Article

Liver tumour immune microenvironment subtypes and neutrophil heterogeneity

https://doi.org/10.1038/s41586-022-05400-x

Received: 7 September 2021

Accepted: 30 September 2022

Published online: 9 November 2022

Check for updates

Ruidong Xue[1,7], Qiming Zhang[2,7], Qi Cao[1,7], Ruirui Kong[1,7], Xiao Xiang[3,7], Hengkang Liu[1], Mei Feng[1], Fangyanni Wang[1], Jinghui Cheng[1], Zhao Li[3], Qimin Zhan[4], Mi Deng[4], Jiye Zhu[3,8]✉, Zemin Zhang[2,5,8]✉ & Ning Zhang[1,4,6,8]✉

The heterogeneity of the tumour immune microenvironment (TIME), organized by various immune and stromal cells, is a major contributing factor of tumour metastasis, relapse and drug resistance[1–3], but how different TIME subtypes are connected to the clinical relevance in liver cancer remains unclear. Here we performed single-cell RNA-sequencing (scRNA-seq) analysis of 189 samples collected from 124 patients and 8 mice with liver cancer. With more than 1 million cells analysed, we stratified patients into five TIME subtypes, including immune activation, immune suppression mediated by myeloid or stromal cells, immune exclusion and immune residence phenotypes. Different TIME subtypes were spatially organized and associated with chemokine networks and genomic features. Notably, tumour-associated neutrophil (TAN) populations enriched in the myeloid-cell-enriched subtype were associated with an unfavourable prognosis. Through in vitro induction of TANs and ex vivo analyses of patient TANs, we showed that CCL4$^+$ TANs can recruit macrophages and that PD-L1$^+$ TANs can suppress T cell cytotoxicity. Furthermore, scRNA-seq analysis of mouse neutrophil subsets revealed that they are largely conserved with those of humans. In vivo neutrophil depletion in mouse models attenuated tumour progression, confirming the pro-tumour phenotypes of TANs. With this detailed cellular heterogeneity landscape of liver cancer, our study illustrates diverse TIME subtypes, highlights immunosuppressive functions of TANs and sheds light on potential immunotherapies targeting TANs.

Primary liver cancer (PLC) has three major histological subtypes—hepatocellular carcinoma (HCC), intrahepatic cholangiocarcinoma (ICC) and combined hepatocellular and intrahepatic cholangiocarcinoma (CHC)[4]. Despite recent progress in immunotherapies[5], our understanding of the baseline TIME landscape in PLC is limited, precluding biomarker identification for better patient stratification. A comprehensive single-cell study covering most cell populations and three major subtypes of PLC with established clinical parameters is needed. Functional contributions of neutrophils in cancer are increasingly recognized[6–10] with both anti-tumour[11,12] and pro-tumour[13,14] roles reported. scRNA-seq has been used to dissect TIME components of PLC[15–22] and neutrophil heterogeneity[23–25], but these studies usually involve antibody-based cell enrichment and are limited in cohort size. Owing to the short lifespan of neutrophils and technical difficulties in handling them, the functional heterogeneity of neutrophils in cancer remains unclear. Here we analysed the cellular landscape of 189 samples collected from patients and mouse models with liver cancer, dissected the TIME subtypes, and investigated the phenotypic and functional heterogeneity of neutrophils in liver cancer.

A large-scale single-cell atlas of liver cancer

To survey the TIME landscape across PLC covering all cell populations, we performed scRNA-seq analysis of 160 samples of 124 treatment-naive patients, including 79 with HCC, 25 with ICC and 7 with CHC (Fig. 1a, Extended Data Fig. 1a and Supplementary Tables 1 and 2). A total of 89 TIME cell clusters were identified among 1,092,172 cells obtained (Fig. 1b,c, Extended Data Figs. 1 and 2, Supplementary Fig. 1, Supplementary Note 1 and Supplementary Table 3). Owing to our large cohort and enrichment-free strategy, we captured more diverse populations and identified a substantial proportion of neutrophils lacking characterization in PLC[15–22] (Extended Data Fig. 1i,j). TIME cell clusters exhibited obvious tissue and cancer type preference, and some were associated with aetiologies (Extended Data Fig. 2b–d). Copy-number analysis showed that most epithelial cells were tumour cells, showing either high hepatic or biliary scores (Extended Data Fig. 1e–h). In contrast to TIME clusters constituting cells across different patients, tumour cell clusters tended to be patient specific. On the basis of PLC subtype composition representing real-world epidemiology, our multifaceted data encompass well-annotated clinical information, a single-cell atlas

[1]Translational Cancer Research Center, Peking University First Hospital, Beijing, China. [2]BIOPIC, Beijing Advanced Innovation Center for Genomics, School of Life Sciences, Peking University, Beijing, China. [3]Beijing Key Surgical Basic Research Laboratory of Liver Cirrhosis and Liver Cancer, Department of Hepatobiliary Surgery, Peking University People's Hospital, Beijing, China. [4]International Cancer Institute, Peking University Health Science Center, Beijing, China. [5]Changping Laboratory, Beijing, China. [6]Yunnan Baiyao Group, Kunming, China. [7]These authors contributed equally: Ruidong Xue, Qiming Zhang, Qi Cao, Ruirui Kong, Xiao Xiang. [8]These authors jointly supervised this work: Jiye Zhu, Zemin Zhang, Ning Zhang. ✉e-mail: gandanwk@vip.sina.com; zemin@pku.edu.cn; zhangning@bjmu.edu.cn

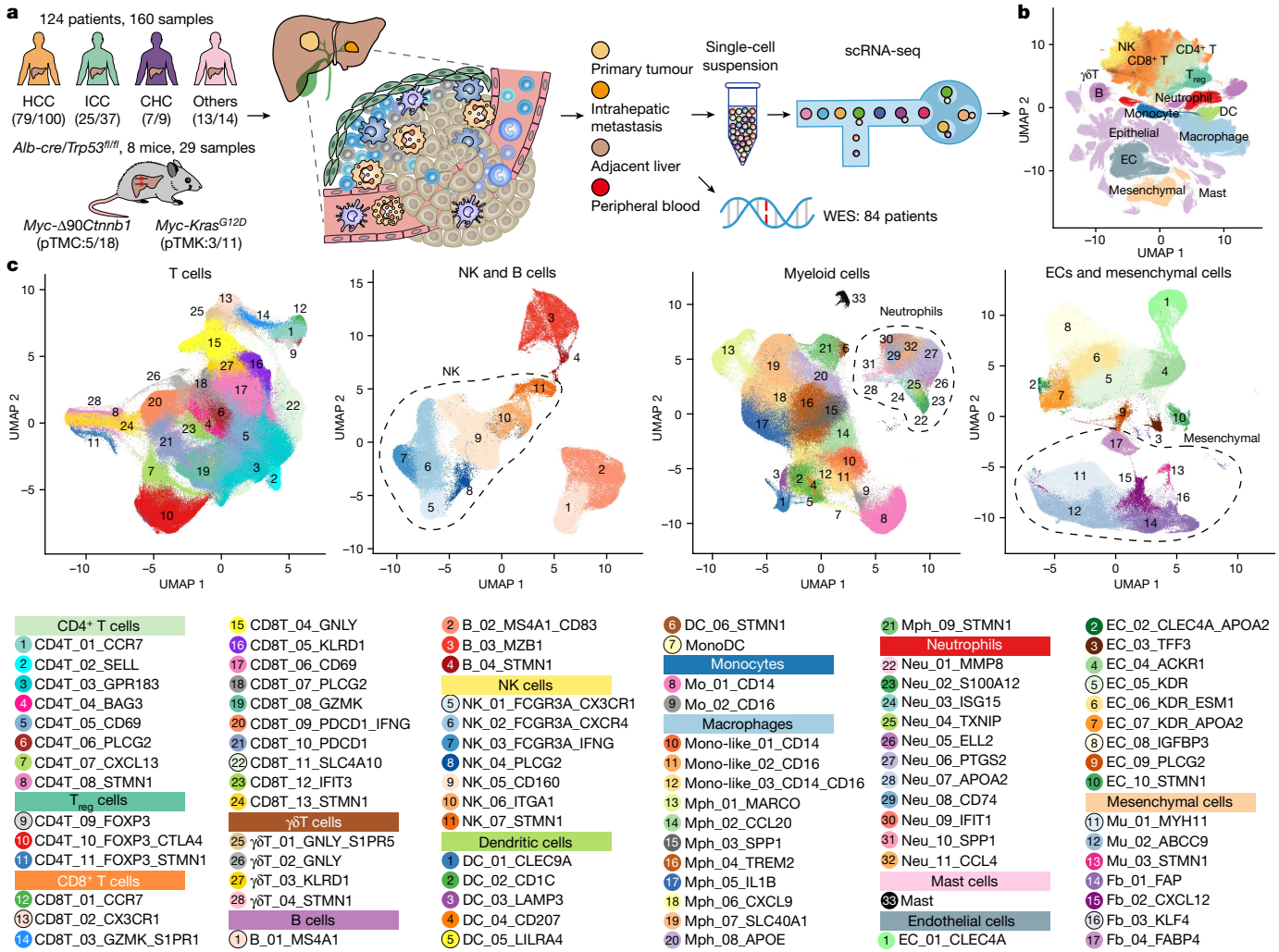


Fig. 1 | The single-cell landscape of 124 patients with liver cancer. a, The experimental workflow. The numbers of cases and samples collected for each cancer type and mouse model are denoted. **b**, Uniform manifold approximation and projection (UMAP) plot showing the major cell types. Dots represent individual cells, and colours represent different cell populations. NK, natural killer cells; T_{reg}, regulatory T cells; DC, dendritic cells; EC, endothelial cells. **c**, UMAP plots showing TIME clusters. To facilitate illustration, cells are grouped into four panels: T cells; natural killer (NK) and B cells; myeloid cells; and endothelial cells (ECs) and mesenchymal cells. Colour code and cluster ID are shown.

with diverse populations and matched genomic profiles, enabling us to examine the cellular heterogeneity landscape of PLC in detail.

Cellular module analyses reveal five TIME subtypes

To investigate TIME subtypes of PLC, we examined co-enrichment patterns of cells from tumour tissues. Hierarchy clustering identified five stable cellular modules (CM1–CM5) (Fig. 2a and Extended Data Fig. 3a). On the basis of the differential enrichment of CM1–CM5, we stratified the patients into five corresponding TIME subtypes (Fig. 2b–e), of which the properties were designated considering four aspects: (1) cell clusters, (2) functional marker gene expression, (3) TIME-related gene signatures[26] and (4) prognostic relevance (Extended Data Fig. 3b–f).

CM1 contained activated myeloid and T cell clusters, including mature dendritic cells enriched in immunoregulatory molecules (DC_03_LAMP3), *CXCL9*$^+$ macrophages (Mph_06_CXCL9), T helper type-1-like cells (CD4T_07_CXCL13) and exhausted T cells (Fig. 1c). High expression of *IFNG*, *GZMB* and *PDCD1*, along with enriched signatures of 'co-activation molecules' and 'checkpoint molecules' suggested that CM1-dominant patients exhibited an immune-activated state, and were therefore designated as TIME-IA (immune activation). The enrichment of Mph_03_SPP1[27] and high *IL1B* expression[28]—both related to immunosuppression—enriched signatures of 'immune suppression by myeloid cells' and 'pro-tumour cytokines', and the association with a worse prognosis collectively suggest immunosuppressive and pro-tumour phenotypes of CM2, and the corresponding patients were therefore designated as TIME-ISM (immune suppressive myeloid).

Stromal cells were enriched in both CM3 and CM4. The enrichment of two stromal clusters (EC_03_TFF3 and Fb_01_FAP), high expression of tumour-activated stromal genes such as *COL1A1*, *MMP11* and *ITGA1*, enriched signatures of 'matrix' and 'cancer-associated fibroblasts' and the association with a worse prognosis led us to designate CM3-dominant patients as TIME-ISS (immune suppressive stromal). By contrast, CM4 contained most endothelial cell and mesenchymal clusters but lacked immune cells. Particularly, the enriched *CXCL12*$^+$ fibroblasts (Fb_02_CXCL12) could exclude T cells from tumour cells[29]. On the basis of these results together, we propose an immune exclusion phenotype (TIME-IE). Unexpectedly, cytotoxic T cells (CD8T_08_GZMK) were also enriched in this cellular module. Using multicolour immunohistochemistry (mIHC), we observed that GZMK$^+$CD8$^+$ T cells mainly localized in the stroma yet were excluded from tumour regions (Fig. 2f), suggesting that these immune-excluded T cells are actually cytotoxic rather than exhausted. CM5 contained liver-resident clusters including residential natural killer cells (NK_05_CD160),

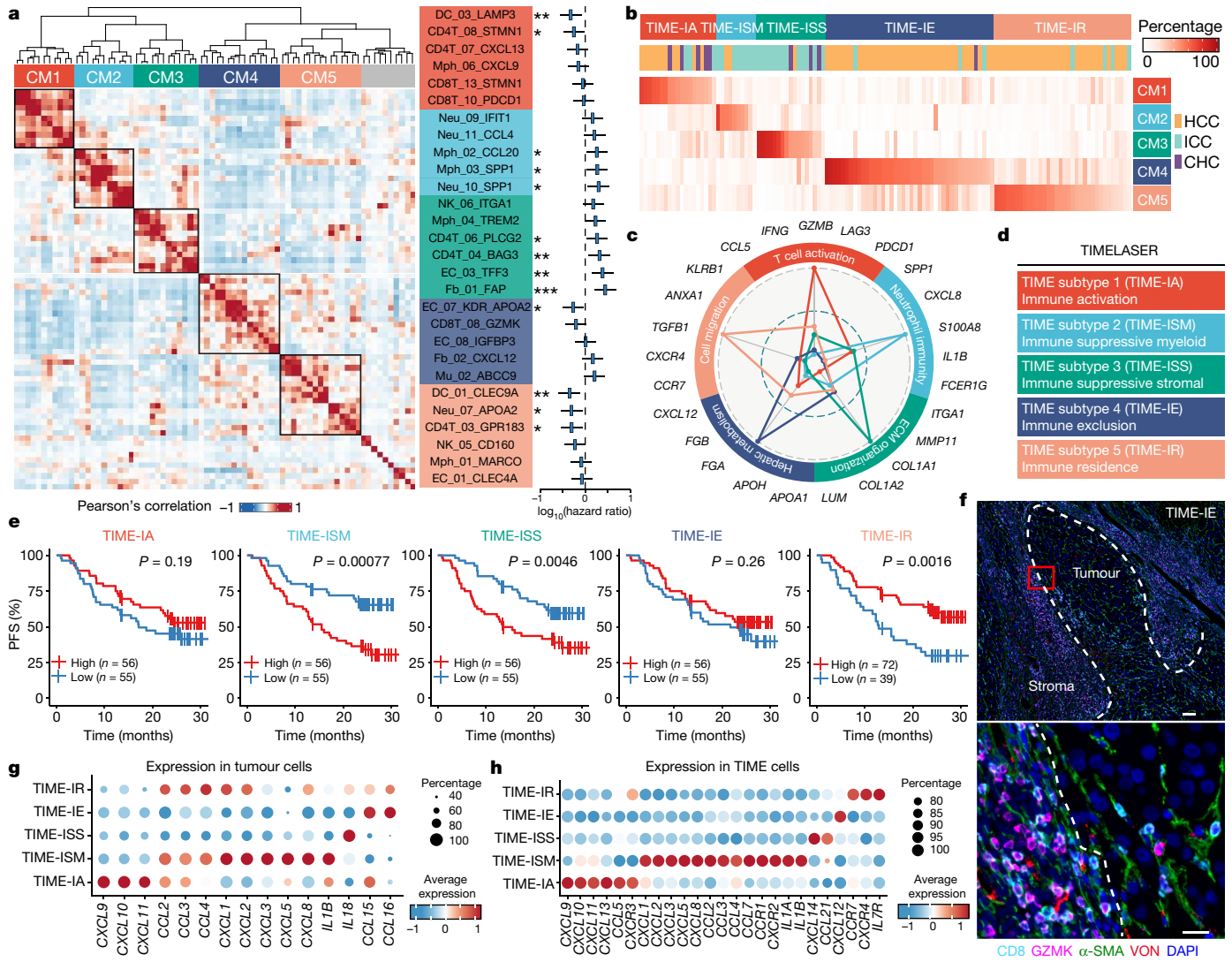


Fig. 2 | Five TIME subtypes of PLC. a, The five cellular modules on the basis of correlations of cell clusters from tumours. Key cell clusters from each cellular module are shown on the right with the forest plot showing the hazard ratio based on progression-free survival. Statistical analysis was performed using log-rank tests; $^{*}P < 0.05$; $^{**}P < 0.01$, $^{***}P < 0.001$. **b**, The percentage of CM1–CM5 across PLC cases. **c**, Radar plot showing marker genes and signalling pathways enriched for CM1–CM5. CMs are denoted by colour. The distance from the dots to the centre of the circle represents the normalized expression of each pathway scaled to 0–1. **d**, Definitions of the five subtypes for TIMELASER phenotypes. **e**, Progression-free survival (PFS) of cases stratified by each cellular module. Statistical analysis was performed using log-rank tests. **f**, Staining of GZMK^{+}CD8^{+} T cells (CD8 and GZMK), fibroblasts (α-SMA) and endothelial cells (VON) in TIME-IE. Scale bars, 100 μm (top) and 20 μm (bottom). **g**,**h**, The expression of specific cytokines, chemokines and receptors in tumour cells (**g**) and TIME cells (**h**) across five TIMELASER subtypes.

Kupffer cells (Mph_01_MARCO) and liver sinusoidal endothelial cells (EC_01_CLEC4A), and was associated with a better prognosis. Thus, CM5-dominant patients were designated as TIME-IR (immune residence).

Taken together, we name this classification scheme TIMELASER, for 'tumour immune microenvironment subtypes at the single-cell resolution including immune activation, suppression, exclusion and residence phenotypes' (Fig. 2d and Extended Data Fig. 5e,f). Survival analysis by assigning each patient into a single-cellular module or by stratifying patients on the basis of each cellular module signature showed consistent results (Fig. 2e and Extended Data Fig. 3e,f), suggesting that our classification is robust and clinically relevant. Reanalysing published scRNA-seq[15–17] and bulk RNA-seq datasets[30–32] revealed similar enrichment of five TIMELASER subtypes across PLC (Extended Data Fig. 4a–e). These results validated our TIMELASER framework and showed that our TIMELASER-derived signatures (Supplementary Table 3) could also be used for bulk data. Moreover, performing co-detection by indexing (CODEX) analysis of representative samples and reanalysing spatial transcriptomes of PLC[33] successfully recapitulated the cellular composition of each subtype, further validating our TIMELASER framework at spatial resolution (Supplementary Note 2 and Extended Data Fig. 4f,g).

TIMELASER subtypes exhibit distinct features

Diverse expression patterns of chemokines and cytokines and their receptors observed in both tumour and TIME cells hint at underlying factors shaping diverse TIMELASER subtypes (Fig. 2g,h, Extended Data Fig. 5a–d and Supplementary Note 3). Concordant chemokine expression patterns of tumour and TIME cells were observed in the TIME-IA and TIME-ISM subtypes, suggesting positive-feedback loops. For example, *CXCL9/10/11–CXCR3* ligand–receptor (L–R) pairs were enriched in TIME-IA, whereas *CXCL1/3/8–CXCR2* L–R pairs were enhanced in TIME-ISM (Fig. 2g,h). These results, along with unique L–R pairs in other

TIME subtypes, suggest that distinct chemokine networks contribute to the organization of TIMELASER subtypes.

Analyses of exome data identified recurrently mutated genes correlated with TIMELASER subtypes, including driver genes such as *TP53*, *CTNNB1*, *KRAS* and *IDH1*[34] (Extended Data Fig. 6a–e and Supplementary Fig. 2). Furthermore, despite the extensive heterogeneity of tumour cells, we characterized eight common gene modules[35] linked with TIMELASER subtypes (Supplementary Note 4 and Extended Data Fig. 6f,g). For example, the cell cycle gene module was enriched in TIME-IA, indicating that these proliferating tumour cells would engage with immune cells. In summary, our TIMELASER subtypes exhibit different chemokine networks, and are associated with distinct somatic alterations and transcriptomic profiles of tumour cells.

