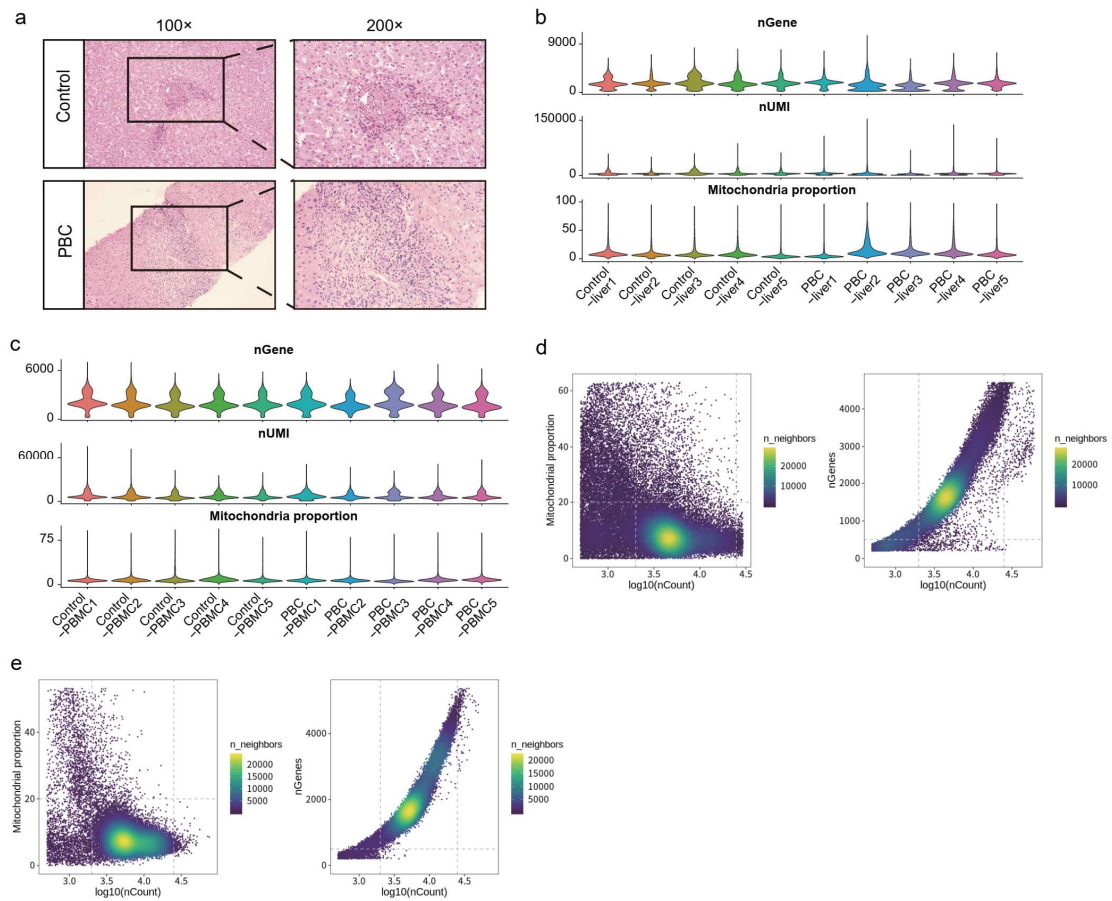
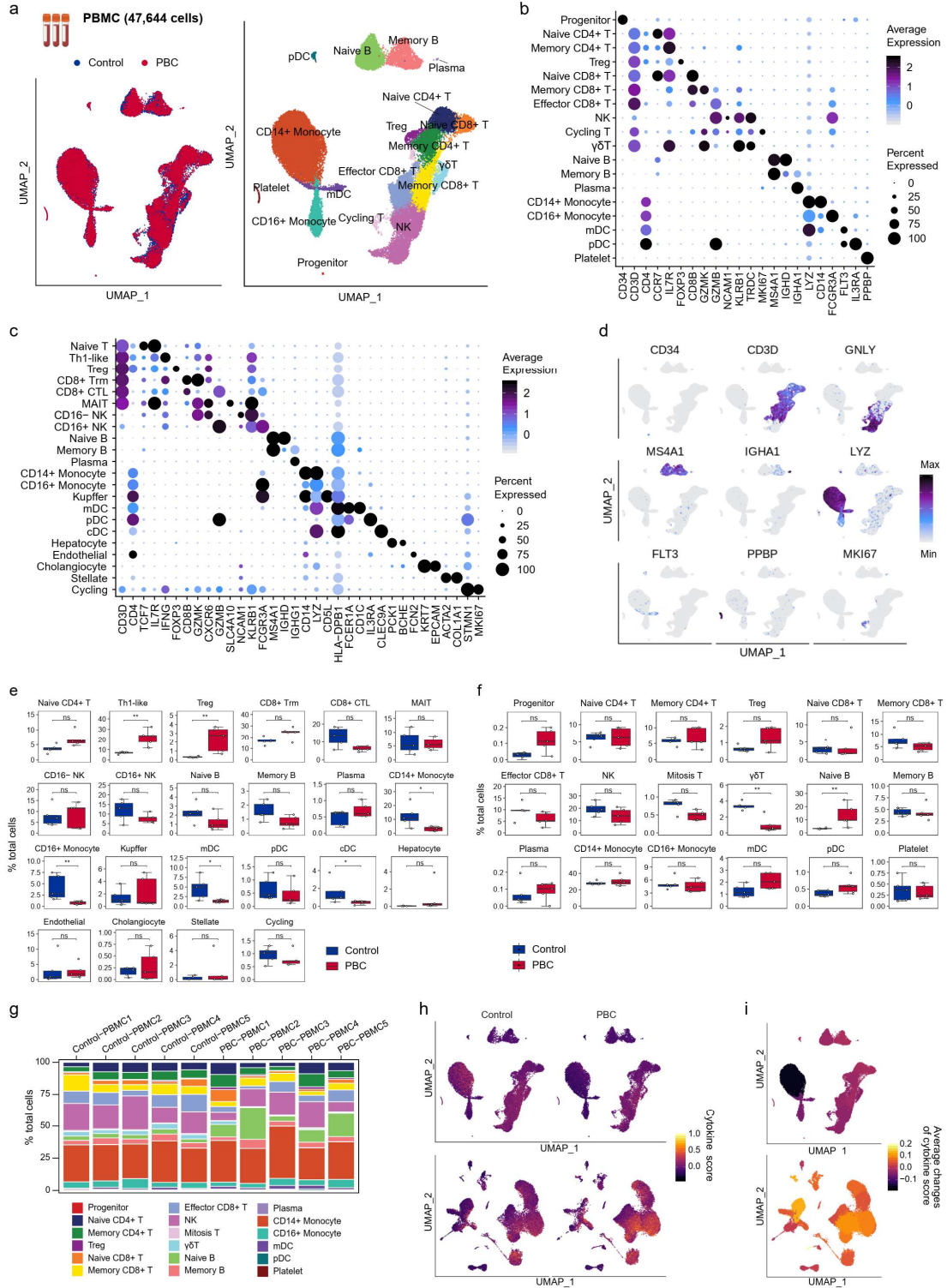
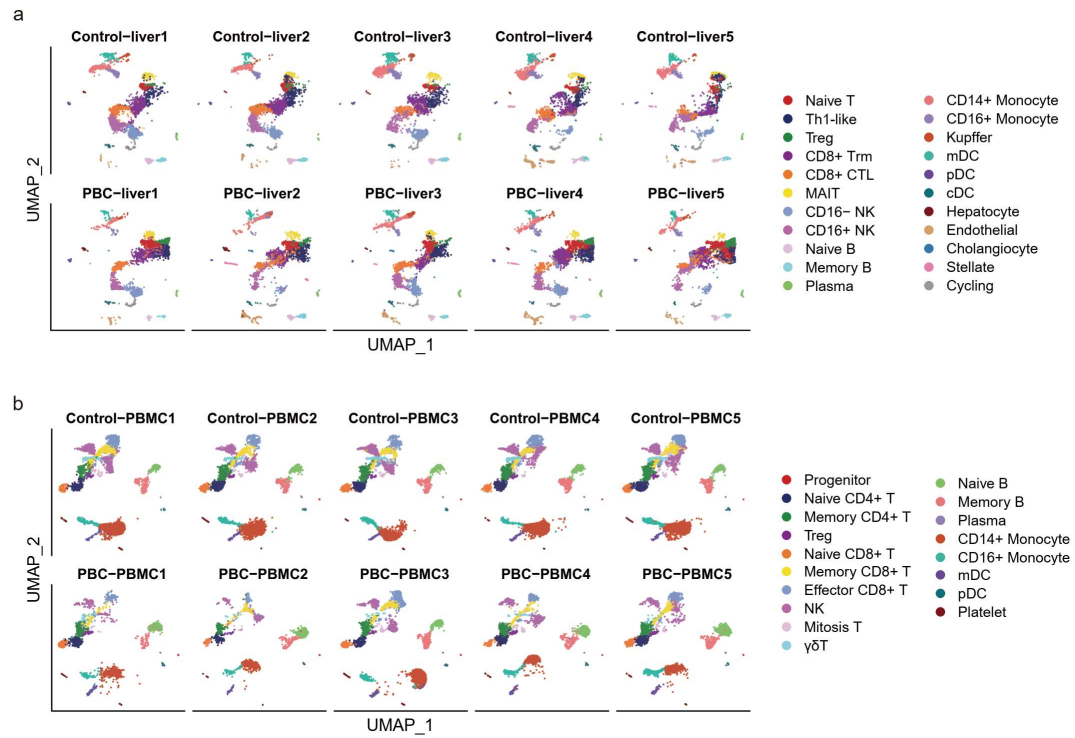




