), colorectal cancer (Li et al., 2017), and kidney cancer (Chung et al., 2017), paving the way for such studies for HCC.

The scRNA-seq technologies in the aforementioned studies were based on either SMART-seq2 (Picelli et al., 2014) or droplet-based platforms (Zheng et al., 2017b). SMART-seq2 provides in-depth coverage for a smaller number of cells, whereas the droplet-based platforms capture a larger scale of cells but with the sacrifice of limited gene coverage (Ziegenhain et al., 2017). Therefore, combined analysis of these two technologies may represent a preferred approach toward achieving an in-depth understanding of the broad landscape. Here we combined full-length and 3′ scRNA-seq technologies to obtain high-quality data for a large collection of CD45$^+$ immune cells from various tissues in HCC patients. For the first time, cell composition, functional states, developmental trajectory, and cellular interactions of immune cells in HCC tumors, adjacent liver, hepatic lymph nodes (LNs), ascites, and blood were explored and compared systematically. Specifically, previously uncharacterized hepatic LNs and ascites were included to dissect the dynamics of immune cells in HCC. Excess accumulation of fluid in the peritoneal cavity, known as ascites (Biecker, 2011; Fukui et al., 2018), is associated with poor prognosis of HCC (Hsu et al., 2013), but the roles of ascites in shaping the tumor immune contexture are still unknown. The combination of the two scRNA-seq technologies and the inclusion of LNs and ascites allowed us to not only characterize the immune composition of HCC with high resolution but also to trace their dynamics.

RESULTS

High-Resolution Immune Landscape of HCC by Integrated Analysis of Full-Length and 3′ scRNA-Seq

To characterize the immune cells in HCC, we applied scRNA-seq methods to study CD45$^+$ cells isolated from tumors and four immune-relevant sites (adjacent liver, hepatic LNs, blood, and ascites) of 16 treatment-naive liver cancer patients (Figures S1A–S1C; Tables S1 and S2). For patient DSN09, both 10x Genomics and SMART-seq2 methods were applied in parallel, giving us the opportunity to evaluate the power of integrated analysis on two types of datasets. Using graph-based clustering (Traag et al., 2019) to analyze the cells from this patient, we identified 20 clusters for 10x data and 22 for SMART-seq2 (Figure 1A). Examination of canonical marker genes revealed major cell populations, including T cells, natural killer (NK) cells, and diverse myeloid-lineage cells in both platforms (Figures S1D and S1E), demonstrating the stability and accuracy of our data. Minor cell populations, though, varied between the two platforms. For instance, the *LAMP3*$^+$ DC, *CD14*$^+$, and *FCGR3A*$^+$ monocyte groups were only identified in SMART-seq2, whereas type 3 innate lymphoid cells (ILCs) were only captured by 10x (Figure 1A). In addition, we observed that SMART-seq2 could help distinguish closely related clusters, possibly by capturing more RNA molecules that contribute to cell type classification (Svensson et al., 2017). For instance, *CD4*$^+$ and *CD8*$^+$ T cell subtypes (*CD8*$^+$*PDCD1*$^+$, *CD8*$^+$*GZMK*$^+$, and *CD8*$^+$*CX3CR1*$^+$ clusters) were readily separated. In contrast, 10x-based T cell clusters were composed of a mixture of *CD4*$^+$ and *CD8*$^+$ T cells, as evidenced by their simultaneous similarities to those defined by SMART-seq2 (Figure S2A). We then integrated the two types of datasets using the Harmony algorithm (Korsunsky et al., 2018). In total, 37 clusters were identified on the combined dataset, with 36 containing cells from both 10x and SMART-seq2, indicating a well-integrated dataset (Figure 2B). Clusters 16, 29, 33, 34, 35, and 36 were dominated by 10x, each containing less than 10 cells from SMART-seq2 (Figure S2B), indicating that such rare cell populations were ''rescued'' by 10x. On the other hand, 129 individual cells initially in the *CLEC9A*$^+$ DCs of 10x clusters were re-assigned to *LAMP3*$^+$ DCs after integration, confirmed by their gene expression (Figure S2C), suggesting that the integration led to identification of rare cell populations that were masked in either platform alone. We also observed that the initial mixture of *CD4*$^+$/*CD8*$^+$ T cells in 10x clusters were subsequently separated after incorporation with SMART-seq2 data (Figure S2B).

We then examined the gene detection sensitivity. At the cluster level, the gene numbers ($\geq$1 read) detected by SMART-seq2 were invariably more numerous than 10x and covered almost all genes detected by 10x (Figure S2D). Saturation analysis revealed that the detected gene number by SMART-seq2 was

Figure 1. Landscape of CD45$^+$ Immune Cells in Liver Cancer

(A) Uniform Manifold Approximation and Projection (UMAP) of SMART-seq2-based (left) and 10x-based (right) single CD45$^+$ cells from patient DSN09. The red line represents 10x-specific cluster, and blue represents SMART-seq2-specific cluster.

(B) UMAP plot showing 37 clusters identified by integrated analysis, colored by cell cluster (top) or platform (bottom).

(C) Saturation curve showing the detected gene number of 10x (top) or SMART-seq2 (bottom). Each dot, colored by platform, represents the detected gene number (y axis) with the given cell number (x axis).

(D) Boxplot showing the detected gene number in specific gene categories of *LAMP3*$^+$ DCs. Each dot represents an individual cell. $p < 2e{-}16$ in all boxplots, Wilcoxon test.

(E) Dot plot showing the comparison of differentially expressed genes (*LAMP3*$^+$ DCs versus others, false discovery rate [FDR] < 0.01, $\log_2$(fold change [FC]) > 0.5) of *LAMP3*$^+$ DCs detected by 10x and SMART-seq2. The x axis and y axis show the $\log_2$(FC) of the detected genes in SMART-seq2 and 10x, respectively.

(F) Enriched GO terms of SMART-seq2- and 10x-specific genes generated in (E).

(G) UMAP projection showing the immune landscape of liver cancer, colored by cluster (left) and tissue (right).

(H) Tissue preference of each cluster estimated by *Ro/e* based on 10x data.

See also Figures S1, S2, S3, and S4 and Tables S1, S2, and S3.

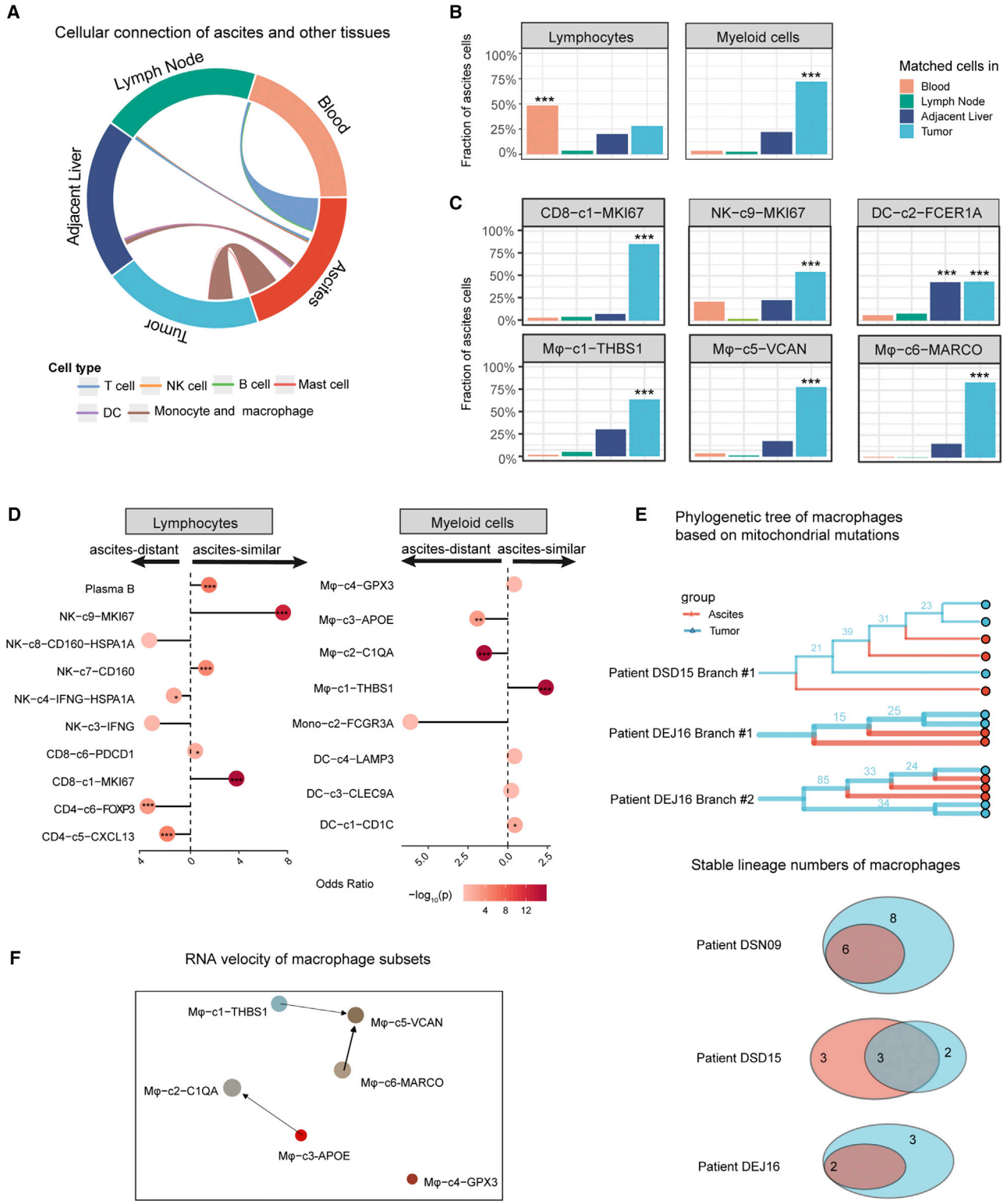


Figure 2. Cellular Relationship of Ascites and Other Tissues

(A) Circos plot showing the proportion of ascites cells matched to cells in other tissues based on expression similarity analysis.

(B) Bar plot showing the cell fraction of lymphocytes and myeloid cells in ascites that aligned to different tissues. ***p < 0.001, Wilcoxon test.

(C) Bar plot showing the cell fraction of subsets of lymphocytes and myeloid cells in ascites that aligned to different tissues. ***p < 0.001, Wilcoxon test.

(legend continued on next page)

much higher than by 10x in all tissues. The gene capturing rate of 20 cells by SMART-seq2 was comparable with that of 1,000 cells by 10x (Figure 1C). Immune-relevant gene categories were also more frequently detected by SMART-seq2 than by 10x, exemplified by *LAMP3*$^+$ DCs ($p < 2e{-}16$, Wilcoxon test) (Figure 1D). For all cell clusters, SMART-seq2 yielded many more differentially expressed genes than 10x. For example, the number of differentially expressed genes detected by the 15 SMART-seq2 cells in *LAMP3*$^+$ DCs (*LAMP3*$^+$ DCs versus other clusters) was 29 times higher than that derived from 129 10x cells (Figure 1E). Notably, transcription factors (TFs) associated with DC development, including *ZBTB46*, were only detected by SMART-seq2 (Figure 1E). These genes also led to additional enriched Gene Ontology (GO) terms. On average, 45 significant GO terms were obtained by SMART-seq2 versus 17 by 10x ($p = 1.12e{-}7$, Wilcoxon test) (Figure S2E). Specifically, for *LAMP3*$^+$ DCs, in addition to the biological processes supported by both platforms, SMART-seq2-specific genes also revealed the enrichment of ''cell-to-cell signaling'' and ''regulation of cell migration,'' providing insight into their potential function (Figure 1F). Taken together, we conclude that the incorporation of SMART-seq2 into the widely used 10x-based scRNA-seq study has a synergistic effect on identifying immune cell types at a higher resolution.

We then integrated the data from all patients, including 66,187 cells generated by 10x and 11,134 cells by SMART-seq2, and performed further analyses (Figure 1G). Clustering analysis revealed major immune cell types in HCC, including T, NK, myeloid, B, plasma B cells, and ILCs (Figures S3A and S3B). T, NK, and myeloid cells were further partitioned through a second round of clustering analysis and yielded 14, 9, and 14 clusters, respectively (Figures S3D, S4A, and S4E).

The immune cell types exhibited different tissue preferences. We quantified tissue enrichment based on the ratio of observed to expected cell numbers in each cluster (*Ro/e*) using the 10x data (Figure 1H and Table S3). Among T cells, CD4-c6-FOXP3 and CD8-c7-PDCD1, corresponding to regulatory T (Treg) cells and exhausted CD8$^+$ T (Tex) cells, respectively, were enriched in tumors, consistent with previous findings (Guo et al., 2018; Zhang et al., 2018; Zheng et al., 2017a). CD8-c1-MKI67, representing proliferative T cells, was also enriched in tumors (Figure 1H). Using RNA velocity, a method inferring precursor-progeny cell dynamics (La Manno et al., 2018), we observed a clear directional flow from proliferative T cells to Tex cells (Figure S3E), consistent with a recent study of melanoma showing that proliferation in CD8$^+$ T cells was most profound during early stages of dysfunction (Li et al., 2019). Tumor-enriched populations like Treg cells are consistently low in both autologous adjacent liver and ''healthy liver'' controls (Figure S3C; MacParland et al., 2018). For NK cell clusters, NK-c3-IFNG, NK-c4-IFNG-HSPA1A, NK-c7-CD160, and NK-c8-CD160-HSPA1A were enriched in tumor compared with other tissues (Figure 1H). Among them, *IFNG*$^+$ NK cells highly expressed circulating NK (cNK) cell markers *FCGR3A*, *CX3CR1*, and the transcription factor T-bet (*TBX21*), whereas *CD160*$^+$ NK cells highly expressed liver-resident NK (lrNK) cell marker *CXCR6* and the transcription factor *EOMES* (Figures S4B–S4D; Male, 2017; Peng and Sun, 2017), indicating the existence of both cNK cells and lrNK cells in HCC. For myeloid cells, 2, 4, 6, and 2 clusters were identified for monocytes, DCs, macrophages, and mast cells, respectively (Figures S4E and S4F). Mφ-c5-VCAN, Mφ-c6-MARCO, and DC-c2-FCER1A were dominantly enriched in ascites (Figure 1H). In summary, by combined analysis of 10x and SMART-seq2 data, we identified multiple immune cell populations with distinct distribution patterns in different tissues of HCC.

Lymphocytes and Myeloid Cells in Ascites Appear to Have Distinct Origins

We investigated the potential origins of cells in ascites by expression alignment (see STAR Methods) to cells in other tissues with the premise that immune cells in ascites were accumulated from other tissues. We observed that lymphocytes significantly resembled cells from blood, whereas myeloid cells mainly mapped to cells of tumors (Figure 2A). Overall, 50.2% (3,314 of 6,603), 24.9% (1,641 of 6,603), and 20.8% (1,375 of 6,603) lymphocytes were assigned to cells in blood, tumors, and adjacent liver, respectively, whereas myeloid cells (79.5%; 2,613 of 3,286) were predominantly assigned to cells found in tumors (Figure 2B). In particular, the ascites-enriched clusters Mφ-c5-VCAN ($p = 0.046$, Fisher's exact test) and Mφ-c6-MARCO ($p < 0.001$) were primarily mapped to tumor-derived cells. Two lymphocyte subsets, NK-c9-MKI67 ($p < 0.001$) and CD8-c1-MKI67 ($p < 0.001$), also had their most similar cells in tumors (Figure 2C).

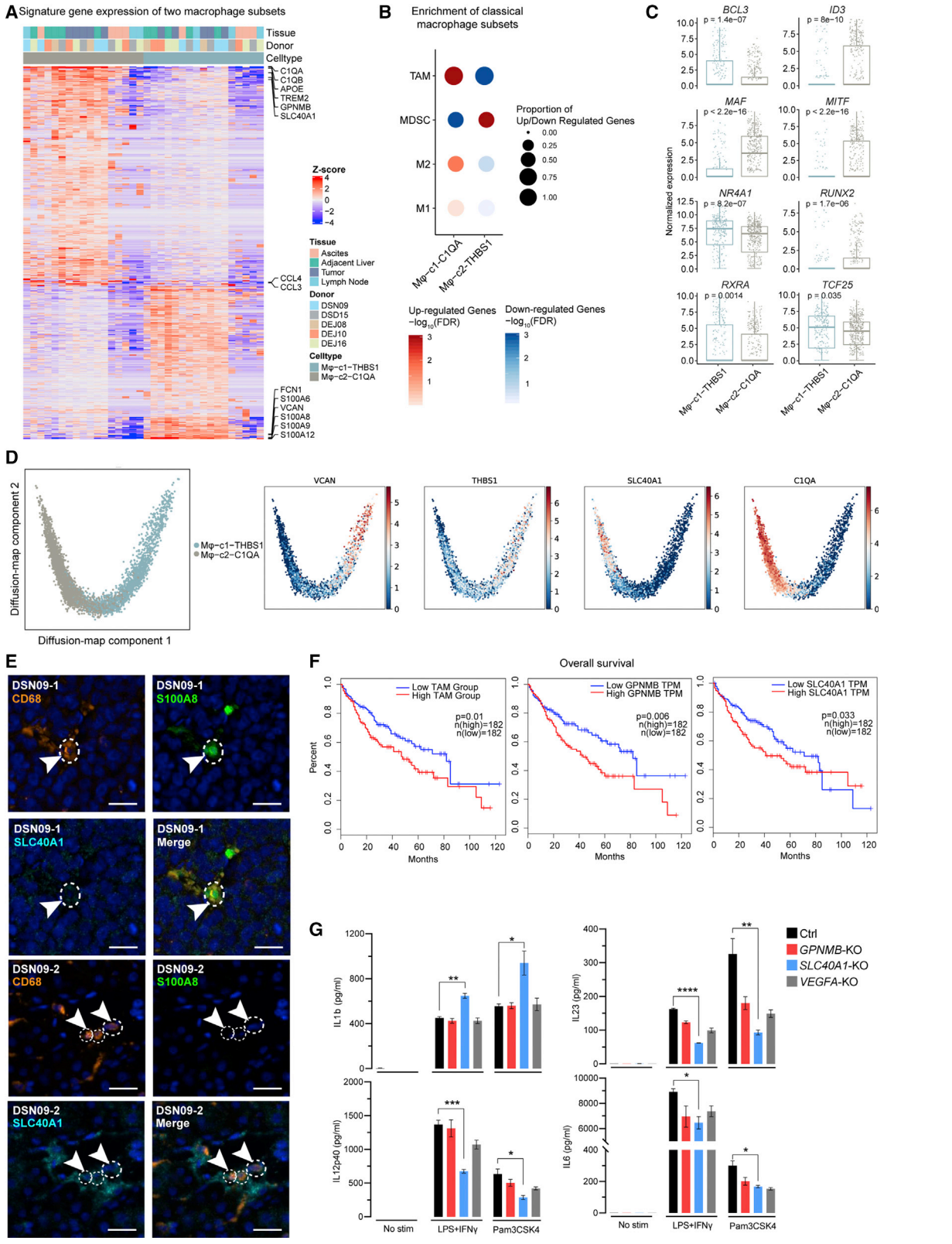
We then performed the reverse analysis to determine which cell types in tumors had the tendency to accumulate in ascites. Individual cells were labeled as ascites-similar or ascites-distant cells based on whether the cells in the tumor could find their best match in ascites. In myeloid cell subsets, Mφ-c1-THBS1 ($p < 0.001$, odds ratio = 0.63, Fisher's exact test) exhibited the highest degree of accumulation in ascites, whereas Mφ-C2-APOE ($p < 0.001$, odds ratio = 0.38) tended to be persistent in tumors (Figure 2D). For lymphocyte subsets, CD8-c1-MKI67, expressing cytotoxic genes (*GZMK*, *GZMA*, *IFNG*, and *TNF*), had a higher tendency to accumulate in ascites ($p < 0.001$, odds ratio = 4.63) (Figure 2D). We identified clonal T cells in CD8-c1-MKI67, sharing the identical T cell receptor (TCR) α-β pair, to span across ascites and tumor (Figure S4G), supporting their migration process between tumors and ascites. We then performed mitochondrial mutation-based lineage tracing following a method using scRNA-seq data (Ludwig et al., 2019). The

(D) Odds ratios and p values based on ascites-similar and ascites-distant cells of tumor-enriched cell clusters. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, Fisher's exact test.

