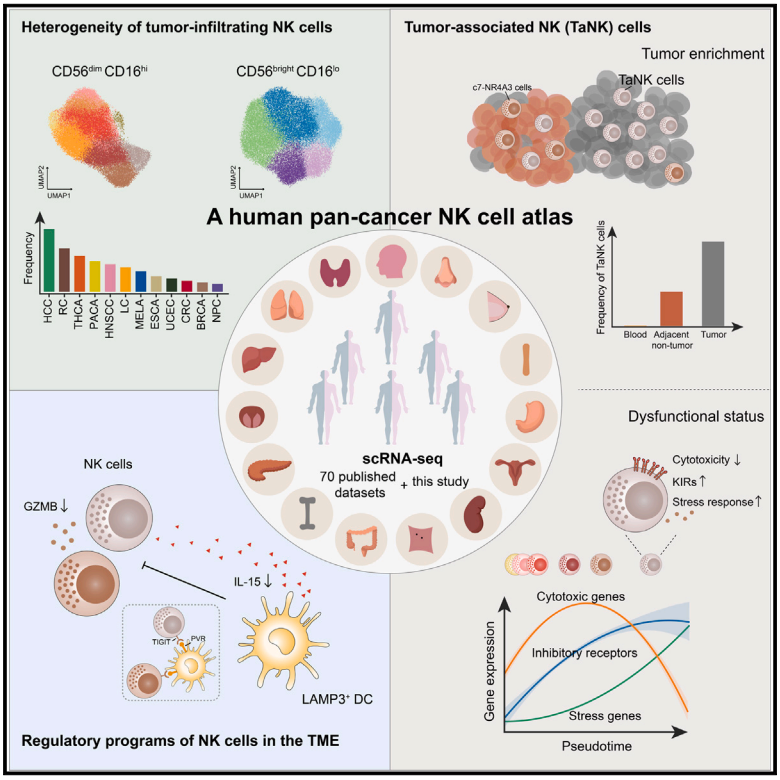
Cell

Resource

A pan-cancer single-cell panorama of human natural killer cells

Graphical abstract



Highlights

- Human NK cells exhibit tumor-type-specific subgroup heterogeneity
- *RGS1* expression is a reliable marker of tissue-infiltrating NK cells
- Transcriptomic features of tumor-associated NK cells indicate impaired cytotoxicity
- LAMP3^{+} DCs appear to regulate NK cell functions in tumors based on spatial analyses

Authors

Fei Tang, Jinhu Li, Lu Qi, ..., Hui Peng, Dongfang Wang, Zemin Zhang

Correspondence

zhuln@pku.edu.cn (L.Z.), huipeng@ustc.edu.cn (H.P.), wangdf19@pku.edu.cn (D.W.), zemin@pku.edu.cn (Z.Z.)

In brief

Integrative single-cell RNA sequencing analyses on natural killer (NK) cells from over 700 patients across 24 tumor types depict shared and tumor-type-specific NK cell features and highlight the potential of specific myeloid cell subpopulations in regulating NK cell anti-tumor function.

Tang et al., 2023, Cell *186*, 4235–4251
September 14, 2023
https://doi.org/10.1016/j.cell.2023.07.034

CellPress

Cell

CellPress
OPEN ACCESS

Resource

A pan-cancer single-cell panorama of human natural killer cells

Fei Tang,[1,7] Jinhu Li,[1,7] Lu Qi,[1,2] Dongfang Liu,[3] Yufei Bo,[1] Shishang Qin,[1] Yuhui Miao,[4] Kezhuo Yu,[1] Wenhong Hou,[4] Jianan Li,[1] Jirun Peng,[3,5] Zhigang Tian,[6] Linnan Zhu,[1,*] Hui Peng,[6,*] Dongfang Wang,[1,*] and Zemin Zhang[1,4,8,*]

[1]Biomedical Pioneering Innovation Center (BIOPIC), Academy for Advanced Interdisciplinary Studies, School of Life Sciences, Peking University, Beijing 100871, China
[2]Changping Laboratory, Beijing 102206, China
[3]Department of Surgery, Beijing Shijitan Hospital, Capital Medical University, Beijing 100038, China
[4]Institute of Cancer Research, Shenzhen Bay Laboratory, Shenzhen 518132, China
[5]Ninth School of Clinical Medicine, Peking University, Beijing 100038, China
[6]The CAS Key Laboratory of Innate Immunity and Chronic Disease, School of Basic Medical Sciences, Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei 230027, China
[7]These authors contributed equally
[8]Lead contact
*Correspondence: zhuln@pku.edu.cn (L.Z.), huipeng@ustc.edu.cn (H.P.), wangdf19@pku.edu.cn (D.W.), zemin@pku.edu.cn (Z.Z.)
https://doi.org/10.1016/j.cell.2023.07.034

SUMMARY

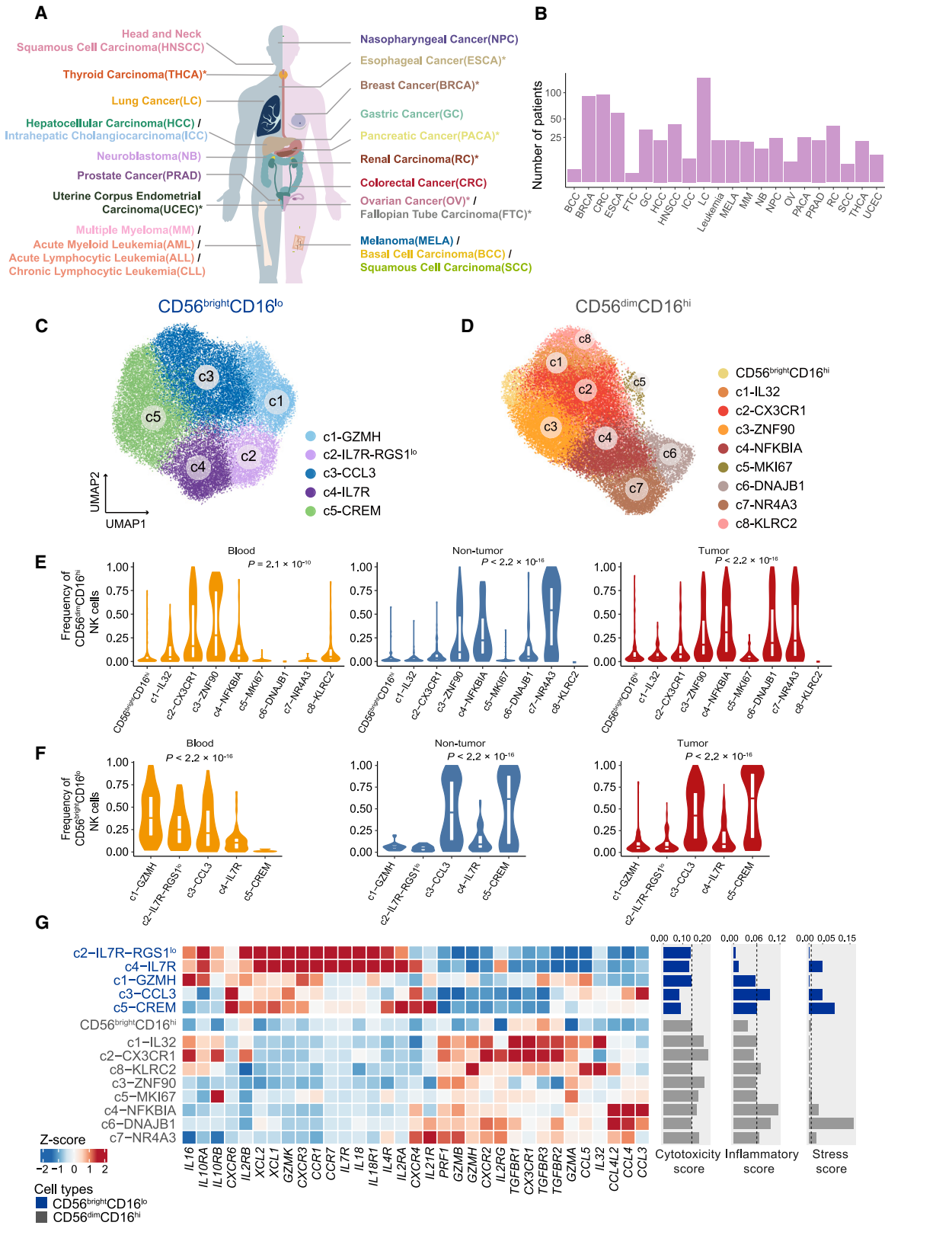
Natural killer (NK) cells play indispensable roles in innate immune responses against tumor progression. To depict their phenotypic and functional diversities in the tumor microenvironment, we perform integrative single-cell RNA sequencing analyses on NK cells from 716 patients with cancer, covering 24 cancer types. We observed heterogeneity in NK cell composition in a tumor-type-specific manner. Notably, we have identified a group of tumor-associated NK cells that are enriched in tumors, show impaired anti-tumor functions, and are associated with unfavorable prognosis and resistance to immunotherapy. Specific myeloid cell subpopulations, in particular LAMP3^{+} dendritic cells, appear to mediate the regulation of NK cell anti-tumor immunity. Our study provides insights into NK-cell-based cancer immunity and highlights potential clinical utilities of NK cell subsets as therapeutic targets.

INTRODUCTION

Although T cell-centric immunotherapies have achieved indisputable clinical successes, the limited number of patients achieving durable response presses for complementary therapeutic strategies.[1] Natural killer (NK) cells, as one important component in the tumor microenvironment (TME), are involved in multiple processes of tumor control, such as direct cell killing and secretion of proinflammatory cytokines.[2,3] A plethora of strategies, characterized with promising properties including their safety and efficacy, have been proposed to harness NK cells for cancer treatment.[4] Of particular note is the remarkable clinical success achieved by chimeric antigen receptor (CAR)-NK cell therapies in lymphoma, myeloma, and leukemia.[5,6] However, NK-cell-based therapies are hindered in solid tumors partly due to the incomplete understanding of tumor-infiltrating NK cells, especially their infiltration to tumor, phenotypic heterogeneities, and dysregulation within the TME.[7,8]

In humans, NK cells can be subdivided into two major groups, CD56^{dim}CD16^{hi} and CD56^{bright}CD16^{lo}, based on the expression levels of CD56 (*NCAM1*) and CD16 (*FCGR3A*).[9,10] The CD56^{dim}CD16^{hi} NK cell population predominantly mediates the killing of target cells by secreting perforin and granzymes,[11] whereas the CD56^{bright}CD16^{lo} NK cell population exhibits immunoregulatory and cytokine-producing capacity.[12] Recently, single-cell RNA sequencing (scRNA-seq) technologies have facilitated the characterization of the heterogeneity of tumor-infiltrating immune cells, providing great opportunities to elucidate the spectrum of tumor-infiltrating NK cell subpopulations.[13] For example, we have identified a *CD160*$^{+}$*HSPA1A*$^{+}$ liver-resident NK cell subset specifically enriched in hepatocellular carcinoma (HCC),[14] and others have reported specialized NK cell populations infiltrated in melanoma with regional variation.[15] In parallel, the distribution and function of human NK cell populations have been investigated in healthy tissues and blood.[16–18] However, for tumor-infiltrating NK cells, the scale of most scRNA-seq analyses is limited, and it is still unclear to what extent they acquire heterogeneous phenotypes in malignant conditions.[19–21]

Compared with CD8^{+} T cells, NK cells serve as alternative sources of cytotoxic activities and combat tumor cells with low mutation load and aberrant expression of major histocompatibility complex (MHC) class I.[22,23] Although the dysfunctional state of CD8^{+} T cells is well-characterized with reduced cytotoxicity



(legend on next page)

and high expression of multiple inhibitory receptors,[24] the NK cell dysfunction has not been studied in detail. Previously, NK cell hypofunction has been reported in HCC, but the immunoregulatory mechanisms have not been analyzed,[25,26] and whether this phenomenon exists in other cancer types is not clear. In addition, although the inhibitory roles of TIGIT and TIM3 have been established in tumor-infiltrating NK cells,[27,28] it is still controversial whether other immune checkpoints, such as PD-1 and CTLA4, play the same role or are even expressed in those cells.[28,29] Furthermore, NK cells are sensitive to immunosuppressive factors of the TME, which may contribute to their dysfunctional phenotype,[4] but how distinct regulatory processes influence the function and abundance of NK cells within the TME of different cancer types remains unclear. Together, these prompted us to conduct a deep investigation of NK cells to elucidate their heterogeneity and dysfunction at the pan-cancer level.

Here, we collected a broad spectrum of published and newly generated scRNA-seq data to construct a comprehensive tumor-infiltrating human NK cell atlas and explored the heterogeneity of NK cells across cancer types and tissues. We uncovered the tumor-infiltrating NK cell state transitions and highlighted the components of the TME that presumably led to the NK cell dysfunction. These data will serve as a rich resource for advancing the understanding of the global properties of NK cells in major cancer types and provide valuable insights into NK-cell-based immunotherapy development.