Neutrophil heterogeneity in liver cancer

The enrichment of multiple neutrophil subsets in TIME-ISM, their association with poor prognosis and their scarcity[36] led us to further examine neutrophils. Using mIHC, we validated the existence of neutrophils in PLC, showing that ICC has significantly more neutrophils than HCC (Extended Data Fig. 7a), consistent with our observation that TIME-ISM is enriched in ICC. A total of 34,307 neutrophils were divided into 11 subsets that exhibited clear tissue separation and cancer-type preference (Fig. 3a,b and Extended Data Fig. 7b–e). Neu_02_S100A12, Neu_03_ISG15 and Neu_04_TXNIP were mainly peripheral blood neutrophils (PBNs) (Extended Data Fig. 7f), whereas Neu_05_ELL2 and Neu_06_PTGS2 were mainly adjacent liver neutrophils (ALNs). All of the other six subsets (Neu_01_MMP8, Neu_07_APOA2, Neu_08_CD74, Neu_09_IFIT1, Neu_10_SPP1 and Neu_11_CCL4) were enriched in tumours and designated as TANs. Developmental trajectory analysis revealed a clear sequential differentiation path from PBNs to ALNs and then to TANs (Fig. 3a and Extended Data Fig. 7e). Notably, a combinatorial high proportion of three TAN subsets from TIME-ISM (Neu_09/10/11, accounting for an average of 86.8% of total TANs) was associated with a worse prognosis (Extended Data Fig. 7g), indicating pro-tumour functions of these TANs.

Distinct gene signatures were observed across these neutrophil subsets (Extended Data Fig. 7h,i). PBNs expressed high levels of secretory vesicle signatures associated with anti-pathogen activities, whereas ALNs and TANs expressed enhanced levels of matrix and chemokine signatures. Analysing regulons of transcription factors (Fig. 3c) revealed higher *SPI1* activity in PBNs, whereas *NFE2L2* and *CREM* were more active in both ALNs and TANs. *MAFG, BHLHE40* and *HES4* were more active in TANs and possibly related to neutrophil reprogramming in tumours (Extended Data Fig. 8a). The activities of these transcription factors were confirmed by accessibility signals detected by assay for transposase accessible chromatin using sequencing (ATAC-seq) analysis of matched patient-derived PBNs, ALNs and TANs (Extended Data Fig. 8b). Collectively, our results present a layered landscape of 11 neutrophil subsets and support that neutrophil differentiation may be orchestrated by transcription factors in a spatiotemporal manner (Supplementary Note 5 and Supplementary Table 3).

To systematically examine the function of TANs, we first induced TANs in vitro by co-culturing human PBNs with three human liver cancer cell lines—HepG2, HCCLM3 and MHCC97H—and a control cell line, HEK293T (Extended Data Fig. 8c–f). Compared with the controls, PBNs co-cultured with various cancer cell lines showed concordant higher expression of pro-angiogenesis and chemokine production signatures, indicating TAN-like phenotypes, and we therefore termed these cells 'in vitro induced TANs'. Signatures of PBN and ALN subsets were downregulated, whereas most TAN subsets were upregulated, with that of Neu_11_CCL4 as the highest (Supplementary Note 6 and Extended Data Fig. 8f), suggesting that this subset is more favoured than others in our co-culture system. These results support the spectrum of neutrophil subsets identified from our scRNA-seq data.

Chemokine secretion and immunosuppression phenotypes of TANs

We next focused on phenotypes and functions of two TAN subsets—Neu_11_CCL4 and Neu_09_IFIT1. *CCL4*$^+$ TANs (Neu_11_CCL4) expressed high levels of chemokine genes *CCL3* and *CCL4*, confirmed by mIHC (Fig. 3d and Extended Data Fig. 8g). In vitro induced TANs also showed elevated *CCL4* expression (Fig. 3e, Extended Data Fig. 8h and Supplementary Fig. 3). We next directly investigated chemokine secretion by ex vivo analysis of patient-derived TANs and non-tumoural neutrophils (that is, PBNs or ALNs, referred to as non-TANs). Compared with non-TANs, TANs showed higher accessibility signals of *CCL4* and CCL4 protein secretion (Fig. 3f,g). *CCL4*$^+$ TANs were predicted to recruit macrophages through *CCL4*–*CCR5* (Extended Data Fig. 8i). Consistently, more autologous monocytes were recruited in the chemotactic assay when co-cultured with TANs (Fig. 3h and Extended Data Fig. 8j). These results validate the chemokine-secreting function of TANs and support that CCL4$^+$ TANs could recruit macrophages.

We also found that TANs showed a marked increase in *CD274* (encoding PD-L1) expression compared with non-TANs, with Neu_09_IFIT1 showing the highest expression (Fig. 3i). *CD274* expression of in vitro induced TANs continuously increased in a time-dependent manner (Fig. 3j). Fluorescence-activated cell sorting (FACS) analysis further showed significantly higher PD-L1 expression in in vitro induced TANs compared with in the controls (Fig. 3k and Extended Data Fig. 8k). ATAC-seq and FACS analyses also revealed higher accessibility signals of *CD274* and PD-L1 expression in patient-derived TANs (Fig. 3l,m). To investigate whether the high PD-L1 expression of TANs would directly inhibit T cell activity, we co-cultured CD8$^+$ T cells with in vitro induced TANs (Extended Data Fig. 9a,b) or patient-derived TANs. CD8$^+$ T cells co-cultured with in vitro induced TANs showed lower protein levels of the T cell cytotoxic marker IFNγ and activation markers CD25 and CD69 (Fig. 3n and Extended Data Fig. 9c). After adding anti-PD-L1 antibodies, the decline in IFNγ in CD8$^+$ T cells was reversed in the PBN-MHCC97H group compared with in the controls (Fig. 3o and Extended Data Fig. 9d), confirming that PD-L1 mediates the suppressive function of TANs. Furthermore, autologous CD8$^+$ T cells co-cultured with human TANs exhibited lower proliferation property (CFSE), and lower levels of IFNγ, GZMB, PRF1 and CD25 (Fig. 3p and Extended Data Fig. 9e). Moreover, mIHC revealed the physical proximity of PD-L1$^+$ neutrophils and PD1$^+$ CD8$^+$ T cells (Fig. 3q), supporting their direct interaction. These results together demonstrate that PD-L1$^+$ TANs suppress cytotoxic CD8$^+$ T cells in PLC.

Notably, two *IFIT1*$^+$ neutrophil subsets enriched in PBNs (Neu_03_ISG15) and TANs (Neu_09_IFIT1) showed distinct PD-L1 expression (Fig. 3i). L–R analyses revealed that Neu_09_IFIT1 cells were more likely to interact with *IFNG*$^+$ lymphocytes (CD8T_13_PDCD1_IFNG and NK_03_FCGR3A_IFNG) through IFNγ–type II IFNR (Supplementary Note 7 and Extended Data Fig. 9f–j). These results indicate that interactions with IFNγ$^+$ cells may contribute to the high PD-L1 expression of Neu_09_IFIT1.

Conserved neutrophil subsets in human and mouse liver cancer

To further examine heterogeneous functions of TANs in vivo, we built two new spontaneous liver cancer mouse models, with the pTMC mice developing mainly HCC and pTMK mice developing mainly ICC (Methods and Extended Data Fig. 10a–c). We performed scRNA-seq analysis of 21 samples that included peripheral blood, tumour-adjacent liver and tumours collected from 6 mice (Extended Data Fig. 10d,e and Supplementary Table 4). A total of 17,780 neutrophils were divided into 12 clusters showing clear tissue specificity and ordered developmental trajectory (Fig. 4a,b and Extended Data Fig. 10f–h). Unbiased cross-species data integration of neutrophil subsets and concordant expression of key signature genes suggested that neutrophils in mouse

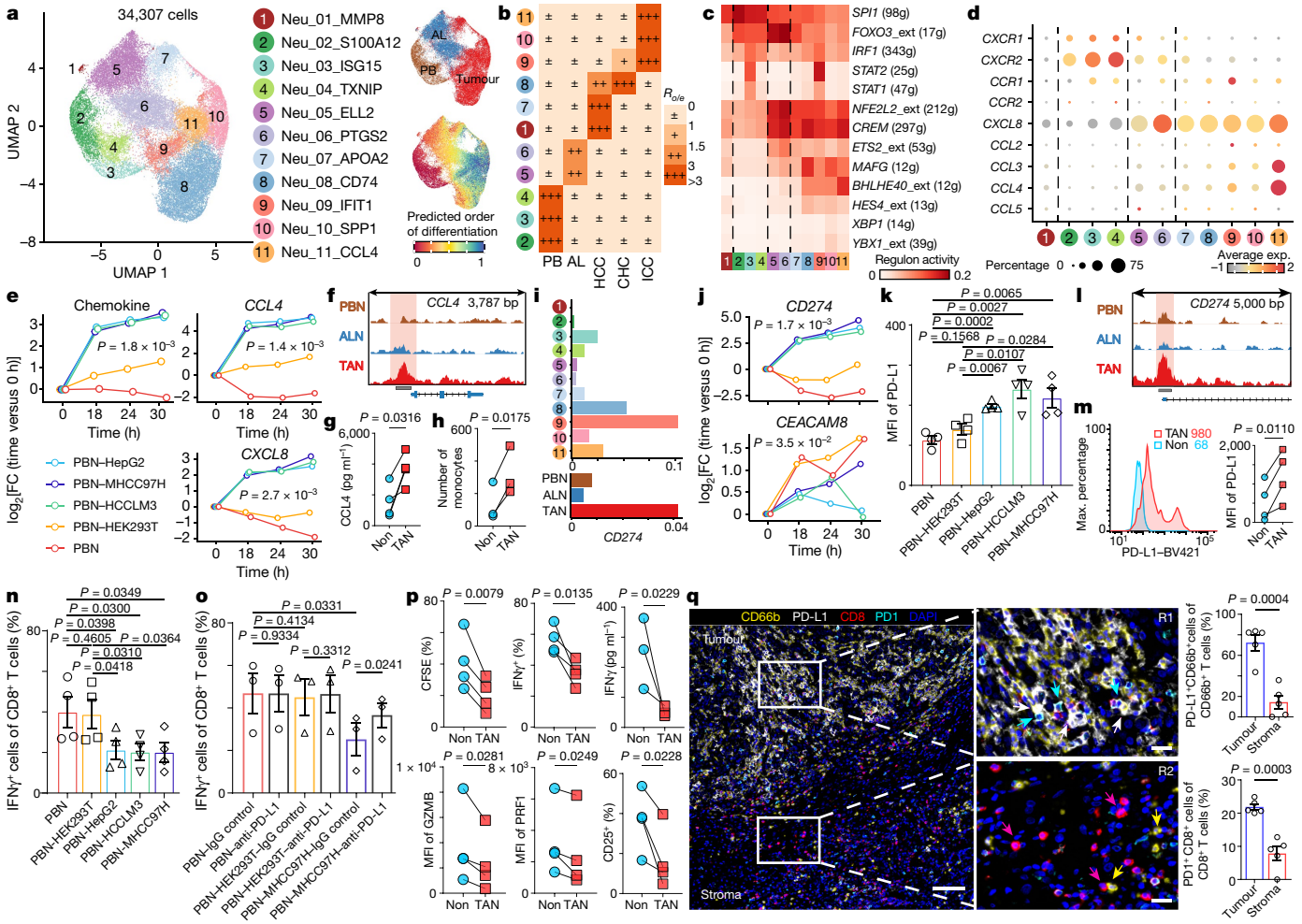


Fig. 3 | Neutrophil heterogeneity and functional validation in humans. **a**, Neutrophil clusters coloured by cluster, tissue source and developmental order. **b**, Tissue preference of neutrophil clusters in humans, revealed by $R_{o/e}$ (ratio of observed cell number to expected cell number). **c**, Transcription factors inferred by SCENIC. The number of target genes for each transcription factor is indicated in parentheses. **d**, Expression (exp.) of chemokines and receptors. **e**, Expression of signatures and genes in in vitro induced TANs. **f**, Normalized ATAC-seq tracks of *CCL4*. The ATAC peak is denoted with the grey line and red shading. **g**, Quantification of CCL4 production using enzyme-linked immunosorbent assay (ELISA). $n = 4$. **h**, Chemotaxis abilities of matched TAN or non-TAN populations on autologous monocytes. $n = 3$. **i**, *CD274* expression. **j**, *CD274* and *CEACAM8* expression as described in **e**. **k**, PD-L1 expression in co-cultured PBNs from **e** examined using FACS. $n = 4$. **l**, ATAC-seq tracks of *CD274* as described in **f**. **m**, PD-L1 expression in matched TAN or non-TAN populations. $n = 4$. MFI, mean fluorescence intensity. **n**, Proportions of IFNγ$^+$CD8$^+$ T cells. $n = 4$. **o**, IFNγ expression in CD8$^+$ T cells co-cultured with different neutrophil–cell line–antibody combinations. $n = 3$. **p**, Comparison of autologous CD8$^+$ T cells co-cultured with matched TANs or non-TANs. $n = 4$, including proliferation (CFSE) and functional marker (CD25, IFNγ, GZMB and PRF1) expression. IFNγ production was further quantified by ELISA ($n = 3$). **q**, Staining of neutrophils (CD66b) and CD8$^+$ T cells. Representative cells are indicated by arrows, including PD-L1$^+$CD66b$^+$ cells (white), PD1$^+$CD8$^+$ T cells (cyan), PD-L1$^-$CD66b$^+$ cells (yellow) and PD1$^-$CD8$^+$ T cells (red). Scale bars, 100 μm (left) and 20 μm (right). The bar plots show the quantification results. $n = 5$. In **g**–**q**, *n* denotes biologically independent samples. For **k**, **n**, **o** and **q**, data are mean ± s.e.m. Statistical analysis was performed using two-sided Student's *t*-tests (**k** and **q**), one-sided Student's *t*-tests (**n**), two-sided paired *t*-tests (**g**, **h**, **m**, **o** and **p**) and two-way ANOVA (**e** and **j**).

and human were largely conserved (Fig. 4c, Extended Data Fig. 10i–k and Supplementary Note 8). Specifically, three TAN subsets from TIME-ISM (Neu_09/10/11) corresponded to mNeu_10/11/12, respectively. Notably, higher *Cd274* expression was observed in mouse TANs, consistent with that in human TANs (Fig. 4d). These results laid the basis for investigating neutrophil-based therapy in our mouse models.

Neutrophil depletion attenuates tumour progression

A collectively pro-tumour phenotype of the TIME-ISM TANs (Neu_09/10/11) led us to examine the therapeutic effect of eliminating those pro-tumour TAN subsets in vivo. As a combinatory in vivo elimination strategy specifically targeting Neu_09/10/11 was not available, we reasoned that neutrophil depletion using anti-Ly6G antibodies[37] might be the most proximate way to mimic such therapy. Neutrophil depletion resulted in significant reductions in liver cancer nodules and tumour weight (Fig. 4e,f). Both the number of TANs and PD-L1 expression of TANs decreased after anti-Ly6G treatment compared with the isotype control (Fig. 4g and Extended Data Fig. 10l–m). IHC analysis confirmed that there were lower numbers of neutrophils and proliferative malignant cells in the Ly6G-blockade group (Fig. 4h and Extended Data Fig. 10n). We further assessed the neutrophil-depletion efficacy by parallel detection of surface and intracellular Ly6G[38] (Extended Data Fig. 10o,p). Analysis of intracellular Ly6G confirmed that about 70% of neutrophils were depleted after Ly6G blockade, consistent with IHC (Fig. 4h and Extended Data Fig. 10n). Furthermore, we observed a 46.6% reduction in infiltrating macrophages (Extended Data Fig. 10n). Although Ly6G blockade did not alter the number of CD8$^+$ T cells, their exhaustion states were relieved as shown by decreased levels of the checkpoint markers PD-1 and TIM3 (Fig. 4g and Extended Data

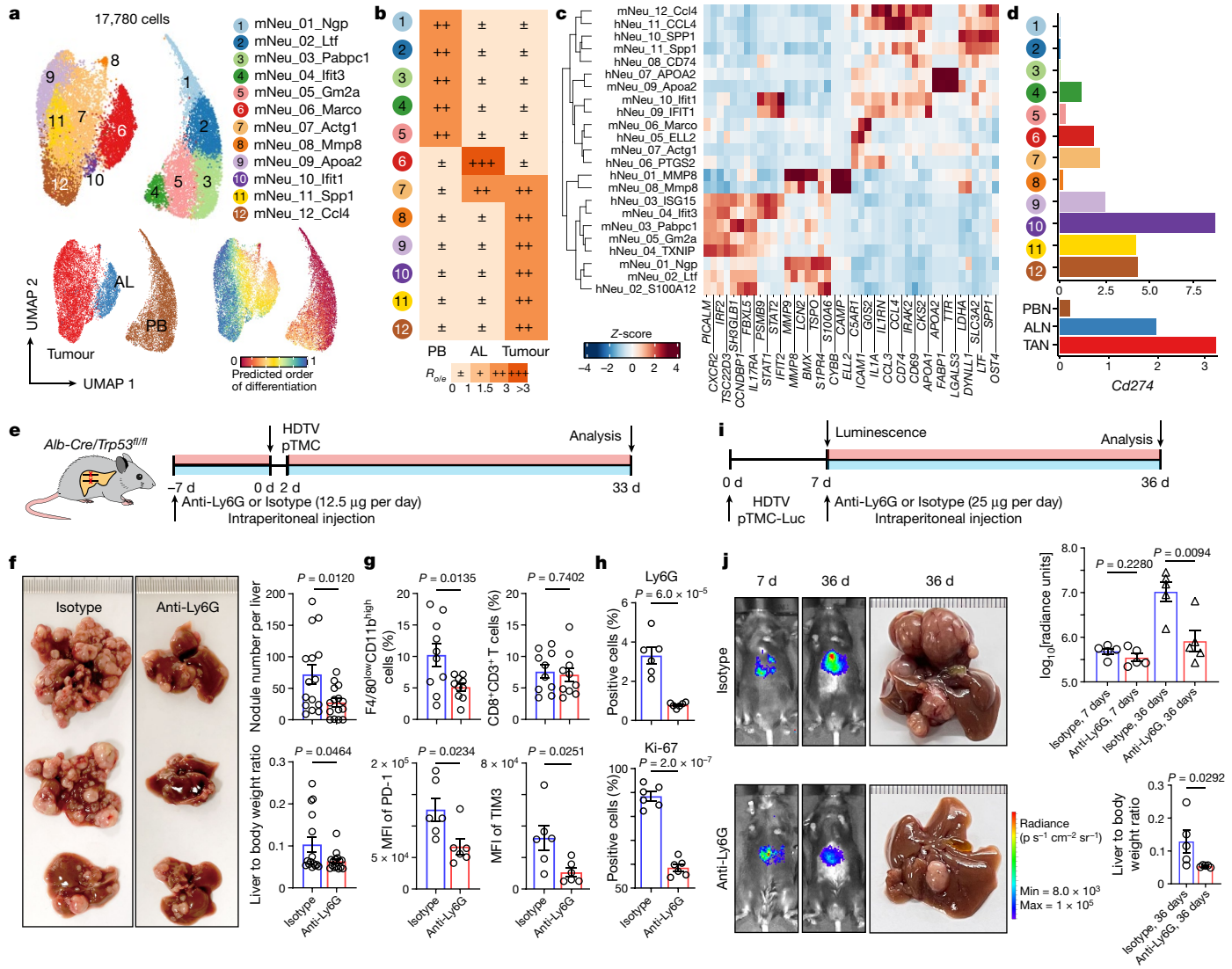


Fig. 4 | Neutrophil heterogeneity and depletion in mouse models. a, UMAP plots showing the neutrophil clusters (top), tissue sources (bottom left) and developmental orders (bottom right) in mice. **b**, The tissue preference of neutrophil clusters in mice revealed by $R_{o/e}$ (ratio of observed cell number to expected cell number). **c**, Heat map comparing representative gene expression across neutrophil clusters in humans and mice. **d**, The expression of *Cd274* in neutrophil clusters (top) and different tissues (bottom) in mice. **e**, Schematic of the anti-Ly6G treatment procedure. **f**, Representative photos of tumours generated in anti-Ly6G and control groups. The ruler tick marks show mm. The bar plots (right) show the nodule numbers per liver and the ratio of liver weight to body weight. $n = 15$. **g**, FACS analyses showing the proportions of macrophages ($n = 10$) and CD8$^+$ T cells ($n = 10$), and the expression of functional markers (PD1 and TIM3; $n = 6$) in tumours of the anti-Ly6G and isotype groups. **h**, The proportions of neutrophils (Ly6G) and proliferating malignant cells (Ki-67) assessed by IHC ($n = 6$). **i**, Schematic of the anti-Ly6G treatment in a therapeutic manner. **j**, Representative images and quantitative results of the tumour load examined by luminescence at 7 days and 36 days in **i**. $n = 5$. Representative photos of tumours at 36 days are also shown. For **f**–**j**, n values denote biologically independent samples; data are mean ± s.e.m. Statistical analysis was performed using two-sided Student's t-tests (**f**–**i** and **j** (top)) and a one-sided Student's t-test (**j**, bottom).