(E) Branches of phylogenetic trees of macrophages in different tissues constructed by mitochondrial mutations. Venn diagrams show several of the number of stable lineages (bootstrap > 30).

(F) RNA velocity based on partition-based graph abstraction (PAGA) projection (Wolf et al., 2019), showing the transition potential from Mφ-c1-THBS1 to Mφ-c5-VCAN.

See also Figure S4 and S5.



(legend on next page)

phylogenetic tree of *MKI67*$^+$ T cells also showed cells in tumor and ascites with common lineages (Figure S4H). In contrast, Treg cells ($p < 0.001$, odds ratio = 0.47) and CD4-c5-CXCL13 ($p < 0.001$, odds ratio = 0.35) were ascites-distant clusters and more likely to be persistent in tumors (Figure 2D). It would be interesting to investigate whether tumors would selectively retain dysfunctional T cells while displacing proliferative cells to ascites.

We next studied the relationship of macrophages in ascites and other tissues. The phylogenetic tree based on mitochondrial mutations revealed a number of macrophages in tumor and ascites sharing common lineages (Figure 2E). Further, RNA velocity analysis revealed that tumor-enriched Mφ-c1-THBS1 exhibited lineage orientation to ascites-enriched Mφ-c5-VCAN (Figure 2F). We next isolated macrophages from different tissues of four additional HCC patients, incubated them with the autologous ascites supernatant, and profiled their transcriptomes. Such incubation did not result in enhanced similarity with ascites macrophages (Figure S5A). Altogether, our data support that specific subsets of lymphocytes and macrophages in ascites could originate from the tumor.

Two Distinct States of Tumor-Enriched Macrophages

Six macrophage clusters were identified in our dataset, of which Mφ-c1-THBS1 and Mφ-c2-C1QA were enriched in tumor tissues based on *Ro/e* (Figure 1H). Although genes upregulated in Mφ-c1-THBS1 were enriched for signatures of myeloid-derived suppressor cells (MDSCs) (Condamine et al., 2016; Zhao et al., 2012), those in Mφ-c2-C1QA simultaneously resembled the signatures for TAMs and M1 and M2 macrophages (Lavin et al., 2017; Figure 3B). The co-existence of M1 and M2 signatures indicated that TAMs were more complex than the classical M1/M2 model, consistent with a previous study (Azizi et al., 2018). A diffusion map of their global transcriptomes showed that Mφ-c1-THBS1 (MDSC-like macrophages) and Mφ-c2-C1QA (TAM-like macrophages) formed a continuum but with distinct expression features (Figure 3D). Specifically, MDSC-like cells highly expressed the S100A family genes *FCN1* and *VCAN*, whereas they expressed low levels of HLA-related genes (Figure 3A). In contrast, TAM-like cells expressed a set of genes found previously in TAMs of lung cancer (Lavin et al., 2017), including *APOE*, *C1QA*, *C1QB*, and *TREM2*. In addition, two additional genes, *SLC40A1*, encoding ferroportin, and *GPNMB*, encoding type I transmembrane glycoprotein, were highly expressed in TAM-like cells (Figure 3A). The transcription factors of these two clusters were distinct, with *ID3*, *MITF*, *RUNX2*, and *MAF* preferentially expressed in TAM-like cells and *BCL3*, *NR4A1*, *RXRA*, and *TCF25* highly expressed in MDSC-like cells (Figure 3C). We further performed multicolor immunohistochemistry (IHC) staining on tumor sections from HCC patients. Among CD68$^+$ macrophages, mutually exclusive signals of S100A8 and SLC40A1 were detected on distinct cells (Figure 3E), supporting the presence of the two distinct macrophage states in HCC.

We then examined the association of gene signatures of macrophage states with the prognosis of The Cancer Genome Atlas (TCGA) liver hepatocellular carcinoma (LIHC). Only the TAM-like signature was associated with a poor prognosis ($p = 0.01$, Cox regression). Two individual genes in the TAM-like signature, *SLC40A1* and *GPNMB*, were also associated with a poor prognosis ($p = 0.033$ and 0.006, respectively) (Figure 3F). To investigate the function of the two genes, we knocked out *SLC40A1* or *GPNMB* in THP-1 monocyte-derived macrophages using the CRISPR-Cas9 system and confirmed the gene deletion efficiency by flow cytometry (fluorescence-activated cell sorting [FACS]) (Figure S5B). We stimulated unedited control, *SLC40A1*-knockout (KO), *GPNMB*-KO, and *VEGFA*-KO THP-1 macrophages with lipopolysaccharide (LPS)+interferon γ (IFNγ) or Pam3CSK4 and measured the production of pro/anti-inflammatory cytokines. In comparison with the control, *GPNMB*-KO produced significantly lower amounts of tumor necrosis factor alpha (TNF-α) under both LPS+IFNγ and Pam3CSK4 conditions (Figure S5C), indicating that GPNMB promotes TNF-α production in macrophages. *SLC40A1*-KO secreted lower amounts of interleukin-23 (IL-23), IL-6, and IL-12p40 but higher amounts of IL-1β (Figure 3G), suggesting that SLC40A1 promoted pro-inflammatory cytokines but suppressed the production of IL1β, a critical regulator of inflammatory responses in the TME.

LAMP3$^+$ DCs: Mature DCs with Potential Migration Capacity in Tumors

The primary function of DCs in cancer immunity is to acquire tumor antigen, migrate to LNs, and activate a *de novo* T cell response (Gardner and Ruffell, 2016; Roberts et al., 2016). We focused on three DC subsets enriched in tumors: DC-c1-CD1C, DC-c3-CLEC9A, and DC-c4-LAMP3 (Figure 1H). DC-c1-CD1C highly expressed *CD1C*, *FCER1A*, and *CLEC10A*, corresponding to cDC2, whereas DC-c3-CLEC9A highly expressed *CLEC9A*, *XCR1*, and *CADM1*, representing cDC1 (Figure 4A;

Figure 3. Two Distinct States of Tumor-Enriched Macrophages

(A) Heatmap showing the signature gene expression of macrophage clusters. Rows represent signature genes. Columns represent individual cells.

(B) Gene enrichment for classical cell type M1, M2, MDSCs and TAMs in comparisons with the macrophage subsets in HCC. Bubble size represents the proportion of upregulated and downregulated genes of clusters based on the gene signature on the y axis. The color of the circle represents the directional FDR. Red, enrichment of upregulated genes in the specific cell subsets; blue, depletion of downregulated genes.

(C) Differentially expressed TFs in the two clusters based on SMART-seq2 (Wilcoxon test).

(D) Diffusion map showing the continuous connection of the two macrophage states (left) and signature gene expression based on 10x (right).

(E) Multicolor IHC staining to validate the two macrophage states, exemplified by patient DSN09. The dotted circle indicates the cell edge. The scale bars represent 20 μm.

(F) Overall survival curves of TCGA LIHC data (Cox regression).

(G) Bar plot showing the cytokine variation in *SLC40A1*-KO, *GPNMB*-KO, and *VEGFA*-KO THP-1 macrophages compared with the unedited control. *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$ (Student's t test).

See also Figure S5.

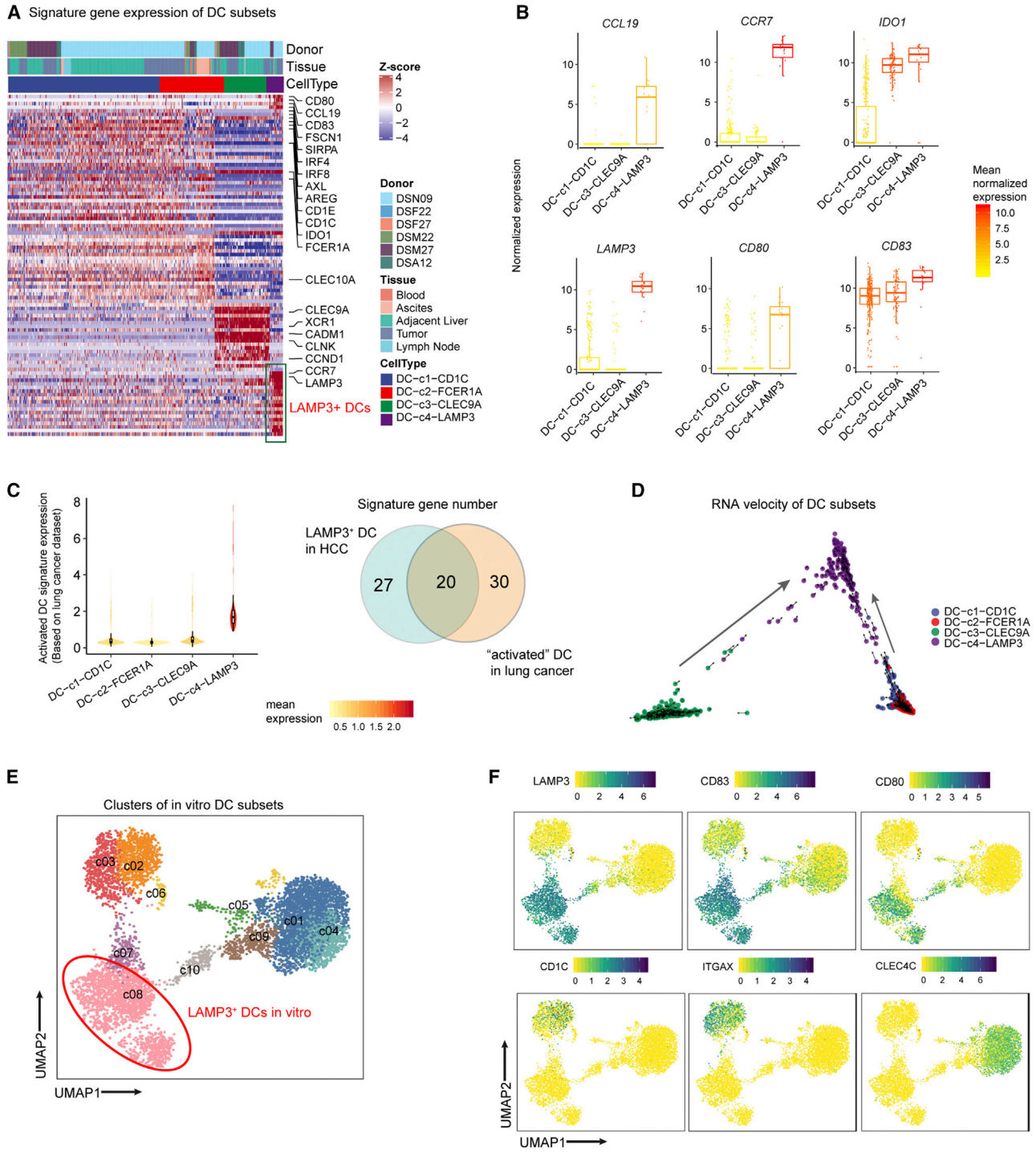


Figure 4. Characteristics and Relationships of DC Subsets

(A) Heatmap showing the gene expression of DCs in HCC based on SMART-seq2. Rows represent signature genes. Columns represent individual cells.

(B) Boxplot showing the signature gene expression of *LAMP3*$^+$ DCs based on SMART-seq2. Each dot represents a single cell.

(C) Expression consistency of the *LAMP3*$^+$ DCs in HCC with the hDC3 in lung cancer (Zilionis et al., 2019). Violin plot: expression of the hDC3 signature in DC subsets in HCC. Venn plot: the number of overlapped signature genes for two datasets.

(D) RNA velocities are visualized on the diffusion map projection of DCs in HCC.

(E) UMAP plot showing the clusters of *in vitro* DCs. Each dot represents a single cell. Color represents cluster.

(F) Signature gene expression of *in vitro* DCs.

See also Figures S5 and S6 and Table S4.

Villani et al., 2017). DC-c4-LAMP3 did not correspond to any classical DC subset *in vivo* based on known markers (Table S4). Cells in DC-c4-LAMP3 expressed the maturation markers *LAMP3*, *CD80*, and *CD83*; the migration marker *CCR7*; and the lymphocyte recirculation chemokines *CCL19* and *CCL21* (Figure 4B). To validate the existence of this population *in vivo*, we performed FACS on five additional HCC patient samples. CD11c$^+$LAMP3$^+$ DCs were identified in tumors, expressing higher CCR7 than cDCs (Figures S5D and S5E). The fraction of LAMP3$^+$ DCs in tumors was higher than that in adjacent liver ($p = 0.0007$, Student's t test) (Figure S5F), consistent with our scRNA-seq data. To examine whether *LAMP3*$^+$ DCs are present in other cancer types, we compared the transcriptome of tumor-infiltrating DCs in our dataset with the transcriptome data of FACS-defined DC subsets infiltrated in breast tumors (Michea et al., 2018). As expected, cDC1 and cDC2 in HCC tumors corresponded well with ones in breast cancer (Figure S5G). Although *LAMP3*$^+$ DCs in HCC did not precisely correspond with DC populations in breast cancer, genes highly expressed in *LAMP3*$^+$ DCs were found in their cDC1, cDC2, and pDC subsets (Figure S5G), suggesting that all of these DC subsets could contain *LAMP3*$^+$ DCs. Compared with the signature genes of scRNA-seq data for lung cancer (Zilionis et al., 2019), we found that *LAMP3*$^+$ DCs highly expressed signature genes derived from the "activated DCs" in lung cancer (Figure 4C; Table S5). Therefore, *LAMP3*$^+$ DCs appear to be present in multiple tumor types and can be detected by scRNA-seq.

Because *LAMP3* has been reported to be associated with DC maturation under CD40L stimulation (de Saint-Vis et al., 1998), we further examined the relationship between *LAMP3*$^+$ DCs and other DC subsets. Naive primary cDC1, cDC2, and pDCs in blood did not express *LAMP3* (Figure S6A). However, upon *in vitro* stimulation of LPS+IFNγ or poly I:C, all three DC subsets upregulated CD83 and LAMP3 (Figure S6B). *In vitro* monocyte-derived DCs (mo-DCs) also upregulated CD83 and LAMP3 upon receiving the maturation stimulus CD40L or LPS. In contrast, incubation with vitamin D, a maturation inhibitor (Malaguarnera et al., 2017; Penna and Adorini, 2000), suppressed CD83 and LAMP3 on mo-DCs (Figure S6C). We then performed scRNA-seq on DCs before and after *in vitro* stimulation to characterize transcriptomic changes of DC subsets during formation of *LAMP3*$^+$ DCs. These DCs could be divided into 10 populations (Figure 4E). Among them, C02 and C03 were identified as cDC2, C01 and C04 as pDCs, and C08 as *LAMP3*$^+$ DCs. We also identified C05 as a newly defined cDC subset, AS-DC (*AXL*$^+$*DAB2*$^+$) (Guilliams et al., 2016; See et al., 2017; Villani et al., 2017; Figures 4F and S6D). *LAMP3*$^+$ DCs comprised a larger population in samples treated with LPS+IFNγ and poly I:C in comparison with untreated samples (Figures 5A–5C), confirming that maturation stimuli induced *LAMP3*$^+$ DCs. Although the number of cDC1 was too small to form a distinct cluster, we examined the existence of cDC1 based on the high expression of *XCR1*, *CLEC9A*, and *CADM1* and observed an increased *LAMP3* level after poly I:C or LPS+IFNγ treatment (Figure S6E). In addition, we found three clusters enriched in the treated samples. C07, corresponding to the transitional stage between cDC2 to *LAMP3*$^+$ DCs, was more prominent in poly-I:C-treated samples, whereas C09 and C10, corresponding to the transitional stage between pDCs and *LAMP3*$^+$ DCs, were more prominent in LPS-treated samples (Figures 5A and 5C). Altogether, these data indicate that LAMP3 expression is a maturation marker of multiple DC subsets.

To further examine the lineage relationship of *LAMP3*$^+$ DCs with other DC subsets, we performed RNA velocity analysis of DC subsets in HCC and observed that 1.5% (22 of 1,465) cDC1 and 0.98% (4 of 410) cDC2 had the potential to transit to *LAMP3*$^+$ DCs (Figure 4D), suggesting that *LAMP3*$^+$ DCs in tumors may originate from both cDC1 and cDC2. For *in vitro* DC subsets after stimulation, RNA velocity analysis showed that cells in transitional stages (C07 and C09) exhibited a strong directional flow toward *LAMP3*$^+$ DCs (Figure 5D), indicating that *in vitro* maturation stimuli induce *LAMP3*$^+$ DC differentiation from cDCs and pDCs but not AS-DCs. Therefore, both *in vivo* and *in vitro* results demonstrated that different primary DC subsets had the potential to become *LAMP3*$^+$ DCs. We then compared the global transcriptome of DCs *in vitro* and in HCC using Spearman correlation. The *LAMP3*$^+$ DCs generated *in vitro* and those in HCC exhibited the highest correlation coefficient (Figure 5E) and shared expression of genes, including *CCR7*, *CCL19*, *FSCN1*, and *CD40* (Figure 5F).

We next examined DC migration potential in HCC using signature genes derived from mouse tissue-migratory cDCs (Miller et al., 2012; Table S5). In comparison with cDC1 and cDC2, *LAMP3*$^+$ DCs exhibited the highest "migration score" (Figure 6A). Using a Transwell migration assay (Figure S6F), we showed that mo-DCs matured with either CD40L+PGE2 or a commercial DC maturation supplement resulted in an increased level of LAMP3, and these matured DCs migrated more readily than immature DCs toward CCL19 (Figures 6B, S6G, and S6H). Conversely, DCs treated with vitamin D failed to migrate (Figure 6B). To test the migration capacity and potential direction of *LAMP3*$^+$ DCs in HCC, we traced the lineage origins of *LAMP3*$^+$ DCs in different tissues by analyzing the mitochondrial mutations. The phylogenetic tree of *LAMP3*$^+$ DCs constructed based on mitochondrial mutations of patient DSN09 identified a small number of cells sharing common lineages in LNs and tumors, more abundant than between other tissues (Figures 6C and 6D). We also performed RNA velocity analysis to examine the migration direction and observed that cDC1 and cDC2 did not show a clear directional flow between tumors and LNs (Figure 6E). In contrast, within *LAMP3*$^+$ DCs, tumor-derived cells exhibited a directional flow toward those in LNs, with a significantly higher tendency than cDC1 and cDC2 ($p = 4.08e{-}06$, odds ratio = 3.03) (Figure 6F), further supporting the potential migration direction. Altogether, we identified *LAMP3*$^+$ DCs in HCC with maturation features. These DCs may originate from cDC1 and cDC2 and exhibit migration capacity toward LNs.