RESULTS

Construction of a human pan-cancer NK cell landscape at single-cell resolution

To construct a comprehensive pan-cancer single-cell transcriptome atlas for tumor-infiltrating NK cells, we first compiled scRNA-seq data from both our newly generated dataset with 47 patients diagnosed with one of the 8 cancer types and 70 additional published datasets (Table S1). These data covered 24 cancer types, including 1,223 samples from 716 patients across the tumor, adjacent non-tumor tissue, peripheral blood, and other tissues such as lymph nodes, and 60 healthy controls (Figures 1A, 1B, S1A, and S1B). After stringent quality control and a combinative strategy of computational gating ($CD3^-CD56^+$/$KLRF1^+$) and unsupervised clustering (Figure S1C; STAR Methods), we obtained a collection of 160,011 high-quality NK cells including 11,963 newly generated NK cells in this study. Notably, it is essential and challenging to build such a large-scale human NK cell atlas, given that within the tumor tissue, NK cells are relatively rare among $CD45^+$ cells (Figure S1D), and their heterogeneity is still poorly understood. Compared with the previous single-cell transcriptome atlas of NK cells constructed with approximately 7,000 NK cells from human blood and spleen,[17] our data scale has expanded >20-fold, representing most major cancer types and multiple tissues.

To unbiasedly define the pan-cancer population structure of NK cells, we integrated scRNA-seq data with minimal batch effects among datasets and performed two rounds of unsupervised clustering (Figure S1E; STAR Methods). The first-round analysis pertained to distinguishing two well-characterized major cell types, $CD56^{bright}CD16^{lo}$ and $CD56^{dim}CD16^{hi}$, based on the high expression of canonical cell markers, *NCAM1* and *FCGR3A*, which corresponded to the previously reported "NK_1" and "NK_2" populations,[17] respectively. The $CD56^{bright}CD16^{lo}$ compartment can be further subdivided into 5 subsets, whereas 9 $CD56^{dim}CD16^{hi}$ subsets were identified during the second-round clustering (Figures 1C and 1D). We did not observe notable differences in the expression of *KLRF1* in either of the two major populations across cancer types or different subsets (Figures S1F and S1G). Notably, previously reported innate lymphoid cell (ILC) signature genes were hardly expressed in these NK cell subsets,[30] further suggesting the purity of our data (Figure S1H). These subsets were all characterized by the high expression of distinct signature genes within each major population (Figures S2A and S2B; Table S2). As expected, previously described subsets were readily identified in our atlas, such as $CD56^{bright}CD16^{lo}$ c5-CREM with high expression of *CREM*[31] (Figure S2B). $CD56^{dim}CD16^{hi}$ c8-KLRC2 NK cells with high expression of *KLRC2* (NKG2C) were regarded adaptive NK cells.[32,33] Our atlas additionally uncovered several underappreciated NK cell subsets with unique transcriptional phenotypes. For example, $CD56^{dim}CD16^{hi}$ c5-MKI67 was distinguished by high expression of proliferation markers like *MKI67* and *STMN1*, and $CD56^{dim}CD16^{hi}$ c6-DNAJB1 specifically expressed genes related to stress response (Figure S2A). Our atlas also captured a low fraction of $CD56^{bright}CD16^{hi}$ NK cells in almost all cancer types, simultaneously expressing high levels of *NCAM1* and *FCGR3A* (Figures S1I and S2D). $CD56^{bright}CD16^{hi}$ NK cells displayed intermediate features between $CD56^{bright}CD16^{lo}$ and $CD56^{dim}CD16^{hi}$ NK cells (Figure S2E), potentially representing developmental intermediates in analogy to mouse $CD27^+CD11b^+$ NK cells, which have also been thought to be a transient maturation stage in mice yet barely detectable by scRNA-seq.[17] We next examined the tissue distribution of all subsets, with distinct tissue enrichment patterns observed, indicating that our integrative analyses can preserve the heterogeneity of different tissues (Figures 1E, 1F, and S2C). To further corroborate the stability of the clustering, NK cells from blood, tumor, and adjacent non-tumor tissues were re-clustered separately, with highly consistent results revealed (Figure S3F). Finally, leveraging the Ratio of Global Unshifted

Figure 1. Pan-cancer single-cell atlas of NK cells and their characteristics

(A) Cancer types involved in the pan-cancer NK cell analysis. *, cancer types with newly generated data.

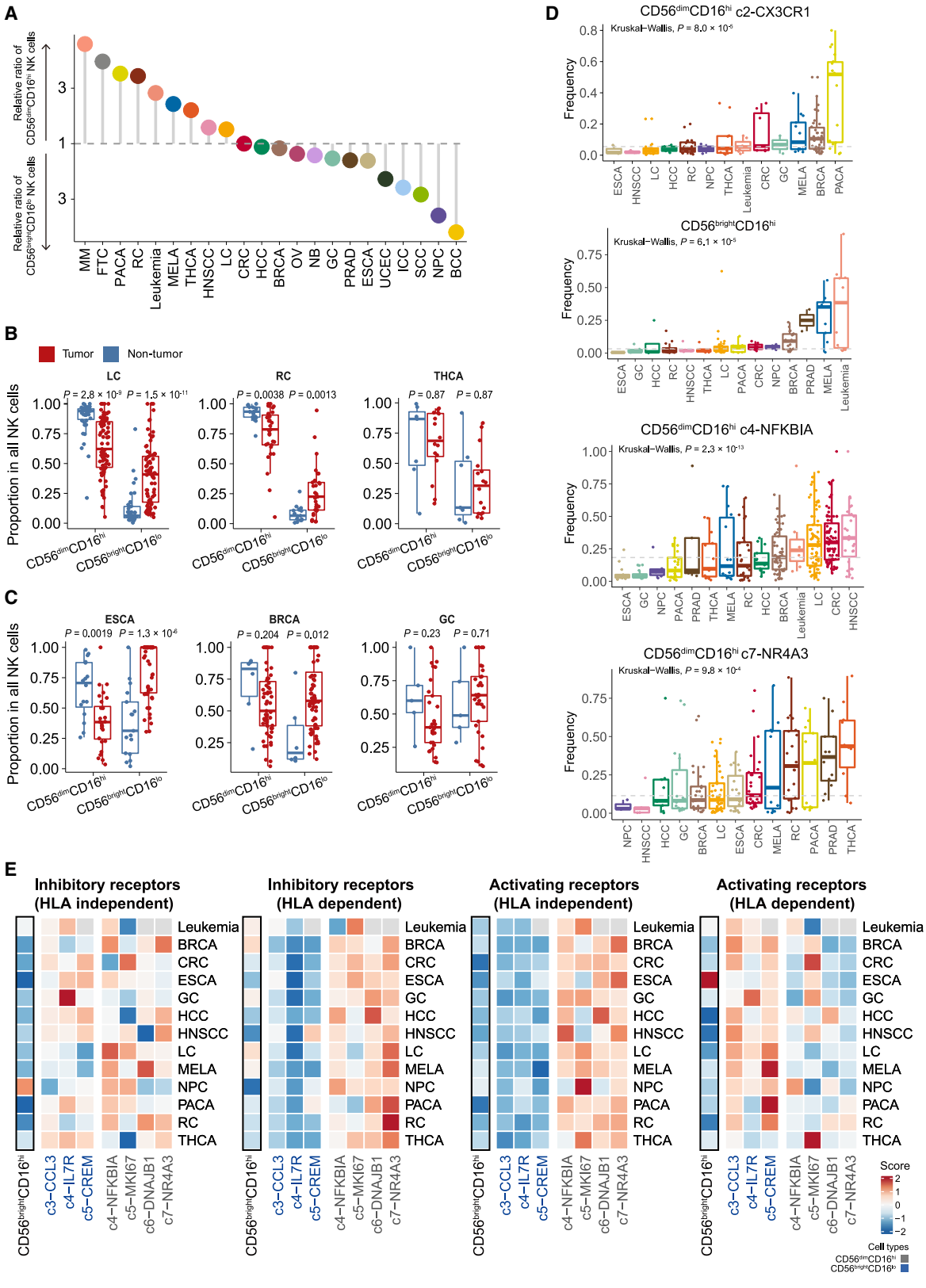
(B) The number of patients across cancer types. The y axis is scaled by a square root transformation.

(C and D) Uniform manifold approximation and projection (UMAP) visualizations of (C) $CD56^{bright}CD16^{lo}$ and (D) $CD56^{dim}CD16^{hi}$ NK cells.

(E and F) Compositions of (E) $CD56^{dim}CD16^{hi}$ and (F) $CD56^{bright}CD16^{lo}$ NK cell subsets across tissues. Kruskal-Wallis test.

(G) The expression pattern of functional genes in NK cells. Score represents the AUCell index of signature genes (STAR Methods). Each dashed line indicates the median of the corresponding score.

See also Figures S1, S2, and S7 and Tables S1, S2, and S7.



(legend on next page)

Entropy (ROGUE) index that can measure the cell cluster purity,[34] we illustrated that all these populations were robust across various cancer types (Figure S1J).

NK cell subsets were involved in different developmental stages as revealed by their expression of common lineage-specific genes. Cells from both the tumor-enriched CD56^{bright}CD16^{lo} c4-IL-7R and the blood-enriched c2-IL-7R-RGS1^{lo}, highly expressing signature of NK cell precursors including *KIT* and *IL7R*,[16,19,20] were mapped at the early developmental stage. By contrast, other immature CD56^{bright}CD16^{lo} subsets exhibited a progressive reduction of the precursor signature and substantial gain of *CD160* expression. The more mature CD56^{dim}CD16^{hi} subsets, especially c6-DNAJB1 and c7-NR4A3, exhibited concomitant upregulation of the killer immunoglobulin-like receptor (KIR) family in addition to *FCGR3A* and *B3GAT1* (CD57) (Figure S2E). NK cell subsets in varied developmental states were concurrent in tumors, indicating that NK cell migration to the tumor may be decoupled with NK cell maturation.

We further inspected the gene expression signature to decipher the functional divergence between different populations (Figures 1G and S2F). CD56^{dim}CD16^{hi} NK cells exhibited high expression of cytotoxic effector genes including perforin (*PRF1*) and most granzyme (*GZMB*, *GZMA*, and *GZMH*) except *GZMK*, which was instead exclusively expressed in CD56^{bright}CD16^{lo} NK cells.[35] CD56^{bright}CD16^{lo} NK cells expressed various cytokine genes such as *IL18*. We also observed that both CD56^{bright}CD16^{lo} c2 and c4 simultaneously expressed *IL18* and its receptor *IL18R1*, implying a potentially crucial role of IL-18-dependent autocrine pathway in these subsets. Strikingly, CD56^{dim}CD16^{hi} NK cell subsets could also exhibit specific expression of certain cytokines, but with a distinct pattern from those CD56^{bright}CD16^{lo} subsets. In particular, the CD56^{dim}CD16^{hi} subset c4-NFKBIA exhibited a relatively high inflammatory score among all NK cell subsets, predominantly expressing *CCL3*, *CCL4*, and *CCL4L2*, indicating their ability to recruit other immune cells such as T cells.[36] Recent studies have associated the recruitment of type 1 conventional dendritic cells (cDC1s) into tumor with NK cells by secreting XCL1, XCL2, and CCL5.[37] Our results identified various NK cell subsets involved in the cDC1 recruitment through cell-type-specific complementary strategies. CD56^{bright}CD16^{lo} c2 and c4 expressed primary levels of *XCL1* and *XCL2*, whereas another cDC1 chemoattractant gene *CCL5* was preferentially produced by both CD56^{bright}CD16^{lo} (c1-GZMH and c3-CCL3) and CD56^{dim}CD16^{hi} (c1-IL-32 and c8-KLRC2) NK cells. Additionally, we depicted the profile of activating and inhibitory receptors, observing clear variations across NK cell subsets (Figure S2G). For example, *KLRC1* was expressed at a higher level in CD56^{bright}CD16^{lo} NK cells than CD56^{dim}CD16^{hi} NK cells, although not specifically, and *LAG3* and *TIGIT* were highly expressed in c8-KLRC2. Notably, compared with other tissue-enriched CD56^{dim}CD16^{hi} subsets (c4-NFKBIA, c5-MKI67, and c7-NR4A3), c6-DNAJB1 NK cells showed the highest stress score and the weakest cytotoxicity, indicating their dramatically different transcriptional and functional phenotypes (Figures 1G and S2F). Taken together, we provide a detailed transcriptome profile of NK cells at a fine-grained subset resolution. Rather than analyzing NK cells as a single population, we dissect subset-specific molecular properties of NK cells and reveal their underappreciated heterogeneities.