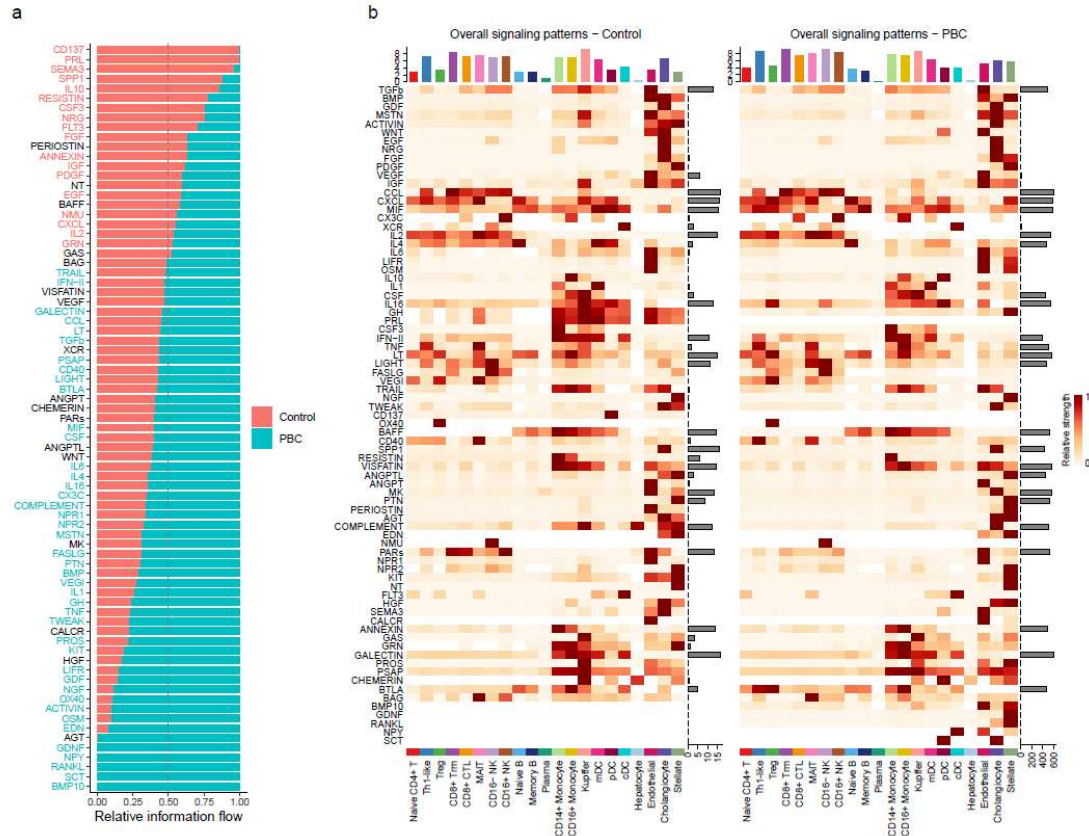
Supplementary Figure. 1 | Representative images of PBC liver and quality control of human scRNA-seq data. **a** Representative hematoxylin-eosin (H&E) staining of normal liver tissue (top) and PBC liver (bottom), the latter shows typical portal inflammation of PBC. **b,c** Violin plots showing the gene number, total UMI counts and mitochondrial gene percentage of liver cells (b) and PBMCs (c) from control and PBC patients. **d,e** Scatter plot showing gene number, UMI number, mitochondrial gene percentage of each cell from the liver (d) or PBMCs (e). Individual point represents one cell and is colored by the number of neighboring points. Dashed lines indicate the filtering threshold for high-quality single-cell data (Gene number ≥ 200 , UMI number $\geq 2,000$ and $\leq 25,000$, mitochondrial gene percentage ≤ 20).