Interaction Relationship of DCs and Lymphocytes

To investigate the interactions between myeloid cells and lymphocytes, we utilized a set of immune-related ligand-receptor (L-R) pairs (Chen and Flies, 2013; Ramilowski et al., 2015; Vento-Tormo et al., 2018) to gain insights into the regulatory relationships among cell clusters. Focusing on cells from tumors, we predicted 97 L-R pairs to mediate interactions between 625

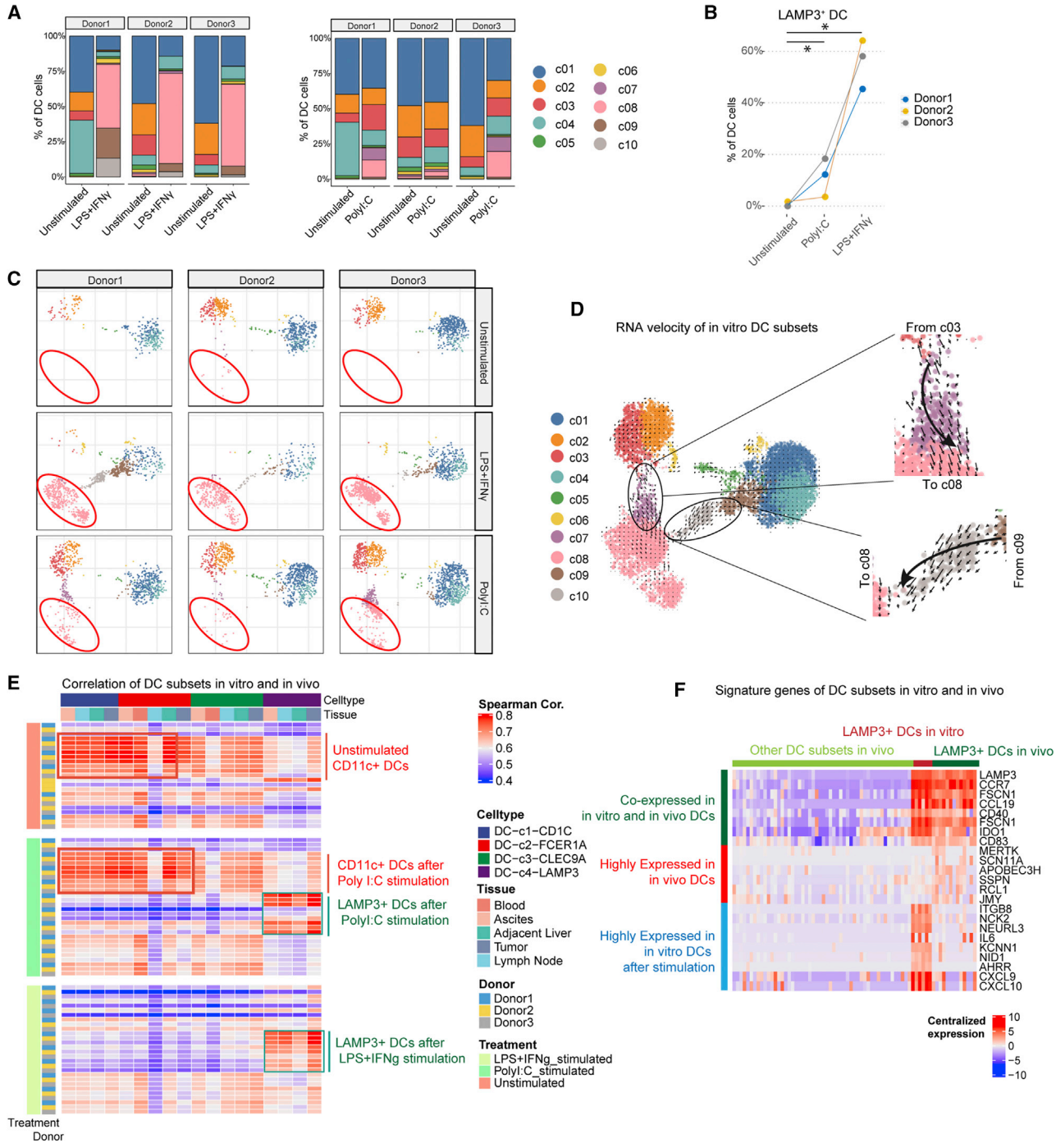


Figure 5. The Relationships of *In Vitro* and *In Vivo* DC Subsets

(A) The proportion of *in vitro* DC clusters. Color represents cluster. The x axis represents different stimuli, and the y axis represents the fraction of clusters.

(B) The percentages of *LAMP3*$^+$ DCs in different stimuli. *$p < 0.05$ (Student's t test).

(C) Disaggregation of the UMAP plots in Figure 4E by donor and stimulus.

(D) RNA velocities are visualized on the UMAP projection in Figure 4E using Gaussian smoothing on a regular grid.

(E) Expression similarity of DCs in HCC and *in vitro*. Rows represent DCs *in vitro*. Columns represent DCs in HCC.

(F) Heatmap showing the gene expression of *LAMP3*$^+$ DCs in HCC and *in vitro*. Rows represent signature genes. Columns represents clusters.

See also Figure S6.

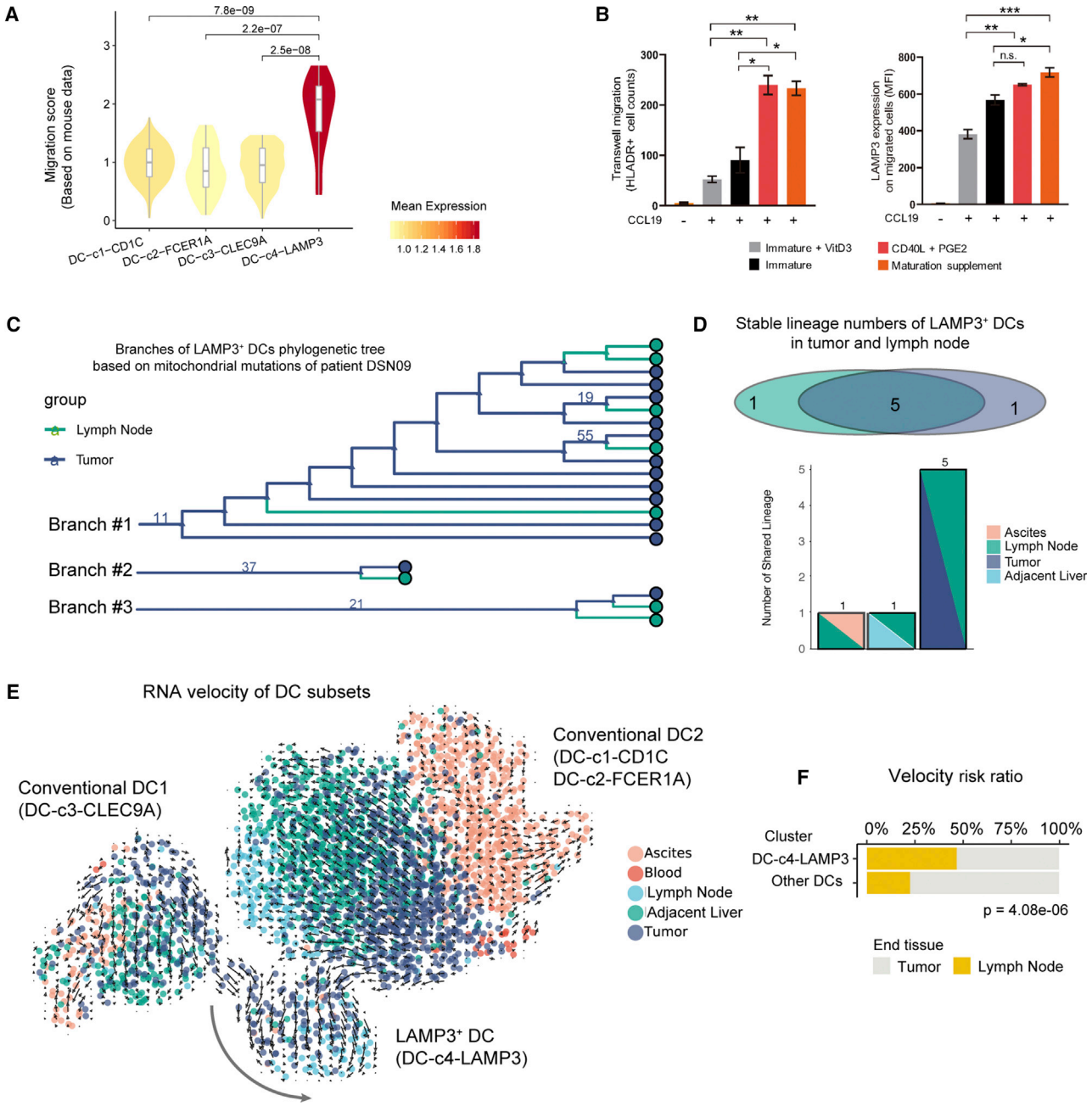


Figure 6. Migration Capacity of *LAMP3*⁺ DCs

(A) Violin plot showing the migration score of DC subsets in HCC (Student's t test).

(B) Bar plot showing the high migration capacity of mature DCs *in vitro*, based on a Transwell assay (left), that highly expressed LAMP3 (right). *p < 0.05, *p < 0.01, and *p < 0.001 (Student's t test).

(C) Branches of the *LAMP3*⁺ DCs mitochondrial phylogenetic tree of patient DSN09.

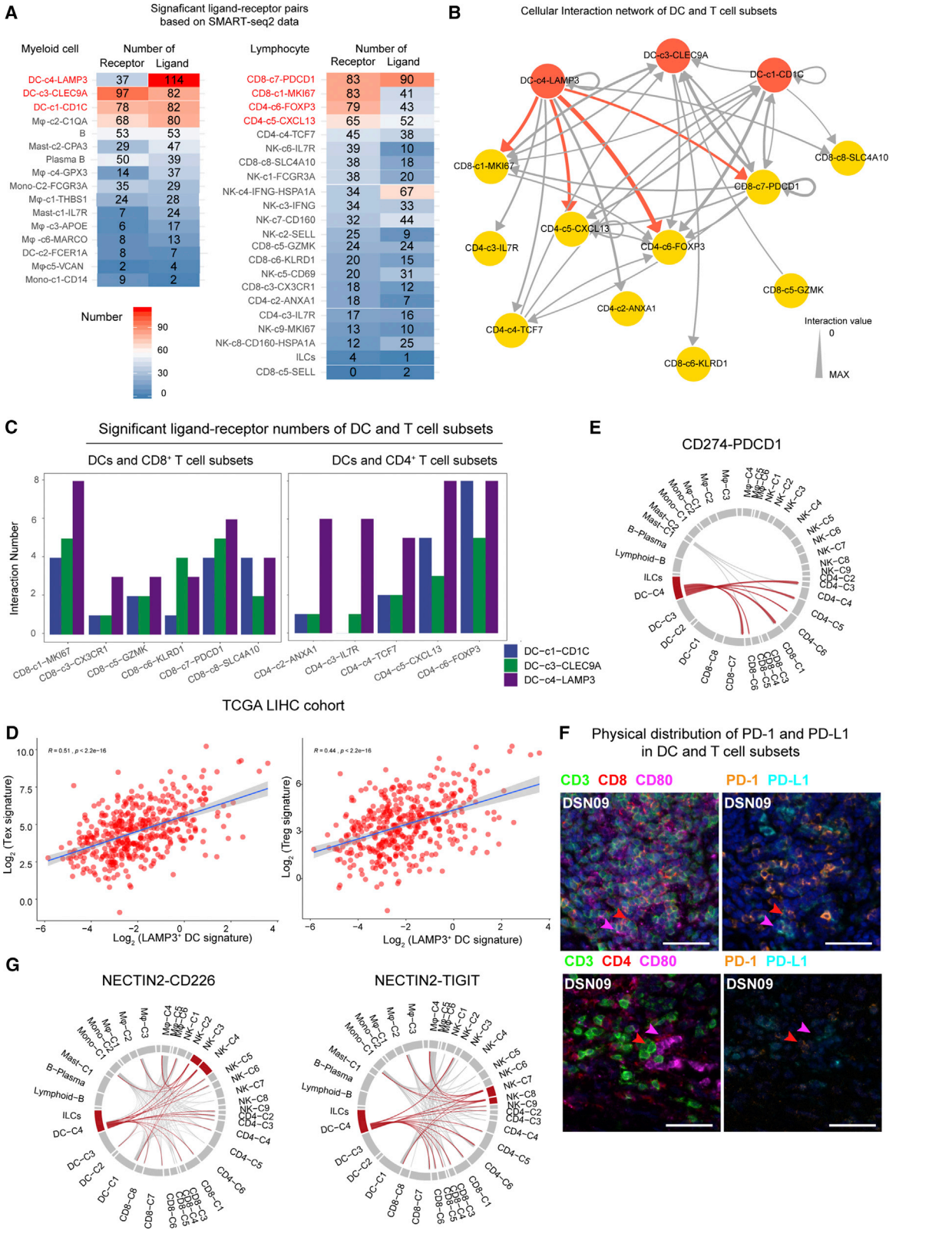
(D) Venn plot showing the stable lineages (bootstrap > 30) in tumors and LNs. The lineage relationship reviewed by phylogenetic trees does not reflect the degree of proliferation The bar plot shows the number of lineages shared in different tissues.

(E) RNA velocities are visualized on the UMAP projection of *in vitro* DCs using Gaussian smoothing on a regular grid.

(F) *LAMP3*⁺ DCs in tumors show a high ratio of RNA velocity links from tumors to LNs (Fisher's exact test).

See also Figure S6 and Table S5.

pairs of cell clusters (Table S6). The SMART-seq2 data revealed more potential L-R pairs because of the higher gene detection sensitivity (Figures 7A and S7A). In line with the important roles of DCs and T cells as immune regulators, DCs harbored the highest ligand numbers, with *LAMP3*⁺ DCs, cDC2, and cDC1 harboring 114, 82, and 82, respectively. T cells harbored the



(legend on next page)

highest receptor numbers in lymphocytes, with Tex cells, CD4-c5-CXCL13, Treg cells, and proliferative T cells harboring 83, 83, 79, and 65, respectively (Figures 7A and 7B).

All three tumor-enriched DC subsets showed strong potential interaction with Tex cells, proliferative T cells, and Treg cells via the co-stimulator CD28/B7 family (CD86-CD28, CD86-CTLA4, ICOSLG-CD28, and ICOLG-CTLA4), and interleukin-15 (IL-15-IL-2RB and IL-15-IL-2RG) (Figure S7B; Table S6). *LAMP3*$^+$ DCs exhibited the highest ligand numbers among the three DC subsets, and the L-R numbers of *LAMP3*$^+$ DCs with T cell subsets were higher than cDC1 and cDC2 (Figure 7C). *LAMP3*$^+$ DCs were predicted to interact with *CD4*$^+$ T cells (CD4-c3-IL-7R, CD4-c5-CXCL13, and CD4-c6-FOXP3) via the CCL19-CCR7 and CCL22-CCR4 axes, suggesting that *LAMP3*$^+$ DCs attracted T cells via migration chemokines (Figure S7B). *LAMP3*$^+$ DCs highly expressed *CD274* (PD-L1) and *PDCD1LG2* (PD-L2) and were predicted to bind to *PDCD1* (PD-1) on central memory T cells (CD4-c3-IL-7R and CD4-c4-TCF7), effector memory T cells (CD8-c4-SELL and CD8-c5-GZMK), Tex cells, Treg cells, and proliferative T cells (Figures 7E and S7B), suggesting that *LAMP3*$^+$ DCs may employ the CD274/PDCD1LG2-PDCD1 axis to regulate multiple types of T cells. Multicolor IHC staining of HCC tumors also showed the physical juxtaposition of PD-1-expressing T cells (CD3$^+$CD4$^+$ or CD3$^+$CD8$^+$) and PD-L1-expressing *LAMP3*$^+$ DCs (CD80$^+$) (Figures 7F and S7D). Based on analysis of 366 TCGA HCC patients, the *LAMP3*$^+$ DC gene signature showed a modest correlation with cytotoxic T cells ($R = 0.26$, $p < 1e{-}07$, Pearson's correlation) but strong correlations with Tex cells ($R = 0.51$, $p < 2.2e{-}16$) and Treg cells ($R = 0.44$, $p < 2.2e{-}16$) (Figures 7D and S7C). Altogether, our data demonstrate that *LAMP3*$^+$ DCs could regulate multiple T cell subsets in HCC tumors via various L-R pairs, including PD-1/PD-L1, and are more likely associated with dysfunction of T cells.

LAMP3$^+$ DCs were also predicted to interact with NK cells via *IL15* and *NECTIN2* (Table S6). *NECTIN2* encodes a membrane ligand, NECTIN, whose interaction with CD226 confers an activating signal but with TIGIT confers an inhibitory signal (Chen and Flies, 2013). Interestingly, *LAMP3*$^+$ DCs were predicted to interact with cNK cells via NECTIN2-CD226 but with lrNK cells via NECTIN2-TIGIT (Figure 7G). This suggests that *LAMP3*$^+$ DCs might regulate distinct NK cell subsets toward opposite directions.

DISCUSSION

Advanced cancer is a systemic disease, and the dynamic response of the immune system at different sites in cancer remains to be completely deciphered. Here we generated transcriptome data by the combination of 10x Genomics and SMART-seq2, covering more than 75,000 individual CD45$^+$ cells of 16 liver cancer patients from multiple immune-relevant tissue sites, providing a rich resource for understanding multi-dimensional characterization of immune cells in HCC. Transcriptome profiling, when augmented by analyses of RNA velocity, mitochondrial mutation-based lineage tracing, and L-R based cell-cell interaction, can lead to a far more dynamic picture, illustrating how various myeloid cells develop within the tumor, cross-talk with lymphocytes, and migrate to LNs or ascites.

Macrophages in tumors have been studied in lung cancer and breast cancer using scRNA-seq data (Azizi et al., 2018; Lavin et al., 2017). Here we identified two distinct macrophage states enriched in HCC tumor tissues. TAM-like macrophages in HCC highly resemble the TAMs identified in lung cancer (Lavin et al., 2017), and enrichment of TAM gene signatures is significantly associated with a survival disadvantage in both LIHC and lung cancer, suggesting this type of tumor-infiltrating TAMs as a potential cellular candidate for therapeutic targeting in multiple types of cancers. TAM-like macrophages in HCC tumors highly express two marker genes, *SLC40A1* and *GPNMB*. *SLC40A1* encodes ferroportin, an iron exporter, and regulates TLR-stimulus-induced pro-inflammatory cytokines, including IL-6, IL-23, and IL-1β, consistent with recent findings of iron metabolism in a polarizing macrophage phenotype in the TME (Mertens et al., 2017; Mora et al., 2019). Based on our observations, we hypothesize that iron metabolism is involved in shaping innate immunity in the TME, but the mechanistic details need further study.

Although much attention has been focused on cDCs (Azizi et al., 2018; Chevrier et al., 2017; Lavin et al., 2017), we found that *LAMP3*$^+$ DCs possess multiple interesting characteristics. First, *LAMP3*$^+$ DCs in tumors appear to be the most active immune-regulators of lymphocytes because they express the highest number of ligands to interact with receptors expressed on T cells and NK cells. It is intriguing that analysis of TCGA data indicates a strong correlation between the *LAMP3*$^+$ DC signature and the Tex or Treg cell signature in LIHC. This contrasts with the well-established functions of mature DCs in priming and activating T cells, implying that DCs in the TME, despite their maturation features, might be related to T cell dysfunction, a mechanism by which a tumor co-opts its environment to evade immune surveillance. Second, *LAMP3*$^+$ DCs are capable of migrating from tumors to hepatic LNs, as evidenced from the same lineages shared by *LAMP3*$^+$ DCs in tumors and LNs, as well as the precursor-progeny relationship supported by RNA

Figure 7. L-R-Based Interaction of DCs and Lymphocytes

(A) The number of ligands and receptors involved in significant L-R pairs of SMART-seq2.

(B) The interactions of DC and T cell clusters in tumors. DCs are filled in red and T cells in yellow. The ligands of DC-c4-LAMP3 are filled in red and others in gray. The arrow width is the sum of interaction values between two clusters. L-R pairs with a value > 10 and $p < 0.01$ are shown.