Tissue heterogeneity of NK cells across cancer types

We next assessed the preference of tumor-infiltrating NK cell population among different cancer types, observing clear discrepancies (Figure 2A; STAR Methods). For example, the immature CD56^{bright}CD16^{lo} NK cells were largely predominant in nasopharyngeal cancer and basal cell carcinoma, whereas the mature CD56^{dim}CD16^{hi} NK cells occupied renal carcinoma and lung cancer, consistent with previous reports[38–42]. Others, such as colorectal cancer and HCC, showed no obvious propensity (Figures 2A, S3A, and S3B). To examine whether the aforementioned tendency could be explained by the organ contexture, we analyzed the intrinsic composition of major NK cell types in adjacent non-tumor tissues and their corresponding changes in tumors. In certain cancer types such as colorectal cancer, the NK cell composition of tumor tissues resembled that of adjacent non-tumor tissues (Figure S3C). As for cancer types including lung cancer and renal carcinoma, despite the conservation of major NK cell types between tumor and adjacent non-tumor tissues, the proportions of CD56^{dim}CD16^{hi} NK cells significantly decreased (Figure 2B). Intriguingly, the dominant NK cell population in breast cancer and esophageal cancer was reversed compared with their adjacent non-tumor tissues (Figure 2C). These observations suggested that the intrinsic organ properties and malignancy-associated factors have compounded effects on shaping the content of NK cell populations.

We further explored the cancer-type specificity for NK cells from the subset perspective and performed unsupervised clustering to stratify the analyzed cancer types by NK cell subset proportions. NK cell subsets such as c2-CX3CR1 showed strong preference in pancreatic cancer, breast cancer, and melanoma. High variabilities were observed across cancer types for certain subsets. For example, c4-NFKBIA and c7-NR4A3 showed dramatically decreased median frequencies from head and neck squamous cell carcinoma and thyroid carcinoma to esophageal cancer and nasopharyngeal cancer (Figure 2D). Although clear variations were observed, cancer types including uterine corpus endometrial carcinoma, basal cell carcinoma, and esophageal cancer were clustered together, all with abundant

Figure 2. Heterogeneity of tumor-infiltrating NK cells across cancer types

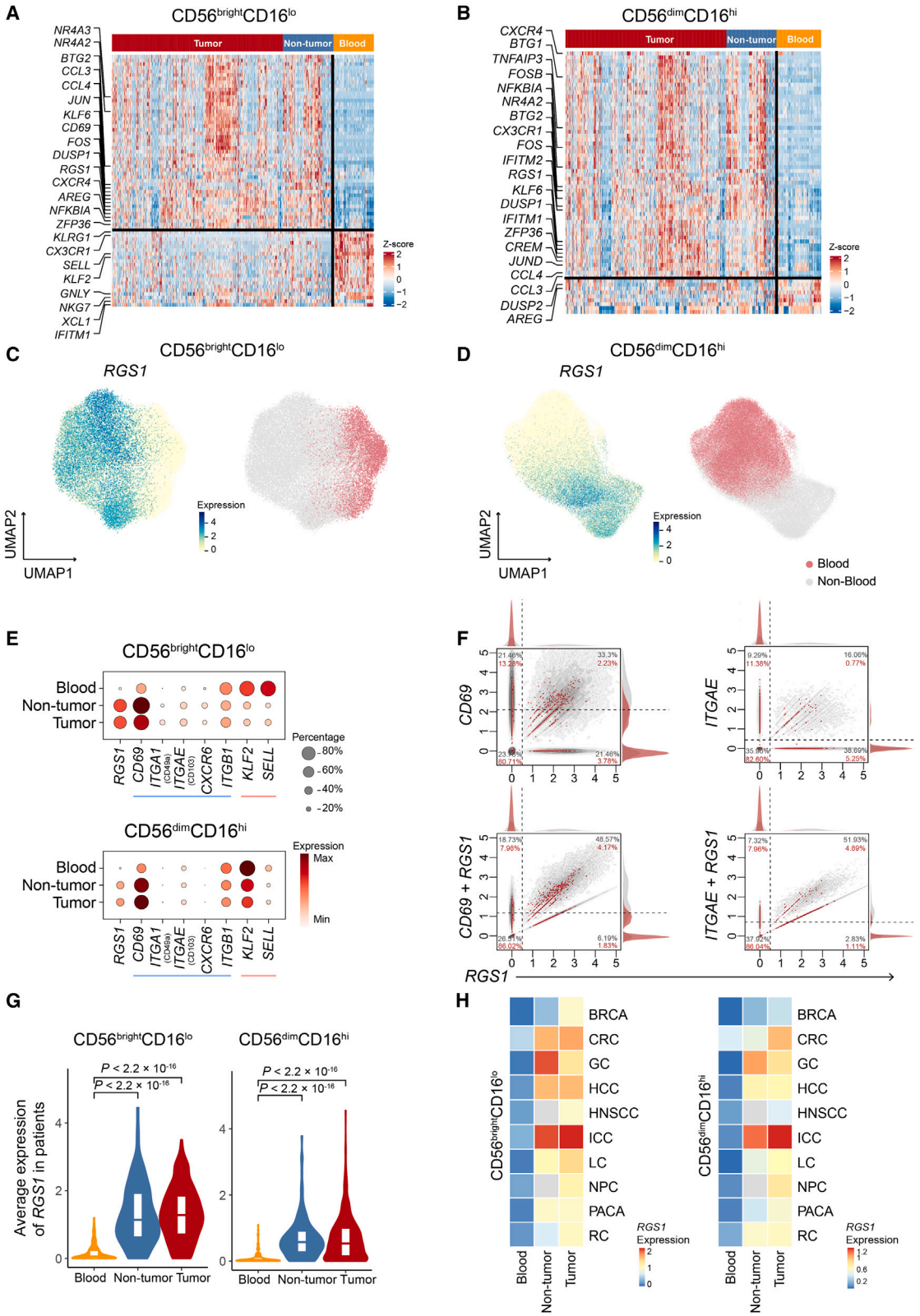
(A) Relative ratios of CD56^{dim}CD16^{hi} and CD56^{bright}CD16^{lo} NK cells in multiple cancer types.

(B and C) Boxplots comparing the proportions of two major NK populations divided by the total NK cell number between tumor and adjacent non-tumor tissues, respectively. Two-sided unpaired Wilcoxon test.

(D) Boxplots showing the frequencies of selected NK cell subsets in tumors. For each subset, only the cancer types with tumor samples > 3 are shown.

(E) Heatmaps showing scores of inhibitory and activating receptors in NK cell subsets (STAR Methods).

See also Figure S3.



(legend on next page)

CD56^{bright}CD16^{hi} c5-CREM and minimal c4-IL-7R NK cells (Figure S3D). Of particular note is that the rare CD56^{bright}CD16^{hi} NK cell subset with the hypomaturation stage as aforementioned was abundant in melanoma and leukemia, especially in the acute myeloid leukemia (AML) subtype (Figures 2D and S3E). Such hypomaturation stage of NK cells has been associated with the reduction of overall survival and relapse-free survival of patients with AML.[43,44] Compared with other NK cell populations, these CD56^{bright}CD16^{hi} cells exhibited distinctive phenotypic and functional shifts in terms of their extremely low activating and inhibitory receptor scores (Figure 2E). Current NK cell-based therapies are focused on augmenting the activation and longevity of NK cells but generally disregard the heterogeneity among cancer types and the suppressive impact of the TME on NK cell cytotoxic functions, which should be considered in future therapeutic strategies.

RGS1 is a hallmark of tissue-infiltrating NK cells

As aforementioned, NK cell components significantly altered from the blood to tissues. Likewise, widespread transcriptional changes were detected between tissue-infiltrating NK cells and their blood counterparts (Figures 3A and 3B; Table S3). Prior studies have identified several markers of tissue-resident NK cells, such as CD69, CD103, CXCR6, and CD49a.[45–47] However, we found that *ITGA1* (CD49a), *ITGAE* (CD103), and *CXCR6* were poorly detected at the single-cell transcriptome level, and *CD69* was expressed widely in NK cells including those from the blood (Figure 3E). Additionally, preferential expression of these markers in NK cell populations of particular tissues has been reported.[16] These motivated us to discover robust NK cell residency markers from the pan-cancer perspective in an unbiased manner.

We selected differentially expressed genes between blood and tissue and further assessed their sensitivity and specificity to distinguish the tissue origin of NK cells. Consequently, *RGS1* (regulator of G protein signaling 1) was pronouncedly recognized, which was exclusively expressed in NK cells within tumor and adjacent non-tumor tissues, but barely detectable in the blood (Figures 3C and 3D). In addition, the expression of *RGS1* was opposite to migration signals including *KLF2* and *SELL* (Figure 3E). Compared with the aforementioned conventional tissue-resident markers, *RGS1* showed much higher sensitivity and specificity (Figures S4A–S4D). We next directly compared the expression patterns of *RGS1* with *ITGAE* and *CD69* as well as their combinations in the blood, observing that *RGS1* alone had the lowest detection rate in the blood, whereas the *RGS1*/*CD69* combination could further increase the positive detection rate in tissues (Figures 3F and S4E). Furthermore, the *RGS1*/*CD69* combination could achieve a higher area under the curve (AUC) value than either alone in distinguishing blood and non-blood NK cells (Figure S4D). Notably, *RGS1* was widely expressed across analyzed patients and cancer types (Figures 3G, 3H, and S4F).

In summary, these characteristics imply the potential role of *RGS1* alone, or its combination with *CD69*, to be a superior marker for tissue-infiltrating NK cells at the transcriptome level. We speculate that the expression of *RGS1* may attenuate the signaling activity of G-proteins,[48] leading to the weakness of NK cell chemotactic migration ability and promoting NK cell residency. Further studies are required to demonstrate the functional mechanism of *RGS1* in NK cells.

Tumor-associated NK cell programs and their characteristics

We next sought to elucidate the specific characteristics of NK cells in tumors. Aside from the tumor-specific enrichment of certain subsets as described above, we found that compared with the adjacent non-tumor tissue, in tumors, the cytokine production of CD56^{bright}CD16^{lo} c3-CCL3 NK cells was lower as indicated by the diminished expression of *CCL3* and *CCL4*, and the expression of *XCL1* and *XCL2* in c5-CREM NK cells decreased in most cancer types (Figures S5A and S5B; Table S5), implicating their functional shifts in tumors. We then identified activated regulons for both tumor-infiltrating CD56^{bright}CD16^{lo} and CD56^{dim}CD16^{hi} NK cell subsets using SCENIC.[49] Importantly, the tumor-enriched c6-DNAJB1 subset exhibited much higher expression of transcription factors such as *KLF6* and *EGR3*, which are associated with the inhibition of cytotoxicity functions[50–52] (Figure S5C).

Using RNA velocity,[53,54] we decoded the transcriptional dynamics of NK cells, observing a clear directional flow from blood-enriched subsets to tumor-infiltrating populations both in CD56^{bright}CD16^{lo} and CD56^{dim}CD16^{hi} NK cells (Figures 4A and S5D). Correspondingly, the expression of *RGS1* was elevated along the velocity flow (Figure S5E). We found that CD56^{dim}CD16^{hi} c6-DNAJB1 NK cells were located at the end of the velocity, thereby inferred as the terminal state (Figure 4A). Notably, CD56^{dim}CD16^{hi} NK cells from the adjacent non-tumor tissue were mainly observed in the uniform manifold approximation and projection (UMAP) area enriched with c7-NR4A3 cells; by contrast, tumor-derived CD56^{dim}CD16^{hi} NK cells were predominant in the UMAP area enriched with c6-DNAJB1 cells

Figure 3. Identification of *RGS1* as a key tissue-infiltrating marker for NK cells

(A and B) Heatmap showing differentially expressed genes for (A) CD56^{bright}CD16^{lo} and (B) CD56^{dim}CD16^{hi} NK cells between blood and non-blood tissue. Rows represent signature genes and columns represent different patients.

(C and D) UMAP plots showing the *RGS1* expression and its tissue distribution in (C) CD56^{bright}CD16^{lo} and (D) CD56^{dim}CD16^{hi} NK cells.

(E) The expression pattern of tissue-resident and migration signals of CD56^{bright}CD16^{lo} and CD56^{dim}CD16^{hi} NK cells in varied tissues. Blue lines mark the tissue-resident signals and red lines mark the migration signals.

(F) The performance of *CD69*, *ITGAE*, and *RGS1* as well as their combinations in distinguishing tissue-derived NK cells based on *in silico* fluorescence-activated cell sorting (FACS). Red dots denote the NK cells derived from blood, and gray dots denote tissue-derived NK cells. Dashed lines indicate the predicted boundary of blood and non-blood NK cells based on the expression level.