Fig. 10l–m). Furthermore, we performed the Ly6G blockade therapeutically in pTMC mice with the luciferase reporter (pTMC-Luc) (Fig. 4i). Ly6G blockade at 36 days after tumour formation (at 7 days) showed substantial reductions in bioluminescence signal and tumour burdens (Fig. 4j). Collectively, neutrophil depletion could attenuate macrophage recruitment and T cell suppression, resulting in tumour inhibition.

To further investigate the neutrophil dynamics during the anti-Ly6G treatment, we performed scRNA-seq of eight samples covering bone marrow, peripheral blood, tumour-adjacent liver tissues and tumours from two mice (Supplementary Fig. 4). Although most TANs diminished after the treatment, mNeu_09_Apoa2 retained and expanded relatively in the tumour. Correspondingly, its human counterpart, Neu_07_APOA2, was associated with favourable prognosis. Furthermore, both human and mouse *APOA2*$^+$ TANs exhibited unique lipid metabolism signatures similar to hepatic lipid-associated macrophages[39,40] and might therefore be lipid-associated neutrophils (Supplementary Note 9). Taken together, neutrophil depletion can alter the TAN composition and attenuate tumour progression in mouse models.

Discussion

On the basis of about 1.3 million cells from human and mouse, our large-scale, sorting-free single-cell analyses delineate a comprehensive cellular landscape of PLC, enabling us to identify five TIMELASER subtypes and decode the neutrophil heterogeneity. The TIMELASER framework covers most cell populations and provides a non-biased stratification of baseline TIME subtypes manifesting spatial resolution.

We speculate that in-depth analysis of these data, along with functional studies, will provide new insights for tumour-TIME and TIME-TIME crosstalk, assist to identify immune cell functions, and guide the identification of biomarkers or targets for immunotherapies[41]. The heterogeneity of neutrophils and their functions in tumorigenesis have been under intense investigation[6–9]. We identified a neutrophil spectrum that is broadly conserved between humans and mice, and clearly show their gene expression, gene signatures and developmental trajectories governed by different sets of transcription factors. Our investigation shows that TANs exhibit a collective pro-tumour phenotype, among which we speculate that the pro-tumour $CCL4^+$, $SPP1^+$ and $PD\text{-}L1^+$ TANs are promising immunotherapy targets, either alone or in combination with immune checkpoint inhibitors. Further exploring the impact of neutrophils on immunotherapies and related confounding clinical factors would offer new opportunities to better understand TAN biology and propose translational research paths for treating liver cancer.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-022-05400-x.

1. Binnewies, M. et al. Understanding the tumor immune microenvironment (TIME) for effective therapy. *Nat. Med.* **24**, 541–550 (2018).
2. Thorsson, V. et al. The immune landscape of cancer. *Immunity* **48**, 812–830 (2018).
3. Li, X. et al. The immunological and metabolic landscape in primary and metastatic liver cancer. *Nat. Rev. Cancer* **21**, 541–557 (2021).
4. Marquardt, J. U., Andersen, J. B. & Thorgeirsson, S. S. Functional and genetic deconstruction of the cellular origin in liver cancer. *Nat. Rev. Cancer* **15**, 653–667 (2015).
5. Finn, R. S. et al. Atezolizumab plus bevacizumab in unresectable hepatocellular carcinoma. *N. Engl. J. Med.* **382**, 1894–1905 (2020).
6. Coffelt, S. B., Wellenstein, M. D. & de Visser, K. E. Neutrophils in cancer: neutral no more. *Nat. Rev. Cancer* **16**, 431–446 (2016).
7. Jaillon, S. et al. Neutrophil diversity and plasticity in tumour progression and therapy. *Nat. Rev. Cancer* **20**, 485–503 (2020).
8. Shaul, M. E. & Fridlender, Z. G. Tumour-associated neutrophils in patients with cancer. *Nat. Rev. Clin. Oncol.* **16**, 601–620 (2019).
9. Ng, L. G., Ostuni, R. & Hidalgo, A. Heterogeneity of neutrophils. *Nat. Rev. Immunol.* **19**, 255–265 (2019).
10. Quail, D. F. et al. Neutrophil phenotypes and functions in cancer: a consensus statement. *J. Exp. Med.* **219**, e20220011 (2022).
11. Cui, C. et al. Neutrophil elastase selectively kills cancer cells and attenuates tumorigenesis. *Cell* **184**, 3163–3177 (2021).
12. Ponzetta, A. et al. Neutrophils driving unconventional T cells mediate resistance against murine sarcomas and selected human tumors. *Cell* **178**, 346–360 (2019).
13. Yang, L. et al. DNA of neutrophil extracellular traps promotes cancer metastasis via CCDC25. *Nature* **583**, 133–138 (2020).
14. Szczerba, B. M. et al. Neutrophils escort circulating tumour cells to enable cell cycle progression. *Nature* **566**, 553–557 (2019).
15. Zhang, Q. et al. Landscape and dynamics of single immune cells in hepatocellular carcinoma. *Cell* **179**, 829–845(2019).
16. Sharma, A. et al. Onco-fetal reprogramming of endothelial cells drives immunosuppressive macrophages in hepatocellular carcinoma. *Cell* **183**, 377–394 (2020).
17. Ma, L. et al. Single-cell atlas of tumor cell evolution in response to therapy in hepatocellular carcinoma and intrahepatic cholangiocarcinoma. *J. Hepatol.* **75**, 1397–1408 (2021).
18. Ma, L. et al. Tumor cell biodiversity drives microenvironmental reprogramming in liver cancer. *Cancer Cell* **36**, 418–430 (2019).
19. Sun, Y. et al. Single-cell landscape of the ecosystem in early-relapse hepatocellular carcinoma. *Cell* **184**, 404–421 (2021).
20. Zheng, C. et al. Landscape of infiltrating t cells in liver cancer revealed by single-cell sequencing. *Cell* **169**, 1342–1356 (2017).
21. Zhang, M. et al. Single-cell transcriptomic architecture and intercellular crosstalk of human intrahepatic cholangiocarcinoma. *J. Hepatol.* **73**, 1118–1130 (2020).
22. Aizarani, N. et al. A human liver cell atlas reveals heterogeneity and epithelial progenitors. *Nature* **572**, 199–204 (2019).
23. Xie, X. et al. Single-cell transcriptome profiling reveals neutrophil heterogeneity in homeostasis and infection. *Nat. Immunol.* **21**, 1119–1133 (2020).
24. Zilionis, R. et al. Single-cell transcriptomics of human and mouse lung cancers reveals conserved myeloid populations across individuals and species. *Immunity* **50**, 1317–1334 (2019).
25. Ballesteros, I. et al. Co-option of neutrophil fates by tissue environments. *Cell* **183**, 1282–1297 (2020).
26. Bagaev, A. et al. Conserved pan-cancer microenvironment subtypes predict response to immunotherapy. *Cancer Cell* **39**, 845–865 (2021).
27. Zhang, L. et al. Single-cell analyses inform mechanisms of myeloid-targeted therapies in colon cancer. *Cell* **181**, 442–459 (2020).
28. Kiss, M. et al. IL1β promotes immune suppression in the tumor microenvironment independent of the inflammasome and gasdermin D. *Cancer Immunol. Res.* **9**, 309–323 (2021).
29. Feig, C. et al. Targeting CXCL12 from FAP-expressing carcinoma-associated fibroblasts synergizes with anti-PD-L1 immunotherapy in pancreatic cancer. *Proc. Natl Acad. Sci. USA* **110**, 20212–20217 (2013).
30. Cancer Genome Atlas Research Network. Comprehensive and integrative genomic characterization of hepatocellular carcinoma. *Cell* **169**, 1327–1341 (2017).
31. Farshidfar, F. et al. Integrative genomic analysis of cholangiocarcinoma identifies distinct IDH-mutant molecular profiles. *Cell Rep.* **18**, 2780–2794 (2017).
32. Xue, R. et al. Genomic and transcriptomic profiling of combined hepatocellular and intrahepatic cholangiocarcinoma reveals distinct molecular subtypes. *Cancer Cell* **35**, 932–947 (2019).
33. Wu, R. et al. Comprehensive analysis of spatial architecture in primary liver cancer. *Sci. Adv.* **7**, eabg3750 (2021).
34. Xiang, X. et al. IDH mutation subgroup status associates with intratumor heterogeneity and the tumor microenvironment in intrahepatic cholangiocarcinoma. *Adv. Sci.* **8**, e2101230 (2021).
35. Wu, S. Z. et al. A single-cell and spatially resolved atlas of human breast cancers. *Nat. Genet.* **53**, 1334–1347 (2021).
36. Ramachandran, P., Matchett, K. P., Dobie, R., Wilson-Kanamori, J. R. & Henderson, N. C. Single-cell technologies in hepatology: new insights into liver biology and disease pathogenesis. *Nat. Rev. Gastroenterol. Hepatol.* **17**, 457–472 (2020).
37. Wculek, S. K. & Malanchi, I. Neutrophils support lung colonization of metastasis-initiating breast cancer cells. *Nature* **528**, 413–417 (2015).
38. Boivin, G. et al. Durable and controlled depletion of neutrophils in mice. *Nat. Commun.* **11**, 2762 (2020).
39. Remmerie, A. et al. Osteopontin expression identifies a subset of recruited macrophages distinct from kupffer cells in the fatty liver. *Immunity* **53**, 641–657 (2020).
40. Guilliams, M. et al. Spatial proteogenomics reveals distinct and evolutionarily conserved hepatic macrophage niches. *Cell* **185**, 379–396 (2022).
41. Zhu, A. X. et al. Molecular correlates of clinical response and resistance to atezolizumab in combination with bevacizumab in advanced hepatocellular carcinoma. *Nat. Med.* **28**, 1599–1611 (2022).

Methods

Patient sample collection

This study was approved by the Research Ethics Committee of both Peking University First Hospital and Peking University People's Hospital. Written informed consent was obtained from each patient. We performed a prospective screen of treatment-naive patients with liver cancer who underwent primary curative resection from March 2019 to January 2020 at Peking University People's Hospital. Fresh tumour and tumour-adjacent liver tissue (AL) samples (at least 2 cm from tumour tissues) were collected within 30 min after the operation. Peripheral blood (PB) samples were collected before the surgery. A total of 124 patients were enrolled and 160 samples were obtained for scRNA-seq, including 79 HCC, 25 ICC, 7 CHC, 2 hepatic haemangioma (HH), 1 adenosquamous carcinoma (ASC), 1 sarcomatoid carcinoma (SAR) and 9 secondary liver cancer (SLC, liver metastases from various primary sites) cases. Focusing on PLC, proportions of HCC, ICC and CHC patients were 69.9%, 22.1% and 6.2%, respectively, consistent with the PLC incidence worldwide. For 14 of 124 patients, AL and PB were collected in addition to tumours. We also performed whole-exome sequencing (WES) for 84 of these patients. No blinding or randomization was performed for the human tumor samples, because this was an observational study. Detailed clinical characteristics and mutational profiles are summarized in Supplementary Tables 1 and 2, respectively.

Mouse models

Trp53$^{fl/fl}$ and *Alb-cre* mice (both C57BL/6) were purchased from the Jackson Laboratory and bred in a pathogen-free environment according to the guidelines of the animal facility in Peking University First Hospital. *Trp53*$^{fl/fl}$ mice were crossed with *Alb-cre* mice to generate the liver conditional *Trp53*-knockout (*Trp53* cKO) mice. 7-week-old male Trp53 cKO mice were used for subsequent experiments. Sleeping beauty transposase (SB100) and transposon pT3-Neo-EF1a-GFP plasmids were purchased from Addgene. cDNA of mouse *Myc* gene was cloned into the transposon vector through the MluI and SpeI restriction enzyme sites, obtaining the pT3-Neo-EF1a-*Myc* plasmid. Next, mutated forms of mouse *Ctnnb1* (*Δ90Ctnnb1*) or mouse *Kras* (*Kras*^{G12D}) were generated by PCR cloning of mouse *Ctnnb1* or *Kras* cDNA. Then, the *Myc* and *Δ90Ctnnb1* transposon plasmid (pT3-EF1a-*Myc*-*Δ90Ctnnb1*, pTMC) was generated through the AscI and NotI restriction sites. Similarly, the *Myc* and *Kras*^{G12D} transposon plasmid (pT3-EF1a-*Myc*-*Kras*^{G12D}, pTMK) was generated. For construction of the pTMC-luciferase plasmid (pTMC-Luc), the luciferase fragment was linked to *Myc* by P2A using In-Fusion cloning. Plasmids for hydrodynamic tail-vein (HDTV) injection were prepared using the EndoFree-Maxi Kit (Qiagen). For HDTV, a total of 30 μg DNA mixture (5:1 ratio of transposon to transposase-encoding plasmid) was suspended in 0.9% saline solution at a final volume equal to 10% of body weight of the mice, and was then injected into 7-week-old male *Trp53*-cKO mice through the tail vein within 5–7 s (ref. [42]). All of the mice were housed in pathogen-free conditions at an ambient temperature 20–26 °C and humidity of 30–70% with a 12 h–12 h light–dark cycle before use. The body weight of mice was monitored twice every week for signs of dynamic tumour growth. The diameter of single tumour was <2 cm. For cKO mouse models, body-weight-matched mice were randomized over the treatment groups, anti-Ly6G and isotype control. For pTMC-Luc mouse model, tumor size was monitored by luminescence signals at day 7 after HDTVi, and tumor-size-matched mice were randomized over the treatment groups, anti-Ly6G and isotype control. No blinding was performed for mouse samples, because this was an observational study. All of the mouse experiments were approved by the Animal Care and Use Committee at Peking University First Hospital.

scRNA-seq analysis of human and mouse samples

Fresh tumour and AL samples were cut into approximately 1 mm^3 pieces in RPMI-1640 medium (Thermo Fisher Scientific) with 10% fetal bovine serum (FBS, Gibco) and enzymatically digested using the MACS tumour dissociation kit (Miltenyi Biotec) for 30 min on a rotor at 37 °C, according to the manufacturer's instructions. After filtering using the 70 μm CellStrainer (BD) in RPMI-1640 medium, the suspended cells were centrifuged at 400*g* for 5 min. After removing the supernatant, cell pellets were resuspended in sorting buffer (PBS supplemented with 2% FBS) after washing twice with PBS. Moreover, 10 ml of fresh PB samples were collected before surgery in EDTA anticoagulant tubes. For PB samples, RBC removal was performed using the ErythroClear kit (STEMCELL) according to the manufacturer's instructions. After collecting single-cell suspensions for different samples, 10 μl of each cell suspension was counted using an automated cell counter (Luna-II, Logos Biosystems) to assess the number of live cells. Throughout the dissociation procedure, cells were maintained on ice whenever possible. The entire procedure was completed in less than 1 h (typically ~45 min) to avoid dissociation-associated artifacts. Cell viability and concentration were then assessed using the Rigel S3 fluorescence cell analyser (Countstar).

To avoid biases introduced by any enrichment steps on the cellular composition of queried samples, the original unsorted single-cell suspensions were directly used for subsequent library construction. Cells were loaded onto the Chromium single cell controller (10x Genomics) to generate single-cell gel beads in the emulsion according to the manufacturer's protocol. scRNA-seq libraries were constructed using Single Cell 3′ Library and Gel Bead Kit v3.1 and sequenced using the NovaSeq 6000 sequencer (Illumina).

scRNA-seq data processing

scRNA-seq data were aligned and quantified using the CellRanger toolkit v.3.1 against the reference genome GRCh38 and GRCm39 for human and mouse samples, respectively. Empty droplets were filtered using the emptyDrops function of the R package dropletUtils v.1.10.3 by assessing whether the RNA content associated with a cell barcode is significantly distinct from the ambient background RNA present within each sample. Cells with FDR < 0.01 (Benjamini–Hochberg-corrected) were selected for further analysis. The quality of cells was assessed based on three metrics: (1) the number of total UMI count per cell (library size) was below 30,000; (2) the number of detected genes was above 500 and below 6,000; (3) the percentage of mitochondrial genes was below 50. As neutrophils showed very low transcript counts as reported[24], the range of detected gene number of neutrophils was set as 100–6,000. Next, we used a cluster-level approach to remove potential doublet cells. In brief, the doublet score was calculated for each cell using doubletCells function of the scran R package v.1.18.7. Cell clusters in each sample were identified by examining the top 50 principal components (PCs) across highly variable genes (HVGs), building neighbour graph by buildSNNGraph function, and then clustering using the cluster_louvain function from the igraph R package v.1.2.9. The median doublet score of each cell cluster was calculated using median-centred MAD-variance normal distribution. Clusters with a median score above the extreme top end of this distribution (Benjamini–Hochberg-corrected $P < 0.1$) were considered as doublets. After quality control, a total of 1,297,609 cells comprising 1,092,172 cells from 160 human samples (124 patients) and 205,437 cells from 29 mouse samples (8 mice) were retained for downstream analysis. Raw counts and $\log_2$ (normalized counts) were computed for each cell. As immune and stromal cells from different patients mixed well, we did not observe obvious batch effect. Gene–cell count matrices from different samples were merged using Seurat (v.3.2.3)[43].

Cell clustering and annotation

To identify major cell types, we used scanpy (v.1.6) Python package[44]. A total of 2,000 HVGs were selected using the highly_variable_genes function, and then the top 50 PCs were calculated using the pca function. We regressed out the effect of percentage of mitochondrial genes

and scaled each gene to unit variance. Nearest neighbourhood graphs were built using the neighbours function, and the community algorithm was applied for clustering using the louvain function (resolution = 1). The dimensionality of each dataset was reduced using UMAP.

We first annotated the 14 major cells types identified in our dataset on the basis of well-known marker genes, including *CD3D*, *CD8A*, *CD4*, *FOXP3*, *TRDC*, *NKG7*, *CD79A* and *MS4A1* for lymphoid lineage ($CD8^+$ T, conventional $CD4^+$ T, T regulatory, γδT, natural killer and B cells); *CD14*, *CD16*, *CD68*, *CD163*, *CD1C*, *LAMP3*, *TPSAB1*, *CSF3R* and *S100A8* for myeloid lineage (monocytes, macrophages, dendritic cells, mast cells and neutrophils); *VWF* and *COL1A1* for stromal cells (endothelial cells and fibroblasts); and *ALB* and *EPCAM* for epithelial cells. Epithelial cells, composed of hepatocytes, cholangiocytes and progenitor cells, were analysed as a whole in cluster analysis. Among these epithelial cells, malignant cells were further distinguished from non-malignant cells by inferring large-scale copy-number variations (CNVs) of each cell using inferCNV (v.1.3.3) R package as described[45]. As non-malignant cells derived from ALs were annotated, we used the average patterns of these cells as a reference for the CNV estimation.

Next, we performed a second round of clustering to further characterize subpopulations of major cell types in the TIME. We converted the scanpy object to Seurat object using the anadata Python package (v.0.7.5) and then clustering using Seurat (v.3.2.3)[43]. To avoid unexpected noise and expression artefacts by dissociation, a total of 1,514 genes associated with mitochondria (50 genes), heat-shock protein (178 genes), ribosome (1,253 genes) and dissociation (33 genes) were excluded (Supplementary Table 1). Owing to variable amount and property of cells in each major cell type, different parameters for clustering were used. For the clustering of T cells, top 20 PCs were selected on the basis of 2,000 HVGs (resolution = 1). For the clustering of natural killer or B cells, the top 10 PCs were selected on the basis of 1,000 HVGs (resolution = 0.6). For monocytes or dendritic cells, the top 10 PCs were selected on the basis of 1,000 HVGs (resolution = 0.8). For macrophages, the top 10 PCs were selected on the basis of 1,500 HVGs (resolution = 1). For endothelial cells or fibroblasts, the top 15 PCs were selected on the basis of 1,000 HVGs (resolution = 1). For neutrophils, the top 8 PCs were selected on the basis of 500 HVGs (resolution = 0.8). Specifically, the resolution of neutrophil clusters was determined on the basis of its biological features. Here we took a scRNA-seq dataset of neutrophils from PB as a reference[23]. The reported three neutrophil subsets (G5a–c) in PB were recapitulated with resolutions of 0.7 and 0.8, with the latter having a better separation of neutrophil clusters.

As a result, we identified 13 $CD8^+$ T, 8 conventional $CD4^+$ T, 3 T regulatory, 4 γδ T, 7 natural killer and 4 B cell clusters for the lymphoid lineage, 5 monocyte, 9 macrophage, 7 DC, 1 mast cell and 11 neutrophil clusters for the myeloid lineage, and 10 endothelial cell and 7 mesenchymal clusters[35,46] for the stromal components. To facilitate data visualization in Fig. 1c, cells were reclustered into four embeddings using Seurat, including (1) T cells, (2) natural killer and B cells, (3) myeloid cells and (4) endothelial cells and mesenchymal cells. Next, we used the FindMarkers or FindAllMarkers function to identify differentially expressed genes (DEGs) with adjusted $P < 0.05$ using Bonferroni correction. Gene Ontology analysis was performed using the clusterProfiler R package (v.3.18.1).