Supplementary Figure. 2 | Marker genes of major cell types and cell subset ratio statistics. **a** UMAP showing the distribution (left) and the annotated cell population (right) of human PBMCs from PBC or control. **b,c** Dot plot displaying expressions of marker genes for each PBMC (b) and liver cell (c) population. **d** UMAP showing the expression of marker genes of major cell types in PBMC scRNA-seq data. **e,f** Box plots showing percentage of each cell subset in total cells in the liver (e) and PBMC (f) scRNA-seq data. The center of the boxplot represents the median, and the bounds of the boxplot represent the 25th and 75th quantiles. Each dot represents one sample. *FDR<0.05, ** FDR<0.01, *** FDR<0.001. **g** Bar plot showing the cell ratio of PBMC cell subsets in each sample. **h** UMAP showing the cytokine score of each cell in PBMC (top) and liver (bottom) data. **i** UMAP showing average changes of the cytokine score of each cell subset in PBMC (top) and liver (bottom) data.



Supplementary Figure. 3 | UMAP showing the scRNA-seq data from human liver (a) and PBMC (b) without batch effects correction.



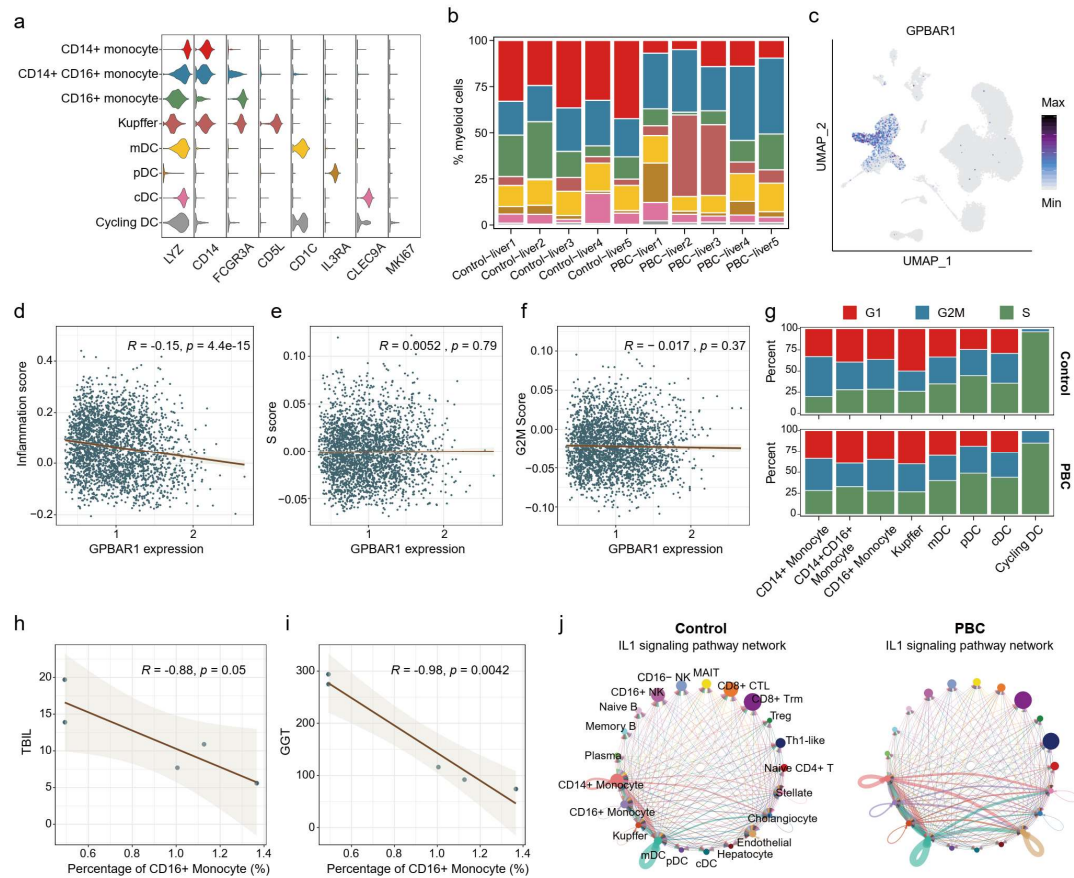
Supplementary Figure. 4 | Cell-cell interaction analysis. **a** Bar plot showing relative strength of information flow/interaction of each signaling pathway in control and PBC. The signaling pathways colored red are enriched in Control and those colored green are enriched in PBC. **b** Heatmap showing the overall signaling pattern (aggregating outgoing and incoming signaling together) of each cell population in control and PBC.