(G) Violin plots showing *RGS1* expression among tissues at the pan-cancer level. Two-sided unpaired Wilcoxon test.

(H) Heatmap showing the mean expression of *RGS1* within the blood, tumor, and adjacent non-tumor tissues across analyzed cancer types.

See also Figure S4 and Table S3.

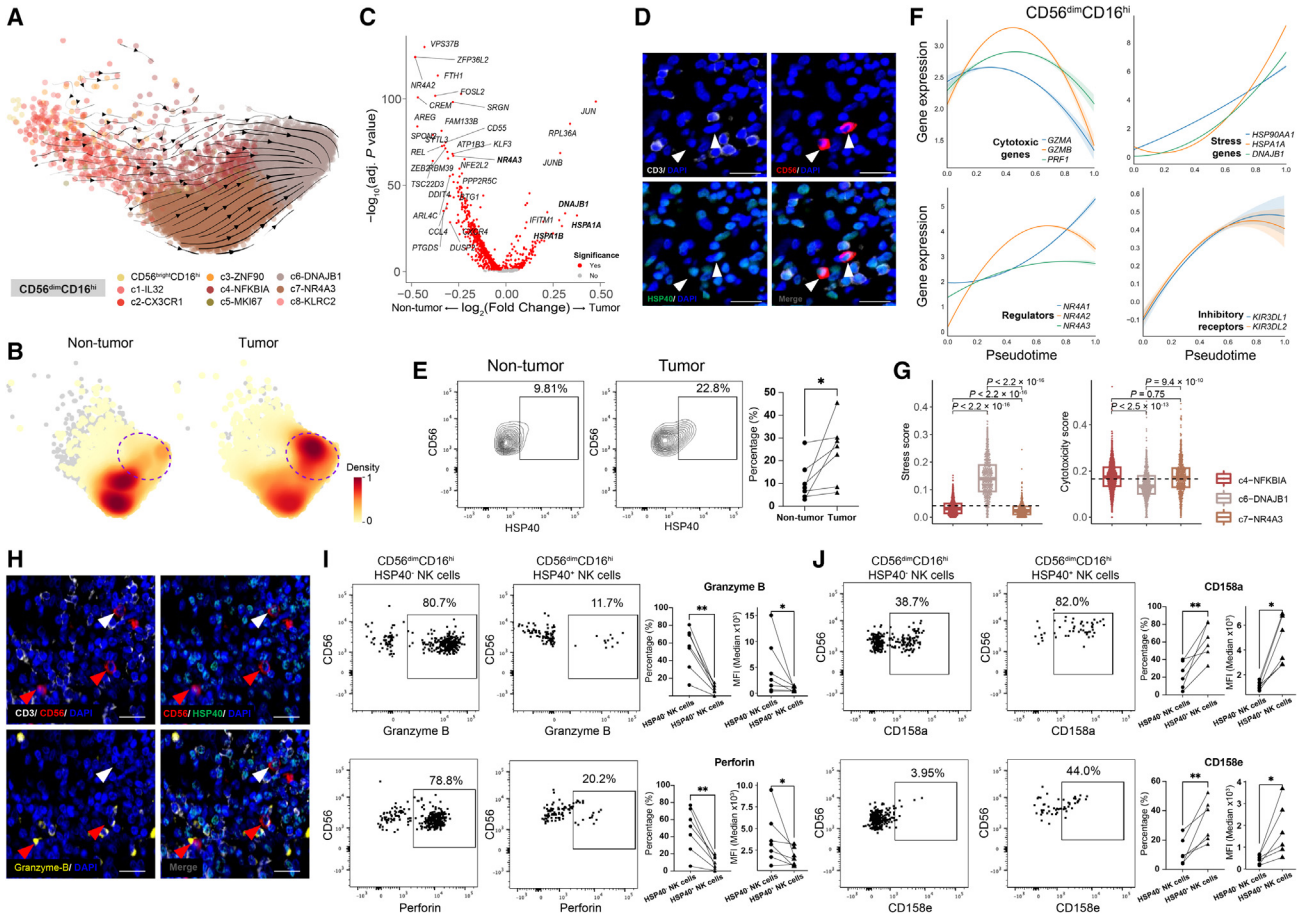


Figure 4. Characteristics of tumor-associated NK cells

(A) RNA velocities overlaid on the UMAP of CD56^{dim}CD16^{hi} NK cells (STAR Methods). Arrows show the RNA velocity field. Dots are colored by CD56^{dim}CD16^{hi} NK cell subsets. Only our newly generated data are used.

(B) Density plots of CD56^{dim}CD16^{hi} NK cells from the tumor and adjacent non-tumor tissue. Dashed lines represent the enriched area of c6-DNAJB1 cells.

(C) Volcano plot showing differentially expressed genes for CD56^{dim}CD16^{hi} NK cells between tumor and adjacent non-tumor tissues. Genes with an adjusted p value < 0.05 are significant. Two-sided unpaired Wilcoxon test.

(D) Representative example of an HCC tumor stained by multiplex immunofluorescence to show TaNK cells (arrows). The scale bar represents 20 μm.

(E) The percentage of CD56^{dim}CD16^{hi} HSP40^{+} NK cells among CD56^{dim}CD16^{hi} NK cells in tumor and adjacent non-tumor regions of liver cancer patients using flow cytometry. *p < 0.05, **p < 0.01, ***p < 0.001, paired t test.

(F) The gene expressions plotted along the pseudotime (STAR Methods).

(G) Boxplots showing signature scores among CD56^{dim}CD16^{hi} c4−NFKBIA, c6-DNAJB1, and c7-NR4A3 cells. Two-sided unpaired Wilcoxon test.

(H) Representative example of an HCC tumor stained by multiplex immunofluorescence to show the expression of granzyme B in HSP40^{+} (white arrows) and HSP40^{-} NK cells (red arrows). The scale bar represents 20 μm.

(I and J) Comparison of (I) cytotoxic granules and (J) inhibitor receptors for CD56^{dim}CD16^{hi} HSP40^{+} and HSP40^{-} NK cells in the tumor region of liver cancer patients using flow cytometry. *p < 0.05, **p < 0.01, ***p < 0.001, paired t test.

See also Figure S5 and Tables S4 and S5.

(Figure 4B). Consistently, markers of tumor-enriched c6-DNAJB1 NK cells such as *DNAJB1* and *HSPA1A* were highly expressed in the tumor-infiltrating CD56^{dim}CD16^{hi} NK cell population (Figure 4C). In addition, we observed similar expression levels of mitochondrial genes in c6-DNAJB1 and c7-NR4A3 NK cells, indicating that the stress phenotype of c6-DNAJB1 NK cells was unrelated to cell quality (Figure S5F). Since c6-DNAJB1 cells were specifically enriched in tumors, we termed this population as tumor-associated NK (TaNK) cells.

The presence of TaNK cells in cancers was further substantiated by multiplex immunofluorescence staining (Figures 4D and S5G). We also validated the tumor enrichment of TaNK cells (CD56^{dim}CD16^{hi} HSP40^{+}) *in vivo* using flow cytometry. In one intrahepatic cholangiocarcinoma and six HCC samples, TaNK cells were identified, and their proportion in tumor-infiltrating CD56^{dim}CD16^{hi} NK cells was higher than that in the matched adjacent liver tissue, consistent with our scRNA-seq data (Figures 4E and S5H; Table S4).

We then used the pseudotime inference analysis[55] to investigate the dynamic of CD56^{dim}CD16^{hi} NK cells and found that TaNK cells increasingly appeared along the inferred pseudotime of CD56^{dim}CD16^{hi} NK cells and were enriched in the terminal stage (Figure S5I), consistent with results of the RNA velocity analysis. To examine the emerging characteristics of TaNK cells, we fitted the gene expression profile to the pseudotime (STAR Methods). Interestingly, CD56^{dim}CD16^{hi} NK cells showed decreased cytotoxicity and elevated expression of inhibitory receptors and stress genes along the transition process (Figure 4F). Of note, the terminal TaNK cells had the lowest cytotoxicity and highest stress scores among all tumor-infiltrating CD56^{dim} CD16^{hi} NK cell subsets. By contrast, the corresponding c7-NR4A3 enriched in adjacent non-tumor tissues was highly cytotoxic (Figure 4G). By performing the multiplex immunofluorescence staining on several cancer types, we observed that TaNK cells exhibited a lower level of GZMB (Figures 4H, S5J, and S5K). Further confirmed by flow cytometry analyses of liver cancer patients, TaNK cells had lower expression of cytotoxic granules (granzyme B and perforin) and higher expression of inhibitory receptors including CD158a (*KIR2DL1*) and CD158e (*KIR3DL1*) compared with the CD56^{dim}CD16^{hi} HSP40^{-} NK cells at the tumor site (Figures 4I and 4J). These results suggest that TaNK cells may be associated with dysfunctional status. Furthermore, we also observed the differential dynamic tendency of the NR4A nuclear receptor family as the pseudotime increases (Figure 4F). c7-NR4A3 NK cells expressed high levels of *NR4A2* and *NR4A3*, whereas TaNK cells highly expressed *NR4A1* (Figure S5L). Intriguingly, *NR4A1* has been identified as a key mediator of the T cell dysfunction[56] and postulated to contribute to restricting the CAR T cell function in solid tumors.[57] In summary, our data suggested that TaNK cells in tumors might be terminally dysfunctional and potentially play critical roles in the TME.

The association of TaNK cells with unfavorable prognosis and immunotherapy resistance

Since NK cells and CD8^{+} T cells exhibit extensive phenotypic and functional similarities,[4,19] we next examined whether immune checkpoint blockade (ICB) therapies targeting CD8^{+} T cells would also impact NK cells. Within tumor-infiltrating NK cells and CD8^{+} T cells, TaNK cells and exhausted T cells (Tex) exhibited a prominent stress state (Figure 5A), suggesting their involvement in the tumor immune response. Both highly expressed a series of inhibitory receptor molecules; however, they held divergent expression profiles over various immune modulatory genes. Conventional immune checkpoint genes such as *PDCD1* and *CTLA4* were barely expressed on TaNK cells (Figure 5A), implying that they are not direct targets of anti-PD-1/CTLA-4 therapies. Thus, TaNK cells may play different roles from Tex cells in the TME and current ICB treatments.

We observed a remarkable discrepancy in the TaNK cell abundance across cancer types (Figure 5B), and tumor stages made marginal impacts on the proportion of TaNK cells (Figure 5C). Notably, in The Cancer Genome Atlas (TCGA) datasets, the high TaNK cell signal in tumors was associated with poor survival for most cancer types (Figures 5D and 5E). We further applied a deep learning-based model to perform deconvolution and cell composition analyses,[58] finding that the high TaNK cell frequency indicated an unfavorable prognosis of cancer patients (Figures S5M and S5N). We additionally examined whether TaNK cells were linked with ICB treatment response by analyzing scRNA-seq data of pretreatment tumors from previous ICB therapy studies of breast cancer and melanoma[15,59] (Figure 5F). Strikingly, a higher proportion of TaNK cells was observed in nonresponsive patients than responsive ones for both cancer types. Further exploiting published bulk data from a wide variety of cancers including melanoma,[60] lung cancer,[61] and metastatic urothelial carcinoma,[62] we validated that nonresponsive patients exhibited stronger TaNK cell signals than responsive patients (Figure 5G).

We speculate that the long-term infiltration could confer the functional state of TaNK cells in tumors, leading to their ineffective killing of malignant cells. The enrichment of TaNK cells is linked to impaired immune responses against the tumor as well as hyposensitivity to current ICB therapy. Our findings reveal the potential role of TaNK cells in tumors and provide a reference to facilitate the rational design of NK cell-based immunotherapies.

Potential mediators in the TME shaping tumor-infiltrating NK cell functions

To gain insights into the regulatory programs of NK cells in the TME, we utilized CellPhoneDB[63] to probe potential cell-cell interactions between NK and other CD45^{+} immune cells, including T and myeloid cells[13,64] (STAR Methods). Compared with T cells, most myeloid cell types except mast cells exhibited strong potential interactions with CD56^{dim}CD16^{hi} NK subsets (Figure 6A). Of particular interest is that TaNK cells were predicted to regulate multiple myeloid cell types via ANXA1, a protein associated with immunosuppression and induction of macrophage reprogramming during inflammatory responses[65,66] (Figure 6B; Table S6). This implied that dysfunctional NK cells might have the potential to suppress proinflammatory macrophages in the TME. To further clarify the role of NK cell-derived ANXA1 in macrophages, we performed multiplex immunofluorescence staining of tumor samples from lung cancer and liver cancer and identified a population of ANXA1^{+} NK cells (Figure S6A). Macrophages close to ANXA1^{+} NK cells were found to exhibit lower expression levels of activation marker CD86 (Figures 6C, 6D, S6B, and S6C) and higher anti-inflammatory marker transforming growth factor β (TGF-β) than those far from ANXA1^{+} NK cells (Figures 6E, 6F, S6D, and S6E).