Experimental and analytical strategies for neutrophils

As neutrophils are very fragile and have a relatively low level of RNA content, both experimental and analytical procedures were improved to capture neutrophils during this study. For scRNA-seq experiments, we kept a minimal hands-on time for the tissue samples. When single-cell suspensions were collected, only the MACS Dead Cell Removal Kit (Miltenyi Biotec) was used to collect viable cells and no FACS enrichment steps were applied, therefore limiting the experimental process from tissue collection (for both tumour and adjacent liver tissues) after surgery to PCR with reverse transcription within 2 h. Prolonging the processing time may cause the failure of neutrophil capture. For data analysis, we set the range of detected UMI as 100–6,000 for neutrophils, while keeping that of other cell populations as 500–6,000 for downstream analysis. A total of 34,307 neutrophils were identified on the basis of the expression of *CSF3R*, *S100A8* and *S100A9*[23,24]. Eleven subsets of human neutrophils were characterized (Fig. 3a and Extended Data Fig. 7b–d) and exhibited clear separation according to the tissue sources of PB, AL and tumour (Fig. 3a,b and Extended Data Fig. 7e), consistent with the previous notion that neutrophils exhibited tissue specificity[6–9]. SingleR (v.1.10.0)[47] was also used to assess the similarity of neutrophil clusters in this study compared to previously reported neutrophil subsets. Neu_02_S100A12, Neu_03_ISG15 and Neu_04_TXNIP were mainly composed of PBNs, matching the reported circulating G5a, G5b and G5c states[23,24] (Extended Data Fig. 7f). Neu_05_ELL2 and Neu_06_PTGS2 were mainly ALNs. All of the other six subsets (Neu_01_MMP8, Neu_07_APOA2, Neu_08_CD74, Neu_09_IFIT1, Neu_10_SPP1 and Neu_11_CCL4) were enriched in tumours and therefore designated as TANs. These TANs were differentially enriched across PLC subtypes, with Neu_01_MMP8 and Neu_07_APOA2 in HCC, and Neu_09_IFIT1, Neu_10_SPP1 and Neu_11_CCL4 in ICC (Fig. 3b).

Calculation of gene signature scores based on scRNA-seq data

Multiple gene signature scores were calculated on the basis of the scRNA-seq data. For each gene signature, individual cells were scored using the AddModuleScore function, which calculated the average expression levels of selected genes at the single-cell level and subtracted by the aggregated expression of control feature sets. Control features were composed of 100 randomly selected genes from each bin where all features were binned into 24 groups based on averaged expression. For malignant cells, hepatic score was calculated based on the expression of 21 hepatocyte-related genes[18] (*ADH1A*, *ADH4*, *AFM*, *AHSG*, *AMBP*, *C4BPB*, *C6*, *CYP2E1*, *CYP4F2*, *F9*, *FGA*, *FGB*, *FGG*, *GC*, *HPX*, *PROC*, *SAA4*, *SERPINA6*, *SERPINC1*, *SERPIND1* and *SLC2A2*). Biliary epithelial score was calculated based on the expression of 13 cholangiocyte-related genes (*KRT14*, *KRT17*, *KRT6A*, *KRT5*, *KRT19*, *KRT8*, *KRT16*, *KRT6B*, *KRT15*, *KRT6C*, *KRTCAP3*, *SFN* and *EPCAM*). For neutrophils, scores for azurophil granule, specific granule, gelatinase granule, secretory vesicle, neutrophil maturation and neutrophil ageing were calculated[23] (Supplementary Table 3). Other functional signatures for neutrophil activation (GO:0042119), neutrophil chemotaxis (GO:0030593), apoptosis (GO:0043065), angiogenesis (GO:0001525), extracellular matrix (GO:0031012), phagocytosis (GO:0006911), type I interferon signalling pathway (GO:0060337) and chemokine activity (GO:008009) were derived from the Gene Ontology database.

Tissue and cancer type enrichment of clusters

To quantify the enrichment of cell clusters across tissues (PB, AL and tumour) and PLC subtypes (HCC, ICC and CHC), we compared the observed and expected cell numbers in each cluster by computing the $R_{o/e}$ value using the epitools (v.0.5-10.1) R package according to the following formula[15]:

$$R_{o/e} = \frac{\text{Observed}}{\text{Expected}},$$

where the expected cell numbers for each combination of cell clusters and tissues were obtained from the χ^2 test. We assumed that one cluster was enriched in a specific tissue or cancer type if $R_{o/e} > 1$.

Identification of cellular modules and TIMELASER subtypes

To examine the potential cellular compositions of different TIME ecosystems in liver cancer, we investigated the co-existence patterns of different TIME cell subpopulations. Pairwise correlation values between the normalized frequency of any two clusters across different tumour samples were calculated using the corr.test function. These values were

then clustered using the pheatmap (v.1.0.12) R package with the ward.D2 cluster method and correlation distance. To avoid potential distortion of clustering due to the low frequency of certain clusters (present in less than 10 tumours), tumour cells fall into 13 PB-enriched clusters (CD4T_01_CCR7, CD4T_09_FOXP3, CD8T_01_CCR7, CD8T_02_CX3CR1, CD8T_03_GZMK_S1PR1, γδT_01_GNLY_S1PR5, NK_01_FCGR3A_CX3CR1, MonoDC, Mo_01_CD14, Mo_02_CD16, Neu_02_S100A12, Neu_03_ISG15 and Neu_04_TXNIP) and 2 additional clusters (Neu_01_MMP8 and Fb_06_FABP3) were excluded from this analysis. As a result, we identified five highly correlated cellular modules. For each patient, cluster-normalized frequencies of clusters from the same cellular module were summed and the most abundant cellular module was designated as the dominant cellular module for this patient. Each cellular module corresponds to a TIMELASER subtype, of which the phenotype was designated based on four aspects: (1) cell populations, (2) marker genes, (3) TIME-related gene signatures as previously defined[26] and (4) prognostic relevance, which combinatorically support the phenotype of our TIME subtypes (Extended Data Fig. 5e,f and Supplementary Table 3).

Classification of TIMELASER subtypes for bulk RNA-seq data

To apply our single-cell based TIMELASER subtypes to published bulk RNA-seq data, we defined gene signatures for each subtype by combining top 8 DEGs of all clusters in the corresponding cellular module (Supplementary Table 3). For each patient, *z*-scores of 5 TIMELASER signatures were computed. First, TIMELASER signature scores were calculated on the basis of the average expression of signature genes, and then subtracted by the aggregated expression of control features. Control features were composed of 100 randomly selected genes from each bin where all features were binned to 24 groups based on averaged expression. Next, *z*-scores of five TIMELASER signatures were calculated by scaling five scores in the same sample. The TIMELASER subtype of each patient was then determined on the basis of the highest signature score across five *z*-scores. For example, we assembled a bulk RNA-seq dataset of 453 patients with PLC collected from three published studies, including TCGA-LIHC (HCC study of TCGA)[30], TCGA-CHOL (ICC study of TCGA)[31] and our previous study of CHC[32]. Classification of TIMELASER subtypes was performed for this large cohort dataset.

To compare our single-cell-based TIMELASER subtypes with the bulk RNA-seq data based molecular functional portrait subtypes[26], we calculated the expression levels of molecular functional portrait signatures for each individual in our dataset (Supplementary Table 3). We first calculated the average expression of a certain gene across TIME cells in each patient using the AverageExpression function. The signature scores were then calculated by the mean expression of involved genes.

WES

DNA was extracted using the DNeasy Blood & Tissue Kit (Qiagen) from fresh-frozen tumour and AL samples. A total of 200 ng to 1 μg DNA was taken from each sample and sheared into fragments of ~300 bp using a Covaris S2 ultrasonicator. The library was constructed using the NEBNext Ultra DNA Library Prep Kit for Illumina and exome regions were captured using Agilent SureSelect All Exon V6. The post-hybridization amplification product was quality-checked and sequenced. Paired-end Illumina reads were aligned to the human genome hg38 (UCSC) using BWA-mem2 (v.2.0pre1) with the default parameters. SAM files were then converted to BAM files and sorted by chromosomal coordinates using Samtools (v.1.10). The Genome Analysis Toolkit (GATK, v.4.1.7.0) was used to remove PCR duplicates and recalibrate the base quality score. Point mutations and indels were identified using Mutect2 (v.4.1.0.0) and VarScan (v.2.4.2). All variants were annotated using ANNOVAR. A series of filtering criteria were applied to the variant candidates: (1) at least 10× coverage was required in the normal sample of each patient bearing at most 1× mutation coverage; (2) at least 10× total coverage was required in tumour samples with at least 3× mutation coverage; (3) variations listed in dbSNP150 were removed unless they were documented in the Catalog of Somatic Mutations in Cancer (COSMIC) database. Finally, point mutations identified by Mutect2 and indels identified by both Mutect2 and VarScan were retained after filtering. All of the variants were annotated using the VEP (v.96; Ensembl Variant Effect Predictor). Tumour ploidy and cellularity were inferred using ABSOLUTE (v.1.0.6). CNVkit (v.0.9.7) was then performed using the default parameters on paired tumour–normal WES data. After segmentation, GISTIC2 (v.2.0.23; Genomic Identification of Significant Targets in Cancer) was applied to identify focal CNVs.

Gene modules of malignant cells

Gene modules of malignant cells were extracted as described previously[35]. For each individual tumour with more than 50 malignant cells, clusters were calculated using Seurat (v.3.2.3) at five resolutions (0.5, 0.8, 1, 1.2, 1.5). For each cluster, the top 200 DEGs were identified and only clusters with more than five tumour cells and more than five DEGs were retained. The DEGs of these clusters were then defined as a gene signature. The redundancy of gene signatures identified from the five resolutions was reduced by a pairwise comparison of gene signatures within each sample. For each pair with a Jaccard index > 0.75, the gene signature with fewer genes was removed. Across all tumours, 1,187 gene signatures were identified. Consensus clustering of the Jaccard similarities between these gene signatures identified eight gene modules. Highly recurrent genes were identified for each gene module and the enriched pathways were calculated using ClusterProfiler (v.3.18.1).

Cell–cell interactions

To investigate cell–cell interactions among clusters from each cellular module, we analysed the L–R pairs using CellphoneDB (v.2.1.7) as described previously[48]. In brief, a $\log_2$-normalized count matrix was subsampled into 500 cells per cluster. Significant L–R pairs were identified after filtering for frequencies below 0.1% or above 2% of all cluster–cluster combinations. For each L–R pair, the total number of this L–R pair across clusters from the same cellular module was counted. Cellular-module-specific L–R pairs were then determined based on the enrichment score by $R_{o/e}$ values ($R_{o/e} > 3$). To identify potential ligands that drive the unique phenotype of Neu_09_IFIT1, we compared the transcriptomic differences between Neu_03_ISG15 and Neu_09_IFIT1, and then used the highly expressed genes in Neu_09_IFIT1 for NicheNet (v.1.1.0) analysis. Genes with $\log_2$[fold change] > 0.2 and adjusted $P < 0.05$ were then used as gene sets of interest. Genes were considered to be expressed when they had non-zero values in at least 10% of the cells in a cell type.

Developmental trajectory

CytoTRACE (v.0.3.3)[49], Monocle (v.2.12)[50], and CellRank (v.1.5.1)[51] were adopted to infer the developmental trajectory of human and mouse neutrophils. CytoTRACE is based on the notion that transcriptional diversity, that is, the number of genes expressed in a cell decreases during differentiation. The $\log_2$-normalized expression matrix was accessed. The predicted orders were projected onto the neutrophil UMAP space. For Monocle 2, we built a new CellDataSet object from cluster-annotated Seurat object using the newCellDataSet function. We used the differentialGeneTest function to derive DEGs from each cluster, and genes with $q < 1 \times 10^{-5}$ were used to order the cells in pseudotime. Dimension reduction was performed using the DDRTree algorithm and then cells were ordered along the trajectory. Moreover, the CytoTRACE scores were also projected on the Monocle trajectory. CellRank was performed to map the cell fate of neutrophil subsets after anti-Ly6G treatment as described.

Regulon network

The regulon network was explored using the R package SCENIC (v.1.1.3)[52], which analysed the co-expression of transcriptional factors and their putative target genes. We built and scored gene regulatory

network using the default parameters. Raw count matrix was used to build co-expression network using the runCorrelation and runGENIE3 functions. Potential regulons based on DNA-motif analysis were selected by RcisTarget and active gene networks were identified by AUCell. Regulon activity for each cell was calculated as the average normalized expression of putative target genes.

Cross-species data integration

Cross-species single-cell data integration was performed using the LIGER v.1.0 workflow[53]. In brief, single-cell datasets of mouse and human neutrophils were preprocessed to produce a raw digital gene expression matrix using createLiger and then normalized. Variable genes were selected and the gene expression was scaled using scaleNotCenter. Shared and species-specific factors were identified through integrative non-negative matrix factorization using optimizeALS. Joint clustering of cells was performed by louvainCluster and then visualized using UMAP.

Survival analysis

Prognostic values of cell clusters and cellular modules were evaluated in our cohort. Kaplan–Meier survival curves were plotted using ggsurvplot function in the R package Survminer v.0.4.9.

IHC and mIHC

Formalin-fixed and paraffin-embedded (FFPE) tissues sectioned to 4 μm were used for histology evaluation of liver tumours in both human and mouse models. Haematoxylin and eosin (H&E) staining was performed for each sample. For IHC and mIHC, tissue slides were deparaffinized with xylene and rehydrated through a graded series of ethanol solutions (100%, 95% and 70%). Then, slides were treated by microwave to induce antigen retrieval using citric acid solution for 15 min. For mouse tumours, primary antibodies for anti-hepatocyte (1:500, ab75677, Abcam), anti-EPCAM (1:200, ab213500, Abcam), anti-Ly6G (1:500, GB11229, Servicebio), anti-Ki-67 (1:500, ab15580, Abcam) and anti-CD68 (1:200, GB113109, Servicebio) were used. Each section was evaluated by 2–3 experienced pathologists. For mIHC analysis of human samples, three panels of primary antibodies were used, including, (1) CD66b (1:1,000, GTX19779, GeneTex) and CCL4 (1:800, ab235961, Abcam); (2) Von (1:100, ab9378, Abcam), α-SMA (1:5,000, ab7817, Abcam), CD8 (1:100, ZA0508, ZSGB), GZMK (1:1,000, ab282703, Abcam); (3) CD8 (1:100, ZA0508, ZSGB), PD1 (1:50, ZM0381, ZSGB), CD66b (as in panel 1) and PD-L1 (1:1,000, ab237726, Abcam). The slides were then incubated with secondary antibodies (1:1,100 μl for each slide; HRP-anti-rabbit IgG, ZSGB, PV-6001; or HRP-anti-mouse IgG, ZSGB, PV-6002) for 10 min at room temperature. After each cycle of staining, heat-induced epitope retrieval was performed to remove all the antibodies including primary antibodies and secondary antibodies. Multiplex immunofluorescence staining was performed using the AlphaTSA Multiplex IHC Kit (AXT36100031, AlphaX). The samples were counterstained for nuclei with DAPI for 10 min and mounted in mounting medium. Multispectral images were scanned with ZEISS AXIOSCAN 7. Cells of interest were quantified using Halo (v.3.4; Indica Labs) or QuPath (v.0.2.0).

CODEX

CODEX was performed on FFPE tissues according to the manufacturer's instructions (Akoya Biosciences)[54]. In brief, 4 μm tissue sections were mounted on poly-L-Lysine-coated coverslips and then deparaffinized and rehydrated. The tissue-retrieval process is the same as for IHC. Tissues were then fixed using prestaining fixing solution and then washed using tissue hydration buffer. For each coverslip, the antibody cocktail (containing β-catenin, CD3e, CD4, CD8, CD11c, CD20, CD31, CD45RO, CD68, E-cadherin, HLA-DR, keratin14, Ki-67, MAC2/galectin-3 and pan-cytokeratin) was then added to the coverslip and staining was performed in a sealed humidity chamber for 3 h. After staining, coverslips were washed for 4 min by staining buffer and fixed in wells containing 1.6% paraformaldehyde for 10 min, followed by three washes in PBS. The coverslips were then incubated in 100% methanol on ice for 5 min, followed by three washes in PBS. Fresh fixative solution was prepared immediately before final fixation, and final fixation was performed at room temperature for 20 min, followed by three washes in PBS. Next, the CODEX reporter plate containing the reporter master mix for every cycle was prepared accordingly. The CODEX multicycle reaction and image acquisition were performed using the Akoya CODEX instrument. During imaging, the tissue was kept in H2 buffer. Hybridization of the fluorescent oligonucleotides was performed in rendering buffer. After imaging, fluorescent oligonucleotides were removed using stripping buffer. Data processing and analysis were performed using CODEX analysis manager and CODEX Multiplex Analysis Manager.

Isolation of immune cells from PB

PB samples (20 ml) were collected from healthy human donors or patients with liver cancer. Density gradient separation was performed with Lymphoprep (STEMCELL, 07861). The layer of peripheral blood mononuclear cells was sent for isolation of $CD8^+$ T cells with anti-CD8 magnetic beads (STEMCELL, 17853), followed by isolation of monocytes with anti-CD14 magnetic beads (STEMCELL, 19359). The bottom layer of erythrocyte/granulocyte pellet was resuspended with Red Cell Lysis Buffer (TIANGEN). Lysis was stopped using RPMI-1640 medium supplemented with 2% FBS, followed by centrifugation at 400*g* for 10 min. Cells were washed twice with PBS and filtered through a 70 μm nylon mesh (FALCON).

Isolation of neutrophils

Neutrophils were isolated from PB, AL and tumours from selected patients. For PB, neutrophils were extracted from the lower layer of the erythrocyte/granulocyte pellet after red blood cells were removed using the ErythroClear kit (STEMCELL). Cells were washed twice with PBS and filtered through a 70 μm nylon mesh (Falcon). Anti-CD66b antibodies (BD, 561650) coupled with magnetic anti-PE microbeads (STEMCELL, 17694) were used to purify neutrophils. For AL and tumours, single-cell suspensions of tissues were collected as descried for scRNA-seq, centrifuged at 300*g* for 5 min and resuspended in 36% Percoll (Sigma, P4937, diluted with PBS), followed by centrifugation at 500*g* for 15 min. Cell pellets were collected and washed twice with PBS. Anti-CD66b antibodies (BD, 561650) coupled with magnetic anti-PE microbeads (STEMCELL, 17694) were further used to purify neutrophils. For survival analysis, PBNs were cultured for 4 days and tested for viability at multiple time points using the Cell Counting Kit-8 (Bestbio). A total of 24.97% and 4.99% of PBNs remained alive after 1 and 3 days, consistent with the lifespan of cultured human neutrophils[55].

Co-culture of PBNs with cell lines

The human embryonic kidney cell line (HEK293T, ATCC number, CRL-3216) and the liver cancer cell line (HepG2, ATCC number, HB-8065) were obtained from American Type Culture Collection (ATCC). Human liver cancer cell lines (HCCLM3 and MHCC97H) were obtained from the Liver Cancer Institute, Zhongshan Hospital, Fudan University (Shanghai, China). All cell lines used in this study were authenticated by applying short tandem-repeat (STR) DNA profiling and tested negative for mycoplasma. All cells were cultured in RPMI1640 medium (Corning) supplemented with 10% fetal bovine serum (FBS) (VISTECH), 100 U ml^{-1} of penicillin and 100 μg ml^{-1} of streptomycin (Hyclone) in a humidified incubator at 37 °C with 5% CO_2. PBNs (1×10^6) were placed in the top insert of a Transwell (0.4 μm, Corning) and tumour cells were placed in the bottom chamber of a 12-well plate and co-cultured for 0 h, 6 h, 12 h, 18 h, 24 h and 30 h. After co-culture, PBNs were sent for qPCR, bulk RNA-seq and FACS analysis, including staining with anti-CD45 (BD, 557833), anti-CD66b (BD, 561650) and anti-PD-L1 (BD, 557924) antibodies.

Co-culture of neutrophils with CD8$^+$ T cells

For the co-culture experiment involving PBNs, cell lines and CD8$^+$ T cells, CD8$^+$ T cells and PBNs were isolated from the same donor at different time points. At day 1, CD8$^+$ T cells were purified and then stimulated with 25 μg ml^{-1} CD3/CD28 T Cell Activator (STEMCELL, 10971) and 50 U ml^{-1} rhIL-2 (STEMCELL, 78036.1) for 3 days. At day 3, PBNs from the same donor were isolated and placed in the lower chamber of a 12-well plate. Tumour cells were placed in the top insert of Transwell (0.4 μm, Corning). After co-culture of PBNs with cell lines for 12 h, CD8$^+$ T cells were added to the bottom chamber at a 1:5 ratio of CD8$^+$ T cells to PBNs and co-cultured for 24 h and 48 h.