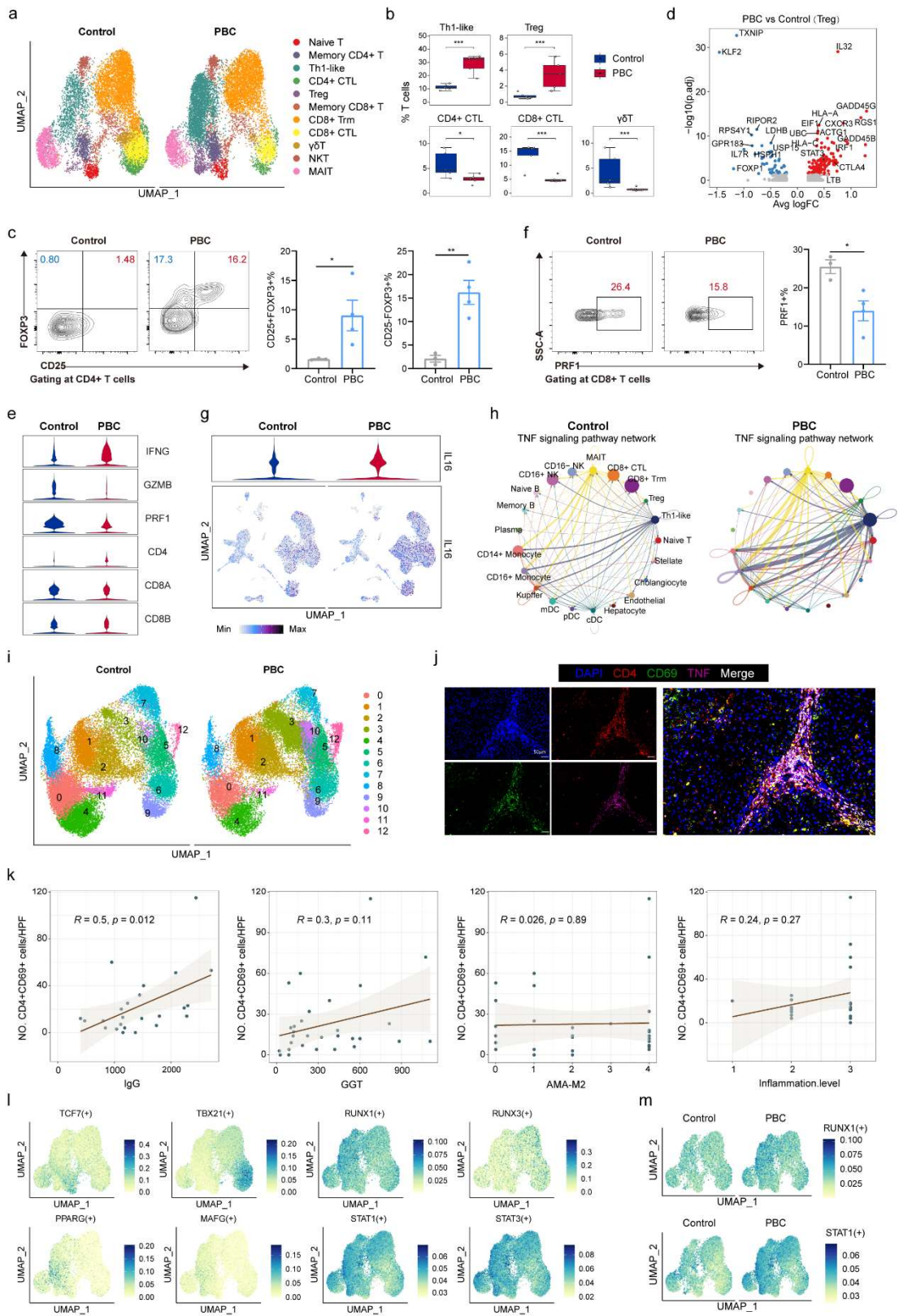
Supplementary Figure. 5 | The inflammatory and cytokine score in each cell subset of the liver and PBMC samples from both PBC patients and the control group. a,b The inflammatory (a) and cytokine score (b) in each cell subset of liver samples from both PBC patients and the control group. c,d The inflammatory (c) and cytokine score (d) in each cell subset of PBMC samples from both PBC patients and the control group. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$



Supplementary Figure. 6 | Characteristics of non-immune cells in the liver of PBC patients. **a** UMAP showing the distribution of annotated non-immune cell subsets from control and PBC groups. **b** UMAP showing the distribution of CX3CR1 and/or GZMB expressing cells in the liver scRNA-seq data. **c** Bubble plot showing the increased signaling pathways between cholangiocytes and other cell types. **d** Schematic diagram depicting the abnormal signal activation in cholangiocytes and activated stellate cells in PBC. Supplementary Fig. 6d created in BioRender. Li, J. (2022) BioRender.com/m28t560.