Notably, among dendritic cell (DC) subsets, LAMP3^{+} DCs, the mature cDCs recently characterized (also called mregDC),[14,67] showed the strongest interaction potential with CD56^{dim}CD16^{hi} NK cells (Figure 6A). The multiplex immunofluorescence analyses showed that LAMP3^{+} DCs were co-localized with NK cells (Figures 6J and S6L). In addition, interactions between them were predicted to be mediated via the IL-15-IL-15 receptor and NECTIN2-TIGIT interaction axes (Figure 6B). Importantly, LAMP3^{+} DCs expressed the highest level of *IL15*, *PVRL2* (*NECTIN2*), and *PVR* among immune populations at the transcriptome level (Figures 6G and S6F). The high expression of IL-15 in LAMP3^{+} DCs was also demonstrated by flow cytometry (Figures 6H, S6G, and S6H; Table S4). IL-15 has been identified as a homeostasis-related cytokine for the longevity maintenance

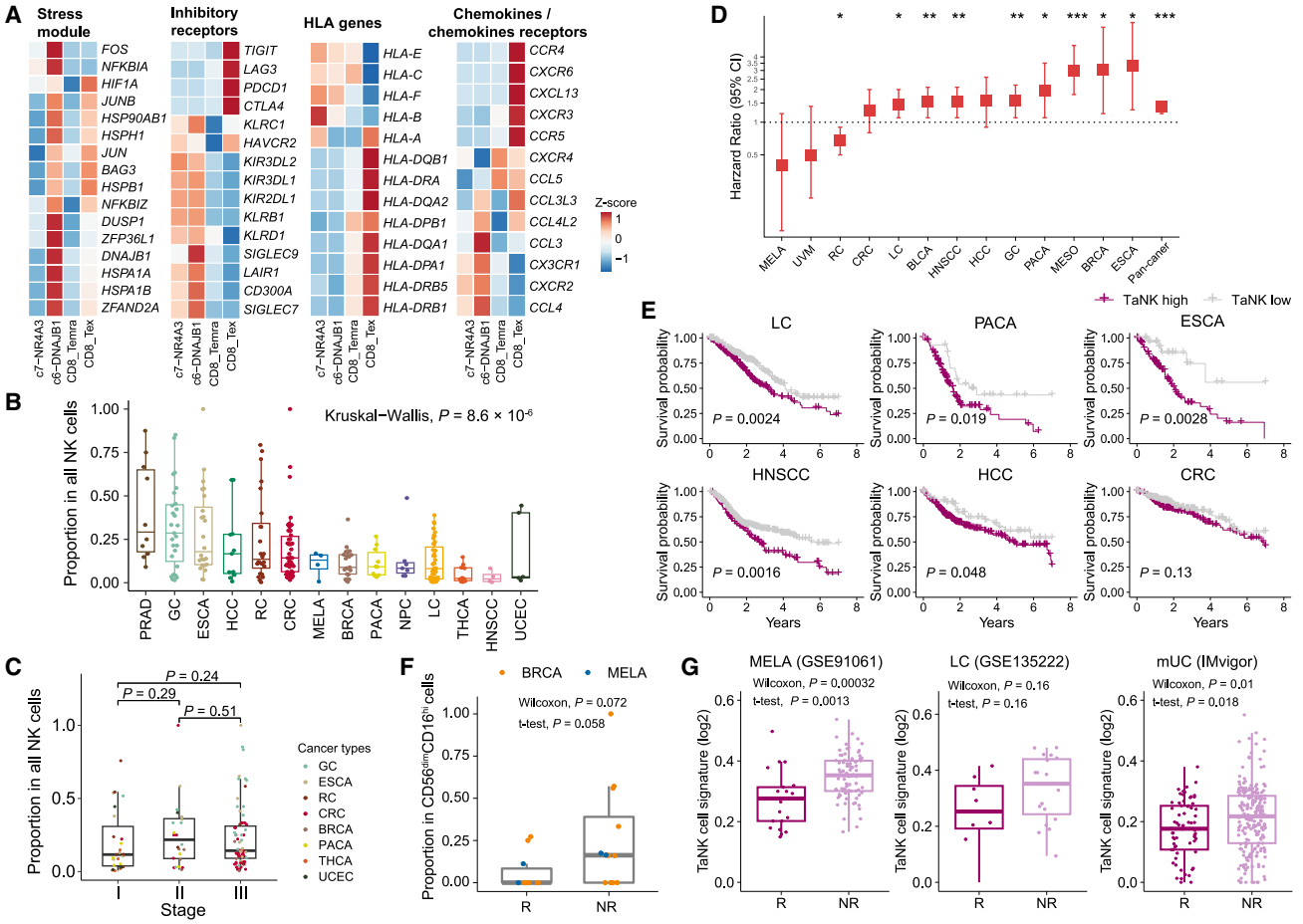


Figure 5. The relationship between TaNK cells and clinical outcomes

(A) Expression patterns of selected genes in specific NK cell and T cell subsets from tumors.

(B) Proportions of CD56^{dim}CD16^{hi} c6-DNAJB1 NK cells in all NK cells across cancer types. Kruskal-Wallis test.

(C) Boxplots comparing the proportion of TaNK cells in all NK cells among different tumor stages. Two-sided unpaired Wilcoxon test.

(D) Forest plot showing the effect of TaNK cells on overall survival. The y axis is scaled by a $\log_{10}$ transformation. *p < 0.05, **p < 0.01, ***p < 0.001. *P* values are adjusted by Benjamini-Hochberg.

(E) Kaplan-Meier plots showing the association of the signature activity of TaNK cells in tumors with prognosis (STAR Methods). +, censored observations; log-rank test.

(F) Boxplots comparing proportions of TaNK cells in CD56^{dim}CD16^{hi} NK cells between non-responders (NRs) and responders (Rs) in ICB therapy datasets (STAR Methods).

(G) Boxplots showing that NR exhibited a higher TaNK cell signal than R in bulk RNA-seq datasets (STAR Methods).

See also Figure S5.

of NK cells and utilized for NK cell infusion and *in vitro* propagation.[4,68] By contrast, TIGIT contributes to suppressing NK cell-mediated immune responses as an inhibitory receptor.[28,69,70] Furthermore, in TCGA datasets, the abundance of LAMP3^{+} DCs was correlated with CD56^{dim}CD16^{hi} NK cells (Figures 6I, S6I, and S6J). We next investigated the specific regulatory process of LAMP3^{+} DCs in tumors and found that tumor-infiltrating LAMP3^{+} DCs exhibited lower expression of *IL15* compared with those in the adjacent non-tumor tissue (Figure S6K), indicating that LAMP3^{+} DCs might have impaired activation effects on CD56^{dim}CD16^{hi} NK cells in the TME. Indeed, NK cells with close physical proximity to LAMP3^{+} DCs expressed granzyme B at a lower level (Figures 6J, 6K, and S6L). Together, our analyses suggested the abnormal regulation of CD56^{dim}CD16^{hi} NK cells by LAMP3^{+} DCs in the TME.

Distinct transcriptome patterns of peripheral blood NK cell subsets

NK cells comprise a sizable proportion of the lymphoid cell compartment in the blood, but their role in the tumor-induced peripheral immune system is relatively opaque. Our atlas contains nine scRNA datasets of blood-derived NK cells from 35 healthy donors (Table S7), enabling us to probe specific alterations of NK cells in the peripheral blood of tumor patients.

We first compared the transcriptome features of circulating NK cells from healthy donors with those from tumor patients.

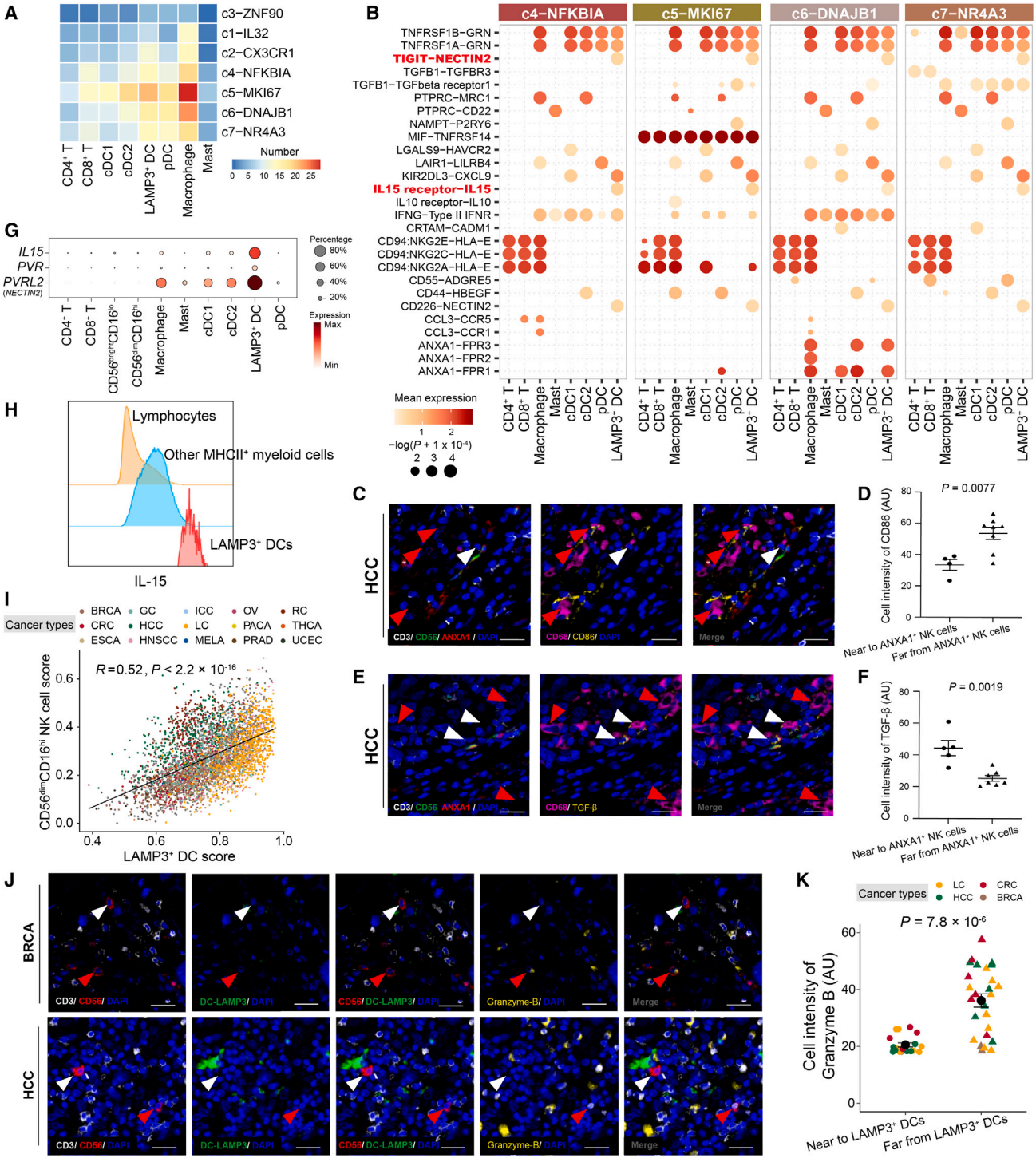


Figure 6. The relationship of LAMP3$^+$ DCs with CD56^{dim}CD16^{hi} NK cells across cancer types

(A) Heatmap showing the number of significant ligand-receptor pairs for each cluster (STAR Methods).

(B) Bubble heatmap showing selected ligand-receptor pairs for interactions of CD56^{dim}CD16^{hi} NK and other immune cell clusters in tumors. Dot size indicates the p value generated by permutation test, and color the mean expression of each ligand-receptor pair.

(C–F) Representative images and quantification of fluorescence intensity of (C and D) CD86 or (E and F) TGF-β expression on macrophages adjacent to and distant from ANXA1$^+$ NK cells in HCC tumors. White arrows represent macrophages adjacent to ANXA1$^+$ NK cells, whereas red arrows far from ANXA1$^+$ NK cells. Scale bar represents 20 μm. Data are represented as mean ± SEM. Two-sided t test.

(legend continued on next page)

Circulating NK cells displayed high similarities among healthy donors from different datasets (Figure S7B). By contrast, for circulating CD56^{bright}CD16^{lo} cells from tumor patients, we observed substantial transcriptome deviations from those in healthy donors; even more dramatic differences were found for circulating CD56^{dim}CD16^{hi} cells in all analyzed cancer types (Figure S7A). Notably, tumor patients exhibited remarkable compositional changes in NK cell subsets, and such patterns appeared to be cancer-type-specific (Figure S7C). For example, the fraction of circulating CD56^{bright}CD16^{lo} c3-CCL3 NK cells increased in colorectal cancer, head and neck squamous cell carcinoma, renal carcinoma, and HCC, but not in other analyzed cancer types.