For the co-culture experiment involving PBNs, ALNs and TANs with CD8$^+$ T cells, PBNs were extracted as described above, whereas ALNs and TANs were purified from single-cell suspensions of tumour and adjacent liver tissues with anti-CD66b antibodies coupled with magnetic anti-PE microbeads from the EasySep PE Selection Kit (STEMCELL). Purified PBNs, ALNs and TANs were directly co-cultured with CD8$^+$ T cells in a 12-well plate at a 1:2.5 ratio of CD8$^+$ T cells to PBNs for 24 h.

After the co-culture, these mixed cells were separated by a BD FACSAria SORP flow cytometer using FACSDiva (v.8.0.1), and the data were analysed using FlowJo (v.10.4). Antibodies against CD45 (BD, 557833), CD3 (BD, 562426), CD8 (BD, 560179) and CD11b (BioLegend, 101256) were used to gate CD8$^+$ T cells and neutrophils. PD-L1 antibodies (BD, 557924) were used to assess the immunosuppression of neutrophils. IFNγ (BD, 557643), GZMB (BD, 561142) and PRF1 (BD, 563762) antibodies were used to assess the cytotoxicity of CD8$^+$ T cells. CD69 (BD, 562884) and CD25 (BD, 563701) antibodies were used to assess the activation status of CD8$^+$ T cells. CFSE (BD, 565082) was used to assess the proliferation of CD8$^+$ T cells. For PD-L1 neutralization, anti-PD-L1 (BE0285, Bio X cell) and the control IgG (BE0086, Bio X cell) were used in the co-culture experiments.

Chemotaxis

PBNs, ALNs or TANs (5×10^5) were suspended in RPMI1640 medium and placed in the bottom chamber of a 12-well plate. Purified CD14$^+$ monocytes (2×10^5) were placed in the top insert of a Transwell (5 μm, Corning) and incubated in macrophage differentiation medium with 100 ng ml^{-1} M-CSF (STEMCELL, 78059). After co-culture of PBNs with monocytes for 48 h, monocytes that migrated and attached to the low surface of the Transwell membrane were fixed with 4% paraformaldyhyde, and stained with 1% crystal violet. The number of migrated monocytes was calculated using Image J (v.1.52k).

RNA isolation and qPCR

Total RNA was isolated using the Trizol RNA Isolation kit (Invitrogen). The Reverse Transcription Reagents kit (TIANGEN) was used for cDNA synthesis from total RNA. qPCR was performed in triplicates using the AriaMx Real-Time PCR System (G8830A). Gene expression of chemokines (*CCL2*, *CCL3*, *CCL4* and *CCL5*) and *CD274* were quantified by the comparative C_t method ($2^{-\Delta\Delta C_t}$) with *GAPDH* as an internal control. The fold change of each gene was calculated at different time points versus 0 h. A list of the primers used for the queried genes is provided in Supplementary Table 5.

Bulk RNA-seq

RNA-seq libraries were constructed using the NEBNext Ultra RNA Library Prep Kit (New England Biolabs) according to the manufacturer's protocol. The library was quality-checked and sequenced using the NovaSeq 6000 sequencer (Illumina). The quality of sequencing reads was evaluated using FastQC. Adaptor sequences and low-quality score bases were trimmed using trimmomatic (v.0.36). These reads were then mapped to human genome reference GRCh38 from Ensembl release 98 using STAR (v.2.5.2b). The fragments per kilobase of exon per million mapped reads (FPKM) values and gene count values were computed using RSEM (v.1.3.1) and DEGs were analysed using the DESeq2 (v.1.24) R package.

ATAC-seq

Fresh neutrophils (1×10^4–5×10^4 cells) isolated from different tissues of patients with liver cancer were immediately sent for bulk ATAC-seq using the TruePrep DNA Library Prep Kit V2 for Illumina (Vazyme, TD501). Raw sequencing reads were trimmed using trimmomatic (v.0.39) and then mapped to the GRCh38 human genome using Bowtie2 (v.2.4.4). PCR duplicates were removed using MarkDuplicates from PicardTools (v.2.23.3). Peaks were called with MACS3 (v.3.0.0a7) and peaks that were found in at least two biological replicates were retained and merged for further analysis. Significantly differentially accessible peaks were identified with adjusted $P < 0.05$, and fold change > 1.5 by DESeq2 (v.1.24). Normalized BigWig files were generated by DeepTools (v.3.5.1) and merged for visualization by pyGenomeTracks (v.3.6).

In vivo neutrophil depletion

The anti-Ly6G antibody (1A8, Bio X Cell) or IgG2a Isotype control (2A3, Bio X Cell) at a dose of 12.5 μg per 100 μl PBS was administered daily through intraperitoneal injection, starting 7 days before HDTV injection of the pTMC plasmid. After 33 days, mice were euthanized by carbon dioxide asphyxiation and the liver tumours were carefully separated from mice. The number of liver tumour nodules was quantified and the ratio of liver weight to body weight was calculated.

To deplete the neutrophils in a therapeutic manner, the pTMC-Luc mouse model was used. Mice were given fresh prepared D-luciferin (150 μg per g) intraperitoneally and incubated for 5 min and imaged using in vivo imaging system. In vivo luciferase bioluminescence signal was detected for an exposure time of 60 s using the Living Image software. At day 7 after the pTMC-Luc HDTV injection, the tumour can be visually detected by the bioluminescence signal, then the anti-Ly6G antibody (1A8, Bio X Cell, BE0075-1) or IgG2a (2A3, Bio X Cell, BE0089) isotype control was injected into mice at a dose of 25 μg per 100 μl PBS on a daily basis. At day 36, the bioluminescence signal was detected, and mice were euthanized. The ratio of liver weight to body weight was calculated. FACS analysis was performed using the following antibodies, CD45 (BioLegend, 103116), CD3ε (BioLegend, 100353), CD8a (BD, 552877), CD11b (BioLegend, 101242), Ly6G (surface, BioLegend, 127618; intracellular, BD, 551461), F4/80 (BioLegend, 123133) to gate CD8$^+$ T cells, neutrophils and macrophages, respectively. Data were analysed using FlowJo (v.10.4) and the gating strategies are shown in Supplementary Fig. 5. PD-L1 antibody (BD, 558091) was used to assess the immunosuppression of neutrophils. PD-1 (CD279, BD, 562523) and TIM3 (BD, 566346) were used to assess the exhaustion of CD8$^+$ T cells. The depletion efficiency of neutrophils was detected by both surface and intracellular Ly6G staining. In brief, the cell suspension was first stained with Ly6G-PE-Cy7 antibodies to cover the surface Ly6G protein. Cells were then fixed and permeabilized and intracellular proteins were stained with Ly6G-PE antibodies.

Statistical analysis

Statistical analyses were performed using GraphPad Prism (v.9.0) (for experimental data), and R (v.3.6.1), RStudio (v.3.5.3) and Python (v.3.7.4) (for sequencing data and matched clinical variables). Comparisons between groups were conducted using χ^2 tests or Fisher's exact test for categorical variables. Student's *t*-tests, Wilcoxon rank-sum tests and ANOVA were used for continuous variables. Paired *t*-tests were used for paired comparisons. Survival analyses were conducted using log-rank tests. $P < 0.05$ was considered to be statistically significant. No statistical methods were used to predetermine the sample size of scRNA-seq libraries. Unless otherwise noted, each experiment was repeated three or more times with biologically independent samples.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Raw sequencing data reported in this paper have been deposited at the Genome Sequence Archive at the National Genomics Data Center (Beijing, China) under the BioProject ID PRJCA007744. The data deposited and made public are compliant with the regulations of the Ministry of Science and Technology of China. To facilitate the use of our data by the wider research community, we developed an interactive web-based tool (http://meta-cancer.cn:3838/scPLC) for analysing and visualizing our single-cell data. Other public data used in this study include reference genomes for human (https://asia.ensembl.org/, GRCh38.p13) and mouse (https://asia.ensembl.org/, GRCm39) and TCGA datasets (https://portal.gdc.cancer.gov/). Source data are provided with this paper.

Code availability

Codes used in this study are available at GitHub (https://github.com/meta-cancer/scPLC).

42. Seehawer, M. et al. Necroptosis microenvironment directs lineage commitment in liver cancer. *Nature* **562**, 69–75 (2018).
43. Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.* **33**, 495–502 (2015).
44. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018).
45. Puram, S. V. et al. Single-cell transcriptomic analysis of primary and metastatic tumor ecosystems in head and neck cancer. *Cell* **171**, 1611–1624 (2017).
46. Muhl, L. et al. Single-cell analysis uncovers fibroblast heterogeneity and criteria for fibroblast and mural cell identification and discrimination. *Nat. Commun.* **11**, 3953 (2020).
47. Aran, D. et al. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat. Immunol.* **20**, 163–172 (2019).
48. Vento-Tormo, R. et al. Single-cell reconstruction of the early maternal-fetal interface in humans. *Nature* **563**, 347–353 (2018).
49. Gulati, G. S. et al. Single-cell transcriptional diversity is a hallmark of developmental potential. *Science* **367**, 405–411 (2020).
50. Qiu, X. et al. Single-cell mRNA quantification and differential analysis with Census. *Nat. Methods* **14**, 309–315 (2017).
51. Lange, M. et al. CellRank for directed single-cell fate mapping. *Nat. Methods* **19**, 159–170 (2022).
52. Aibar, S. et al. SCENIC: single-cell regulatory network inference and clustering. *Nat. Methods* **14**, 1083–1086 (2017).
53. Welch, J. D. et al. Single-cell multi-omic integration compares and contrasts features of brain cell identity. *Cell* **177**, 1873–1887 (2019).
54. Schürch, C. M. et al. Coordinated cellular neighborhoods orchestrate antitumoral immunity at the colorectal cancer invasive front. *Cell* **182**, 1341–1359 (2020).
55. Fan, Y. et al. Targeting multiple cell death pathways extends the shelf life and preserves the function of human and mouse neutrophils for transfusion. *Sci. Transl. Med.* **13**, eabb1069 (2021).

Acknowledgements We thank Y. Guo, C. Shan and J. Ren from National Center for Protein Sciences at Peking University for FACS and CODEX assistance. This work is jointly supported by National Natural Science Foundation of China (81988101, 82173035, 82030079, 81972656, 81802813, 81902401, 81972735 and 81872508), the National Science and Technology Major Project of China (2018ZX10723204), Beijing Natural Science Foundation (7212108), Changping Laboratory, the Michigan Medicine and PKU-HSC JI for Translational and Clinical Research (BMU2020JI005) and Sino-Russian Math Center in PKU.

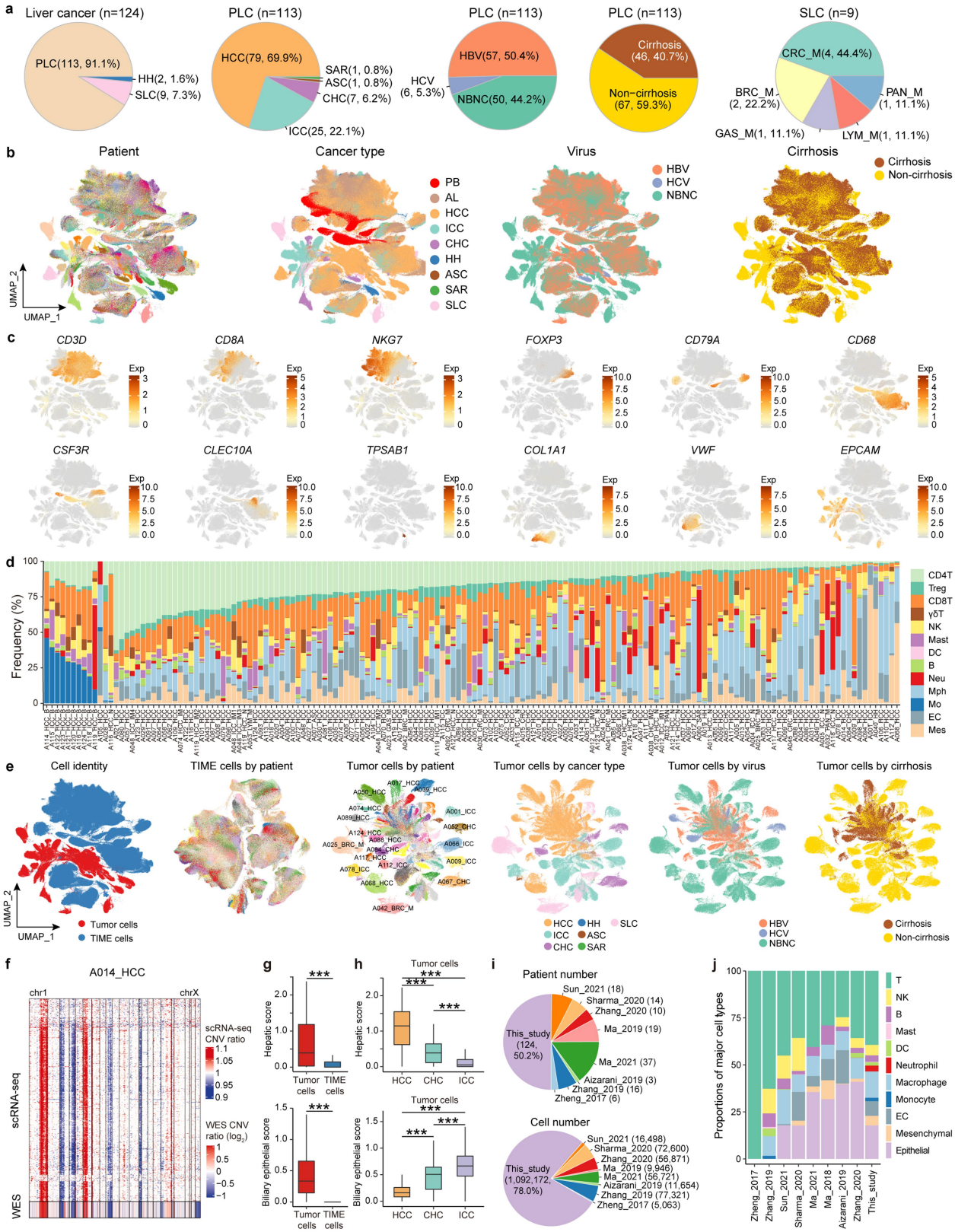
Author contributions R.X., Z.Z. and N.Z. conceived and designed the project. R.X., X.X., Z.L. and J.Z. collected the human samples and clinical information. X.X., Z.L. and J.Z. performed pathological examination. R.X. and X.X. performed the scRNA-seq experiments. Q.C., Q. Zhang and R.X. performed bioinformatic analyses. Q. Zhang, R.X. and Q.C. performed IHC, mIHC and CODEX experiments. R.K., R.X., M.F. and F.W. performed functional experiments of neutrophils. R.K. and R.X. constructed the mouse models and analysed the in vivo data. R.X., Q. Zhang, Q.C., R.K., X.X., H.L., Q. Zhan, M.D., J.Z., Z.Z. and N.Z. discussed and interpreted the data. Q.C., R.X. and J.C. built the online website. R.X., Q. Zhang, Q.C. and R.K. wrote the manuscript with help from Z.Z. and N.Z.; Z.Z., J.Z. and N.Z. supervised the project.

Competing interests Z.Z. is a founder of Analytical BioSciences and is a consultant for InnoCare Pharma and ArsenalBio. N.Z. is the CSO of Yunnan Baiyao Group. The other authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-022-05400-x.

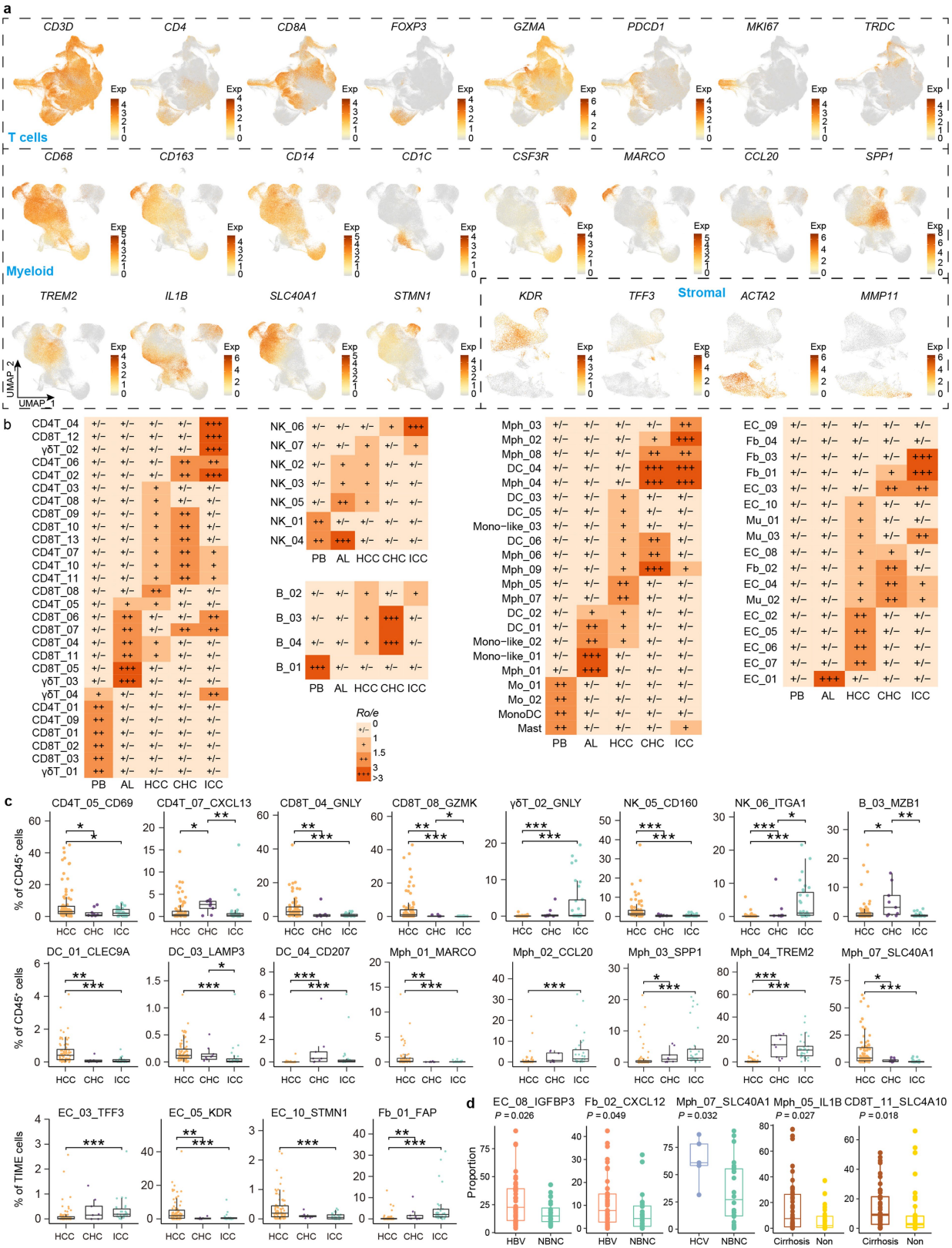
Correspondence and requests for materials should be addressed to Jiye Zhu, Zemin Zhang or Ning Zhang.