(C) The number of L-R pairs from DC to T cell subsets. Pairs with value > 10 and $p < 0.01$ were counted.

(D) Correlation of the *LAMP3*$^+$ DC signature with Tex or Treg cells based on TCGA LIHC data, normalized by *CD45* expression. Each dot represents a patient (Pearson's correlation analysis).

(E) The predicted interactions mediated by CD274-PDCD1.

(F) Multicolor IHC staining of PD-L1 and PD-1 expression in *LAMP3*$^+$ DCs and T cells, exemplified by patient DSN09. The scale bar represents 30 μm.

(G) The predicted interactions mediated by NECTIN2-CD226/TIGIT.

See also Figure S7 and Table S6.

velocity. Further supported by the observation that cells in CD8-c3-CX3CR1 and CD8-c5-GZMK in LNs share the same TCR with T cells in tumor or adjacent liver tissues (Figure S6I), it appears that *LAMP3*$^+$ DCs could migrate from tumors to LNs and prime T cell migration to tumors to exert effector functions. Importantly, we found that *LAMP3*$^+$ DCs exist not only in HCC but also in breast cancer and lung cancer (Figures 4C and S5G; Michea et al., 2018; Zilionis et al., 2019). Altogether, *LAMP3*$^+$ DCs may represent a common DC subset in tumors, arising from cDCs through maturation, with a unique capacity to regulate lymphocytes in the TME through cross-talk and migrate to LNs, correlating with dysfunctional T cells.

Analyses of multiple tissues in combination with computational lineage tracing reveals potential migration patterns of tumor-infiltrating immune cells to other sites, implying that the cellular TME may shape the immune response of distant sites. Our analyses demonstrate that tumor-infiltrating *LAMP3*$^+$ DCs can migrate or home to the hepatic LNs. A disease-related function of the hepatic artery LNs is not completely surprising because this LN station is known to be involved in liver-related immune pathology such as hepatitis B and C (Soresi et al., 2003) and primary biliary cirrhosis (Sadamoto et al., 1998). Although it is not a predominant site for malignant cell metastasis (Yuki et al., 1990; Zeng et al., 2005), hepatic artery LNs can serve as one of the draining LNs for tumor-associated DCs in HCC based on our data. We also described the cellular relationship of tumors and ascites in HCC for the first time. We found that ascites of the HCC is a unique tissue type that harbors multiple myeloid cells and lymphocytes from tumors. The myeloid cells include Mφ-c5-VCAN, Mφ-c6-MARCO, and DC-c2-FCER1A. In contrast, cDC1 appears to be tumor resident, whereas *LAMP3*$^+$ DCs migrate from tumors to LNs but not to ascites. Ascites also harbors lymphoid cells, including CD8-c1-MKI67 and NK-c9-MKI67, which implies a potential tumor immunosuppression process with the retention of Tex and Treg cells within tumors while allowing proliferative T cells, with high granzyme expression, to migrate outside of the tumor. Such a possibility will need further study. The clinical relevance of this finding needs to be addressed, correlating the phenotype of ascites-derived immune cells with the occurrence of ascites or survival of the patients. Because performing paracentesis to obtain ascites fluid is far simpler and safer than performing repeat liver biopsies, the potential use of ascites-derived myeloid cells as HCC biomarkers should be further explored.

In summary, our comprehensive characterization of immune cells from different tissue sites reveals the dynamic nature of immune cells in the cancer setting. Our analysis uncovers differential lineage and migratory relationships of myeloid and lymphoid cells in the TME, LNs, and ascites. To facilitate use of our data for the wide research community, we developed an interactive web-based tool (http://cancer-pku.cn:3838/HCC/) for analyzing and visualizing our single-cell data for one or multiple user input genes. Our data can be a valuable resource for further investigation to gain deeper biological insights that will lead to novel therapeutic targets and biomarkers of response for current immunotherapies for liver cancer.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- LEAD CONTACT AND MATERIALS AVAILABILITY
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Human Specimens
- METHOD DETAILS
  - Sample collection and single cell processing
  - Flow cytometry of primary PBMC
  - *In vitro* stimulation of DC subsets
  - Single cell RNA-seq process for patients
  - Single cell RNA-seq process for *in vitro* DCs
  - *In vitro* stimulation of macrophages
  - Bulk RNA and DNA isolation and sequencing
  - Multi-color immunohistochemistry
  - Genetic knockout in THP-1 cells using CRISPR-Cas9
  - Multiplex cytokine assay using THP-1 cells
  - Dendritic cells migration assay
  - Single cell RNA-seq data processing
  - Processing bulk RNA-seq data
  - Dimension reduction and unsupervised clustering
  - Comparing DCs in different cancer types
  - Cell similarity analysis across tissues
  - TCR analysis
  - Lineage tracing based on mitochondrial mutation
  - RNA velocity-based cell fate tracing
  - Cell-cell interaction analysis
  - TCGA data analysis
- DATA AND CODE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.cell.2019.10.003.

ACKNOWLEDGMENTS

We thank the BIOPIC high-throughput sequencing facility and the Computing Platform of the Center for Life Science for their contributions. We thank L. Hang and H. Yang for sample preparation and F. Wang and X. Zhang for assistance with FACS. We thank F. Ginhoux, H. Comeau, L. Zhang, X. Guo, and C. Zheng for discussions. This project was supported by the Beijing Advanced Innovation Centre for Genomics at Peking University, the Beijing Key Clinical Specialty Construction Project (2018), the National Natural Science Foundation of China (31530036 and 91742203), SLS-Qidong Innovation Fund, and Boehringer-Ingelheim.

AUTHOR CONTRIBUTIONS

Conceptualization, Z.Z., X.R., Q.Z., and Y. He; Resources, N.L., W.Z., Y.M., M.X., and J.P.; Methodology, Z.Z., X.R., and Y. He; Investigation, Q.Z., R.G., Y. Han, and X.H.; Formal Analysis, Y. He, Q.Z., Y. Han, and G.Z.; Validation, S.P.K., S.J.P., G.W.P., M.M., D.K., S.C., C.H., S.P., A.B.V., and Q.Z.; Writing – Original Draft, Q.Z., Y. He, X.R., and Z.Z.; Writing – Review & Editing, Q.Z., Y. He, K.L., X.R., M.B., and Z.Z.

DECLARATION OF INTERESTS

K.L., S.P.K., S.J.P., G.W.P., M.M., D.K., S.C., C.H., S.P., and A.B.V. are employees of Boehringer-Ingelheim. M.B. is a founder of Vacthera. Z.Z. is a

founder of Analytical Bioscience and an advisor for InnoCare. All financial interests are unrelated to this study. Other authors declare no competing interests.

Received: January 28, 2019
Revised: June 14, 2019
Accepted: October 3, 2019
Published: October 31, 2019

REFERENCES

Azizi, E., Carr, A.J., Plitas, G., Cornish, A.E., Konopacki, C., Prabhakaran, S., Nainys, J., Wu, K., Kiseliovas, V., Setty, M., et al. (2018). Single-Cell Map of Diverse Immune Phenotypes in the Breast Tumor Microenvironment. Cell *174*, 1293–1308.e36.

Biecker, E. (2011). Diagnosis and therapy of ascites in liver cirrhosis. World J. Gastroenterol. *17*, 1237–1248.

Binnewies, M., Mujal, A.M., Pollack, J.L., Combes, A.J., Hardison, E.A., Barry, K.C., Tsui, J., Ruhland, M.K., Kersten, K., Abushawish, M.A., et al. (2019). Unleashing Type-2 Dendritic Cells to Drive Protective Antitumor CD4(+) T Cell Immunity. Cell *177*, 556–571.e16.

Bray, N.L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. Nat. Biotechnol. *34*, 525–527.

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, L.A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J. Clin. *68*, 394–424.

Chen, L., and Flies, D.B. (2013). Molecular mechanisms of T cell co-stimulation and co-inhibition. Nat. Rev. Immunol. *13*, 227–242.

Chevrier, S., Levine, J.H., Zanotelli, V.R.T., Silina, K., Schulz, D., Bacac, M., Ries, C.H., Ailles, L., Jewett, M.A.S., Moch, H., et al. (2017). An Immune Atlas of Clear Cell Renal Cell Carcinoma. Cell *169*, 736–749.e18.

Chung, W., Eum, H.H., Lee, H.O., Lee, K.M., Lee, H.B., Kim, K.T., Ryu, H.S., Kim, S., Lee, J.E., Park, Y.H., et al. (2017). Single-cell RNA-seq enables comprehensive tumour and immune cell profiling in primary breast cancer. Nat. Commun. *8*, 15081.

Condamine, T., Dominguez, G.A., Youn, J.I., Kossenkov, A.V., Mony, S., Alicea-Torres, K., Tcyganov, E., Hashimoto, A., Nefedova, Y., Lin, C., et al. (2016). Lectin-type oxidized LDL receptor-1 distinguishes population of human polymorphonuclear myeloid-derived suppressor cells in cancer patients. Sci. Immunol. *1*, aaf8943.

de Saint-Vis, B., Vincent, J., Vandenabeele, S., Vanbervliet, B., Pin, J.J., Aït-Yahia, S., Patel, S., Mattei, M.G., Banchereau, J., Zurawski, S., et al. (1998). A novel lysosome-associated membrane glycoprotein, DC-LAMP, induced upon DC maturation, is transiently expressed in MHC class II compartment. Immunity *9*, 325–336.

El-Khoueiry, A.B., Sangro, B., Yau, T., Crocenzi, T.S., Kudo, M., Hsu, C., Kim, T.Y., Choo, S.P., Trojan, J., Welling, T.H.R., et al. (2017). Nivolumab in patients with advanced hepatocellular carcinoma (CheckMate 040): an open-label, non-comparative, phase 1/2 dose escalation and expansion trial. Lancet *389*, 2492–2502.

Engblom, C., Pfirschke, C., and Pittet, M.J. (2016). The role of myeloid cells in cancer therapies. Nat. Rev. Cancer *16*, 447–462.

Erard, N., Knott, S.R.V., and Hannon, G.J. (2017). A CRISPR Resource for Individual, Combinatorial, or Multiplexed Gene Knockout. Mol. Cell *67*, 1080.

Forner, A., Reig, M., and Bruix, J. (2018). Hepatocellular carcinoma. Lancet *391*, 1301–1314.

Fukui, H., Kawaratani, H., Kaji, K., Takaya, H., and Yoshiji, H. (2018). Management of refractory cirrhotic ascites: challenges and solutions. Hepat. Med. *10*, 55–71.

Gardner, A., and Ruffell, B. (2016). Dendritic Cells and Cancer Immunity. Trends Immunol. *37*, 855–865.

Georgoudaki, A.M., Prokopec, K.E., Boura, V.F., Hellqvist, E., Sohn, S., Östling, J., Dahan, R., Harris, R.A., Rantalainen, M., Klevebring, D., et al. (2016). Reprogramming Tumor-Associated Macrophages by Antibody Targeting Inhibits Cancer Progression and Metastasis. Cell Rep. *15*, 2000–2011.

Gnjatic, S., Bronte, V., Brunet, L.R., Butler, M.O., Disis, M.L., Galon, J., Hakansson, L.G., Hanks, B.A., Karanikas, V., Khleif, S.N., et al. (2017). Identifying baseline immune-related biomarkers to predict clinical outcome of immunotherapy. J. Immunother. Cancer *5*, 44.

Guilliams, M., Dutertre, C.A., Scott, C.L., McGovern, N., Sichien, D., Chakarov, S., Van Gassen, S., Chen, J., Poidinger, M., De Prijck, S., et al. (2016). Unsupervised High-Dimensional Analysis Aligns Dendritic Cells across Tissues and Species. Immunity *45*, 669–684.

Guo, X., Zhang, Y., Zheng, L., Zheng, C., Song, J., Zhang, Q., Kang, B., Liu, Z., Jin, L., Xing, R., et al. (2018). Global characterization of T cells in non-small-cell lung cancer by single-cell sequencing. Nat. Med. *24*, 978–985.

Hackl, H., Charoentong, P., Finotello, F., and Trajanoski, Z. (2016). Computational genomics tools for dissecting tumour-immune cell interactions. Nat. Rev. Genet. *17*, 441–458.

Hsu, C.Y., Lee, Y.H., Huang, Y.H., Hsia, C.Y., Su, C.W., Lin, H.C., Lee, R.C., Chiou, Y.Y., Lee, F.Y., Huo, T.I., and Lee, S.D. (2013). Ascites in patients with hepatocellular carcinoma: prevalence, associated factors, prognostic impact, and staging strategy. Hepatol. Int. *7*, 188–198.

Korsunsky, I., Fan, J., Slowikowski, K., Zhang, F., Wei, K., Baglaenko, Y., Brenner, M., Loh, P.-R., and Raychaudhuri, S. (2018). Fast, sensitive, and accurate integration of single cell data with Harmony. bioRxiv. https://doi.org/10.1101/461954.

La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastriti, M.E., Lönnerberg, P., Furlan, A., et al. (2018). RNA velocity of single cells. Nature *560*, 494–498.

Lambrechts, D., Wauters, E., Boeckx, B., Aibar, S., Nittner, D., Burton, O., Bassez, A., Decaluwé, H., Pircher, A., Van den Eynde, K., et al. (2018). Phenotype molding of stromal cells in the lung tumor microenvironment. Nat. Med. *24*, 1277–1289.

Lavin, Y., Kobayashi, S., Leader, A., Amir, E.D., Elefant, N., Bigenwald, C., Remark, R., Sweeney, R., Becker, C.D., Levine, J.H., et al. (2017). Innate Immune Landscape in Early Lung Adenocarcinoma by Paired Single-Cell Analyses. Cell *169*, 750–765.e17.

Li, H., Courtois, E.T., Sengupta, D., Tan, Y., Chen, K.H., Goh, J.J.L., Kong, S.L., Chua, C., Hon, L.K., Tan, W.S., et al. (2017). Reference component analysis of single-cell transcriptomes elucidates cellular heterogeneity in human colorectal tumors. Nat. Genet. *49*, 708–718.

Li, H., van der Leun, A.M., Yofe, I., Lubling, Y., Gelbard-Solodkin, D., van Akkooi, A.C.J., van den Braber, M., Rozeman, E.A., Haanen, J., Blank, C.U., et al. (2019). Dysfunctional CD8 T Cells Form a Proliferative, Dynamically Regulated Compartment within Human Melanoma. Cell *176*, 775–789.e18.

Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923–930.

Llovet, J.M., Ricci, S., Mazzaferro, V., Hilgard, P., Gane, E., Blanc, J.F., de Oliveira, A.C., Santoro, A., Raoul, J.L., Forner, A., et al.; SHARP Investigators Study Group (2008). Sorafenib in advanced hepatocellular carcinoma. N. Engl. J. Med. *359*, 378–390.

Ludwig, L.S., Lareau, C.A., Ulirsch, J.C., Christian, E., Muus, C., Li, L.H., Pelka, K., Ge, W., Oren, Y., Brack, A., et al. (2019). Lineage Tracing in Humans Enabled by Mitochondrial Mutations and Single-Cell Genomics. Cell *176*, 1325–1339.e22.

Lun, A.T., Bach, K., and Marioni, J.C. (2016). Pooling across cells to normalize single-cell RNA sequencing data with many zero counts. Genome Biol. *17*, 75.

Lun, A.T.L., Riesenfeld, S., Andrews, T., Dao, T.P., Gomes, T., and Marioni, J.C.; participants in the 1st Human Cell Atlas Jamboree (2019). EmptyDrops: distinguishing cells from empty droplets in droplet-based single-cell RNA sequencing data. Genome Biol. *20*, 63.

MacParland, S.A., Liu, J.C., Ma, X.Z., Innes, B.T., Bartczak, A.M., Gage, B.K., Manuel, J., Khuu, N., Echeverri, J., Linares, I., et al. (2018). Single cell RNA sequencing of human liver reveals distinct intrahepatic macrophage populations. Nat. Commun. *9*, 4383.

Malaguarnera, L., Marsullo, A., Zorena, K., Musumeci, G., and Di Rosa, M. (2017). Vitamin D_3 regulates LAMP3 expression in monocyte derived dendritic cells. Cell. Immunol. *311*, 13–21.

Male, V. (2017). Liver-Resident NK Cells: The Human Factor. Trends Immunol. *38*, 307–309.

McCarthy, D.J., Campbell, K.R., Lun, A.T., and Wills, Q.F. (2017). Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. Bioinformatics *33*, 1179–1186.

Mertens, C., Mora, J., Ören, B., Grein, S., Winslow, S., Scholich, K., Weigert, A., Malmström, P., Forsare, C., Fernö, M., et al. (2017). Macrophage-derived lipocalin-2 transports iron in the tumor microenvironment. OncoImmunology *7*, e1408751.

Michea, P., Noël, F., Zakine, E., Czerwinska, U., Sirven, P., Abouzid, O., Goudot, C., Scholer-Dahirel, A., Vincent-Salomon, A., Reyal, F., et al. (2018). Adjustment of dendritic cells to the breast-cancer microenvironment is subset specific. Nat. Immunol. *19*, 885–897.

Miller, J.C., Brown, B.D., Shay, T., Gautier, E.L., Jojic, V., Cohain, A., Pandey, G., Leboeuf, M., Elpek, K.G., Helft, J., et al.; Immunological Genome Consortium (2012). Deciphering the transcriptional network of the dendritic cell lineage. Nat. Immunol. *13*, 888–899.

Mora, J., Mertens, C., Meier, J.K., Fuhrmann, D.C., Brüne, B., and Jung, M. (2019). Strategies to Interfere with Tumor Metabolism through the Interplay of Innate and Adaptive Immunity. Cells *8*, E445.

Peng, H., and Sun, R. (2017). Liver-resident NK cells and their potential functions. Cell. Mol. Immunol. Published online September 18, 2017. https://doi.org/10.1038/cmi.2017.72.

Penna, G., and Adorini, L. (2000). 1 Alpha,25-dihydroxyvitamin D3 inhibits differentiation, maturation, activation, and survival of dendritic cells leading to impaired alloreactive T cell activation. J. Immunol. *164*, 2405–2411.

Peranzoni, E., Lemoine, J., Vimeux, L., Feuillet, V., Barrin, S., Kantari-Mimoun, C., Bercovici, N., Guérin, M., Biton, J., Ouakrim, H., et al. (2018). Macrophages impede CD8 T cells from reaching tumor cells and limit the efficacy of anti-PD-1 treatment. Proc. Natl. Acad. Sci. USA *115*, E4041–E4050.

Picelli, S., Faridani, O.R., Björklund, A.K., Winberg, G., Sagasser, S., and Sandberg, R. (2014). Full-length RNA-seq from single cells using Smart-seq2. Nat. Protoc. *9*, 171–181.

Pijuan-Sala, B., Griffiths, J.A., Guibentif, C., Hiscock, T.W., Jawaid, W., Calero-Nieto, F.J., Mulas, C., Ibarra-Soria, X., Tyser, R.C.V., Ho, D.L.L., et al. (2019). A single-cell molecular map of mouse gastrulation and early organogenesis. Nature *566*, 490–495.

Puram, S.V., Tirosh, I., Parikh, A.S., Patel, A.P., Yizhak, K., Gillespie, S., Rodman, C., Luo, C.L., Mroz, E.A., Emerick, K.S., et al. (2017). Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer. Cell *171*, 1611–1624.e24.

Ramilowski, J.A., Goldberg, T., Harshbarger, J., Kloppmann, E., Lizio, M., Satagopam, V.P., Itoh, M., Kawaji, H., Carninci, P., Rost, B., and Forrest, A.R. (2015). A draft network of ligand-receptor-mediated multicellular signalling in human. Nat. Commun. *6*, 7866.

Ray, K. (2017). Liver cancer: Nivolumab: checkmate for hepatocellular carcinoma? Nat. Rev. Gastroenterol. Hepatol. *14*, 326.