Next, we focus on the CD56^{dim}CD16^{hi} c8-KLRC2 adaptive NK cells, which were enriched in certain cancer types such as colorectal cancer and gastric cancer (Figure S7D). Adaptive NK cells have been viewed as an attractive source of CAR NK cells, due to their effector characteristics of augmented cytokine response and intrinsic resistance to the immunosuppressive effects.[4] Of particular interest is that these cells specifically expressed MHC class II genes, compared with other circulating NK cells based on our data (Figure S7E). We additionally examined functional shifts of these NK cells in tumor patients and found that compared with those in healthy donors, patient-derived adaptive NK cells had significantly higher expression of functional genes and MHC class II genes (Figures S7F and S7H), implying their highly activated state in tumor patients. We further confirmed the high expression of MHC class II molecules on circulating NK cells from HCC patients compared with those from healthy donors by flow cytometry (Figure S7G). Accordingly, genes upregulated in patient-derived adaptive NK cells were involved in pathways such as positive regulation of the immune effector process (Figure S7I). Taken together, our analyses revealed that circulating NK cells were involved in the systematic change of the peripheral immune environment during tumor progression.

DISCUSSION

In this study, we collected a wide variety of NK cells covering 24 cancer types and systematically explored the unappreciated complexity of tumor-infiltrating NK cells. It is indeed challenging to simultaneously achieve high purity and completeness of the assembled NK cell atlas due to the heterogeneity of datasets from different labs and the similar transcriptional phenotype among NK cells, effector T cells, and other ILCs. Our computational strategy can ensure a pure NK cell atlas with a minimal effect on the downstream proportion comparisons and other analyses. In particular, we confirmed that well-characterized human ILC signature genes were hardly expressed in all identified NK cell subsets (Figure S1H). However, any computational strategy may inevitably leave certain NK cells out, and the effects of our strategy on NK cells derived from different tissues and developmental stages as well as the potential induced bias should be explored further.

Based on our high-dimensional scRNA-seq data on a large scale, we separately explored CD56^{bright}CD16^{lo} and CD56^{dim}CD16^{hi} subsets, which corresponded to the previously reported NK_1 and NK_2,[17] respectively, and revealed that tumor-infiltrating NK cells were structured with heterogeneous populations accompanying phenotypic variation and functional diversity. These NK cells appear to involve in extensive anti-tumor responses such as direct killing of cancerous cells, secretion of proinflammatory cytokines, and recruitment of other immune components (Figures 7A and 7B). NK cell populations exhibited substantial cancer-type preferences, which were associated with both intrinsic organ properties and factors from the TME. Particularly, the reduction of CD56^{dim}CD16^{hi} NK cells in tumors observed for most cancer types represents a potential mechanism of tumor escaping from the NK cell immunosurveillance. Previously, higher NK cell activity has been associated with the response to trastuzumab, an anti-HER2 antibody, for breast cancer patients,[71] which inspires us that the function of tumor-infiltrating NK cells may affect the clinical efficacy of antibody drugs through antibody-dependent cellular cytotoxicity.

Facilitating NK cell infiltration in solid malignancies has been a key focus of developing therapeutic NK cell products. We recognized *RGS1* as a key marker of tissue-infiltrating NK cells at the transcriptome level (Figure 7C). Analogously, *RGS1* is highly correlated with T cell function and tissue residency,[72,73] but the functional effect on NK cells is still unclear and needs further investigation. It is still difficult to accurately discriminate whether *RGS1* marks all NK cells entering tissues or only *bona fide* tissue-resident NK cells.

We identified a tumor-enriched NK cell subset in potentially dysfunctional states, named TaNK cells (Figure 7D). Analogous to the exhaustion of T cells,[24,74] the dysfunction of NK cells suggests the impairment of natural cytotoxicity.[75,76] Notably, although TaNK cells are not always the dominant component of tumor-infiltrating NK cells in all cancers, the enrichment of TaNK cells in tumors is robust in various cancers including those dominated by CD56^{bright}CD16^{lo} NK cells. The high abundance of these cells was related to unfavorable prognosis and immunotherapy resistance in multiple cancer types, implicating their roles in clinical settings. We speculated that TaNK cell enrichment may reflect or affect the tumor immune responses in the

(G) The expression patterns of *IL15*, *PVR*, and *PVRL2* in major immune populations based on scRNA-seq data. Dot size represents percentage, and color the mean expression.
(H) A representative plot of IL-15 expression in the tumor tissue of an HCC patient, analyzed by flow cytometry.
(I) Scatterplots showing correlations of CD56^{dim}CD16^{hi} NK cells with LAMP3^{+} DCs in the TCGA dataset (STAR Methods). Pearson correlation test.
(J) Multiplex immunofluorescence staining to show the co-localization of LAMP3^{+} DCs (DC-LAMP3^{+}) and NK cells (CD3^{-}CD56^{+}). The scale bar represents 20 μm.
(K) Quantification of the fluorescence intensity for granzyme B in NK cells near to LAMP3^{+} DCs (n = 19) or far from LAMP3^{+} DCs (n = 26) from (D) using the Halo v3.3 image analysis platform (Indica Labs). Data are represented as mean ± SEM. Two-sided unpaired Wilcoxon test.
(A), (B), and (G) are plotted using our newly generated data.
See also Figure S6 and Tables S4 and S6.

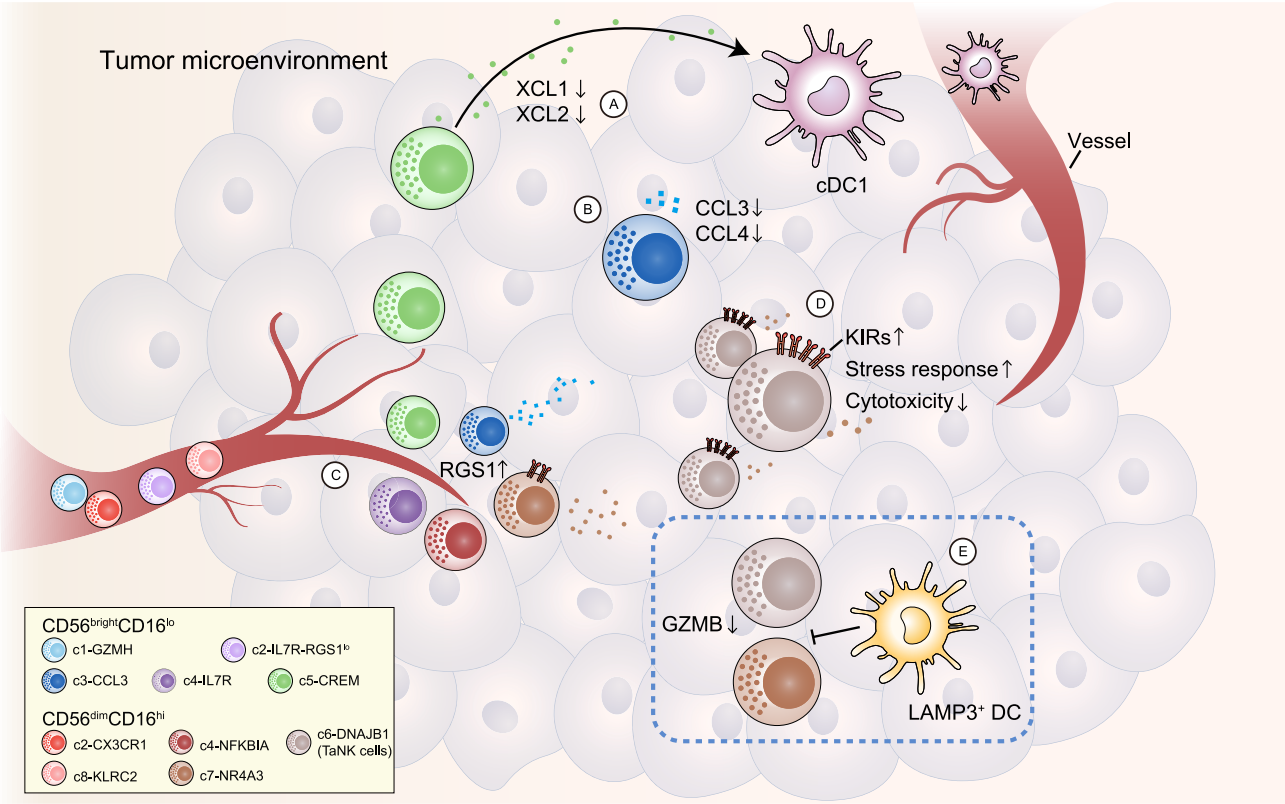


Figure 7. Summary of NK cell features and dynamics in this study

Sketch map showing the phenotypic shifts of different tumor-infiltrating NK cell subsets

(A) CD56^{bright}CD16^{lo} c5-CREM NK cells downregulated *XCL2* and *XCL1* in tumors, possibly leading to attenuated cDC1 recruitment.

(B) CD56^{bright}CD16^{lo} c3-CCL3 NK cells exhibited reduced cytokine production in tumors, including *CCL3* and *CCL4*.

(C) In contrast to circulating NK cells, tumor-infiltrating NK cells consistently expressed *RGS1*.

(D) c6-DNAJB1 NK cells, exhibiting higher expression of KIRs, elevated stress response, and reduced cytotoxicity, were considered the potentially dysfunctional state of tumor-infiltrating CD56^{dim}CD16^{hi} NK cells.

(E) Cellular interactions with LAMP3^{+} DCs shape tumor-infiltrating CD56^{dim}CD16^{hi} NK cell functions.

TME, although those cells may be not the direct target of ICB therapies. Further exploration of larger cohorts is expected to corroborate the functional roles of TaNK cells in both tumor progression and ICB treatment. We also explored the mechanism of how these NK cells are affected in tumors. We identified internal regulators for TaNK cells, including *NR4A1*, which has been reported to inhibit effector T cell differentiation and play critical roles in T cell exhaustion.[56,57] External factors from other cell populations in the TME including myeloid-derived suppressor cells, Tregs, and tumor-associated macrophages can also suppress the anti-tumor function of NK cells.[4,77] Our analyses and experimental evidence support myeloid cells as core mediators of NK cells. Specifically, LAMP3^{+} DCs can serve as a crucial regulator and potentially inhibit CD56^{dim}CD16^{hi} NK cell function in the TME, whereas further direct functional validation is still imperative (Figure 7E). Overall, our results provide cues to maintain the anti-tumor activity of NK cells *in vivo* via searching for cell-intrinsic and TME-associated factors.

Emerging studies have unveiled the alteration of systemic immunity during tumor progression.[78] Our study revealed a CD56^{dim}CD16^{hi} c8-KLRC2 adaptive NK cell population that is characterized by the upregulation of certain proinflammatory cytokines and MHC class II genes in the peripheral blood of tumor patients. Although the mechanism underlying their phenotypic shifts in tumor patients remains unclear, a possible explanation is that the malignancy-induced release of cytokines might result in an inflammatory context of peripheral blood and then activate NK cells accompanied by MHC class II molecule expressions. We expect future studies to pay more attention to the role of adaptive NK cells in the systemic immunity and elucidate the connection between circulating NK cells and intra-tumoral immune responses.

In summary, our comprehensive analyses enhance the current understanding of NK cells from a pan-cancer view, illuminating insights into NK cell population structures as well as tumor-induced local and systemic NK cell responses. To facilitate the usage of our data for the wide research community, an interactive portal (http://pan-nk.cancer-pku.cn/) has been developed for analyzing and visualizing our single-cell data. We envision that our large-scale data can further promote the

application of NK cell-based immunotherapy to more cancer patients.

Limitations of the study

We lack detailed clinical information on patients from public datasets. Further investigating the functional and compositional variances of NK cells in specific contexts including different tumor stages, metastatic positions, and expression status of MHC class I class genes, is challenging in our study. Epigenetic characteristics of TaNK cells are still unexplored, and the functional validation for these cells is limited due to the difficulty of *in vitro* culturing. Estimations about the clinical significance of TaNK cells are not conditioned on other possibly confounding immune cell types. Nevertheless, this first pan-cancer single-cell NK cell atlas captures the complex tumor-infiltrating NK cell characteristics and informs future developmental directions of NK cell-based immunotherapy strategies that can be personalized for the maximal clinical benefit.