Peer review information *Nature* thanks Andres Hidalgo, Alexander Swarbrick and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

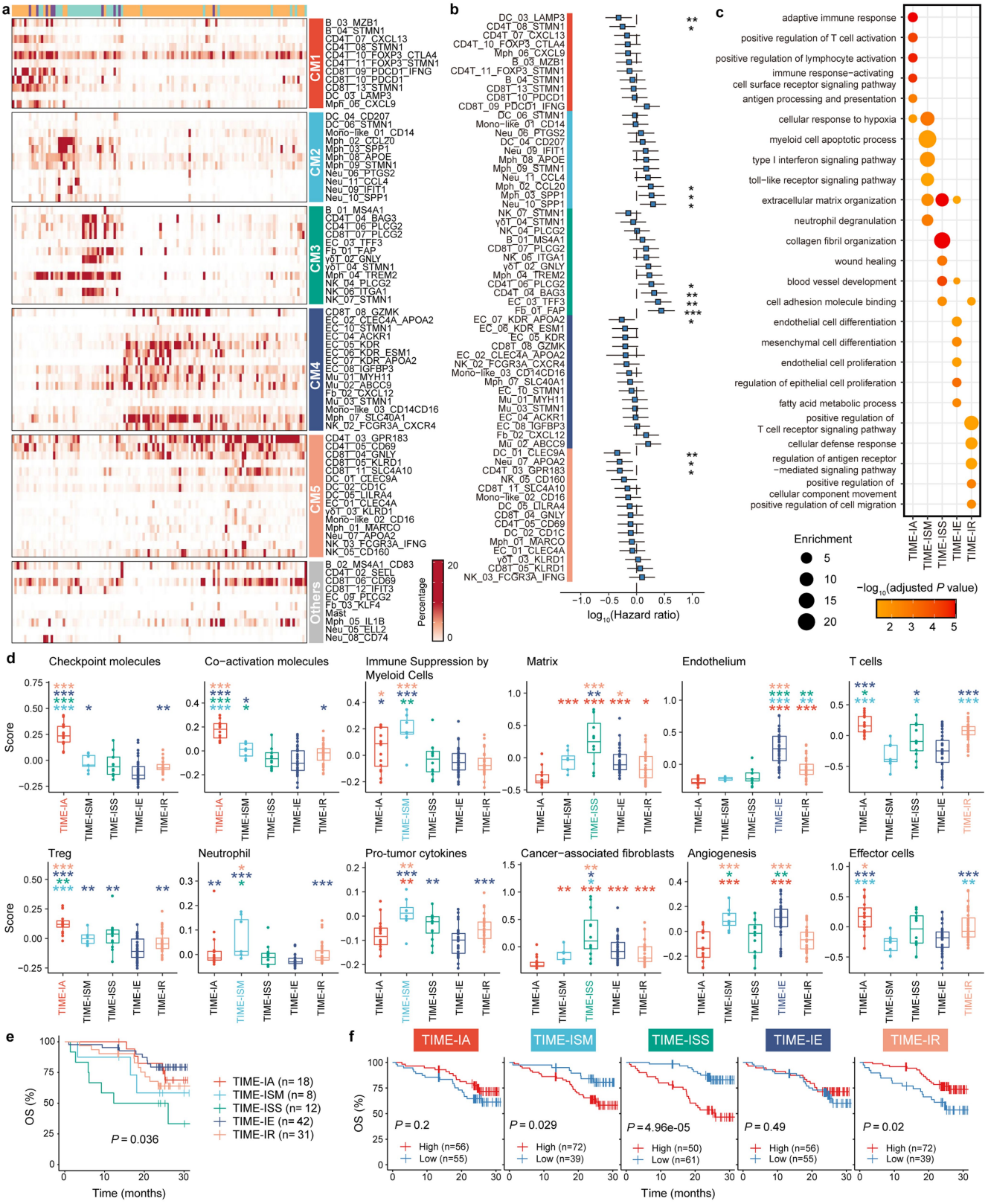


Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | Patient cohort and cluster information. a, Pie charts showing the composition of cancer types in our cohort. HCC, hepatocellular carcinoma; ICC, intrahepatic cholangiocarcinoma; CHC, combined hepatocellular and cholangiocarcinoma; HH, hepatic hemangioma; ASC, adenosquamous carcinoma; SAR, sarcomatoid carcinoma; SLC, secondary liver cancer. CRC_M, liver metastasis from colorectal cancer, PAN_M, liver metastasis from pancreatic cancer, LYM_M, liver metastasis from lymphoma, GAS_M, liver metastasis from gastric cancer, BRC_M, liver metastasis from breast cancer. **b**, UMAP plots showing the distribution of patients, cancer types, viruses and liver cirrhosis states. Dots represent individual cells. PB, peripheral blood; AL, adjacent liver; HBV, hepatitis B virus, HCV, hepatitis C virus, NBNC, double negative of HBV and HCV. **c**, UMAP plots showing expression of canonical marker genes of major cell populations including T cells (*CD3D, CD8A, FOXP3*), NK cells (*NKG7*), B cells (*CD79A*), macrophages (*CD68*), neutrophils (*CSF3R*), dendritic cells (*CLEC10A*), mast cells (*TPSAB1*), fibroblasts (*COL1A1*), endothelial cells (*VWF*), and epithelial cells (*EPCAM*). **d**, Stacked barplot showing the distribution of major cell types in each sample. **e**, UMAP plots showing the distribution of cell identities for tumour cells and TIME cells. Tumour cells were further coloured by patient, cancer type, virus, and cirrhosis. **f**, CNV profiles inferred from scRNA-seq data for each cell and from matched bulk exome data in the sample A014_HCC. **g**, Boxplots showing hepatic scores and biliary epithelial scores in tumour (n = 193,877 cells) and TIME cells (n = 898,295 cells). Cells are from 124 patients. **h**, Boxplots showing hepatic scores and biliary epithelial scores in tumour cells of different PLC subtypes (HCC, n = 96,211 cells from 79 cases, ICC, n = 52,345 cells from 25 cases, CHC, n = 15,493 cells from 7 cases). Cells are from 111 patients. **i**, Pie charts showing the patient number (top) and cell number (bottom) of our study and published single cell studies for PLC. Colours represent different studies. **j**, Stacked barplot showing proportions of major cell populations among different studies. Colours represent major cell populations. In **g**-**h**, n denotes individual cells. Two-sided Wilcoxon rank-sum test is used. For boxplots, centre line shows median, box limits indicate upper and lower quartiles, and whiskers extend 1.5 times the interquartile range, while data beyond the end of the whiskers are outlying points that are plotted individually. ***, $P < 0.001$.

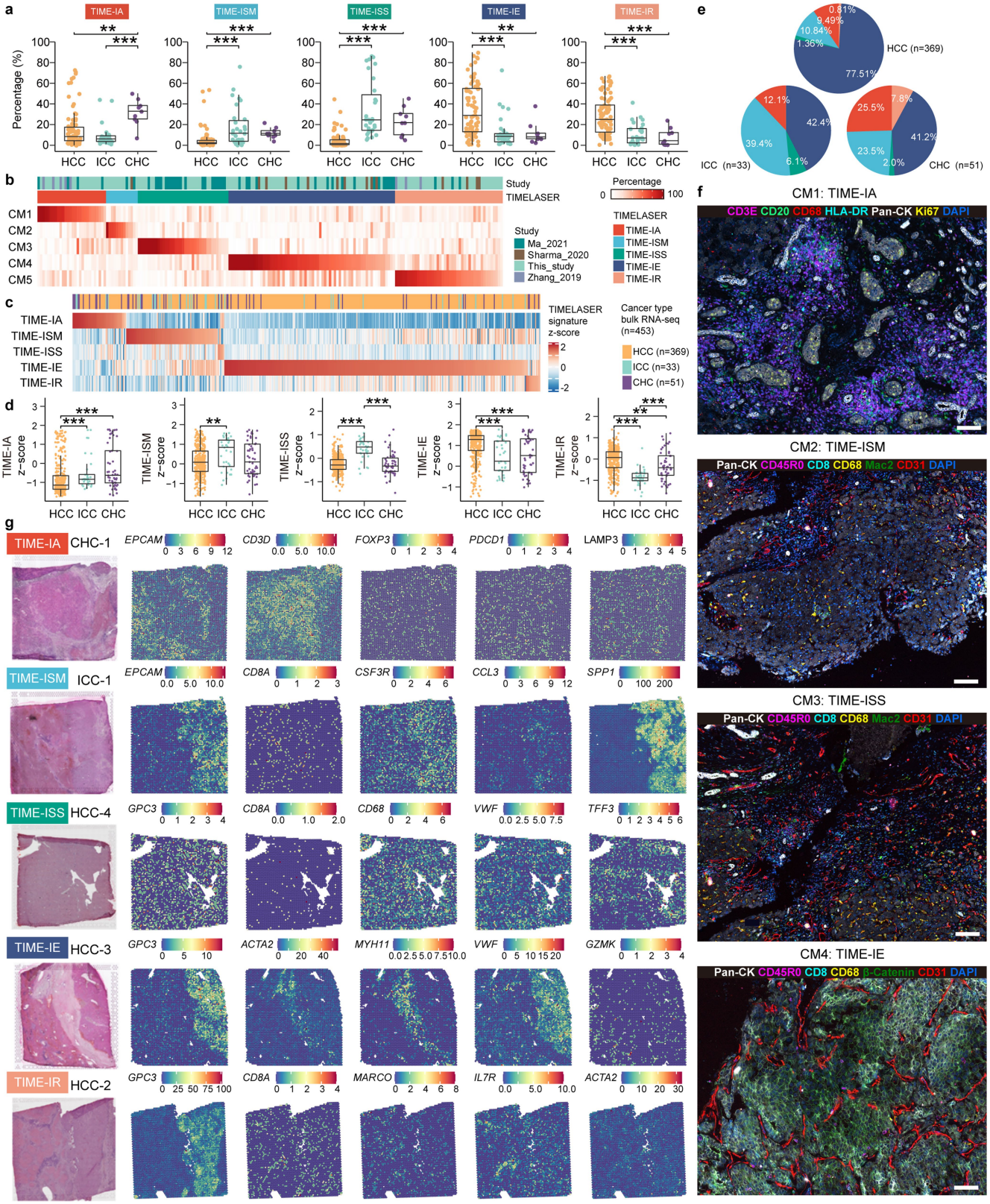


Extended Data Fig. 2 | Gene expression and tissue preference of 89 TIME cell clusters. a, UMAP plots showing the expression of canonical marker genes for clusters in each major cell population. Normalized expression level was abbreviated as Exp. **b**, Heatmap showing tissue preferences of clusters in each major cell population revealed by $R_{o/e}$. **c**, Boxplots showing proportions of several tumour-enriched TIME clusters divided by PLC subtypes. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$. (HCC, $n = 79$ cases, ICC, $n = 25$ cases, CHC, $n = 7$ cases). **d**, Boxplots showing proportions of several cell clusters associated with virus or cirrhosis. (HBV, $n = 57$ cases, HCV, $n = 6$ cases, NBNC, $n = 50$ cases; cirrhosis, $n = 46$ cases, non-cirrhosis, $n = 67$ cases). In **c**-**d**, n denotes biologically independent samples. Two-sided Wilcoxon rank-sum test is used. For boxplots, centre line shows median, box limits indicate upper and lower quartiles, and whiskers extend 1.5 times the interquartile range, while data beyond the end of the whiskers are outlying points that are plotted individually.



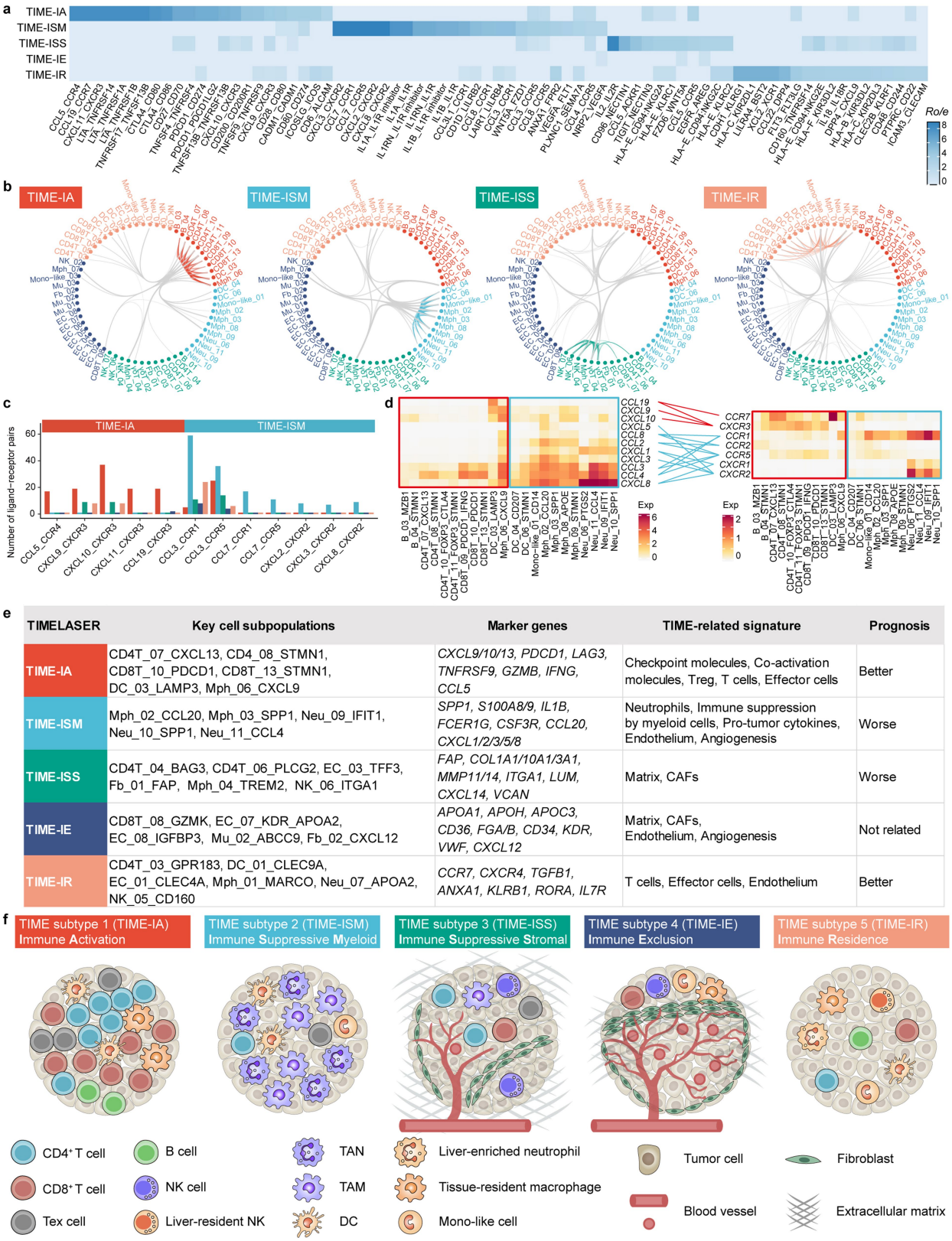
Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Clusters, signatures, and prognosis of five TIMELASER subtypes. a, Heatmap showing frequencies of TIME cell clusters in 5 CMs. **b**, Forest plot showing the clinical relevance of clusters in each CM revealed by $\log_{10}$(hazard ratio) based on PFS. Cox regression. Log-rank test. **c**, Dot heatmap showing enriched pathways across TIMELASER subtypes. Benjamini–Hochberg-adjusted hypergeometric test. **d**, Boxplots showing the expression of given signatures in different TIMELASER subtypes. Signature scores of TIMELASER subtypes with overhead asterisk are significantly higher than that of subtypes with corresponding asterisk colour. Wilcoxon rank-sum test, two-sided. (TIME-IA, $n = 18$ cases, TIME-ISM, $n = 8$ cases, TIME-ISS, $n = 12$ cases, TIME-IE, $n = 42$ cases, TIME-IR, $n = 31$ cases, n denotes biologically independent patients). For boxplots, centre line shows median, box limits indicate upper and lower quartiles, and whiskers extend 1.5 times the interquartile range, while data beyond the end of the whiskers are outlying points that are plotted individually. **e**, Overall survival (OS) with each patient assigned to a single CM. Log-rank test. **f**, OS of cases stratified by each TIMELASER module. Log-rank test. In **b** and **d**, *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.



Extended Data Fig. 4 | See next page for caption.

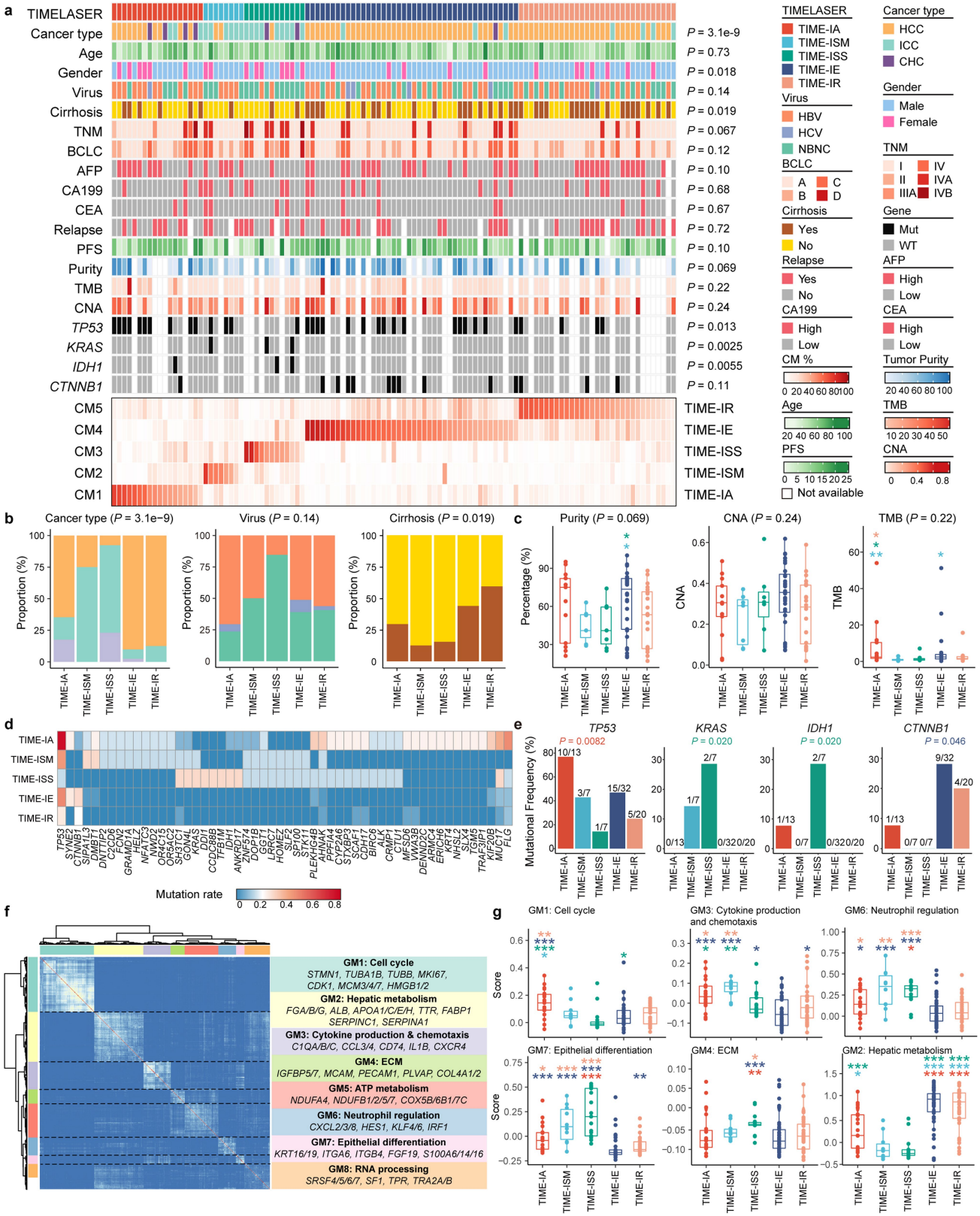
Extended Data Fig. 4 | Validation of five TIMELASER subtypes. a, Boxplots showing the percentage of TIMELASER modules across 3 PLC subtypes. (HCC, $n = 79$ cases, ICC, $n = 25$ cases, CHC, $n = 7$ cases). **b**, Heatmap showing the percentage of CM1–5 across tumours in our cohort and three published scRNA-seq cohorts. **c**, Expression of signature genes for the five TIMELASER subtypes in 453 published liver cancer bulk RNA-seq data. **d**, Boxplot showing z-scores of signature genes for five TIMELASER subtypes in different cancer types. Colours represents HCC (orange, $n = 369$ cases), ICC (green, $n = 33$ cases) and CHC (purple, $n = 51$ cases). **e**, Pie charts showing the proportion of TIMELASER subtypes in **c**. **f**, Representative CODEX results showing four different TIMELASER subtypes. For each sample, only six representative antibodies staining are displayed in the figure along with DAPI. Scale bar, 500 μm. **g**, Validation of TIMELASER by a published spatial transcriptomic study of liver cancer. H&E staining and the corresponding spatial feature plots of different marker genes of cell types are shown in different samples. In **a** and **d**, n denotes biologically independent samples. Two-sided Wilcoxon rank-sum test is used. For boxplots, centre line shows median, box limits indicate upper and lower quartiles, and whiskers extend 1.5 times the interquartile range, while data beyond the end of the whiskers are outlying points that are plotted individually. **, $P < 0.01$; ***, $P < 0.001$.