Ringelhan, M., Pfister, D., O'Connor, T., Pikarsky, E., and Heikenwalder, M. (2018). The immunology of hepatocellular carcinoma. Nat. Immunol. *19*, 222–232.

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. *43*, e47.

Roberts, E.W., Broz, M.L., Binnewies, M., Headley, M.B., Nelson, A.E., Wolf, D.M., Kaisho, T., Bogunovic, D., Bhardwaj, N., and Krummel, M.F. (2016). Critical Role for CD103(+)/CD141(+) Dendritic Cells Bearing CCR7 for Tumor Antigen Trafficking and Priming of T Cell Immunity in Melanoma. Cancer Cell *30*, 324–336.

Sadamoto, T., Joplin, R., Keogh, A., Mason, A., Carman, W., and Neuberger, J. (1998). Expression of pyruvate-dehydrogenase complex PDC-E2 on biliary epithelial cells induced by lymph nodes from primary biliary cirrhosis. Lancet *352*, 1595–1596.

Salmon, H., Idoyaga, J., Rahman, A., Leboeuf, M., Remark, R., Jordan, S., Casanova-Acebes, M., Khudoynazarova, M., Agudo, J., Tung, N., et al. (2016). Expansion and Activation of CD103(+) Dendritic Cell Progenitors at the Tumor Site Enhances Tumor Responses to Therapeutic PD-L1 and BRAF Inhibition. Immunity *44*, 924–938.

See, P., Dutertre, C.A., Chen, J., Günther, P., McGovern, N., Irac, S.E., Gunawan, M., Beyer, M., Händler, K., Duan, K., et al. (2017). Mapping the human DC lineage through the integration of high-dimensional techniques. Science *356*, eaag3009.

Soresi, M., Bonfissuto, G., Magliarisi, C., Riili, A., Terranova, A., Di Giovanni, G., Bascone, F., Carroccio, A., Tripi, S., and Montalto, G. (2003). Ultrasound detection of abdominal lymph nodes in chronic liver diseases. A retrospective analysis. Clin. Radiol. *58*, 372–377.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics *30*, 1312–1313.

Svensson, V., Natarajan, K.N., Ly, L.H., Miragaia, R.J., Labalette, C., Macaulay, I.C., Cvejic, A., and Teichmann, S.A. (2017). Power analysis of single-cell RNA-sequencing experiments. Nat. Methods *14*, 381–387.

Tang, Z., Kang, B., Li, C., Chen, T., and Zhang, Z. (2019). GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. Nucleic Acids Res. *47* (W1), W556–W560.

Tirosh, I., Izar, B., Prakadan, S.M., Wadsworth, M.H., 2nd, Treacy, D., Trombetta, J.J., Rotem, A., Rodman, C., Lian, C., Murphy, G., et al. (2016). Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science *352*, 189–196.

Traag, V.A., Waltman, L., and van Eck, N.J. (2019). From Louvain to Leiden: guaranteeing well-connected communities. Sci. Rep. *9*, 5233.

Vento-Tormo, R., Efremova, M., Botting, R.A., Turco, M.Y., Vento-Tormo, M., Meyer, K.B., Park, J.E., Stephenson, E., Polański, K., Goncalves, A., et al. (2018). Single-cell reconstruction of the early maternal-fetal interface in humans. Nature *563*, 347–353.

Villani, A.C., Satija, R., Reynolds, G., Sarkizova, S., Shekhar, K., Fletcher, J., Griesbeck, M., Butler, A., Zheng, S., Lazo, S., et al. (2017). Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. Science *356*, eaah4573.

Wagner, J., Rapsomaniki, M.A., Chevrier, S., Anzeneder, T., Langwieder, C., Dykgers, A., Rees, M., Ramaswamy, A., Muenst, S., Soysal, S.D., et al. (2019). A Single-Cell Atlas of the Tumor and Immune Ecosystem of Human Breast Cancer. Cell *177*, 1330–1345.e18.

Wolf, F.A., Angerer, P., and Theis, F.J. (2018). SCANPY: large-scale single-cell gene expression data analysis. Genome Biol. *19*, 15.

Wolf, F.A., Hamey, F.K., Plass, M., Solana, J., Dahlin, J.S., Göttgens, B., Rajewsky, N., Simon, L., and Theis, F.J. (2019). PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. Genome Biol. *20*, 59.

Yuki, K., Hirohashi, S., Sakamoto, M., Kanai, T., and Shimosato, Y. (1990). Growth and spread of hepatocellular carcinoma. A review of 240 consecutive autopsy cases. Cancer *66*, 2174–2179.

Yusufi, N., Mall, S., Bianchi, H.O., Steiger, K., Reder, S., Klar, R., Audehm, S., Mustafa, M., Nekolla, S., Peschel, C., et al. (2017). In-depth Characterization of a TCR-specific Tracer for Sensitive Detection of Tumor-directed Transgenic T Cells by Immuno-PET. Theranostics *7*, 2402–2416.

Zeng, Z.C., Tang, Z.Y., Fan, J., Qin, L.X., Ye, S.L., Zhou, J., Sun, H.C., Wang, B.L., and Wang, J.H. (2005). Consideration of role of radiotherapy for lymph node metastases in patients with HCC: retrospective analysis for prognostic factors from 125 patients. Int. J. Radiat. Oncol. Biol. Phys. *63*, 1067–1076.

Zhang, L., Yu, X., Zheng, L., Zhang, Y., Li, Y., Fang, Q., Gao, R., Kang, B., Zhang, Q., Huang, J.Y., et al. (2018). Lineage tracking reveals dynamic relationships of T cells in colorectal cancer. Nature *564*, 268–272.

Zhao, F., Hoechst, B., Duffy, A., Gamrekelashvili, J., Fioravanti, S., Manns, M.P., Greten, T.F., and Korangy, F. (2012). S100A9 a new marker for monocytic human myeloid-derived suppressor cells. Immunology *136*, 176–183.

Zheng, C., Zheng, L., Yoo, J.K., Guo, H., Zhang, Y., Guo, X., Kang, B., Hu, R., Huang, J.Y., Zhang, Q., et al. (2017a). Landscape of Infiltrating T Cells in Liver Cancer Revealed by Single-Cell Sequencing. Cell *169*, 1342–1356.e16.

Zheng, G.X., Terry, J.M., Belgrader, P., Ryvkin, P., Bent, Z.W., Wilson, R., Ziraldo, S.B., Wheeler, T.D., McDermott, G.P., Zhu, J., et al. (2017b). Massively parallel digital transcriptional profiling of single cells. Nat. Commun. *8*, 14049.

Ziegenhain, C., Vieth, B., Parekh, S., Reinius, B., Guillaumet-Adkins, A., Smets, M., Leonhardt, H., Heyn, H., Hellmann, I., and Enard, W. (2017). Comparative Analysis of Single-Cell RNA Sequencing Methods. Mol. Cell *65*, 631–643.e4.

Zilionis, R., Engblom, C., Pfirschke, C., Savova, V., Zemmour, D., Saatcioglu, H.D., Krishnan, I., Maroni, G., Meyerovitz, C.V., Kerwin, C.M., et al. (2019). Single-Cell Transcriptomics of Human and Mouse Lung Cancers Reveals Conserved Myeloid Populations across Individuals and Species. Immunity *50*, 1317–1334.e10.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-Human CD45 FITC (FACS)	eBioscience	Cat #11-0459-42
Anti-Human CD3 eFlour 450 (FACS)	eBioscience	Cat #48-0037-41
Anti-Human CD4 FITC (FACS)	eBioscience	Cat #11-0048-41
Anti-Human CD8a APC (FACS)	eBioscience	Cat #17-0086-41
Anti-Human CD25 PE (FACS)	eBioscience	Cat #12-0259-42
Anti-Human CD68 PE-eFlour 610 (FACS)	eBioscience	Cat #61-0689-42
7-AAD Viability Staining Solution (FACS)	eBioscience	Cat #00-6993-50
Anti-CD68 antibody (IHC)	Abcam	Cat #ab213363
Anti-SLC40A1 antibody (IHC)	Abcam	Cat #ab78066
Anti-S100A8 antibody (IHC)	Abcam	Cat #22506
Anti-CD80 antibody (IHC)	Abcam	Cat #ab134120
Anti-CD3 antibody (IHC)	Abcam	Cat #ab11089
Anti-CD4 antibody (IHC)	Abcam	Cat #ab133616
Anti-CD8 antibody (IHC)	Abcam	Cat #ab17147
Anti-PD1 antibody (IHC)	Abcam	Cat #ab137132
Anti-PD-L1 antibody (IHC)	Abcam	Cat #ab213524
BUV737-CD45RA (FACS)	BD Biosciences	Cat #612846
BV421-CD303 (FACS)	Biolegend	Cat #354212
BV510-CD123 (FACS)	Biolegend	Cat #306022
BV786 HLA-DR (FACS)	BioLegend	Cat #307642
BV650 CD3 (FACS)	Biolegend	Cat #317324
BV650 CD16 (FACS)	Biolegend	Cat #302042
BV650 CD19 (FACS)	Biolegend	Cat #302238
BV650 CD56 (FACS)	Biolegend	Cat #318343
BV711 CD141 (FACS)	BD Biosciences	Cat #563155
FITC CD14 (FACS)	Biolegend	Cat #325603
PerCP-Cy5.5 CD83 (FACS)	Biolegend	Cat #305320
APC CLEC9A (FACS)	Biolegend	Cat #353805
PE-dazzle CD45 (FACS)	Biolegend	Cat #304051
PE-Cy7 CCR7 (FACS)	Biolegend	Cat #353225
Clec9A (FACS)	Biolegend	Cat #353804
PE-dazzle CD45 (FACS)	Biolegend	Cat #304051
APC-Cy7 CD1c (FACS)	Biolegend	Cat #331520
CD3 PE-Cy7 (Lin) (FACS)	BD Biosciences	Cat #557749
CD14 PE- Cy7 (Lin) (FACS)	BD Biosciences	Cat #561385
CD16 PE-Cy7 (Lin) (FACS)	BD Biosciences	Cat #560716
CD19 PE-Cy7 (Lin) (FACS)	BD Biosciences	Cat #557835
CD20 PE-Cy7 (Lin) (FACS)	BD Biosciences	Cat #560735
CD56 PE-Cy 7 (Lin) (FACS)	BD Biosciences	Cat #557747
HLA-DR APC (FACS)	BD Biosciences	Cat #560744
CD123 BV605 (FACS)	BD Biosciences	Cat #564197
CD11c BUV395 (FACS)	BD Biosciences	Cat #563787
CD141 BB515 (FACS)	BD Biosciences	Cat #565084

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
CD1c AF700 (FACS)	Biolegend	Cat #331530
DC-LAMP PE (FACS)	BD Biosciences	Cat #558126
Isotype Control PE (FACS)	BD Biosciences	Cat #555749
CD83-APC-Cy7 (FACS)	Biolegend	Cat #305339
Anti-SLC40A1 (FACS)	Novus	Cat #NBP2-75923F
Anti-GPNMB (FACS)	Thermo Fisher	Cat #12-9838-42
Anti-VEGFA (FACS)	BD Biosciences	Cat #555036
Anti-mouse IgG-AF488 (FACS)	Biolegend	Cat #405319
Biological Samples		
Human monocyte TPH1 cell	ATCC®	TIB-202
Critical Commercial Assays		
Tumor Dissociation Kit human	Miltenyi Biotec	Cat #130-095-929
Fixation/Permeabilization Solution Kit	BD Biosciences	Cat #554714
SureSelectXT Target Enrichment System for Illumina Paired-End Multiplexed Sequencing Library Kit	Aglient	Cat #G9701
TruePrep DNA Library Prep Kit V2 for Illumina	Vazyme Biotech	Cat #TD503
Chromium Single Cell 3′ Library and Bead kit	10x Genomics	Cat #PN-120237
Chromium Single Cell 3′ Chip Kit v2	10x Genomics	Cat #PN-120236
Chromium i7 Multiplex Kit	10x Genomics	Cat #PN-120262
Hiseq 3000/4000 SBS kit	Illunima	Cat #FC-410-1003
Hiseq 3000/4000 PE cluster kit	Illunima	Cat #PE-410-1001
Deposited Data		
Data files for single-cell RNA sequencing (raw data)	This paper	HRA000069
Data files for single-cell RNA sequencing (processed data)	This paper	EGA S00001003449
Data files for bulk RNA sequencing (processed data)	This paper	EGA S00001003449
Data files for bulk exome sequencing (processed data)	This paper	EGA S00001003449
Oligonucleotides		
Primer: GAPDH Forward: TTGGCTACAGCAACAGGGTG	This paper	N/A
Primer: GAPDH Reverse: TCTACATGGCAACTGTGAGGAG	This paper	N/A
GPNMB sgRNA1: ATGAAAGACCTTCTGCTTACATGAGGGAGC	This paper	N/A
GPNMB sgRNA2: CACTGCGAAACCTTACCTAAAGAAGGGGTG	This paper	N/A
SLC40A1 sgRNA1: ACAGAATGTTTCAGTCATCCTGTGTGGAAT	This paper	N/A
SLC40A1 sgRNA2: AGAGCAGAACGTACTCCACGCACATGGATA	This paper	N/A
Software and Algorithms		
Scater	McCarthy et al., 2017	https://github.com/davismcc/scater
Scran	Lun et al., 2016	https://github.com/MarionLab/scran
Scanpy v1.4	Wolf et al., 2018	https://github.com/theislab/scanpy.git
Harmony	Korsunsky et al., 2018	https://github.com/pardeike/Harmony
DropletUtils	Lun et al., 2019	https://github.com/MarioniLab/DropletUtils
Scvelo		https://github.com/theislab/scvelo
Kallisto v0.44	Bray et al., 2016	https://pachterlab.github.io/kallisto
Velocyto.py	La Manno et al., 2018	https://github.com/velocyto-team/velocyto.py
RAxML v8.0	Stamatakis, 2014	https://cme.h-its.org/exelixis/web/software/raxml/index.html
Cellranger v2.3	10x Genomics	https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/ahta-is-cell-ranger

LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Zemin Zhang (zemin@pku.edu.cn).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Human Specimens

Sixteen patients who were pathologically diagnosed with liver cancer, including thirteen males and three females, were enrolled in this study after approved by the Ethics Committee of Beijing Shijitan Hospital, Capital Medical University. The samples of six of the patients (P1202, P0407, P0205, P0508, P0322 and P1116) were collected and analyzed in our previous study (Zheng et al., 2017a). All patients in this study provided written informed consent for sample collection and data analyses. Their ages ranged from 26 to 84, with a median age of 55. Patient DSD15 was pathologically diagnosed with intrahepatic cholangiocarcinoma (ICC), while other fifteen patients were diagnosed with hepatocellular carcinoma (HCC). Twelve of the patients were HBV-positive based on the HBsAg test. None of the patients was treated with chemotherapy or radiation prior to tumor resection. The stages of these patients were classified according to the guidance of AJCC version 8. Among all the patients, four were diagnosed as stage I, six as stage II, four as stage III and two as stage IV. The available clinical characteristics of these patients are summarized in Table S1. Their peripheral blood, tumors, adjacent liver tissues, lymph nodes (common hepatic artery lymph nodes) and ascites were obtained for the subsequent immune cell isolation. Tumor tissues were dissected about 2 cm from tumor edge, while the adjacent liver tissues were at least 2 cm from the matched tumors. For patients DEJ08, DEJ10 and DEJ16, because of the large tumor size, we collected two spatial sites within one tumor, with one site close to the tumor edge and thus labeled as tumor-edge and the other close to tumor core and labeled as tumor-core.

METHOD DETAILS

Sample collection and single cell processing

Fresh tumor, adjacent liver tissue and lymph node samples were cut into approximately 1 mm^3 pieces in the RPMI-1640 medium (Invitrogen) with 10% fetal bovine serum (FBS; Sciencell), and enzymatically digested with MACS tumor dissociation kit (Miltenyi Biotec) for 30 min on a rotor at 37°C, according to the manufacturer's instructions. After filtered by 70 μm Cell-Strainer (BD) in the RPMI-1640 medium (Invitrogen), the suspended cells were centrifuged at 400 g for 5 min. After removing the supernatant, the pelleted cells were suspended in red blood cell lysis buffer (Solarbio) and incubated on ice for 2 min to lyse red blood cells. The cell pellets were re-suspended in sorting buffer (PBS supplemented with 2% FBS) after washing twice with PBS (Invitrogen).

Peripheral blood mononuclear cells (PBMCs) and ascites were isolated using HISTOPAQUE-1077 (Sigma-Aldrich) solution as previously described (Zheng et al., 2017a). Briefly, 5 mL fresh peripheral blood or ascites was collected prior to surgery in EDTA anticoagulant tubes and subsequently layered onto HISTOPAQUE-1077. After centrifugation, immune cells remained at the plasma-HISTOPAQUE-1077 interface and were transferred to a new tube, with twice washing using PBS. These immune cells were re-suspended with sorting buffer.

Flow cytometry of primary PBMC

Human PBMC was obtained after ficoll centrifugation. 1-2 million cells were stained with zombie yellow (Biolegend Cat# 423103) according to the manual. After blocking with Fc Receptor Blocking Solution (Biolegend Cat# 422301), cell surface staining was performed in FACS buffer containing antibody cocktails (CD45, CD11c, CD45RA, CD1c, CD141, CLEC9A, CD123, CD303, CCR7, CD83, HLA-DR, CD3, CD19, CD56, CD16, CD14) on ice for 1 hour. After washing twice with FACS buffer, the cells were fixed using BD Cytofix/cytoperm solution (BD Bioscience Cat# 554722) 20 min on ice. Cells were washed with BD perm/wash buffer (BD Bioscience Cat# 554723), intracellular blocking with mouse IgG and Miltenyi FcR blocking, then intracellular staining was performed using PE-anti-LAMP3 (BD Biosciences Cat# 558126), or PE-isotype control (BD Biosciences Cat# 555749) for 1 hour on ice. Cells were washed twice with perm/wash buffer and then analyzed on the BD Fortessa.

In vitro stimulation of DC subsets

Fresh blood was obtained from healthy volunteers. PBMCs were isolated from heparinized fresh blood by standard density gradient centrifugation with Ficoll-Paque Plus (GE Healthcare). Subsequently, Pan-DCs were obtained by negative selection using Human Pan-DC pre-enrichment kit (Stemcell technologies). DCs were cultured in CellGenix GMP DC medium (CellGenix), and stimulated with LPS-EB Ultrapure (Invivogen) and INF-γ (Peprtotech); both at 100 ng/ml or with Poly I:C (Invivogen) at 1μg/ml. The cells were analyzed 20 hour and 12 hour post stimulation for FACS and scRNA-seq respectively. Prior to scRNA-seq, the dead cells were removed using dead cell removal kit (Miltenyi).

For membrane staining, cells (2×10^5) were incubated with conjugated mAbs for 30 min at 4°C. For intracellular staining, Fixation/Permeabilization Solution Kit (BD Biosciences) was used. Intracellular staining was performed as per manufacturer's protocol.

Briefly, the cells were fixed with fixation solution for 20 min at 4°C. The cells were then incubated with conjugated antibody or respective isotype control for 1 hour at 4°C. FACS analyses were performed using LSRFortessa (Becton Dickinson).

For flow cytometry analyses, the following mAbs were used: SP34-2 (CD3-PE-Cy7), M5E2 (CD14-PE-Cy7), 3G8 (CD16-PE-Cy7), SJ25C1 (CD19-PE-Cy7), 2H7 (CD20-PE-Cy7), B159 (CD56-PE-Cy7), G46-6 (HLA-DR-APC), 7G3 (CD123-BV605), B-ly6 (CD11c-BUV395), 1A4 (CD141-BB515), I10-1112 (CD208-PE) were purchased from BD Biosciences. L161 (CD1c-AF700) was purchased from Biolegend.