STAR★METHODS

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

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Human participants
- METHOD DETAILS
  - Sample collection and flow cytometry
  - Multiplex immunofluorescent staining
  - Single cell RNA library preparation and sequencing
  - Single-cell RNA-seq data preprocessing
  - Batch effect correction and unsupervised clustering
  - The proportion and tissue distribution of NK cell subsets
  - Calculation of signature score
  - Definition of cytotoxicity, inflammatory and stress gene sets
  - Definition of HLA-dependent and -independent receptors
  - Hierarchical clustering of cancer types
  - Specificity and sensitivity assessment
  - RNA velocity analysis
  - Pseudotime trajectory inference
  - SCENIC analysis
  - Cell-cell interaction analysis by CellPhoneDB
  - Analysis of immunotherapy datasets
  - TCGA RNA-seq data analysis
  - Survival analysis
  - Comparison with human peripheral blood for NK cells
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Statistical analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cell.2023.07.034.

ACKNOWLEDGMENTS

We thank Y. Zhang, Y. He, S. Cheng, Q. Zhang, C. Li, Z. Li, and X. Chen for discussions. Part of the analysis in this study was performed on the High Performance Computing Platform of the Center for Life Science (Peking University). This project was supported by funding from the National Natural Science Foundation of China (81988101, 31991171, 91959000, 592259205, 62203019, and 92159305) and Beijing Municipal Science and Technology Commission (Z221100007022002).

AUTHOR CONTRIBUTIONS

Z.Z., D.W., and L.Z. designed this study. F.T., J.L., and S.Q. performed data analysis. Z.Z., L.Z., H.P., and Z.T. contributed to the experimental design. L.Q., L.Z., and Y.B. performed the experiments. Y.M., K.Y., and J.L. provided experimental methods. D.L. and J.P. provided clinical samples. F.T., J.L, L.Z., H.P., D.W., and Z.Z. wrote the manuscript, with all authors contributing to providing feedback.

DECLARATION OF INTERESTS

Z.Z. is a founder of Analytical Bioscience and also serves on the Advisory Board of *Cell*. All financial interests are unrelated to this study.

Received: November 15, 2022
Revised: March 28, 2023
Accepted: July 28, 2023
Published: August 21, 2023

REFERENCES

1. Waldman, A.D., Fritz, J.M., and Lenardo, M.J. (2020). A guide to cancer immunotherapy: from T cell basic science to clinical practice. Nat. Rev. Immunol. *20*, 651–668. https://doi.org/10.1038/s41577-020-0306-5.
2. Huntington, N.D., Vosshenrich, C.A.J., and Di Santo, J.P. (2007). Developmental pathways that generate natural-killer-cell diversity in mice and humans. Nat. Rev. Immunol. *7*, 703–714. https://doi.org/10.1038/nri2154.
3. Freud, A.G., and Caligiuri, M.A. (2006). Human natural killer cell development. Immunol. Rev. *214*, 56–72. https://doi.org/10.1111/j.1600-065X.2006.00451.x.
4. Myers, J.A., and Miller, J.S. (2021). Exploring the NK cell platform for cancer immunotherapy. Nat. Rev. Clin. Oncol. *18*, 85–100. https://doi.org/10.1038/s41571-020-0426-7.
5. Li, Y., Hermanson, D.L., Moriarity, B.S., and Kaufman, D.S. (2018). Human iPSC-derived natural killer cells engineered with chimeric antigen receptors enhance anti-tumor activity. Cell Stem Cell *23*, 181–192.e5. https://doi.org/10.1016/j.stem.2018.06.002.
6. Zhang, L., Meng, Y., Feng, X., and Han, Z. (2022). CAR-NK cells for cancer immunotherapy: from bench to bedside. Biomark. Res. *10*, 12. https://doi.org/10.1186/s40364-022-00364-6.
7. Habif, G., Crinier, A., André, P., Vivier, E., and Narni-Mancinelli, E. (2019). Targeting natural killer cells in solid tumors. Cell. Mol. Immunol. *16*, 415–422. https://doi.org/10.1038/s41423-019-0224-2.
8. Melaiu, O., Lucarini, V., Cifaldi, L., and Fruci, D. (2019). Influence of the tumor microenvironment on NK cell function in solid tumors. Front. Immunol. *10*, 3038.
9. Lanier, L.L., Le, A.M., Civin, C.I., Loken, M.R., and Phillips, J.H. (1986). The relationship of CD16 (Leu-11) and Leu-19 (NKH-1) antigen expression on human peripheral blood NK cells and cytotoxic T lymphocytes. J. Immunol. *136*, 4480–4486.

10. Nagler, A., Lanier, L.L., Cwirla, S., and Phillips, J.H. (1989). Comparative studies of human FcRIII-positive and negative natural killer cells. J. Immunol. *143*, 3183–3191.

11. Jacobs, R., Hintzen, G., Kemper, A., Beul, K., Kempf, S., Behrens, G., Sykora, K.W., and Schmidt, R.E. (2001). CD56bright cells differ in their KIR repertoire and cytotoxic features from CD56dim NK cells. Eur. J. Immunol. *31*, 3121–3127. https://doi.org/10.1002/1521-4141(2001010)31:10<3121::AID-IMMU3121>3.0.CO;2-4.

12. Cooper, M.A., Fehniger, T.A., Turner, S.C., Chen, K.S., Ghaheri, B.A., Ghayur, T., Carson, W.E., and Caligiuri, M.A. (2001). Human natural killer cells: a unique innate immunoregulatory role for the CD56(bright) subset. Blood *97*, 3146–3151. https://doi.org/10.1182/blood.V97.10.3146.

13. Zheng, L., Qin, S., Si, W., Wang, A., Xing, B., Gao, R., Ren, X., Wang, L., Wu, X., Zhang, J., et al. (2021). Pan-cancer single-cell landscape of tumor-infiltrating T cells. Science *374*, abe6474. https://doi.org/10.1126/science.abe6474.

14. Zhang, Q., He, Y., Luo, N., Patel, S.J., Han, Y., Gao, R., Modak, M., Carotta, S., Haslinger, C., Kind, D., et al. (2019). Landscape and dynamics of single immune cells in hepatocellular carcinoma. Cell *179*, 829–845.e20. https://doi.org/10.1016/j.cell.2019.10.003.

15. de Andrade, L.F., Lu, Y., Luoma, A., Ito, Y., Pan, D., Pyrdol, J.W., Yoon, C.H., Yuan, G.C., and Wucherpfennig, K.W. (2019). Discovery of specialized NK cell populations infiltrating human melanoma metastases. JCI Insight *4*. https://doi.org/10.1172/jci.insight.133103.

16. Dogra, P., Rancan, C., Ma, W., Toth, M., Senda, T., Carpenter, D.J., Kubota, M., Matsumoto, R., Thapa, P., Szabo, P.A., et al. (2020). Tissue determinants of human NK cell development, function, and residence. Cell *180*, 749–763.e13. https://doi.org/10.1016/j.cell.2020.01.022.

17. Crinier, A., Milpied, P., Escalière, B., Piperoglou, C., Galluso, J., Balsamo, A., Spinelli, L., Cervera-Marzal, I., Ebbo, M., Girard-Madoux, M., et al. (2018). High-dimensional single-cell analysis identifies organ-specific signatures and conserved NK cell subsets in humans and mice. Immunity *49*, 971–986.e5. https://doi.org/10.1016/j.immuni.2018.09.009.

18. Björkström, N.K., Ljunggren, H.G., and Michaëlsson, J. (2016). Emerging insights into natural killer cells in human peripheral tissues. Nat. Rev. Immunol. *16*, 310–320. https://doi.org/10.1038/nri.2016.34.

19. Abel, A.M., Yang, C., Thakar, M.S., and Malarkannan, S. (2018). Natural killer cells: development, maturation, and clinical utilization. Front. Immunol. *9*, 1869.

20. Vivier, E., Artis, D., Colonna, M., Diefenbach, A., Di Santo, J.P., Eberl, G., Koyasu, S., Locksley, R.M., McKenzie, A.N.J., Mebius, R.E., et al. (2018). Innate lymphoid cells: 10 years on. Cell *174*, 1054–1066. https://doi.org/10.1016/j.cell.2018.07.017.

21. Jacquelot, N., Seillet, C., Vivier, E., and Belz, G.T. (2022). Innate lymphoid cells and cancer. Nat. Immunol. *23*, 371–379. https://doi.org/10.1038/s41590-022-01127-z.

22. Dhatchinamoorthy, K., Colbert, J.D., and Rock, K.L. (2021). Cancer immune evasion through loss of MHC class I antigen presentation. Front. Immunol. *12*, 636568.

23. Freeman, A.J., Vervoort, S.J., Ramsbottom, K.M., Kelly, M.J., Michie, J., Pijpers, L., Johnstone, R.W., Kearney, C.J., and Oliaro, J. (2019). Natural killer cells suppress T cell-associated tumor immune evasion. Cell Rep. *28*, 2784–2794.e5. https://doi.org/10.1016/j.celrep.2019.08.017.

24. McLane, L.M., Abdel-Hakeem, M.S., and Wherry, E.J. (2019). CD8 T cell exhaustion during chronic viral infection and cancer. Annu. Rev. Immunol. *37*, 457–495. https://doi.org/10.1146/annurev-immunol-041015-055318.

25. Sun, C., Sun, H., Zhang, C., and Tian, Z. (2015). NK cell receptor imbalance and NK cell dysfunction in HBV infection and hepatocellular carcinoma. Cell. Mol. Immunol. *12*, 292–302. https://doi.org/10.1038/cmi.2014.91.

26. Cai, L., Zhang, Z., Zhou, L., Wang, H., Fu, J., Zhang, S., Shi, M., Zhang, H., Yang, Y., Wu, H., et al. (2008). Functional impairment in circulating and intrahepatic NK cells and relative mechanism in hepatocellular carcinoma patients. Clin. Immunol. *129*, 428–437. https://doi.org/10.1016/j.clim.2008.08.012.

27. Bi, J., and Tian, Z. (2019). NK cell dysfunction and checkpoint immunotherapy. Front. Immunol. *10*, 1999.

28. Zhang, Q., Bi, J., Zheng, X., Chen, Y., Wang, H., Wu, W., Wang, Z., Wu, Q., Peng, H., Wei, H., et al. (2018). Blockade of the checkpoint receptor TIGIT prevents NK cell exhaustion and elicits potent anti-tumor immunity. Nat. Immunol. *19*, 723–732. https://doi.org/10.1038/s41590-018-0132-0.

29. Judge, S.J., Dunai, C., Aguilar, E.G., Vick, S.C., Sturgill, I.R., Khuat, L.T., Stoffel, K.M., Van Dyke, J., Longo, D.L., Darrow, M.A., et al. (2020). Minimal PD-1 expression in mouse and human NK cells under diverse conditions. J. Clin. Invest. *130*, 3051–3068. https://doi.org/10.1172/JCI133353.

30. Björklund, Å.K., Forkel, M., Picelli, S., Konya, V., Theorell, J., Friberg, D., Sandberg, R., and Mjösberg, J. (2016). The heterogeneity of human CD127+ innate lymphoid cells revealed by single-cell RNA sequencing. Nat. Immunol. *17*, 451–460. https://doi.org/10.1038/ni.3368.

31. Smith, S.L., Kennedy, P.R., Stacey, K.B., Worboys, J.D., Yarwood, A., Seo, S., Solloa, E.H., Mistretta, B., Chatterjee, S.S., Gunaratne, P., et al. (2020). Diversity of peripheral blood human NK cells identified by single-cell RNA sequencing. Blood Adv. *4*, 1388–1406. https://doi.org/10.1182/bloodadvances.2019000699.

32. Schlums, H., Cichocki, F., Tesi, B., Theorell, J., Beziat, V., Holmes, T.D., Han, H., Chiang, S.C.C., Foley, B., Mattsson, K., et al. (2015). Cytomegalovirus infection drives adaptive epigenetic diversification of NK cells with altered signaling and effector function. Immunity *42*, 443–456. https://doi.org/10.1016/j.immuni.2015.02.008.

33. Costa-García, M., Ataya, M., Moraru, M., Vilches, C., López-Botet, M., and Muntasell, A. (2019). Human cytomegalovirus antigen presentation by HLA-DR+ NKG2C+ adaptive NK cells specifically activates polyfunctional effector memory CD4+ T lymphocytes. Front. Immunol. *10*, 687.

34. Liu, B., Li, C., Li, Z., Wang, D., Ren, X., and Zhang, Z. (2020). An entropy-based metric for assessing the purity of single cell populations. Nat. Commun. *11*, 3155. https://doi.org/10.1038/s41467-020-16904-3.

35. Bade, B., Boettcher, H.E., Lohrmann, J., Hink-Schauer, C., Bratke, K., Jenne, D.E., Virchow, J.C., Jr., and Luttmann, W. (2005). Differential expression of the granzymes A, K and M and perforin in human peripheral blood lymphocytes. Int. Immunol. *17*, 1419–1428. https://doi.org/10.1093/intimm/dxh320.