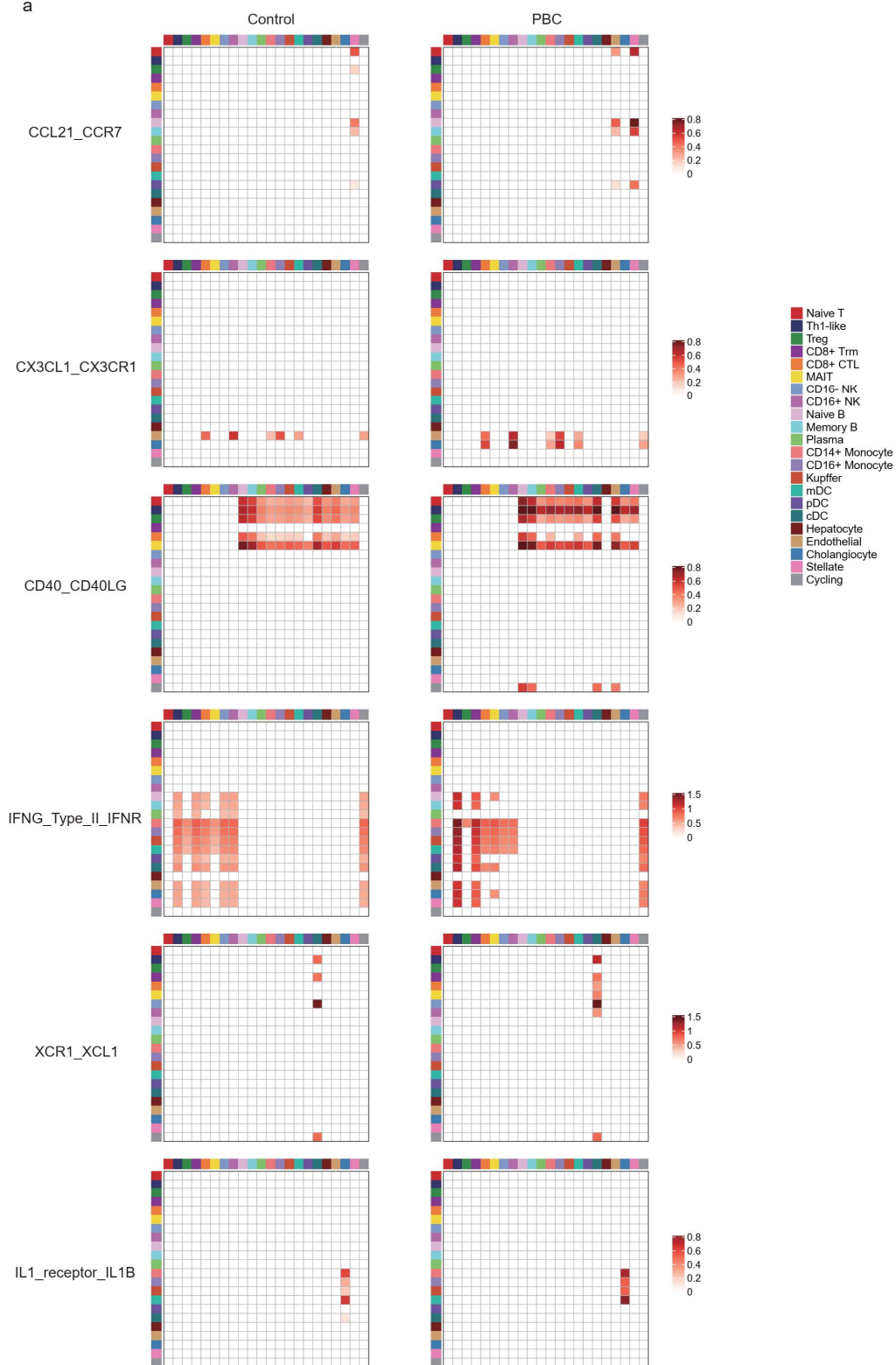


Supplementary Figure. 7 | Characteristics of myeloid cells in the liver of PBC patients. **a** Violin plot showing the expression of marker genes in subpopulations of myeloid cells in the liver. **b** Stacked bar plot showing percentages of myeloid cell subsets in each sample. Colors represent cell subpopulations shown in figure 3a. **c** UMAP showing the distribution of GPBAR1 expressing cells in the liver scRNA-seq data. **d-f** Scatter plot with linear regression line showing the correlation between the inflammatory score (d) S score (e), G2M score (f) and GPBAR1 expression in monocyte. S and G2M scores indicate the possibility of cells in the S and G2M phase, respectively. **g** Bar plot showing the percentage of each cell cycle phase in each myeloid cell subset. **h,i** Scatter plot with linear regression line showing the correlation between TBIL (h) or GGT (i) and the percentage of CD16⁺ monocyte in scRNA-seq data. **j** Circle diagrams showing the cell-cell interaction difference of IL1 signaling pathways between PBC and control.

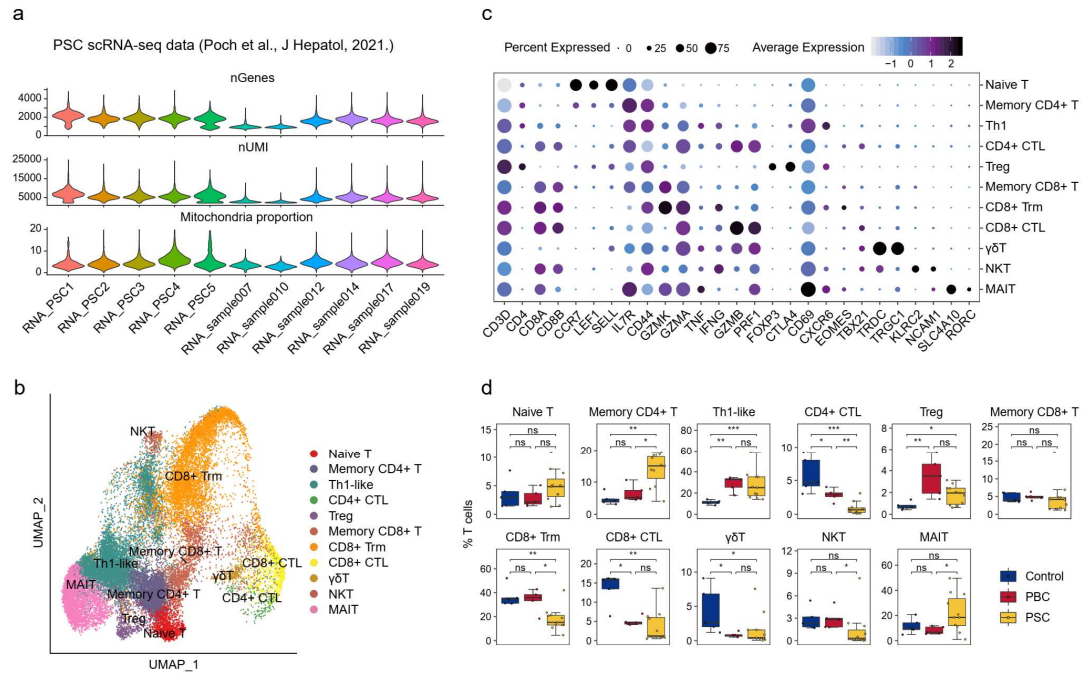


Supplementary Figure. 8 | Characteristics of hepatic T cells in PBC patients. **a** UMAP showing distribution of annotated T cell subsets from control and PBC groups. **b** Box plots showing percentages of T cell subsets in the liver. **c** Representative flow plots and summarized data of proportion of CD25⁺FOXP3⁺ CD4⁺ T cells and CD25⁺FOXP3⁺ CD4⁺ T cells from control group patients and PBC patients' liver biopsy samples (control $n=3$, PBC $n=4$), data are presented as means $\pm$ s.d. Significance was determined by two-tailed Student's t-test, $*P<0.05$, $**P<0.01$, $***P<0.001$, $****P<0.0001$. Source data are provided as a Source Data file including exact P . **d** Scatter plot showing the differentially expressed genes in Treg cells from PBC group relative to Control group. Red dots represented up-regulated genes. Blue dots represented down-regulated genes. **e** Violin plots showing the expression of IFNG, GZMB, PRF1, CD4, CD8A, CD8B in all T cells in control and PBC liver data. **f** Representative flow plots and summarized data of proportion of PRF1⁺CD8⁺ T cells from control group patients and PBC patients' liver biopsy samples (control $n=3$, PBC $n=4$), data are presented as means $\pm$ s.d. Significance was determined by two-tailed Student's t-test, $*P<0.05$, $**P<0.01$, $***P<0.001$, $****P<0.0001$. Source data are provided as a Source Data file including exact P . **g** Violin plot (top) and UMAP (bottom) showing the expression of IL16 in total liver cells from control and PBC. **h** Circle diagrams showing the cell-cell interaction difference of the TNF signaling pathway between control and PBC. **i** UMAP showing the distribution of control and PBC T cells from combined liver and PBMC T cells. **j** Representative immunofluorescent images ($n=5$) of DAPI (blue), CD4 (red), CD69 (green), and TNF (purple) in the liver of PBC patients. **k** Scatter plot with linear regression line showing the correlation between the percentage of CD4⁺CD69⁺ cells in the liver of PBC patients and pathological features (GGT, IgG, and Inflammation level). GGT: γ -glutamyl transferase. **l** UMAP showing AUC score of TCF7(+), TBX21(+), RUNX1(+), RUNX3(+), PPARG(+), MAFG(+), STAT1(+), STAT3(+). **m** UMAP showing AUC score of RUNX1(+) and STAT1(+) in control and PBC groups.