e

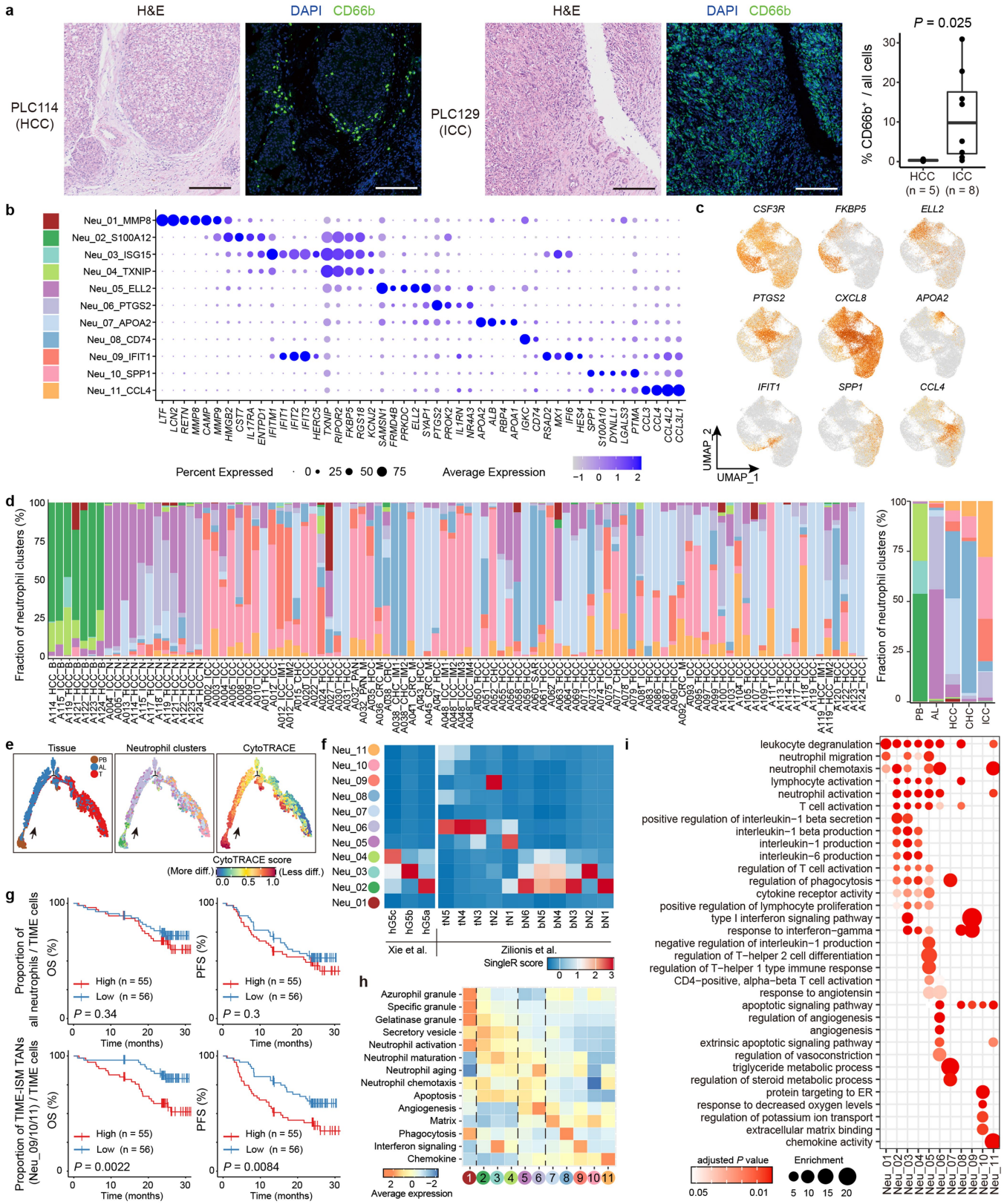
TIMELASER	Key cell subpopulations	Marker genes	TIME-related signature	Prognosis
TIME-IA	CD4T_07_CXCL13, CD4_08_STMN1, CD8T_10_PDCD1, CD8T_13_STMN1, DC_03_LAMP3, Mph_06_CXCL9	*CXCL9/10/13, PDCD1, LAG3, TNFRSF9, GZMB, IFNG, CCL5*	Checkpoint molecules, Co-activation molecules, Treg, T cells, Effector cells	Better
TIME-ISM	Mph_02_CCL20, Mph_03_SPP1, Neu_09_IFIT1, Neu_10_SPP1, Neu_11_CCL4	*SPP1, S100A8/9, IL1B, FCER1G, CSF3R, CCL20, CXCL1/2/3/5/8*	Neutrophils, Immune suppression by myeloid cells, Pro-tumor cytokines, Endothelium, Angiogenesis	Worse
TIME-ISS	CD4T_04_BAG3, CD4T_06_PLCG2, EC_03_TFF3, Fb_01_FAP, Mph_04_TREM2, NK_06_ITGA1	*FAP, COL1A1/10A1/3A1, MMP11/14, ITGA1, LUM, CXCL14, VCAN*	Matrix, CAFs	Worse
TIME-IE	CD8T_08_GZMK, EC_07_KDR_APOA2, EC_08_IGFBP3, Mu_02_ABCC9, Fb_02_CXCL12	*APOA1, APOH, APOC3, CD36, FGA/B, CD34, KDR, VWF, CXCL12*	Matrix, CAFs, Endothelium, Angiogenesis	Not related
TIME-IR	CD4T_03_GPR183, DC_01_CLEC9A, EC_01_CLEC4A, Mph_01_MARCO, Neu_07_APOA2, NK_05_CD160	*CCR7, CXCR4, TGFB1, ANXA1, KLRB1, RORA, IL7R*	T cells, Effector cells, Endothelium	Better

Extended Data Fig. 5 | L-R networks and feature summary of five TIMELASER subtypes. a, Heatmap showing $R_{o/e}$ enrichment values of TIMELASER-specific L-R pairs. **b**, Chord diagrams showing the interactions within each TIMELASER subtype mediated by specific L-R pairs. Line width is proportional to interaction intensity and coloured by TIMELASER subtypes. **c**, Barplots showing the number of ligand-receptor pairs significantly enriched in TIME-IA and TIME-ISM modules. **d**, Heatmap showing the expression of chemokines and the corresponding receptors in TIME-IA and TIME-ISM patients. Exp, normalized mean expression. **e**, Summary of key features across TIMELASER subtypes. **f**, Schematic for five TIMELASER subtypes. Selected cell populations are shown for each TIMELASER subtype with tumour cells as background. Tex, exhausted T cell; NK, nature killer cell; TAN, tumour-associated neutrophil; TAM, tumour-associated macrophage; DC, dendritic cell.



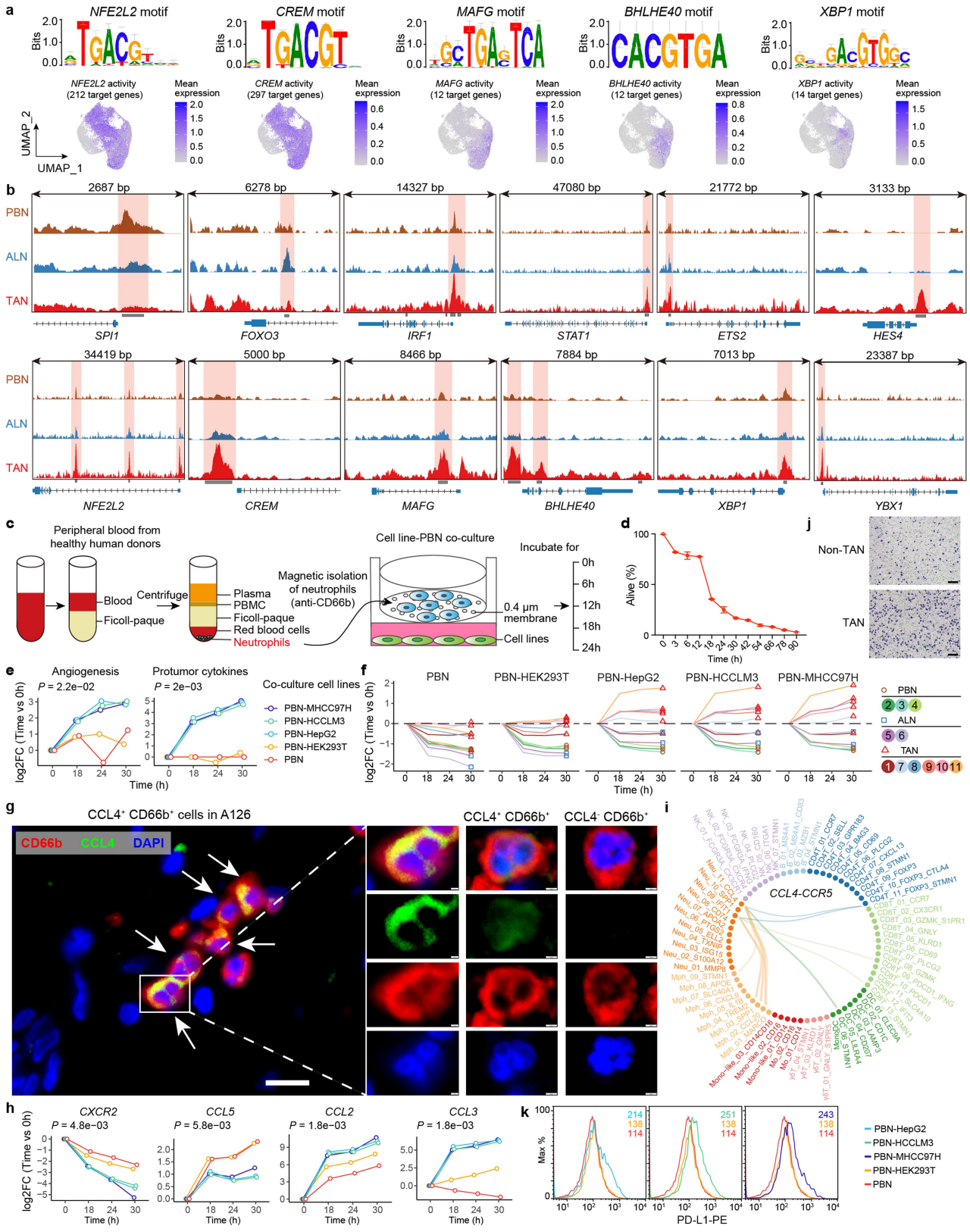
Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | Mutational landscape and GMs of malignant cells. **a**, Heatmap showing frequencies of five TIMELASER subtypes across 111 PLC patient samples. Detailed clinical and molecular attributes of individual tumour samples are annotated. P values to the right indicate significant non-random distributions for each attribute. Chi-square test is used for categorical variables. Two-way ANOVA test is used for continuous variables. **b**, Stacked barplots showing the distribution of cancer types, virus and cirrhosis state across TIMELASER subtypes. Chi-square test. **c**, Boxplots showing the distribution of tumour purity, CNA and TMB inferred by WES data across TIMELASER subtypes. Two-way ANOVA test is used for comparison of multiple groups. Two-sided Wilcoxon rank-sum test is used for comparison between any two groups. **d**, Heatmap showing the mutational rate of somatic mutations enriched in different TIMELASER subtypes. **e**, Barplots showing mutational frequencies of *TP53*, *KRAS*, *IDH1*, and *CTNNB1* in different TIMELASER subtypes. Colours represent different TIMELASER subtypes. One-sided Fisher's exact test. Tests are performed between the denoted TIME subtype (*P* value colour coded) and a combination of all others. **f**, Heatmaps showing the eight common gene modules (GMs) extracted from tumour cells. **g**, Boxplots showing the distributions of signature scores of GMs across tumours stratified into five TIMELASER subtypes. Overhead asterisk is significantly higher than that of subtypes with corresponding asterisk colour. Wilcoxon rank-sum test, two-sided. In **c** and **e**, (TIME-IA, $n = 13$ cases, TIME-ISM, $n = 7$ cases, TIME-ISS, $n = 7$ cases, TIME-IE, $n = 32$ cases, TIME-IR, $n = 20$ cases). In **g**, (TIME-IA, $n = 18$ cases, TIME-ISM, $n = 8$ cases, TIME-ISS, $n = 12$ cases, TIME-IE, $n = 42$ cases, TIME-IR, $n = 31$ cases). In **c, e**, and **g**, *n* denotes biologically independent patients. For boxplots, centre line shows median, box limits indicate upper and lower quartiles, and whiskers extend 1.5 times the interquartile range, while data beyond the end of the whiskers are outlying points that are plotted individually. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.



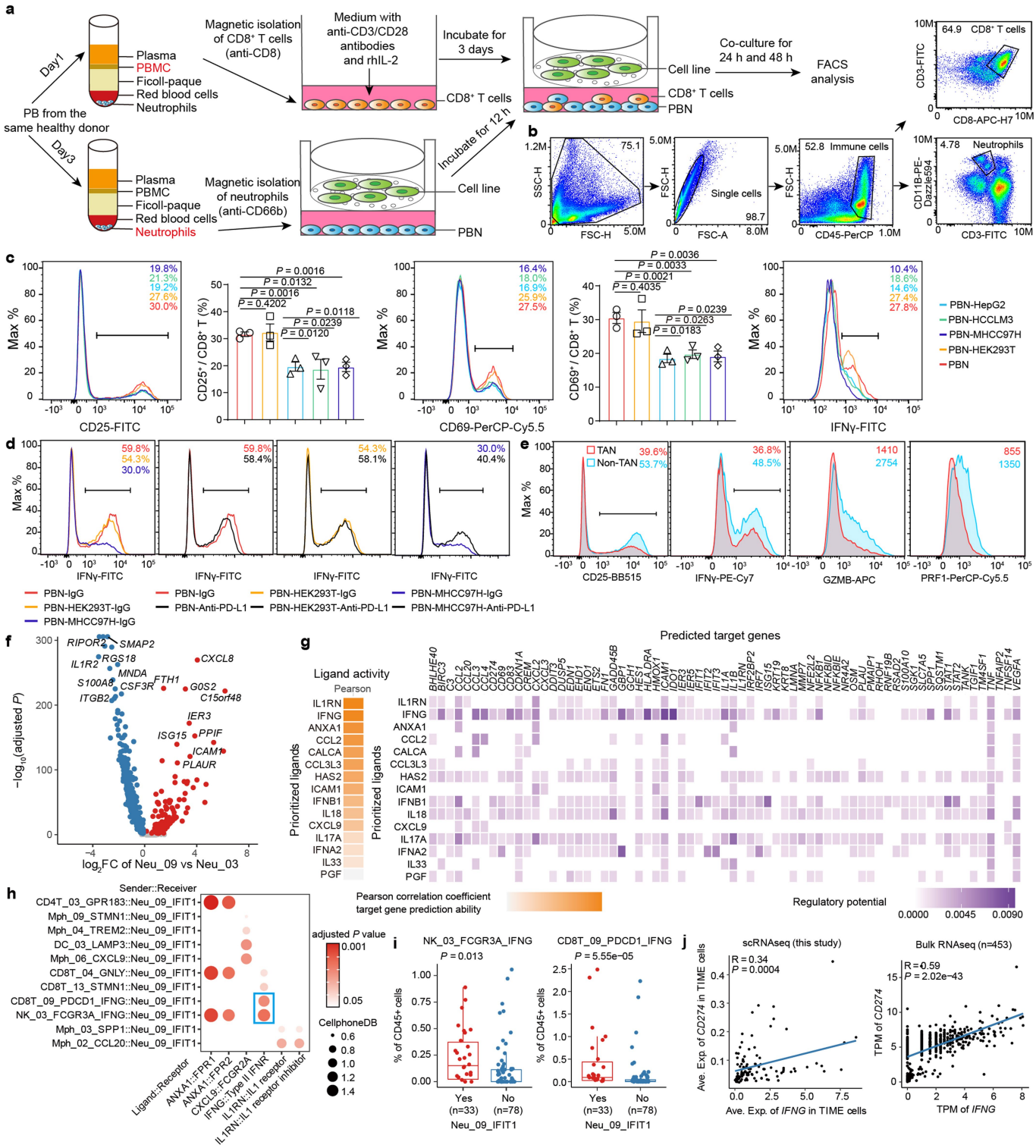
Extended Data Fig. 7 | See next page for caption.

Extended Data Fig. 7 | Neutrophil heterogeneity in human PLC. a, H&E and IHC plots showing the neutrophil frequencies in HCC and ICC patients. Scale bar, 20 μm. Boxplot to the right is the quantitative result. Student's t-test, two sided. (HCC, $n = 5$ cases, ICC, $n = 8$ cases, n denotes biologically independent samples.) In the boxplot, centre line shows median, box limits indicate upper and lower quartiles, and whiskers extend 1.5 times the interquartile range, while data beyond the end of the whiskers are outlying points that are plotted individually. **b**, Dot heatmap showing the row-scaled expression of marker genes for neutrophil clusters. **c**, UMAP plots showing the expression of typical marker genes for neutrophil subsets. Exp, normalized expression. **d**, Distribution of neutrophil clusters by patient. **e**, Monocle trajectories of neutrophils coloured by tissues (left), cluster identities (middle) and CytoTRACE scores (right). Each dot represents a single cell. Cell orders are inferred based on the expression of the most variable genes across neutrophil clusters. **f**, Heatmap showing similarity scores of peripheral blood neutrophil clusters from *Xie et al.* and lung cancer neutrophil clusters from *Zillionis et al.* compared with liver cancer neutrophil clusters inferred by singleR. **g**, OS and PFS of patients stratified by the proportion of all neutrophils and neutrophil clusters (Neu_09/10/11) in TIME-ISM. Log-rank test. **h**, Average expression of classic neutrophil scores and TAN-specific gene scores in neutrophil clusters. **i**, Gene ontology analysis showing the enrichment of specific pathways in neutrophil clusters. Benjamini–Hochberg-adjusted hypergeometric test.



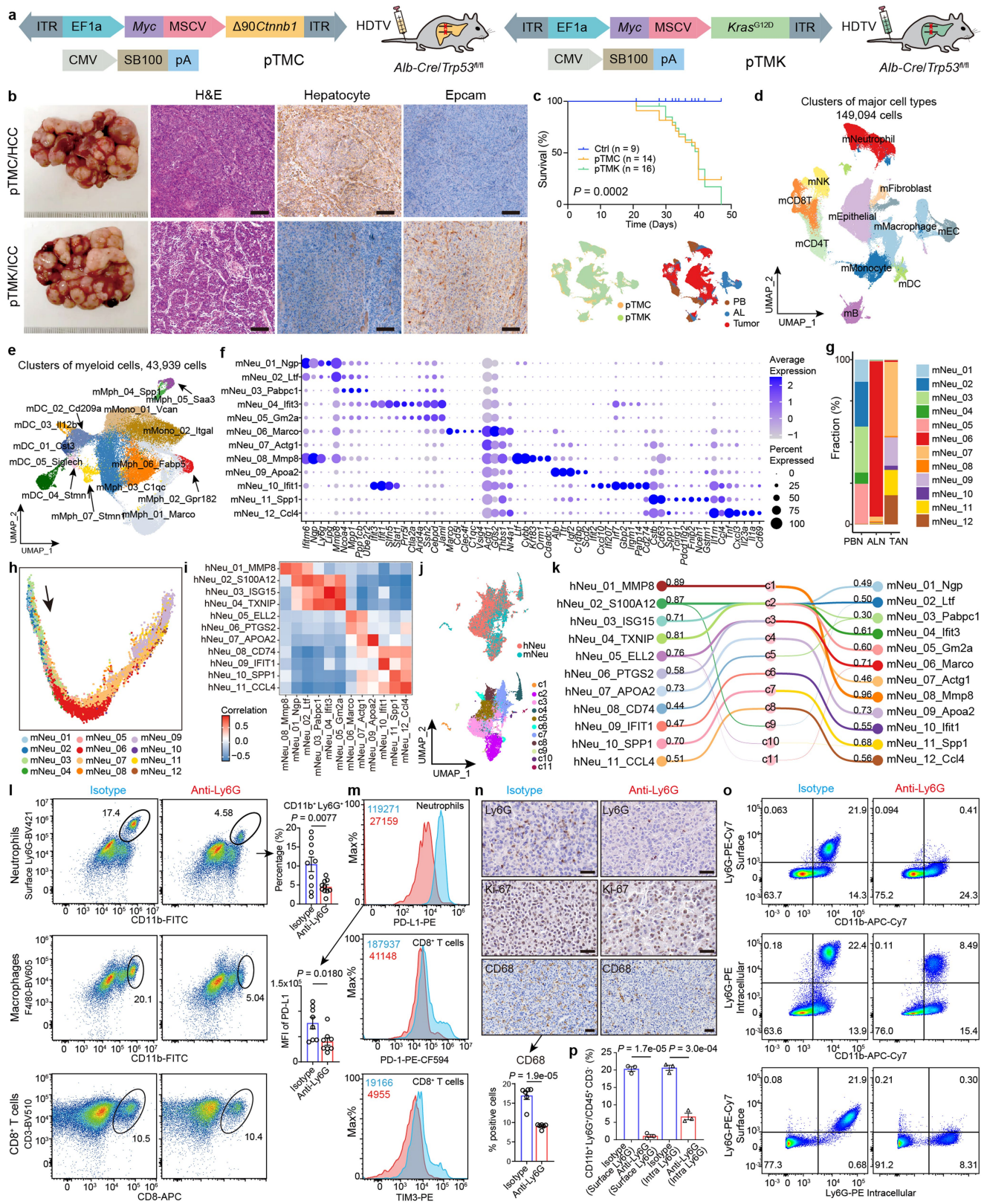
Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | Transcription factors, mIHC and in vitro validation of neutrophil clusters. a, UMAP plots showing regulon activities of five representative transcription factors for specific neutrophil clusters. Binding motifs of these transcription factors are shown on the top. **b**, Normalized ATAC-seq sequencing tracks of selected transcription factor loci in matched PBN, ALN, and TAN isolated from the same patient. ATAC peaks detected by MACS3 are denoted with the grey box above the gene body and highlighted with light red shading. **c**, Workflow of co-culture experiments of PBNs with or without cell line (liver cancer cell line HepG2, HCCLM3, and MHCC97H, control cell line HEK293T). **d**, Survival curve of PBN in culture condition ($n = 3$, n denotes biologically independent samples). Data are presented as mean values ± SEM. **e**, Expression of TAN-related signatures in PBNs co-cultured with or without different cell lines for 0 h, 18 h, 24 h, and 30 h. **f**, Expression of gene signatures of different neutrophil subsets in PBNs co-cultured with or without different cell lines. **g**, White arrows mark $CCL4^+CD66b^+$ neutrophils, with one cell highlighted by the four enlarged panels on the right. Middle panels show another representative $CCL4^+CD66b^+$ neutrophil while right panels show a represen tative CCL4−$CD66b^+$ neutrophil. Scale bars are 20 μm and 2 μm. **h**, Expression of selected genes in PBNs co-cultured with or without different cell lines for 0 h, 18 h, 24 h, and 30 h. **i**, Chord diagrams showing interactions between neutrophils and other cell types mediated by *CCL3-CCR1* and *CCL4-CCR5*. Line width is proportional to interaction intensity, coloured by cell types with receptors. **j**, Crystal violet staining of migrated monocytes co-cultured with matched TAN or non-TAN. Scale bar, 100 μm. **k**, FACS analysis showing the PD-L1 expression of PBNs co-cultured with or without different cell lines for 24 h. Two-way ANOVA test is used for **e** and **h**.