Single cell RNA-seq process for patients

Based on FACS analysis, single $CD45^+$ cells (anti-human CD45, HI30, eBiocience) of the samples of patients DSF22, DSF27, DSM22, DSM27, DSA12, DSN09, DSD15, DEJ08, DEJ10 and DEJ16 were sorted into 1.5 mL low binding tubes (Eppendorf) with 50 mL sorting buffer for droplet-based scRNA-seq, or into wells of 96-well plates (Axygen) for plated-based scRNA-seq, prepared with lysis buffer which contained 1 μL 10 mM dNTP mix (Fermentas), 1 μL 10 μM Oligo dT primer, 1.9 μL 1% Triton X-100 (Sigma) plus 0.1 μL 40 U/μl RNase Inhibitor (Takara). The sealed plates were stored frozen at −80°C.

For droplet-based scRNA-seq, single cells were processed through the GemCode Single Cell Platform using the GemCode Gel Bead, Chip and Library Kits (10x Genomics, Pleasanton) as per the manufacturer's protocol (Zheng et al., 2017a). The loaded cell numbers of DSN09, DSD15, DEJ08 and DEJ10 were 10,000 for each sample, and the cell numbers of DEJ16 was 5,000 (Table S2). The cells were then partitioned into Gel Beads in Emulsion in the GemCode instrument, where cell lysis and barcoded reverse transcription of RNA occurred, followed by amplification, shearing and 3′ adaptor and sample index attachment. Libraries were sequenced on an Illumina Hiseq 4000. The cost per cell was ∼$3, including reagents and sequencing.

For plated-based scRNA-seq, reverse transcription and transcriptome amplifications were performed following SMART-seq2 protocol. The amplified cDNA products were purified with 1x Agencourt XP DNA beads (Beckman), Following the first round of beads purification, the cDNA of each single cell was quantified with qPCR of *GAPDH*, and fragment analysis (using fragment analyzer AATI). The DNA products with high quality were further cleaned with 0.5x Agencourt XP DNA beads (Beckman) to eliminate short fragments (less than 500 bp). The concentration of each sample was then quantified with Qubit HsDNA kits (Invitrogen). The libraries were constructed with the TruePrep DNA Library Prep Kit V2 for Illumina (Vazyme Biotech). Constructed libraries were analyzed by an Illumina Hiseq 4000 with 150 bp paired-end reads. The cost per cell was ∼$60, including reagents and sequencing.

Single cell RNA-seq process for *in vitro* DCs

Cells were centrifuged, resuspended in HBSS/0.04% BSA buffer and washed twice. Cells were then filtered by a 30 μm cell strainer and stored on ice while awaiting processing. Viability (> 95%), concentration and aggregate ratio (< 3%) were determined using the NucleoCounter NC-200 (ChemoMetec, Allerod, Denmark).

Single-cell libraries were generated via the Chromium Controller and the Single Cell 3′ Reagent Kit v3 (10x Genomics, Pleasanton, CA) according to the manufacturer's instructions. In total 8,700 cells were loaded onto the Single Cell Chip B for encapsulation into droplets, aiming to capture ∼5000 cells. cDNA was amplified by 12 PCR cycles and a total of 100 ng cDNA was used for library preparation (12 PCR cycles in the index PCR). Amplified libraries were purified using the recommended double-sided size selection (0.6x and 0.8x), as well as an additional final clean-up step using 1x SPRISelect Beads (Beckman Coulter, Brea, CA) to ensure full removal of primer and adaptor dimers. Qualitative and quantitative assessments of the final libraries were performed on a 12-channel Fragment Analyzer with the High Sensitivitiy NGS Fragment 1-6000 bp assay (Agilent, Santa Clara, CA), and the Qubit 3.0 Fluorometer with the 1x dsDNA High Sensitivity kit (Life Technologies, Carlsbad, CA), respectively. Libraries with an average final concentration of ∼30 nM and fragment size of ∼450 bp were then normalized, pooled and sequenced on the NextSeq 500 using a 150 cycle NextSeq 500 High Output Kit v2.5 (Illumina, San Diego, CA) with the following run set-up parameters: 28/8/0/91, Read1/i7/i5/Read2.

In vitro stimulation of macrophages

$CD68^+$ (anti-human CD68, eBioscience) macrophages were isolated based on FACS from tumor, adjacent liver, ascites and blood of 4 additional HCC patients. Cells were incubated with autologous ascites supernatant plus 10% FBS, and then then subject to RNA-seq at 0, 24 and 48 hours based on SMART-seq2 protocol. For the 0 hour, 3 replicates of 200 cells from each sample were collected into lysis buffer based on FACS. For 24 hours and 48 hours, 3 replicates of 10-50 cells from each sample were collected into lysis buffer by mouth pipette. Constructed libraries were analyzed by an Illumina Hiseq 4000 with 150 bp paired-end reads.

Bulk RNA and DNA isolation and sequencing

Genomic DNA of tissue samples were extracted using the QIAamp DNA Mini Kit (QIAGEN) according to the manufacturer's specification. The concentrations of DNA were quantified using NanoDrop instrument (Thermo) and the qualities of DNA were evaluated with agarose gel electrophoresis. Exon libraries were constructed using the SureSelectXT Target Enrichment System for Illumina Paired-End Multiplexed Sequencing Library kit (Agilent). Samples were sequenced on the Illumina Hiseq 4000.

For bulk RNA analysis, small fragments of tissues were first stored in RNAlater RNA stabilization reagent (QIAGEN) after surgical resection and kept on ice to avoid RNA degradation. RNA was extracted using the RNeasy Mini Kit (QIAGEN) according to the manufacturer's specification. The concentrations of RNA were quantified using the NanoDrop instrument (Thermo) and the qualities of RNA were evaluated with fragment analyzer (AATI). Libraries were constructed using NEBNext Poly(A) mRNA Magnetic Isolation

Module kit (NEB) and NEBNext Ultra RNA Library Prep Kit for Illumina Paired-end Multiplexed Sequencing Library (NEB). Samples were sequenced on the Illumina Hiseq 4000.

Multi-color immunohistochemistry

Human tissue specimens were provided by Beijing Shijitan Hospital under an approved Institutional Review Board protocol. The specimens were collected within 30 min after the tumor resection and fixed in formalin for 48 hour. Dehydration and embedding in paraffin was performed following routine methods. These paraffin blocks were cut into 5 μm slides and adhered on the slides glass. Then the paraffin sections were placed in the 70°C paraffin oven for 1 hr before deparaffinized in xylene and then rehydrated in 100%, 90%, 70% alcohol successively. Antigen was retrieved by critic acid buffer (pH 6.0) in the 95°C water bath for 20 min. Endogenous peroxidase was inactivated by incubation in 3% H_2O_2 for 15 min. Following a preincubation with 10% normal goat serum to block nonspecific sites for 30 min, the sections were incubated with primary antibodies in a humidified chamber at 4°C overnight. The primary antibodies and IHC metrics used in the validation of two distinct macrophage state were: rabbit anti-human CD68 (Abcam, 1:2000), rabbit anti-human SLC40A1 (Abcam, 1:200) and mouse anti-human S100A8 (Abcam, 1:500). The primary antibodies and IHC metrics used in the validation of the potential physical interaction (co-localization) between *LAMP3*$^+$ DCs and T cell subsets were: CD3, CD4, CD8, CD80, PD-1 and PD-L1. After the sections were washed with PBS twice for 5 min, the antigenic binding sites were visualized using the GTVisionTMIIDetection System/Mo&Rb according to the manufacturer's protocol.

Genetic knockout in THP-1 cells using CRISPR-Cas9

Human monocyte THP-1 cells (ATCC® TIB-202) were transduced with lentivirus encoding Streptococci pyogenus Cas9 (Cas9) by the spinfection method to generate THP1-Cas9 cells. Briefly, cells were cultured for two weeks in RPMI 1640 (Thermo Fisher) medium supplemented with 10% fetal bovine serum (FBS), 100 U/ml penicillin and 100 μg/ml streptomycin, 2 mM L-glutamine, 12.5 mM HEPES (Thermo Fisher) and 0.05 mM β-mercaptoethanol (GIBCO). On the day of transduction, 2.5 × 10^5 cells were seeded in each well of a 24-well plate in 500 μL of culture medium. Lentiviral particles were diluted to achieve MOI of 1.0 in 500 μL culture medium supplemented with 8 μg/ml polybrene (EMD Millipore) and mixed with THP1 cells. To spinfect, cells were centrifuged at 800 g for 45 min at 32°C. After 16 −18 hours of incubation at 37°C post spinfection, cells were selected with 10 μg/ml blasticidin for 5 days. To knockout a specific gene, cells were transduced with dual single-guide RNA (sgRNA) CRoatan constructs targeting *GPNMB*, *SLC40A1* and *VEGFA* (Erard et al., 2017). Un-transduced cells were eradicated by maintaining cells under 2 μg/ml puromycin selection for 5 days.

Knockout efficiency was determined by FACS. To get the highest possible expression of GPNMB, SLC40A1 and VEGFA at the protein level, THP-1 cells were differentiated into macrophages by 150 nM PMA treatment for 24 hours and then stimulated with 20 ng/ ml lipopolysaccharide (LPS) plus 20 ng/ml γ-interferon (IFNγ) for 24 hours. The genetic knockout of surface proteins, GPNMB and SLC40A1, was measured using a standard surface staining protocol for FACS analysis. To measure knockout efficiency of secreted factor VEGFA, stimulated THP-1 cells were treated with Golgi Plug and Golgi Stop (BD Biosciences) for 4 hours and intracellular staining was performed using the BD-PermFix protocol. Anti-mouse IgG-AF488 was used as the secondary antibody to stain anti-VEGFA.

Multiplex cytokine assay using THP-1 cells

Gene edited and THP1-Cas9 (control) cells seeded at the density of 1 × 10^5 cells in 96-well flat bottom plates were differentiated into macrophages by 150 nM PMA treatment for 24 hours. Then cells were stimulated with either 100 ng/ml TLR agonist Pam3CSK4 or 20 ng/ml lipopolysaccharide (LPS) plus 20 ng/ml γ-interferon (IFNγ) for 24 hours, and supernatants from each condition were collected for analysis. Secreted chemokines and cytokines in these supernatants were quantified using the LEGENDplexTM Human Macrophage (13-plex) Panel kit (BioLegend) according to the manufacturer's instructions.

Dendritic cells migration assay

Migratory properties of immature and mature mo-DCs was measured using transwell chemotactic assay with a polycarbonate filter of 5-micron pore size in 24-well transwell chambers (Corning Costar, Cambridge, MA). ImmunoCult-ACF Dendritic Cell medium (StemCell Technologies) containing 250 ng/ml CCL19 (R&D Systems, #361-MI-025) or medium alone as a control for spontaneous migration was added to the lower chamber. Approximately 0.5-1 × 10^5 mo-DCs in 100 uL were added to the upper chamber and were incubated for 3 hours at 37°C. A 600 uL aliquot of the cells that migrated to the bottom chamber was collected. The surface antigens of migrated cells were stained with LIVE/DEAD Fixable Green dead cell stain (1:2000, ThermoFisher), HLA-DR-APC (Biolegend, clone# L243) and CD83-APC-Cy7 (Biolegend, clone# HB15e). Then cells were subjected to intracellular staining using BD-PermFix protocol and CD208-PE (1:10; BD Biosciences, clone# I10-1112). Cells were counted using BD CountBright absolute counting beads and immuno-phenotyped by FACS using BD Canto II.

Single cell RNA-seq data processing

Droplet-based sequencing data were aligned and quantified using the Cell Ranger Single-Cell Software Suite (version 2.3, 10x Genomics) against the GRCh38 human reference genome. To call real cells from empty droplets, we used the emptyDrops() function from R package dropletUtils (Lun et al., 2019), which assesses whether the RNA content associated with a cell barcode is significantly

distinct from the ambient background RNA present within each sample. Cells with $p < 0.01$ (Benjamini–Hochberg-corrected) were considered for further analysis. Quality of cells were then assessed based on three metrics step by step: (1) The number of total UMI counts per cell (library size); (2) The number of detected genes per cell; (3) The proportion of mitochondrial gene counts. Low-quality cells were filtered if the library size or the number of detected genes (in log10 scale) was below the medians of all cells minus 3 × the median absolute deviation. Cells were filtered out if the proportion of mitochondrial gene counts was higher than the median of all cells minus 3 × the median absolute deviation.

We then tried to identify doublets cells following methods as previous described (Pijuan-Sala et al., 2019). Briefly, a doublet score was computed for each cell by applying the 'doubletCells' function (scran R package) to each 10x sample separately. We next identified clusters of cells in each sample by computing the first 50 principal components across all genes, building a shared nearest-neighbor graph (10 nearest neighbors; 'buildSNNGraph' function; scran R package), and applying the Louvain clustering algorithm ('cluster_louvain' function; igraph R package; default parameters) to it. Only HVGs (calculated separately for each sample) were used for the clustering. This procedure was repeated in each identified cluster to break the data into smaller clusters, ensuring that small regions of high doublet density were not clustered with large numbers of singlets. For each cluster, the median doublet score was considered as a summary of the scores of its cells, as clusters with a high median score were likely to contain mostly doublets. Doublet calls were made in each sample by considering a null distribution for the scores using a median-centered MAD-variance normal distribution, separately for each sample. The MAD estimate was calculated only on values above the median to avoid the effects of zero-truncation, as doublet scores cannot be less than zero. All cells in clusters with a median score at the extreme upper end of this distribution (Benjamini–Hochberg-corrected $p < 0.1$) were labeled as doublets. After removing doublet cells, a total of 66,187 cells were retained for downstream analysis.

The size factor of each cell was computed using a pooling strategy implemented in the R function computeSumFactors. Normalized counts were then computed by dividing the counts for each cell by the size factor for that cell. A log2 transformation was applied to normalized counts.

SMART-seq2 sequencing data were aligned with STAR, using genome reference hg38 and annotation downloaded from NCI GDC (Genomic Data Commons) (https://api.gdc.cancer.gov/data/254f697d-310d-4d7d-a27b-27fbf767a834).Gene-specific read counts were calculated using featureCounts (Liao et al., 2014) with parameters '-M -O–fraction -Q 30'. Quality of cells were then assessed based on three metrics step by step: (1) The number of total counts per cell (library size); (2) The number of detected genes per cell; (3) The proportion of mitochondrial gene counts. Low-quality cells were filtered if the library size or the number of detected genes was below the median of all cells minus 3 × the median absolute deviation. We also removed all cells with the proportion of mitochondrial gene counts higher than 10%. The size factor for each cell was computed using a pooling strategy implemented in the R function computeSumFactors in scran package. Normalized counts/expression were then computed by dividing the counts for each cell by the size factor for that cells. A log2 transformation was applied to normalized counts. A total of 7,099 cells were retained for downstream analysis.

Processing bulk RNA-seq data

The bulk RNA expression was quantified by kallisito (version 0.44) with the same annotated transcriptome used for droplet-based data. Gene expression were calculated by aggregating transcript expression (TPM) belonging to the same gene, and the TPM of gene were then converted to read counts implemented in the R function tximport.

For the mini-bulk SMART-seq2 data of *in vitro* stimulation of macrophages, same processed as scRNA SMART-seq2 data were applied. Briefly, reads were aligned with STAR, using genome reference hg38 and annotation downloaded from NCI GDC (Genomic Data Commons) (https://api.gdc.cancer.gov/data/254f697d-310d-4d7d-a27b-27fbf767a834).Gene-specific read counts were calculated using featureCounts with parameters '-M -O–fraction -Q 30'.

Dimension reduction and unsupervised clustering

To integrate cells into a shared space from different platforms for unsupervised clustering, we used the harmony algorithm (Korsunsky et al., 2018) to integrate two datasets (SMART-seq2 and 10x). To detect the most variable genes used for harmony algorithm, we first fitted a regression trend between the variance of the $\log_2$ transformed normalized counts and the abundance of each gene using the trendVar function in scran. Next, we used the decomposeVar function from scran package to compute the biological variation for each gene (Lun et al., 2016). Genes were ordered based on their biological variation. We used genes with their biological variation > 0 and FDR < 0.01 across all samples within the same platform as informative genes for integration. We calculate a PCA matrix with 50 components using such informative genes by MultiBatchPCA() function implemented in scran package. We then feed this PCA matrix into HarmonyMatrix() function implemented in Harmony R package. As harmony can integrate over multiple covariates, we set donor and platform as two technical covariates for correction. And the corresponding theta parameter were set as 2 and 4 respectively. The batch-corrected shared space output by harmony then used to build nearest neighbor graph using scanpy. Such nearest neighbor graph was used to find clusters by leiden community detection algorithm (Traag et al., 2019), as leiden algorithm is able to yield communities that are guaranteed to be connected compared to louvain algorithm. To identify cluster-specific markers genes, differential expression testing was performed using the findMarkers function in the scran R package within the same platform. This function per-

formed pairwise Welch t tests between groups of cells, after blocking on uninteresting factors of variation (here were set as donor). Genes with the minimum rank across all pairwise comparison less than 30 were considered as cluster-specific markers. Clusters were annotated based on the expression of known maker genes.

The first round of clustering identified six major cell types including T and NK cells, myeloid cells, B cells, plasma B cells and type three ILCs. A second round of clustering was performed on T and NK cells to separate T cells and NK cells. To identify clusters within each major cell type, we performed a third round of clustering on T, NK and myeloid cells respectively. The procedure of the second and third round of clustering is the same as first round, starting from unfiltered expression matrix, including finding HVGs, calculating PCA matrix and performing integration analyses by Harmony.

For dimension reduction, we calculated UMAP, diffusion-map and force-directed graph using batch-corrected shared space output by Harmony (Korsunsky et al., 2018). All reduced dimensions were calculated by scanpy (Wolf et al., 2018). Partition-based graph abstraction (PAGA) were calculated by scanpy.

Comparing DCs in different cancer types

To compare our scRNA-based dendritic cell transcriptome with FACS-based breast cancer dendritic cell transcriptome (Michea et al., 2018), we aggregated counts across all cells for each gene in each DC cluster within in each sample to obtain 'pseudo-bulk' samples. For instance, assuming there were 10 cells in the tumor of Donor 1 were identified as *LAMP3*$^{+}$ DC cells, for each gene, counts of these 10 cells would be summed up. We then use the EdgeR package to compute log CPM of each gene as normalized gene expression.

For breast cancer dendritic cell transcriptome, we downloaded the original sequence fastq file from SRA PRJNA380940. The same procedure as SMART-seq2 data analysis was applied to the downloaded data to generate counts table. Log CPM of each gene were calculated as normalized gene expression.

To identify gene signatures associated with particular type of dendritic cell, we performed pairwise comparison across different types of dendritic cells using the Genewise Negative Binomial Generalized Linear Models with Quasi-likelihood Tests (glmQLFit) of EdgeR R package. Only differentially expressed genes with FDR of < 0.01 and logFC > 0compared to all the other types of dendritic cells were considered as DC cell type-specific signatures.

Pairwise comparison across scRNA-based pseudo-bulk data is also performed to get differentially expressed genes between each two types of dendritic cells. A self-contained test implemented in mroast function in R limma package was performed to test whether any gene signatures are enriched in the differentially expressed genes (or tend to be differentially expressed) (Ritchie et al., 2015).

Cell similarity analysis across tissues

To quantify the cell similarities among different tissues, we projected the combined cell-by-gene expression matrix onto a shared low-dimensional PC space. We used the union of HVG sets identified within each tissue for PCA projection. For each cell within each tissue, we query for its nearest neighbor using the R function queryKNN() among cells of all other tissues in the low-dimensional space spanned by the top 50 PCs. The potential origin of each cell was estimated by the tissue origin of its nearest neighbor cell. To determine the statistical significance of tissue origin for each cell type, we performed permutation test by randomly shuffling the tissue labels for all the nearest neighbor cells 1000 times.