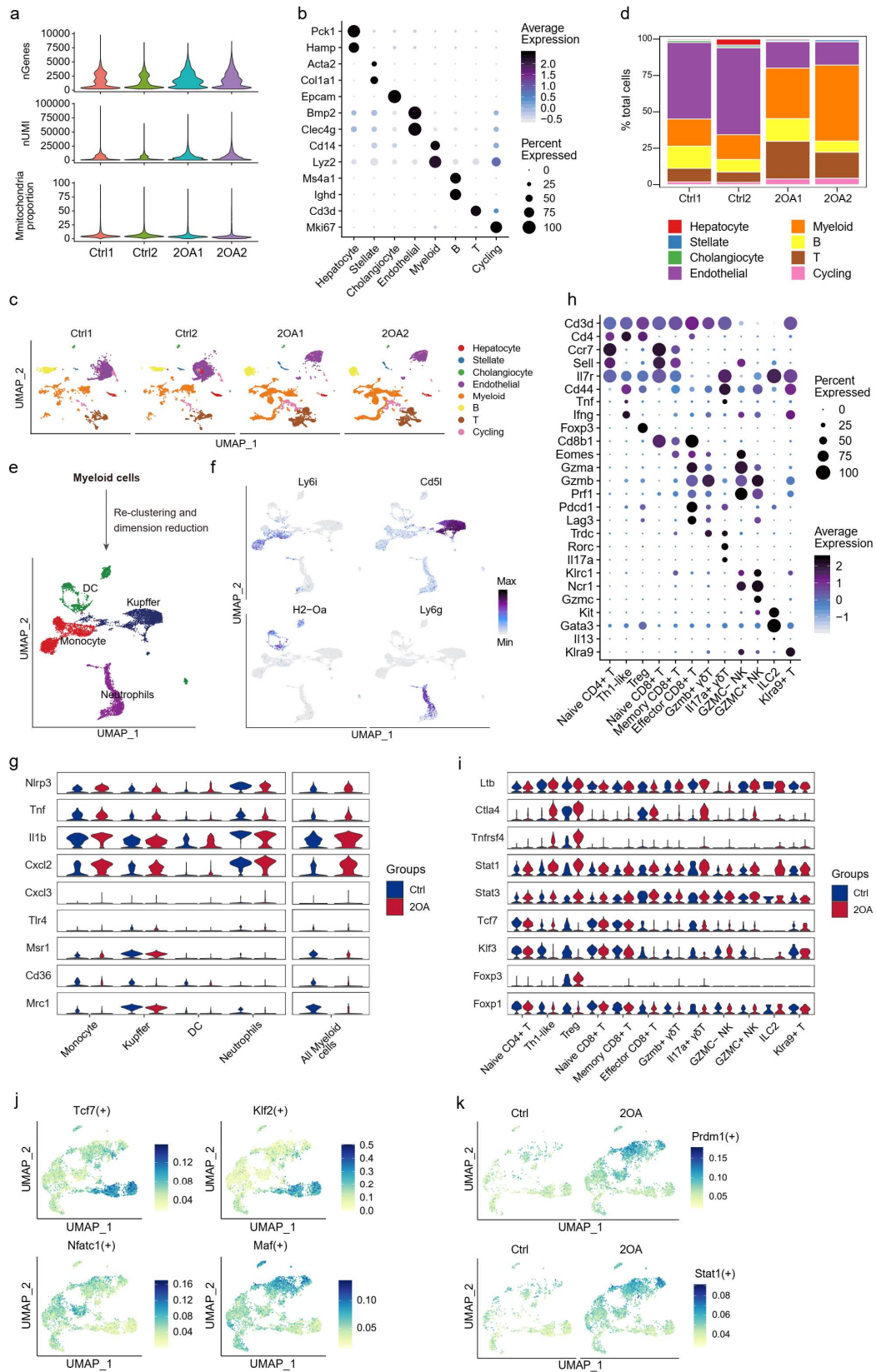
a



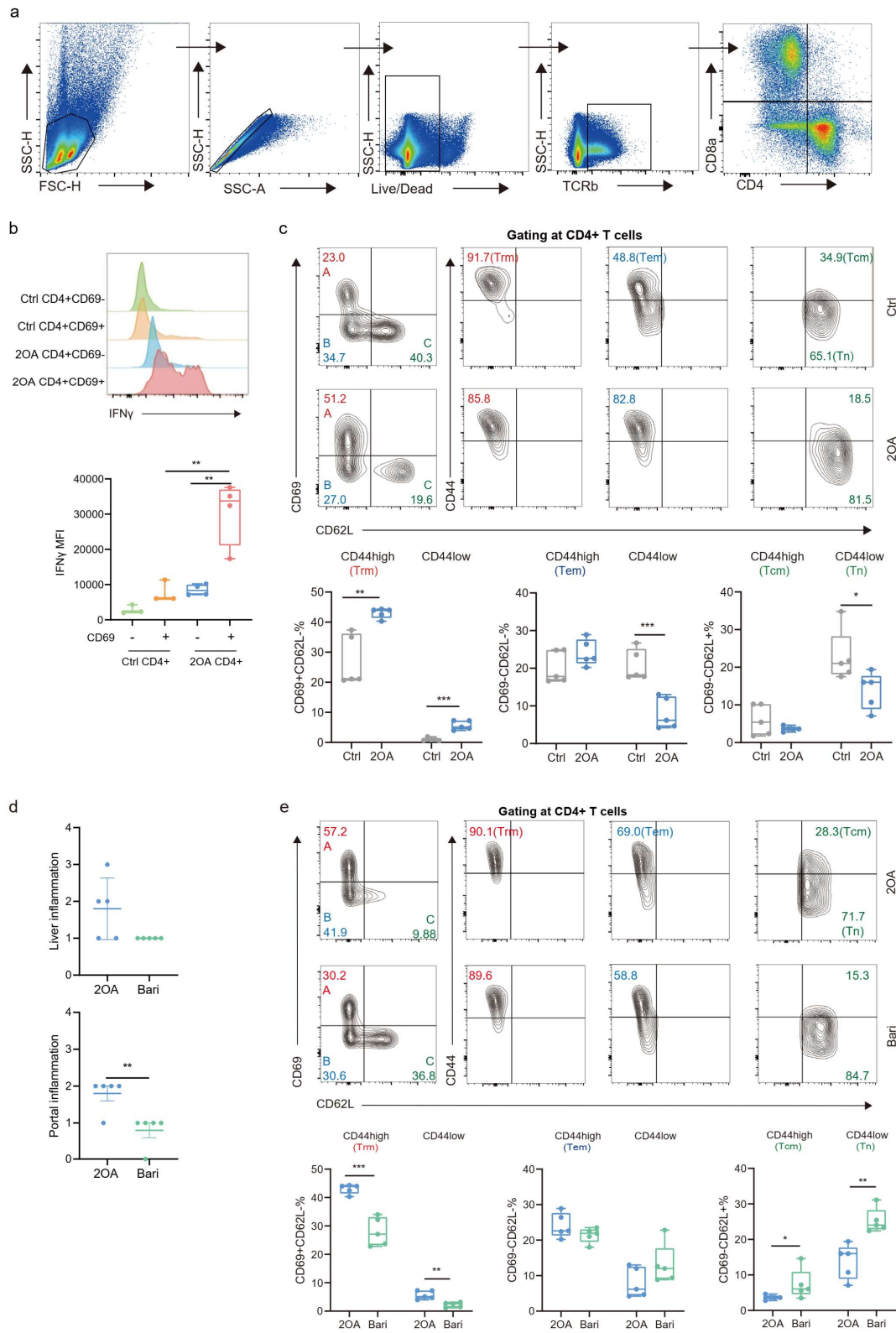
Supplementary Figure. 9 | The cell-cell communication analysis results of human liver scRNA-seq data using CellPhoneDB. Heatmap showing the significant receptor-ligand pairs involved in cell-cell interactions between the control and PBMC groups. The gradient red color represents the mean expression of each ligand-receptor pair in the cell-cell interaction pairs. The top and left annotation bars indicate the cell subsets in liver scRNA-seq data.