Extended Data Fig. 9 | See next page for caption.

Extended Data Fig. 9 | Co-culture experiment of cell line-PBN-CD8$^+$ T cells and analyses of two *IFIT*$^+$ neutrophil subsets. a, Experimental workflow. **b**, Gating strategy separating neutrophils from CD8$^+$ T cells in the bottom chamber of co-culture system in **a**. **c**, FACS analysis showing the expression of CD25 ($n = 3$), CD69 ($n = 3$), and IFNγ ($n = 4$) in PBNs co-cultured with different cell lines. Student's t-test, one-sided. Data are presented as mean values ± SEM. **d**, FACS analysis showing the expression of IFNγ in CD8$^+$ T cells when anti-PD-L1 or the IgG control is added to the co-culture system. **e**, FACS analysis showing the expression of CD25, IFNγ, GZMB, and PRF1 in CD8$^+$ T cells co-cultured with matched TAN or non-TAN isolated from patients with liver cancer. **f**, Volcano plot showing differentially expressed genes between Neu_03_ISG15 and Neu_09_IFIT1. Benjamini-Hochberg adjusted Wilcoxon rank-sum test, two-sided. **g**, Heatmap showing the predicted ligand activity by NicheNet on genes highly expressed in Neu_09_IFIT1. Pearson correlation indicates the ability of each ligand to predict the target genes, and better predictive ligands are thus ranked higher. **h**, Dot heatmap showing the selected ligand-receptor pairs between different cell populations and Neu_09_IFIT1. Benjamini-Hochberg adjusted permutation test. **i**, Boxplots showing the proportion of two *IFNG*$^+$ populations between patients with or without Neu_09_IFIT1. Wilcoxon rank-sum test, two-sided. (Yes, $n = 33$ cases, No, $n = 78$ cases). For boxplots, centre line shows median, box limits indicate upper and lower quartiles, and whiskers extend 1.5 times the interquartile range, while data beyond the end of the whiskers are outlying points that are plotted individually. **j**, Pearson correlation between the expression of *CD274* and *IFNG* in TIME cells in this study (left) or in the collected bulk RNA-seq datasets (right). In **c** and **i**, *n* denotes biologically independent samples.



Extended Data Fig. 10 | See next page for caption.

Extended Data Fig. 10 | scRNA-seq and functional analyses of mouse models. **a**, Schematic of liver cancer mouse models. Intrahepatic delivery of the transposable vectors pTMC (encoding *Myc* and *Δ90Ctnnb1*) or pTMK (encoding *Myc* and $Kras^{G12D}$) via HDTV in *Alb-Cre* × $Trp53^{fl/fl}$ mice. **b**, Representative photos, H&E, and IHC staining of HCC and ICC mouse models. Rulers in the photo show a minimum unit of mm. Scale bar on the staining slides is 20 μm. **c**, Survival curve of liver cancer mouse model. Log-rank test. **d**, UMAP plot showing major cell types of mice with liver cancer. Dots represent individual cells, and colours represent the major cell populations. mILC: innate lymphoid cells, mNeu: neutrophils, mMph: macrophages, mMono: monocytes, mEC: endothelial cells; mFb: fibroblasts, mEpithelial: hepatocytes, biliary cells and progenitors; the first letter m indicates mouse clusters. The two small UMAP plots show the distribution of mouse models (left) and tissue types (right). **e**, UMAP plot showing myeloid clusters including 5 DC, 2 monocyte and 7 macrophage clusters for liver cancer mouse models. **f**, Dot heatmap showing the row-scaled expression of typical marker genes for neutrophil clusters in mice. **g**, Stacked barplot showing the fraction of 12 mouse neutrophil subsets across PB, AL, and tumour. **h**, The trajectory path of mouse neutrophil clusters inferred by Monocle2. Each dot represents a single cell. Cell orders are inferred from the expression of the most variable genes. The trajectory direction is determined by biological prior. **i**, Heatmap showing Pearson's correlations across neutrophil clusters in human and mouse. **j**, UMAP plots showing the integration of mouse and human neutrophil clusters. **k**, Sankey plot showing the similarities of the joint clusters, mouse tissue isolated neutrophil clusters, and human sample isolated neutrophil clusters. **l**, FACS analysis on neutrophil, macrophage, and $CD8^+$ T cell populations in isotype and anti-Ly6G groups. The right barplot shows the decreased neutrophil number in anti-Ly6G group ($n = 10$). **m**, FACS analyses and coloured histogram showing reduced PD-L1 expression in TANs and reduced PD-1 and TIM3 expression in tumour-infiltrated $CD8^+$ T cells of the anti-Ly6G group compared with isotype control. The left barplot shows the decreased PD-L1 expression of neutrophils in anti-Ly6G group ($n = 8$). **n**, IHC of CD68 in tumour regions of mice treated with isotype control or anti-Ly6G antibody ($n = 6$). **o**, FACS analysis showing the expression of surface and intracellular Ly6G in the isotype control and anti-Ly6G treatment groups. **p**, Bar plot showing the statistical analysis of FACS results ($n = 3$). In **l**-**p**, *n* denotes biologically independent samples, data are presented as mean values ± SEM, and two-sided Student's t-test is used.

nature portfolio

Corresponding author(s): Ning Zhang, Zemin Zhang, Jiye Zhu

Last updated by author(s): Sep 7, 2022

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	scRNA-seq data were collected by Chromium single cell controller (10x Genomics) with built-in software. FACSDiva v8.0.1 was used for collecting FACS data. Multispectral images were scanned with ZEISS AXIOSCAN 7.
Data analysis	Codes used in this study are provided at https://github.com/meta-cancer/scPLC. All scRNA-seq data were processed with CellRanger v3.1. R v3.6.1, Rstudio v3.5.3 and the following R packages were used: dropletUtils v1.10.3, scran v1.18.7, igraph v1.2.9, Seurat v3.2.3, inferCNV v1.3.3, clusterProfiler v3.18.1, epitools v0.5-10.1, pheatmap v1.0.12, NicheNet v1.1.0, CellphoneDB v2.1.7, CytoTRACE v0.3.3, Monocle v2.12, SCENIC v1.1.3, SciBet v1.0, SingleR v1.10.0, LIGER v1.0, and Survminer v0.4.9. FastQC v0.11.8, STAR v2.5.2b, RSEM v1.3.1, and DESeq2 v1.24 were used for bulk and scRNA-seq analysis. Python v3.7.4 and python packages: anadata v0.7.5, scanpy v1.6, and CellRank v1.5.1 were also used. WES data were analyzed with BWA-mem2 v2.0pre1, Samtools v1.10, GATK v4.1.7.0, VarScan v2.4.2, Mutect2 v4.1.0.0, ANNOVAR, VEP v96, ABSOLUTE v1.0.6, CNVkit v0.9.7, and GISTIC2 v2.0.23. ATAC-seq data were analyzed with trimmomatic v0.39, Bowtie2 v2.4.4, PicardTools v2.23.3, MACS3 v3.0.0a7, DEseq2 v1.24, DeepTools v3.5.1, and pyGenomeTracks v3.6. FlowJo v10.4 was used for FACS analysis. Halo v3.4 , QuPath v0.2.0, and Image J v1.52k software were used for image analysis. GraphPad Prism v9.0 was used for statistical analysis for data collected from functional experiments.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Raw sequence data reported in this paper have been deposited in the Genome Sequence Archive in National Genomics Data Center (Beijing, China), under the BioProject ID: PRJCA007744. The data deposited and made public are compliant with the regulations of the Ministry of Science and Technology of China. To facilitate usage of our data for the wide research community, we developed an interactive web-based tool (http://meta-cancer.cn:3838/scPLC) for analyzing and visualizing our single-cell data. Other public data used in this study includes, reference genomes for human (https://asia.ensembl.org/, GRCh38.p13) and mouse (https://asia.ensembl.org/, GRCm39), and TCGA datasets (https://portal.gdc.cancer.gov/).

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were used to predetermine sample size of scRNA-seq libraries. We performed a prospective screen of treatment-naive liver cancer patients underwent primary curative resection from March 2019 to January 2020 at Peking University People's Hospital. A total of 124 patients were enrolled and 160 samples were sent for scRNA-seq, including 79 HCC, 25 ICC, 7 CHC, 2 hepatic hemangioma (HH), 1 adenosquamous carcinoma (ASC), 1 sarcomatoid carcinoma (SAR) and 9 secondary liver cancer (SLC, liver metastases from various primary sites) cases. Detailed clinical characteristics were summarized in Supplementary Table 1. All samples that passed the QC of single cell libraries were included. For functional validation experiments, three or more biologically independent human or mouse samples were used to reach the requirements of statistical analyses.
Data exclusions	There is no data that were excluded from the analyses.
Replication	For scRNA-seq analysis, there is no replication for the human tumor samples. For experimental validations, each experiment was repeated three or more times (n denotes biologically independent samples, and is denoted in the legend for each figure).
Randomization	No randomization was performed for the human tumor samples because this is an observational study. For cKO mouse models, body weight-matched mice were randomized over the treatment groups, anti-Ly6G and isotype control. For pTMC-Luc mouse model, tumor size were monitored by luminescence signals at day 7 after HDTVi, and tumor size-matched mice were randomized over the treatment groups, anti-Ly6G and isotype control.
Blinding	Blinding was not considered appropriate for this study because this is an observational study. Our analyses and results were based on the cancer type of samples determined by pathologists.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

Antibodies for flow cytometry
Source Catalog number Clone name Lot number Dilution
Anti-CD45 Antibody (PerCP), Mouse Monoclonal Sino Biological. Inc 10086-MM05-C MM05 PK09MY2006 1:25
BD Pharmingen™ APC-Cy™7 Mouse Anti-Human CD45 BD 557833 2D1 1232778 1:100
BD Pharmingen™ FITC Mouse Anti-Human CD3 BD 555332 UCHT1 6285760 1:100
BD Horizon™ BV421 Mouse Anti-Human CD3 BD 562426 UCHT1 1152667 1:100
BD Horizon™ BV421 Mouse Anti-Human CD2 BD 562639 RPA-2.10 1005982 1:100
BD Pharmingen™ APC-H7 Mouse anti-Human CD8 BD 560179 SK1 259953 1:100
BD Horizon™ BV510 Mouse Anti-Human CD8 BD 563919 SK1 1012142 1:100
BD Horizon™ BV786 Mouse Anti-Human CD25 BD 563701 M-A251 9081958 1:100
BD Horizon™ BB515 Mouse Anti-Human CD25 BD 564468 2A3 0072491 1:100
BD Horizon™ BV421 Mouse Anti-Human CD69 BD 562884 FN50 7040967 1:100
BD Pharmingen™ PE-Cy™7 Mouse Anti-Human CD69 BD 557745 FN50 9311621 1:100
Recombinant Anti-IFN gamma Antibody (PE), Rabbit Monoclonal Sino Biological. Inc 11725-R003-P R003 HR11OC3001 1:10
BD Pharmingen™ PE-Cy™7 Mouse Anti-Human IFN-γ BD 557643 B27 1046592 1:100
BD Pharmingen™ PE Mouse Anti-Human Granzyme B BD 561142 GB11 1109119 1:100
BD Pharmingen™ PerCP-Cy™5.5 Mouse Anti-Human Perforin BD 563762 δG9 1216609 1:100
BD Pharmingen™ PE Mouse Anti-Human CD274 BD 557924 MIH1 7096869 1:100
BD Pharmingen™ Alexa Fluor® 647 Mouse Anti-Human CD66b BD 561645 G10F5 1123620 1:100
BD Pharmingen™ PE Mouse Anti-Human CD66b BD 561650 G10F5 1158304 1:100
PE/Dazzle™ 594 anti-mouse/human CD11b Antibody BioLegend 101256 M1/70 B276558 1:100
Brilliant Violet 711™ anti-mouse/human CD11b Antibody BioLegend 101242 M1/70 B345610 1:100
FITC anti-mouse/human CD11b Antibody BioLegend 101206 M1/70 B260639 1:100
APC/Cyanine7 anti-mouse CD45 Antibody BioLegend 103116 30-F11 B297466 1:100
Brilliant Violet 510™ anti-mouse CD3ε Antibody BioLegend 100353 145-2C11 B340821 1:100
Recombinant Anti-CD8 alpha Antibody (APC), Rabbit Monoclonal Sino Biological. Inc 50389-R208-A R208 HS14MY1303 1:10
BD Pharmingen™ PE-Cy™7 Rat Anti-Mouse CD8a BD 552877 53-6.7 1152381 1:100
Brilliant Violet 605™ anti-mouse F4/80 Antibody BioLegend 123133 BM8 B284701 1:100
BD Horizon™ BV421 Rat Anti-Mouse LY-6G BD 562737 1A8 9162683 1:100
BD Pharmingen™ PE Rat Anti-Mouse Ly-6G BD 551461 1A8 0337169 1:100
PE/Cyanine7 anti-mouse Ly-6G Antibody BioLegend 127618 1A8 B351626 1:100
BD Horizon™ PE-CF594 Hamster Anti-Mouse CD279 (PD-1) BD 562523 J43 8255909 1:100
BD Pharmingen™ PE Rat Anti-Mouse CD274 BD 558091 MIH5 1005002 1:100
BD Pharmingen™ PE Mouse Anti-Mouse CD366 (TIM-3) BD 566346 5D12/TIM-3 8305644 1:100
InVivo Mab rat IgG2a isotype control (Isotype1) Bio X Cell BE0089 2A3 815022F1
InVivoMab anti-mouse Ly6G Bio X Cell BE0075-1 1A8 737721M1
InVivoMab anti-human PD-L1 (B7-H1) Bio X Cell BE0285 29E.2A3 804721J2
InVivoMab mouse IgG2b isotype control Bio X Cell BE0086 MPC-11 779420O1

Antibodies for histological analyses
Anti-CD66 antibody GeneTex GTX19779 80H3 821901081 1:1000
Anti-CCL4/MIP-1 beta antibody Abcam ab235961 EP521Y GR3273495-7 1:800
Recombinant Anti-PD-L1 antibody Abcam ab237726 CAL10 GR3300061-4 1:1000
anti-human CD8 ZSGB-BIO ZA0508 SP16 21096408 1:100
anti-human PD1 ZSGB-BIO ZM0381 UMAB199 21102842 1:50
Recombinant Anti-Granzyme K antibody Abcam ab282703 EPR24601-164 GR3398831-2 1:1000
Anti-alpha smooth muscle Actin antibody Abcam ab7817 1A4 GR3425194-7 1:5000
Anti-Von Willebrand Factor antibody Abcam ab9378 NA GR3333649-2 1:100
Anti-mouse Ly6G servicebio GB11229 1A8 AC2101177B 1:500
Anti-Ki67 antibody Abcam ab15580 NA GR3426431-1 1:500
Anti-CD68 antibody servicebio GB113109 NA c62921012803 1:200
Recombinant Anti-EpCAM antibody Abcam ab213500 EPR20532-222 NA 1:1000
Anti-Hepatocyte Specific Antigen antibody Abcam ab75677 OCH1E5 GR3371755-1 1:1
HRP-anti-rabbit IgG ZSGB PV-6001 NA 2181D1217 1:1
HRP-anti-mouse IgG ZSGB PV-6002 NA 2131D1129 1:1

Validation

All antibodies used in this study are commercially available. They are validated by the vendors for the specific assay and species used, with the validation reports available on the vendor's website. All antibodies were titrated to determine the optimal working concentration. Isotype controls were used to gate on cells staining with the antibody of interest.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

The human embryonic kidney cell line (HEK293T, ATCC number, CRL-3216) and liver cancer cell line (HepG2, ATCC number, HB-8065) were obtained from American Type Culture Collection (ATCC). Human liver cancer cell lines (HCCLM3 and MHCC97H) were obtained from the Liver Cancer Institute, Zhongshan Hospital, Fudan University (Shanghai, China).

Authentication

Cell lines used in this study (HEK293T, HepG2, HCCLM3, and MHCC97H) were authenticated by applying short tandem-repeat (STR) DNA profiling.

Mycoplasma contamination	All cell lines tested negative for Mycoplasma.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified lines were used in this study.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Trp53fl/fl and Alb-Cre mice (both C57Bl/6) were purchased from the Jackson Laboratory and bred in a pathogen-free environment per guidelines of the animal facility in Peking University First Hospital. Trp53fl/fl mice were crossed with Alb-Cre mice to generate the liver conditional Trp53 knockout (Trp53 cKO) mice. 7-week-old male Trp53 cKO mice were used for experiments. All mice were housed in pathogen free conditions at an ambient temperature 20-26°C and humidity of 30-70% with a 12:12 hour light:dark cycle prior to use. Body weight of mice was monitored twice every week for signs of dynamic tumor growth. The diameter of single tumor was < 2cm.
Wild animals	This study did not involve wild animals.
Field-collected samples	This study did not involve field-collected samples.
Ethics oversight	All mouse experiments were approved by the Animal Care and Use Committee at Peking University First Hospital.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about studies involving human research participants

Population characteristics	A total of 124 patients were enrolled and 160 samples were sent for scRNA-seq, including 79 HCC, 25 ICC, 7 CHC, 2 hepatic hemangioma (HH), 1 adenosquamous carcinoma (ASC), 1 sarcomatoid carcinoma (SAR) and 9 secondary liver cancer (SLC, liver metastases from various primary sites) cases. Detailed clinical characteristics were summarized in Supplementary Table 1.
Recruitment	We performed a prospective screen of treatment-naive liver cancer patients underwent primary curative resection from March 2019 to January 2020 at Peking University People's Hospital. All patients with surgical samples available for analyses were included. Written informed consent was obtained from each patient. No self-selection biased was anticipated.
Ethics oversight	This study was approved by the Research Ethics Committee of both Peking University First Hospital and Peking University People's Hospital.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Fresh tumor and AL samples were cut into approximately 1 mm3 pieces in the RPMI-1640 medium (ThermoFisher) with 10% fetal bovine serum (FBS, Gibco) 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. The enzymatic reaction was stopped using cold medium and tissue suspension was filtered through a 70 µm nylon mesh (FALCON).
Instrument	BD FACSAria SORP
Software	FACSDiva v8.0.1 and FlowJo v10.4
Cell population abundance	Abundance of the cell population in the sorted samples were indicated in Supplementary Fig. 5.

Gating strategy

Intact cells were gated according to the FSC-A and SSC-A. Doublets were excluded by the FSC-H and FSC-A. Dead cells were excluded based on DAPI staining. Gating Strategies for specific cell populations were shown in Supplementary Fig. 5. Sort purities were routinely confirmed, as assessed by post-sort measurements of respective target cell populations. Purity of post sort populations was >90%.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.