TCR analysis

The TCR sequences for each single cell from the Smart-seq2 data were assembled by the TraCeR method (Yusufi et al., 2017), leading to the identification of the CDR3 sequence, the rearranged TCR genes and their expression abundance (transcripts per million, TPM). Only in-frame TCR alpha-beta pairs were considered to define the dominant TCR of a single cell. Non-T cells were negative control here. Each unique dominant alpha-beta pair was defined as a clonotype. If one clonotype was present in at least two cells, this clonotype would be considered clonal, and the number of cells with such dominant alpha-beta pair indicated the degree of clonality of the clonotype.

Lineage tracing based on mitochondrial mutation

For SMART-seq2 data, mitochondrial mutations of each cell were determined following Ludwig et al. (2019). Specifically, for each mitochondrial genome position, the allele frequency (AF) of a base *b* and the total coverage of this position was computed. For 10x data, we use VarTrix (https://github.com/10XGenomics/vartrix) to calculate the alternative allele frequency and the coverage of each position in the mitochondrial chromosome.

To build phylogenetic trees based on the mitochondrial mutations, we retained mitochondrial genome position with coverage > 20 and at least 85% cells within the interest cell populations expressed. Because of heteroplasmy of mitochondria genome, a cutoff of alternative allele frequency > 0.1 was set to identify subclone mutations in the mitochondrial genome. All bases were concatenated into a unique DNA sequence for each cell. Multiple alignment was then constructed by aligning all sequences by position. After removing identical sequences, we calculated phylogenetic trees using RAxML with parameters '-f a -p 1000 -x 319 -m GTRCAT'. and estimated stable cell lineages (bootstrap value > 30). The lineage relationship reviewed by phylogenetic trees does not reflect the degree of proliferation. Outgroup of the phylogenetic trees were set using the original reference sequence from hg38.

RNA velocity-based cell fate tracing

To perform the RNA velocity analysis, the spliced reads and unspliced reads were recounted by the velocyto python package based on previous aligned bam files of scRNA-seq data. The calculation of RNA velocity values for each gene in each cell and embedding RNA velocity vector to low-dimension space were done by following the scvelo python pipeline.

We calculated the velocity-based cell transition matrix by transition_matrix() function from scvelo, of which the element was the Pearson correlation coefficient between the velocity vector and cell state difference vectors of the column cell as previously described (La Manno et al., 2018). We estimated the destination of a cell by identifying the highest correlation value. Then Fisher's exact test was performed on 2x2 cluster-by-cluster or cluster-by-tissue contingency tables to test the fate destinations of interested cell clusters. To infer the migration directions of macrophages and dendritic cells we first constructed partition-based graph abstraction for macrophage population and dendritic cell populations respectively, and then oriented edges among cell populations using RNA velocity information as previously described (Wolf et al., 2019).

Cell-cell interaction analysis

The cell-cell interaction analysis was based on the expression of immune-related receptors and ligands. The gene list contained 168 pairs of well-annotated receptors and ligands, including cytokines, chemokines and co-stimulators (Table S6; Chen and Flies, 2013). Cell types that had at least 5 cells and occupied $\geq$ 10% of immune cells from either tumor or adjacent liver tissues were considered. We estimated the potential interaction between two cell types mediated by a specific ligand-receptor pair by the product of the average expression levels of the ligand in one cell type and the corresponding receptor in the other cell type. To examine the statistical significance of the estimated interaction intensity, permutations were applied on the cell type tags of individual cells for 1000 times, and the p value was estimated by the number of permutations that had interacting intensity larger than the real value. Adjusted p value by Bonferroni correction was calculated for multiple testing correction across the hundreds of ligand-receptor pairs. If a pair of ligand and receptor had a value of interacting intensity larger than 10 (in SMART-seq2 data; in 10x data, the cut-off is 0.38, the same percentage as SMART-seq2 data), and an adjusted p value less than 0.01 between two cell types, we defined this ligand-receptor pair as a potential molecular axis mediating interactions between the two cell types. For a given pair of ligand and receptor, cell types with the average expression level of either the ligand or the receptor less than 1 ($\log_2$(Normalized Counts) < 1) were filtered.

TCGA data analysis

The TCGA hepatocellular carcinoma (LIHC) data were used to test the correlation of selected genes and patient survival. The gene expression data and the clinical data were downloaded from UCSC Xena (http://xena.ucsc.edu/). The feature genes used for MDSC-like and TAM-like states were based on differentially expressed genes (FDR < 0.01, $\log_2$ (FC) > 2) of MDCS- / TAM-like macrophages versus other macrophage subsets. To correct the effect of $CD45^+$ cell levels within each sample, the expression of selected genes in tumor were divided by the expression of *PTPRC* (CD45). The statistical analysis was performed by GEPIA2 (Tang et al., 2019). For correlation analysis, *LAMP3*$^+$ DC signatures were based on differentially expressed genes among all DC subsets (Top Rank $\leq$ 30). Treg and T exhausted signatures were based on Zheng et al. (2017a). We calculated the mean of the expression (TPM) for all signature genes as signature score. To correct the effect of $CD45^+$ cell levels within each sample, signature score was divided by that of the expression of *PTPRC* (CD45). Spearman correlation between signatures were calculated cor() function in R.

DATA AND CODE AVAILABILITY

The accession numbers for the sequencing raw data and processed data in this paper are GSA (Genome Sequence Archive in BIG Data Center, Beijing Institute of Genomics, Chinese Academy of Sciences): HRA000069 and EGA: EGAS00001003449 respectively. Analysis of such HCC data can also be found at http://cancer-pku.cn:3838/HCC/.

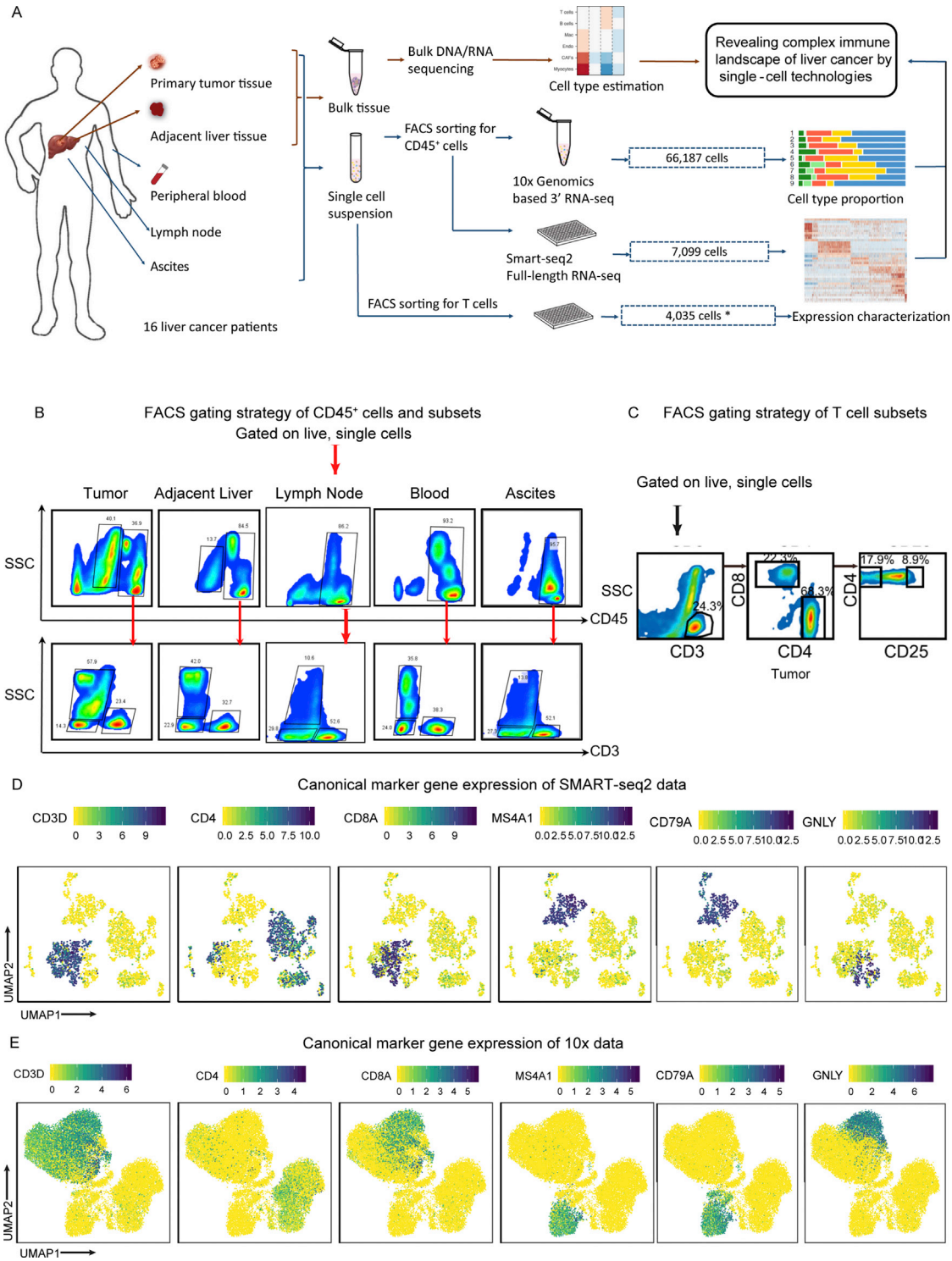


Figure S1. Study Design and Basic Information of the Single-Cell RNA-Seq Data Generated by Two Platforms, Related to Figure 1

(A) Scheme of the overall study design. Two technologies (droplet-based 10x Genomics and plate-based SMART-seq2) of scRNA-seq were applied to CD45$^+$ cells and T cell subsets derived from blood, tumor, adjacent liver tissue, LN and ascites, and the output data of two technologies were integrated and combined for downstream analysis. *, the data have been published in our previous study (Zheng et al., 2017a).

(legend continued on next page)

(B) FACS gating strategy of $CD45^+$ cells sorting. For 10x scRNA-seq, we collected cells based on CD45 antibody (upper row). For SMART-seq2 scRNA-seq, we collected cells based on CD45, CD3 antibodies and cell size information (bottom row).
(C) FACS gating strategy of T cell subsets sorting.
(D) and (E) show the canonical marker genes expression for cell type annotation of patient DSN09 based on SMART-seq2 and 10x data respectively.

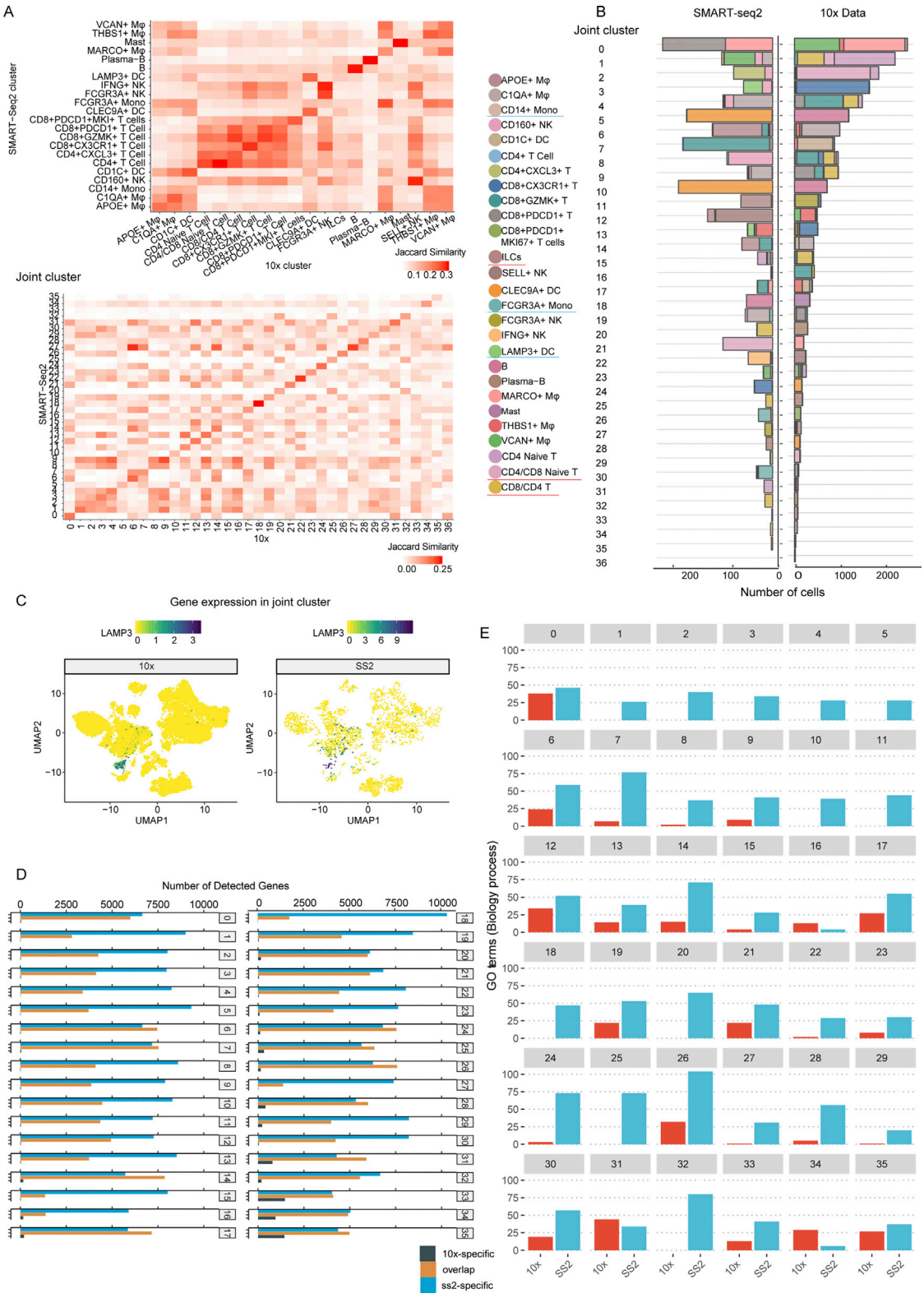


Figure S2. Comparison and Integration of Full-Length and 3′ Single-Cell RNA-Seq Data, Related to Figure 1

(A) Heatmap showing the Jaccard similarity of SMART-seq2 and 10x data in separate cluster (upper), and SMART-seq2 and 10x data in joint clusters (bottom). Jaccard similarity was calculated using top 500 cell type specific markers.

(legend continued on next page)

(B) Bar plots showing the distribution of SMART-seq2 and 10x clusters in joint clusters. Red lines represent 10x-specific clusters, and blue lines represent SMART-seq2-specific clusters.
(C) *LAMP3* expression of joint clusters.
(D) Bar plots showing number of detected genes in each cluster.
(E) Gene ontology (Biology process) term numbers generated by the differentially expressed genes of 10x and SMART-seq2 data respectively.

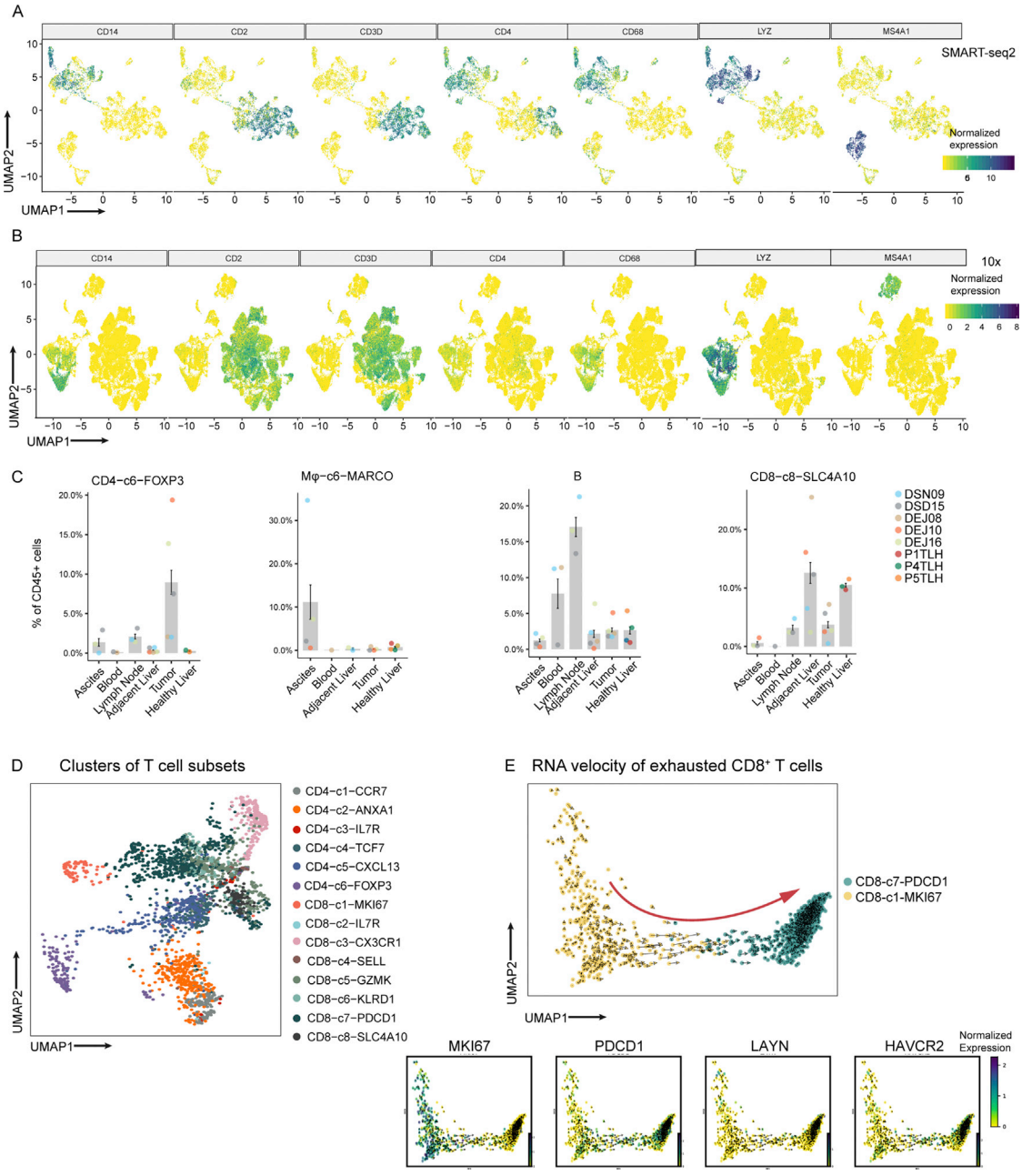


Figure S3. Cluster Characterization of the Global Landscape and T Cells Based on 10x and SMART-seq2 Data, Related to Figure 1