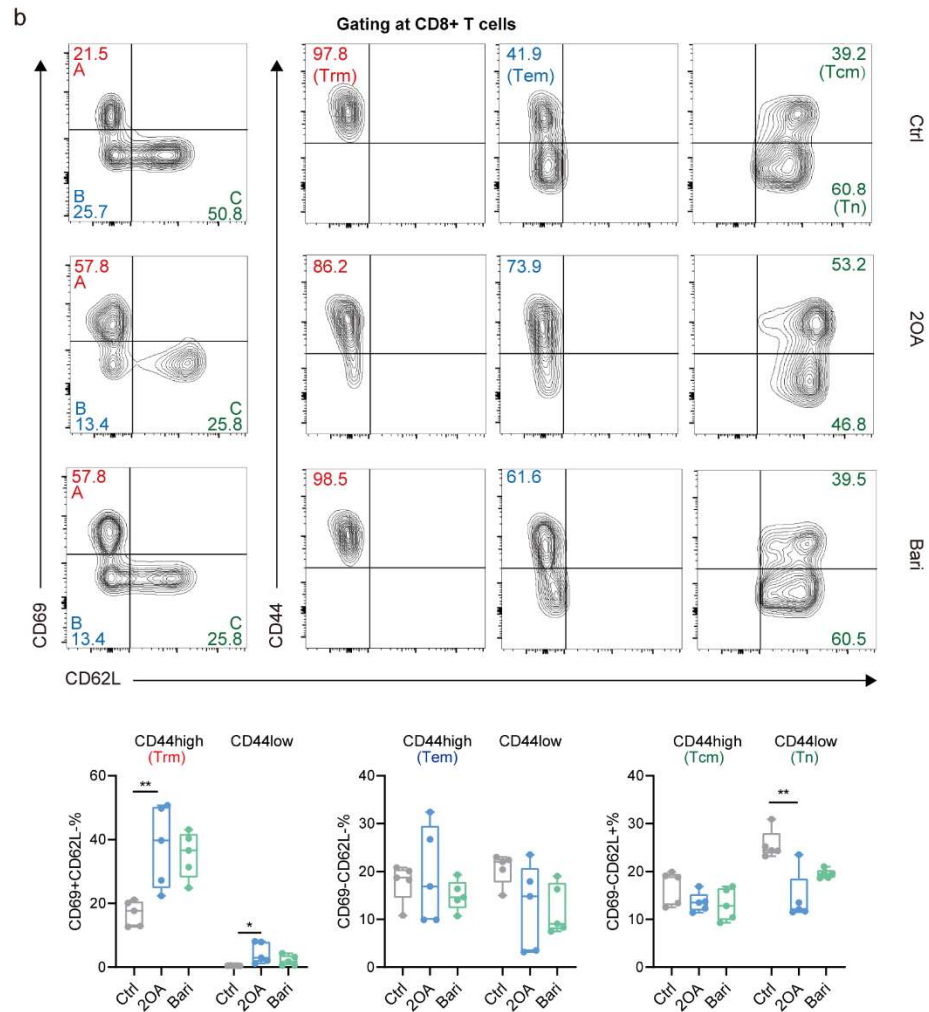
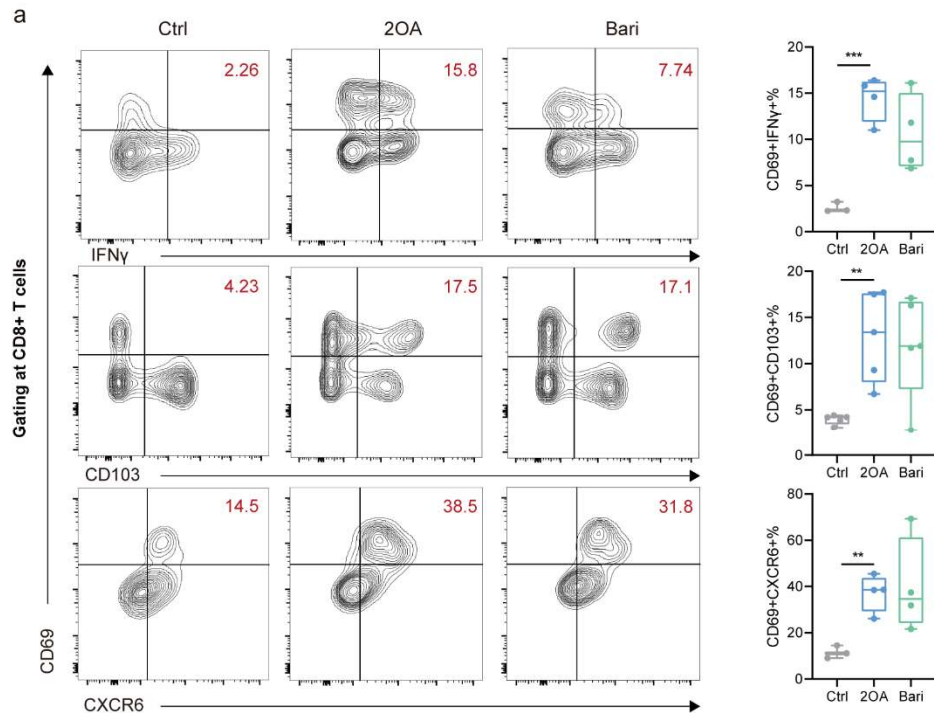
Supplementary Figure. 10 | Comparison of transcriptional features between hepatic T cells derived from patients with PBC and PSC. **a** Violin plots showing the gene number, total UMI counts and mitochondrial gene percentage of liver PSC patients. The scRNA-seq data of PSC from Poch et al., J Hepatol, 2021. **b** UMAP showing the distribution of annotated T cell subsets from PSC patients. **c** Bubble plot showing the expression of marker genes in T cell subsets. **d** Box plots showing the cell number proportion of each T cell subset in each group.



Supplementary Figure. 11 | Characteristics of immune cells in the liver of PBC mouse model analyzed by scRNA-seq. **a** Violin plots showing the gene number, total UMI counts and mitochondrial gene percentage of liver cells from control and PBC mice. **b** Bubble plot showing the expression of marker genes of major cell subsets of mouse scRNA-seq data. **c** UMAP showing the scRNA-seq data from mouse liver without batch effects correction. **d** Bar plot showing the cell ratio of main cell types in different samples of mouse scRNA-seq data. **e** UMAP showing the distribution of annotated myeloid cell subsets. **f** UMAP projection showing expression of Ly6i, Cd51, H2-Oa and Ly6g. **g** Violin plot showing the expression of Nlrp3, Tnf, Il1b, Cxcl12, Cxcl3, Tlr4, Msr1, Cd36, Mrc1 in myeloid cell subsets from control and 2OA groups. **h** Bubble plot showing the expression of marker genes of T cell subsets of mouse scRNA-seq data. **i** Violin plot showing the expression of Ltb, Ctla4, Tnfrsf4, Stat1, Stat3, Tcf7, Klf3, Foxp3, Foxp1 in T cell subsets from control and 2OA groups. **j** UMAP showing AUC score of Tcf7(+), Klf2(+), Nfatc1(+), Maf(+). **k** UMAP showing AUC score of Prdm1(+) Stat1(+) in control and 2OA-BSA groups.



Supplementary Figure. 12 | Identification and characterization of Th1-like cells in the liver of PBC mice and effects of baricitinib on cholangitis in PBC mice. **a** flow cytometry gating strategies for intrahepatic CD8a⁺ and CD4⁺ T cells. **b** Expression level of IFN- γ in CD4⁺CD69⁻ or CD4⁺CD69⁺ T cells in control and 2OA-BSA groups' liver samples ($n=3$ Ctrl, $n=4$ 2OA) using Representative density plots (left) and box plots (right). Source data are provided as a Source Data file including exact *P*. **c** Flow cytometry analysis (top graphs) and percentages (bottom graphs) of Trm (CD69⁺CD62L⁻CD44^{high}), Tem (CD69⁻CD62L⁻CD44^{high}), Tcm (CD69⁻CD62L⁺CD44^{high}), and Tn (CD69⁻CD62L⁺CD44^{low}) cells among CD4⁺ T cells from the liver of Ctrl and 2OA-BSA mice ($n=5$ Ctrl, $n=5$ 2OA). Source data are provided as a Source Data file including exact *P*. **d** Liver inflammation score and the portal inflammation score from H&E-stained liver biopsies from 2OA-BSA and Bari group mice ($n=5$ 2OA, $n=5$ Bari). Source data are provided as a Source Data file including exact *P*. **e** Flow cytometry analysis of Trm (CD69⁺CD62L⁻CD44^{high}), Tem (CD69⁻CD62L⁻CD44^{high}), Tcm (CD69⁻CD62L⁺CD44^{high}), and Tn (CD69⁻CD62L⁺CD44^{low}) (top graphs) and percentages of Trm, Tem, Tcm, and Tn cells (bottom graphs) among CD4⁺ T cells from the liver of 2OA-BSA and Bari mice ($n=5$ 2OA, $n=5$ Bari). Data are presented as means $\pm$ s.d. Significance was determined by two-tailed Student's t-test, * $P<0.05$, ** $P<0.01$. Source data are provided as a Source Data file including exact *P*. In b,c, e box plots show median (center line), the upper and lower quantiles (box), and the range of the data (whiskers).



Supplementary Figure. 13 | Flow cytometry revealed the characteristics of liver resident CD8⁺ T cells in PBC mouse model. **a** Flow cytometry analysis (left graphs) and percentage (right graphs) of CD69⁺CXCR6⁺ cells ($n=3$ Ctrl, $n=4$ 2OA, $n=4$ Bari), CD69⁺CD103⁺ cells ($n=5$ Ctrl, $n=5$ 2OA, $n=5$ Bari), and CD69⁺IFN- γ ⁺ cells ($n=3$ Ctrl, $n=4$ 2OA, $n=4$ Bari) among CD8⁺ T cells from the liver of Ctrl, 2OA-BSA and Bari mice. Source data are provided as a Source Data file including exact *P*. **b** Flow cytometry analysis (top graphs) and percentages (bottom graphs) of Trm (CD69⁺CD62L⁻CD44^{high}), Tem (CD69⁻CD62L⁻CD44^{high}), Tcm (CD69⁻CD62L⁺CD44^{high}), and Tn (CD69⁻CD62L⁺CD44^{low}) cells among CD8⁺ T cells from the liver of Ctrl, 2OA and Bari mice ($n=5$ Ctrl, $n=5$ 2OA, $n=5$ Bari). Data are presented as means $\pm$ s.d. Significance was determined by two-tailed Student's t-test, * $P<0.05$, ** $P<0.01$, *** $P<0.001$. Box plots show median (center line), the upper and lower quantiles (box), and the range of the data (whiskers). Source data are provided as a Source Data file including exact *P*.