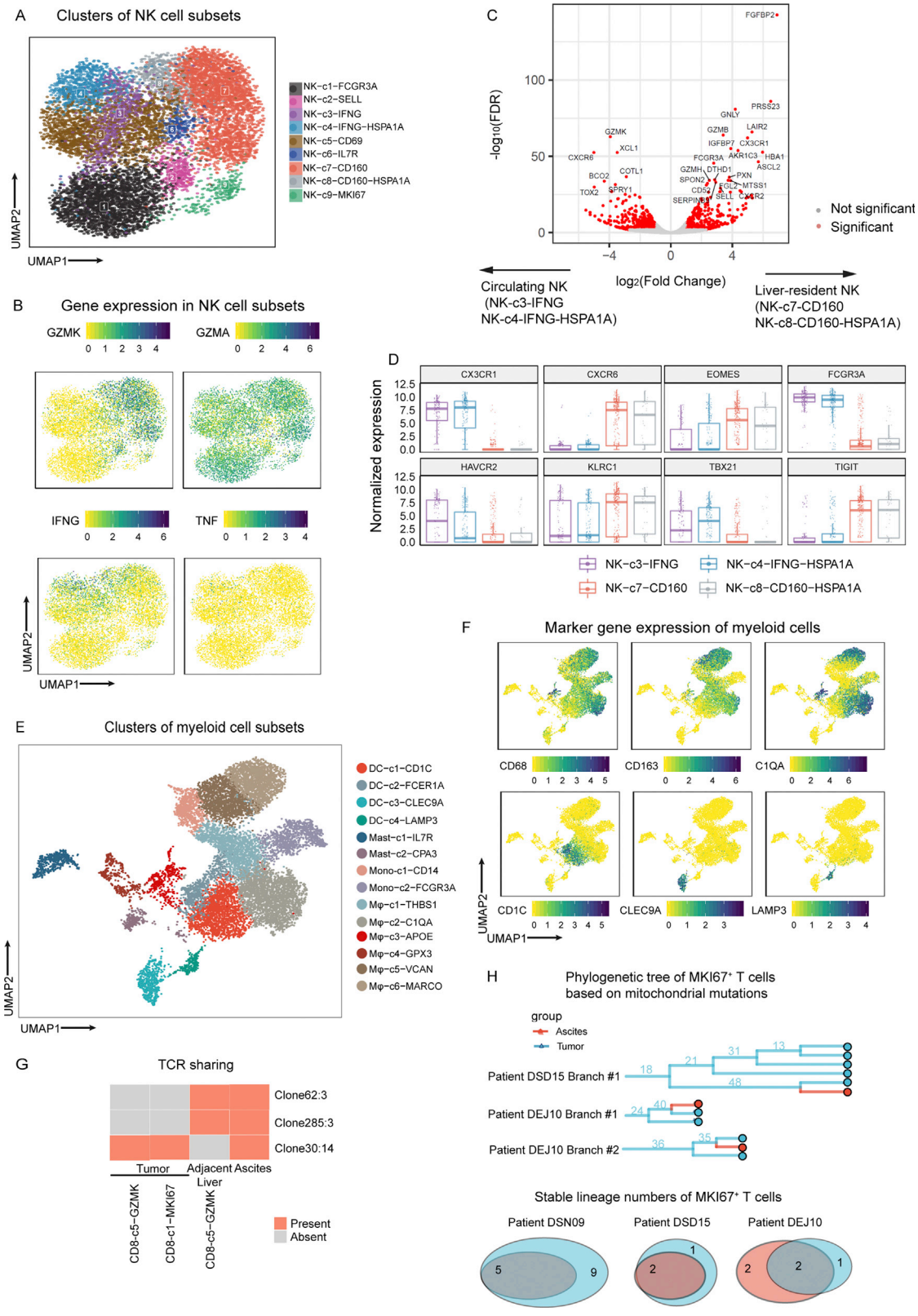
(A) Expression of canonical marker genes for cell type annotation based on SMART-seq2 data.

(B) Expression of canonical marker genes for cell type annotation based on 10x data.

(C) Bar plots showing different tissue enriched populations, together with "healthy liver" controls (P1TLH, P4TLH and P5TLH) (MacParland et al., 2018).

(D) The UMAP projection of T cells based on SMART-seq2 data. Each dot represents an individual cell.

(E) RNA velocities are visualized on the UMAP projection of CD8-c1-MKI67 and CD8-c7-PDCD1 populations, colored by clusters.



(legend on next page)

Figure S4. Cluster Characterization of NK and Myeloid Cells Based on SMART-seq2 and 10x Data, Related to Figures 2 and 3

(A) The UMAP projection of NK cells, showing the formation of 9 sub-clusters. Each dot represents an individual cell, colored according to cell cluster number.

(B) The UMAP plot showing the expression of cytotoxic genes in NK cell populations (*GZMK, GZMA, IFNG* and *TNF*).

(C) Differentially expressed genes between circulating NK cells (NK-c3-IFNG and NK-c4-IFNG-HSPA1A) and liver-resident NK cells (NK-c7-CD160 and NK-c8-CD160-HSPA1A). Red dots are significant genes (log_2 (FC) > 1, FDR < 0.01).

(D) Boxplot of marker gene expression in cNK and lrNK clusters.

(E) The UMAP projection of myeloid cell clusters, showing the formation of 14 clusters. Each dot represents an individual cell, colored according to cell cluster number.

(F) The UMAP plots showing the expression of signature genes in myeloid cell populations.

(G) TCR sharing between cell clusters in tumor or adjacent liver tissues, and cells in ascites. Rows represent different clonotypes, and columns represent different cell clusters or tissues.

(H) Branches of phylogenetic trees of proliferative T cells in ascites and tumor constructed by mitochondrial mutations. Venn diagram shows the number of stable lineages (Bootstrap > 30).

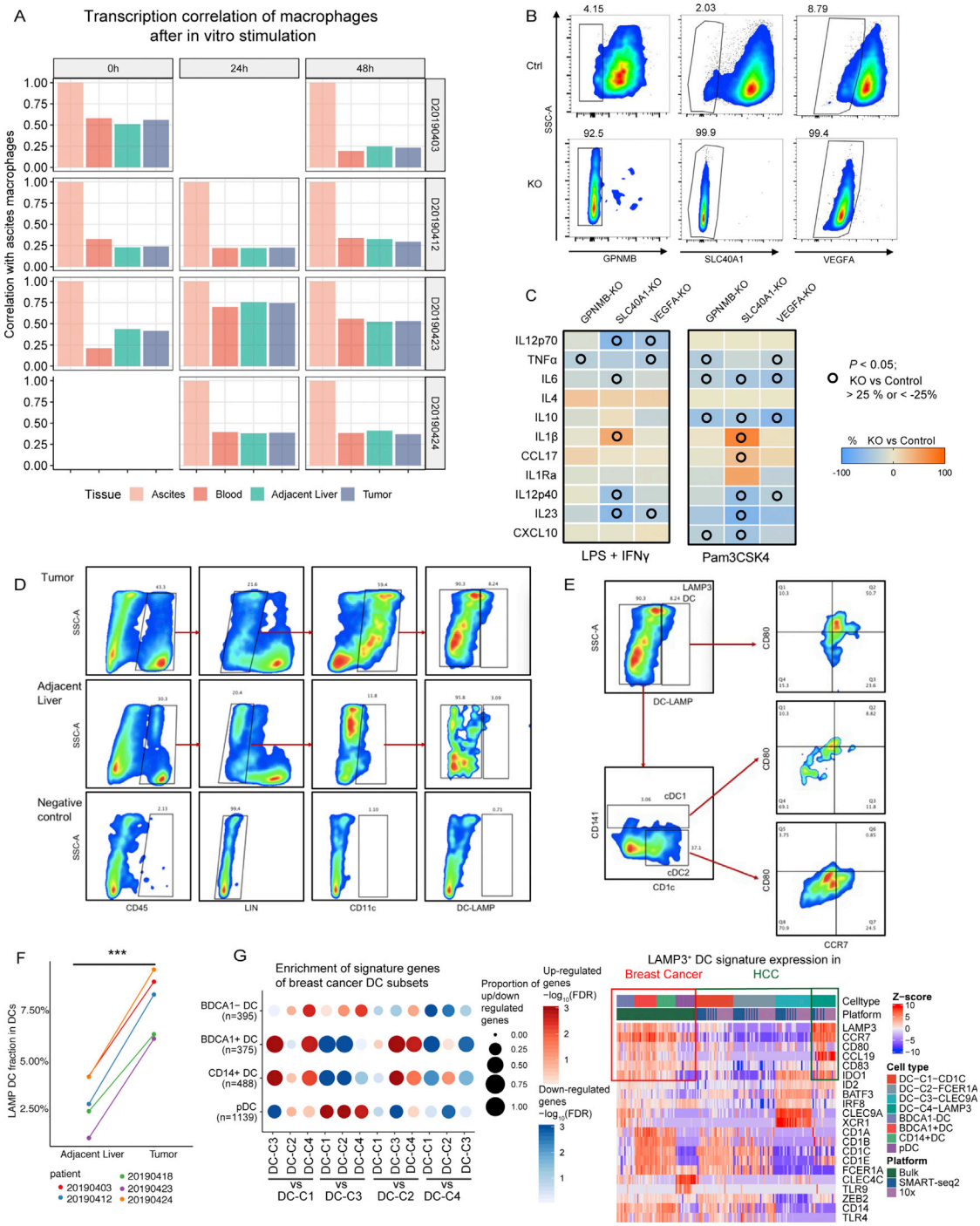


Figure S5. CRISPR Knockout of THP-1 Cells and Validation of the Existence of *LAMP3*$^+$ Population in HCC, Related to Figures 2, 3, and 4

(A) Bar plot showing the expression similarity (measured by Spearman correlation) of macrophages from different tissues to those from ascites with *in vitro* stimulation by autologous ascites supernatant.

(B) FACS analysis showing the knockout efficacy of three genes *SLC40A1*, *GPNMB* and *VEGFA* in THP-1 cells.

(C) Heatmap showing the cytokine alteration after CRISPR knockout in (B), compared with unedited-control (Student's t test).

(D) FACS gating strategy of DC subsets from tumor and adjacent liver tissues of HCC patients.

(E) FACS analysis showing the higher expression of CCR7 and CD80 in LAMP3$^+$ DCs than cDC1 and cDC2.

(F) Dot plot showing the higher proportion of LAMP3$^+$ DCs in tumors than adjacent liver tissues. *p < 0.05; **p < 0.01; ***p < 0.001; Student's t test.

(legend continued on next page)

(G) The left plot shows the enrichment of up/downregulated genes of different DC subset pairs in our data in those DC subsets generated based on FACS (Michea et al., 2018). The bubble size represents the proportion of genes in the gene signature (y axis) expressed in the corresponding cell subsets (x axis), and the color bar represents the p value of the enrichment/depletion. The right heatmap shows the signature genes of each cluster of the two datasets (HCC and breast cancer).

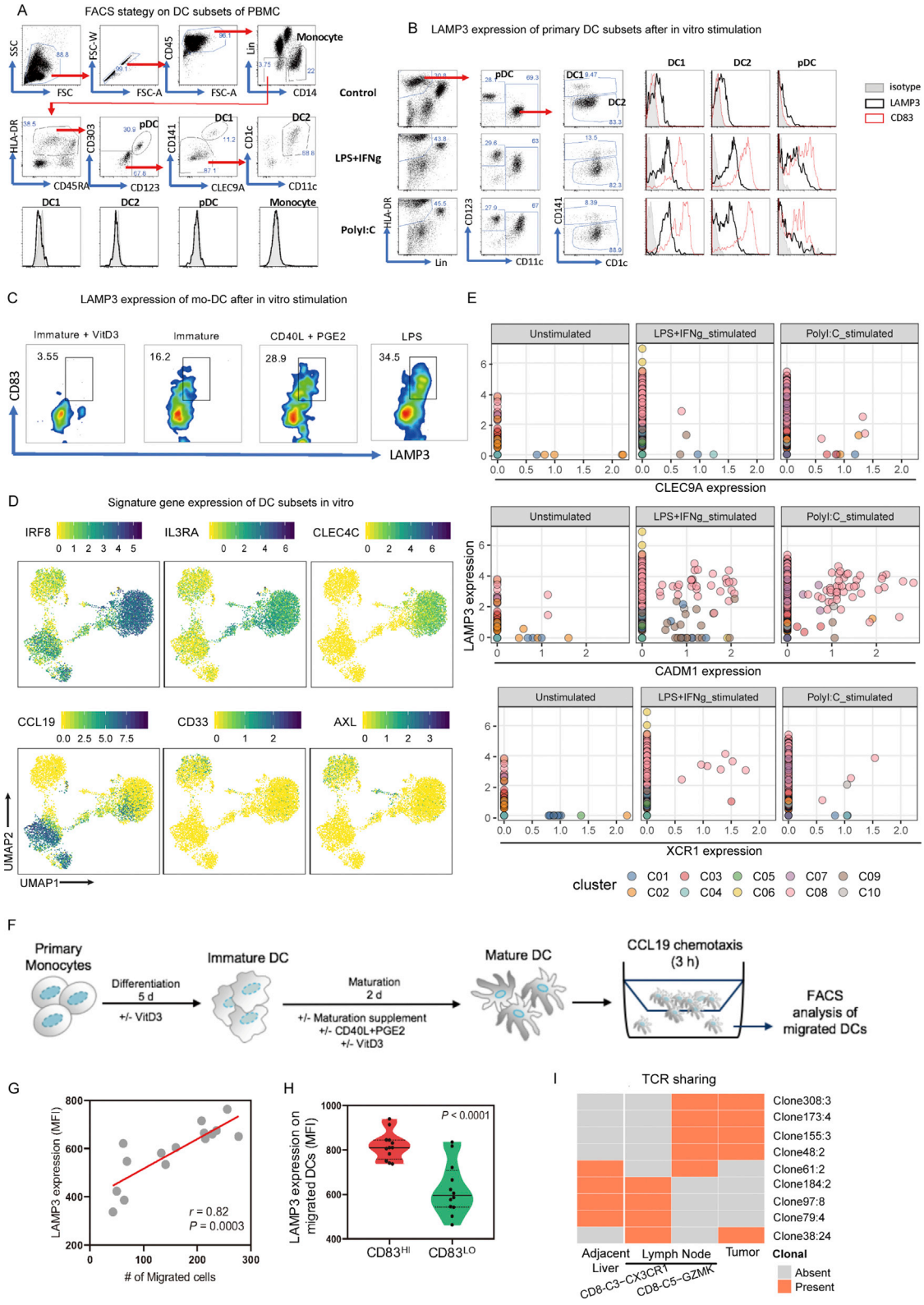


Figure S6. *In Vitro* Experiment Showing the Maturation and Migration Properties of DC Subsets, Related to Figure 4, 5, and 6

(A) FACS gating strategy of DC subsets from PBMCs of healthy donors, showing no expression of *LAMP3*.

(B) Histograms showing the expression of LAMP3 in different DC subsets after *in vitro* stimulations (LPS+IFNγ or PolyI:C).

(legend continued on next page)

(C) LAMP3 expression in mo-DC subsets under different stimuli conditions.
(D) UMAP plots showing the expression of signature genes based on literature (Villani et al., 2017) in PBMC-derived DCs with/without stimulations.
(E) The expression relationship of *LAMP3* with cDC1 marker genes *CLEC9A*, *XCR1* and *CADM1* from *in vitro* DC scRNA-seq.
(F) Transwell experiment design.
(G) The LAMP3 expression is positively associated with the number of migrated cells in the migration assay of (F) (Pearson's correlation analysis).
(H) LAMP3 is highly expressed in $CD83^{hi}$ migrated cells (Student's t test).
(I) TCR sharing between cell clusters in tumor and adjacent liver tissues and cells in LN. Rows represent different clonotypes, and columns represent different cell clusters or tissues.

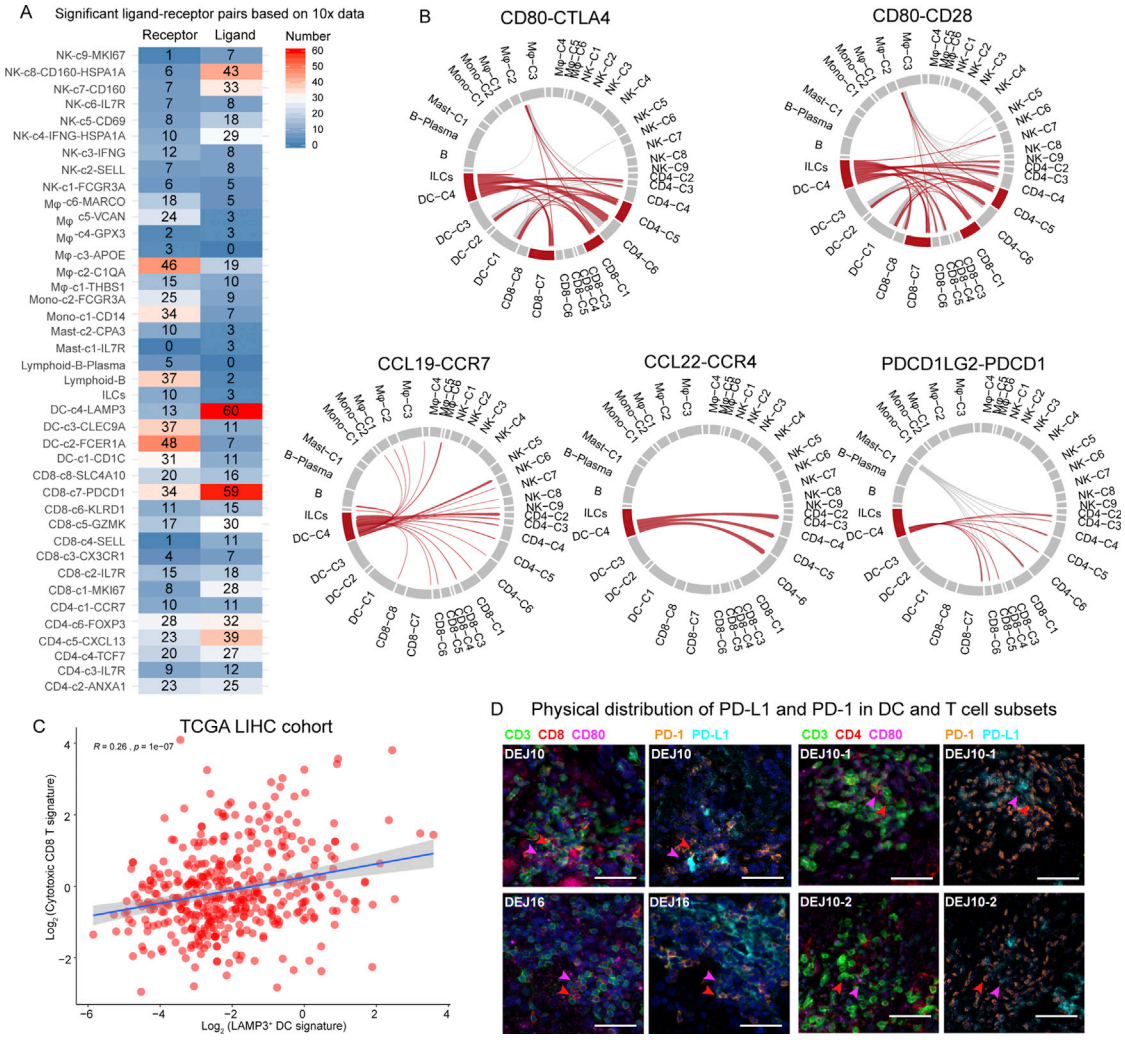


Figure S7. Predicted Cell-Cell Interactions Based on Ligand-Receptor Interactions, Related to Figure 7

(A) Number of ligands and receptors involving in significant ligand-receptor pairs for each cell cluster based on 10x data.

(B) Circos plots showing the predicted cell-cell interactions of CD80-CTLA4 and CD80-CD28, CCL19-CCR7, CCL22-CCR4, PDCD1LG2-PDCD1 axes among different clusters.

(C) Correlation of the *LAMP3*$^{+}$ DC gene signature with that of cytotoxic T cells based on TCGA LIHC data, normalized by *CD45* expression. Each dot represents a patient (Pearson's correlation analysis).

(D) Multicolor IHC staining to validate the interaction pairs of PD-L1 and PD-1 in *LAMP3*$^{+}$ DCs and *CD8*$^{+}$/*CD4*$^{+}$ T cells, exemplified by patient DEJ10 and DEJ16. The scale bar represents 30 μm.