36. Castellino, F., Huang, A.Y., Altan-Bonnet, G., Stoll, S., Scheinecker, C., and Germain, R.N. (2006). Chemokines enhance immunity by guiding naive CD8+ T cells to sites of CD4+ T cell–dendritic cell interaction. Nature *440*, 890–895. https://doi.org/10.1038/nature04651.

37. Böttcher, J.P., Bonavita, E., Chakravarty, P., Blees, H., Cabeza-Cabrerizo, M., Sammicheli, S., Rogers, N.C., Sahai, E., Zelenay, S., and Reis e Sousa, C. (2018). NK cells stimulate recruitment of cDC1 into the tumor microenvironment promoting cancer immune control. Cell *172*, 1022–1037.e14. https://doi.org/10.1016/j.cell.2018.01.004.

38. Pazina, T., MacFarlane, A.W., Bernabei, L., Dulaimi, E., Kotcher, R., Yam, C., Bezman, N.A., Robbins, M.D., Ross, E.A., Campbell, K.S., et al. (2021). Alterations of NK cell phenotype in the disease course of multiple myeloma. Cancers *13*. https://doi.org/10.3390/cancers13020226.

39. Ziblat, A., Iraolagoitia, X.L.R., Nuñez, S.Y., Torres, N.I., Secchiari, F., Sierra, J.M., Spallanzani, R.G., Rovegno, A., Secin, F.P., Fuertes, M.B., et al. (2021). Circulating and tumor-infiltrating NK cells from clear cell renal cell carcinoma patients exhibit a predominantly inhibitory phenotype characterized by overexpression of CD85j, CD45, CD48 and PD-1. Front. Immunol. *12*, 681615. https://doi.org/10.3389/fimmu.2021.681615.

40. Wang, H., Zhang, X., Wald, D., Gong, S., Wang, W., and Wang, J. (2022). Downregulation of CD16 surface expression on NK cells is a favorable prognostic factor in acute myeloid leukemia. Blood *140*, 11834–11835. Suppl 1. https://doi.org/10.1182/blood-2022-168947.

41. Le, D.T., Huynh, T.R., Burt, B., Van Buren, G., Abeynaike, S.A., Zalfa, C., Nikzad, R., Kheradmand, F., Tyner, J.J., and Paust, S. (2021). Natural killer

cells and cytotoxic T lymphocytes are required to clear solid tumor in a patient-derived xenograft. JCI Insight *6*. https://doi.org/10.1172/jci.insight.140116.

42. Mele, D., Pessino, G., Trisolini, G., Luchena, A., Benazzo, M., Morbini, P., Mantovani, S., Oliviero, B., Mondelli, M.U., and Varchetta, S. (2022). Impaired intratumoral natural killer cell function in head and neck carcinoma. Front. Immunol. *13*, 997806. https://doi.org/10.3389/fimmu.2022.997806.
43. Scoville, S.D., Nalin, A.P., Chen, L., Chen, L., Zhang, M.H., McConnell, K., Beceiro Casas, S., Ernst, G., Traboulsi, A.A.-R., Hashi, N., et al. (2018). Human AML activates the aryl hydrocarbon receptor pathway to impair NK cell development and function. Blood *132*, 1792–1804. https://doi.org/10.1182/blood-2018-03-838474.
44. Chretien, A.S., Fauriat, C., Orlanducci, F., Galseran, C., Rey, J., Bouvier Borg, G., Gautherot, E., Granjeaud, S., Hamel-Broza, J.F., Demerle, C., et al. (2017). Natural killer defective maturation is associated with adverse clinical outcome in patients with acute myeloid leukemia. Front. Immunol. *8*, 573. https://doi.org/10.3389/fimmu.2017.00573.
45. Freud, A.G., Mundy-Bosse, B.L., Yu, J., and Caligiuri, M.A. (2017). The broad spectrum of human natural killer cell diversity. Immunity *47*, 820–833. https://doi.org/10.1016/j.immuni.2017.10.008.
46. Sojka, D.K., Plougastel-Douglas, B., Yang, L., Pak-Wittel, M.A., Artyomov, M.N., Ivanova, Y., Zhong, C., Chase, J.M., Rothman, P.B., Yu, J., et al. (2014). Tissue-resident natural killer (NK) cells are cell lineages distinct from thymic and conventional splenic NK cells. eLife *3*, e01659. https://doi.org/10.7554/eLife.01659.
47. Peng, H., Jiang, X., Chen, Y., Sojka, D.K., Wei, H., Gao, X., Sun, R., Yokoyama, W.M., and Tian, Z. (2013). Liver-resident NK cells confer adaptive immunity in skin-contact inflammation. J. Clin. Invest. *123*, 1444–1456. https://doi.org/10.1172/JCI66381.
48. Druey, K.M., Blumer, K.J., Kang, V.H., and Kehrl, J.H. (1996). Inhibition of G-protein-mediated MAP kinase activation by a new mammalian gene family. Nature *379*, 742–746. https://doi.org/10.1038/379742a0.
49. Aibar, S., González-Blas, C.B., Moerman, T., Huynh-Thu, V.A., Imrichova, H., Hulselmans, G., Rambow, F., Marine, J.C., Geurts, P., Aerts, J., et al. (2017). SCENIC: single-cell regulatory network inference and clustering. Nat. Methods *14*, 1083–1086. https://doi.org/10.1038/nmeth.4463.
50. Safford, M., Collins, S., Lutz, M.A., Allen, A., Huang, C.T., Kowalski, J., Blackford, A., Horton, M.R., Drake, C., Schwartz, R.H., et al. (2005). Egr-2 and Egr-3 are negative regulators of T cell activation. Nat. Immunol. *6*, 472–480. https://doi.org/10.1038/ni1193.
51. Li, S., Miao, T., Sebastian, M., Bhullar, P., Ghaffari, E., Liu, M., Symonds, A.L.J., and Wang, P. (2012). The transcription factors Egr2 and Egr3 are essential for the control of inflammation and antigen-induced proliferation of B and T cells. Immunity *37*, 685–696. https://doi.org/10.1016/j.immuni.2012.08.001.
52. Cao, Z., Sun, X., Icli, B., Wara, A.K., and Feinberg, M.W. (2010). Role of Krüppel-like factors in leukocyte development, function, and disease. Blood *116*, 4404–4414. https://doi.org/10.1182/blood-2010-05-285353.
53. La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastriti, M.E., Lönnerberg, P., Furlan, A., et al. (2018). RNA velocity of single cells. Nature *560*, 494–498. https://doi.org/10.1038/s41586-018-0414-6.
54. Bergen, V., Lange, M., Peidli, S., Wolf, F.A., and Theis, F.J. (2020). Generalizing RNA velocity to transient cell states through dynamical modeling. Nat. Biotechnol. *38*, 1408–1414. https://doi.org/10.1038/s41587-020-0591-3.
55. Haghverdi, L., Büttner, M., Wolf, F.A., Buettner, F., and Theis, F.J. (2016). Diffusion pseudotime robustly reconstructs lineage branching. Nat. Methods *13*, 845–848. https://doi.org/10.1038/nmeth.3971.
56. Liu, X., Wang, Y., Lu, H., Li, J., Yan, X., Xiao, M., Hao, J., Alekseev, A., Khong, H., Chen, T., et al. (2019). Genome-wide analysis identifies NR4A1 as a key mediator of T cell dysfunction. Nature *567*, 525–529. https://doi.org/10.1038/s41586-019-0979-8.
57. Chen, J., López-Moyado, I.F., Seo, H., Lio, C.-W.J., Hempleman, L.J., Sekiya, T., Yoshimura, A., Scott-Browne, J.P., and Rao, A. (2019). NR4A transcription factors limit CAR T cell function in solid tumours. Nature *567*, 530–534. https://doi.org/10.1038/s41586-019-0985-x.
58. Menden, K., Marouf, M., Oller, S., Dalmia, A., Magruder, D.S., Kloiber, K., Heutink, P., and Bonn, S. (2020). Deep learning–based cell composition analysis from tissue expression profiles. Sci. Adv. *6*, eaba2619. https://doi.org/10.1126/sciadv.aba2619.
59. Zhang, Y., Chen, H., Mo, H., Hu, X., Gao, R., Zhao, Y., Liu, B., Niu, L., Sun, X., Yu, X., et al. (2021). Single-cell analyses reveal key immune cell subsets associated with response to PD-L1 blockade in triple-negative breast cancer. Cancer Cell *39*, 1578–1593.e8. https://doi.org/10.1016/j.ccell.2021.09.010.
60. Riaz, N., Havel, J.J., Makarov, V., Desrichard, A., Urba, W.J., Sims, J.S., Hodi, F.S., Martín-Algarra, S., Mandal, R., Sharfman, W.H., et al. (2017). Tumor and microenvironment evolution during immunotherapy with nivolumab. Cell *171*, 934–949.e16. https://doi.org/10.1016/j.cell.2017.09.028.
61. Jung, H., Kim, H.S., Kim, J.Y., Sun, J.M., Ahn, J.S., Ahn, M.J., Park, K., Esteller, M., Lee, S.H., and Choi, J.K. (2019). DNA methylation loss promotes immune evasion of tumours with high mutation and copy number load. Nat. Commun. *10*, 4278. https://doi.org/10.1038/s41467-019-12159-9.
62. Balar, A.V., Galsky, M.D., Rosenberg, J.E., Powles, T., Petrylak, D.P., Bellmunt, J., Loriot, Y., Necchi, A., Hoffman-Censits, J., Perez-Gracia, J.L., et al. (2017). Atezolizumab as first-line treatment in cisplatin-ineligible patients with locally advanced and metastatic urothelial carcinoma: a single-arm, multicentre, phase 2 trial. Lancet *389*, 67–76. https://doi.org/10.1016/S0140-6736(16)32455-2.
63. Efremova, M., Vento-Tormo, M., Teichmann, S.A., and Vento-Tormo, R. (2020). CellPhoneDB: inferring cell–cell communication from combined expression of multi-subunit ligand–receptor complexes. Nat. Protoc. *15*, 1484–1506. https://doi.org/10.1038/s41596-020-0292-x.
64. Cheng, S., Li, Z., Gao, R., Xing, B., Gao, Y., Yang, Y., Qin, S., Zhang, L., Ouyang, H., Du, P., et al. (2021). A pan-cancer single-cell transcriptional atlas of tumor infiltrating myeloid cells. Cell *184*, 792–809.e23. https://doi.org/10.1016/j.cell.2021.01.010.
65. McArthur, S., Juban, G., Gobbetti, T., Desgeorges, T., Theret, M., Gondin, J., Toller-Kawahisa, J.E., Reutelingsperger, C.P., Chazaud, B., Perretti, M., et al. (2020). Annexin A1 drives macrophage skewing to accelerate muscle regeneration through AMPK activation. J. Clin. Invest. *130*, 1156–1167. https://doi.org/10.1172/JCI124635.
66. Rhys, H.I., Dell'Accio, F., Pitzalis, C., Moore, A., Norling, L.V., and Perretti, M. (2018). Neutrophil microvesicles from healthy control and rheumatoid arthritis patients prevent the inflammatory activation of macrophages. EBiomedicine *29*, 60–69. https://doi.org/10.1016/j.ebiom.2018.02.003.
67. Maier, B., Leader, A.M., Chen, S.T., Tung, N., Chang, C., LeBerichel, J., Chudnovskiy, A., Maskey, S., Walker, L., Finnigan, J.P., et al. (2020). A conserved dendritic-cell regulatory program limits antitumour immunity. Nature *580*, 257–262. https://doi.org/10.1038/s41586-020-2134-y.
68. Ranson, T., Vosshenrich, C.A.J., Corcuff, E., Richard, O., Müller, W., and Di Santo, J.P. (2003). IL-15 is an essential mediator of peripheral NK-cell homeostasis. Blood *101*, 4887–4893. https://doi.org/10.1182/blood-2002-11-3392.
69. Stanietsky, N., Simic, H., Arapovic, J., Toporik, A., Levy, O., Novik, A., Levine, Z., Beiman, M., Dassa, L., Achdout, H., et al. (2009). The interaction of TIGIT with PVR and PVRL2 inhibits human NK cell cytotoxicity. Proc. Natl. Acad. Sci. USA *106*, 17858–17863. https://doi.org/10.1073/pnas.0903474106.
70. Bi, J., Zhang, Q., Liang, D., Xiong, L., Wei, H., Sun, R., and Tian, Z. (2014). T-cell Ig and ITIM domain regulates natural killer cell activation in murine acute viral hepatitis. Hepatology *59*, 1715–1725. https://doi.org/10.1002/hep.26968.

71. Beano, A., Signorino, E., Evangelista, A., Brusa, D., Mistrangelo, M., Polimeni, M.A., Spadi, R., Donadio, M., Ciuffreda, L., and Matera, L. (2008). Correlation between NK function and response to trastuzumab in metastatic breast cancer patients. J. Transl. Med. *6*, 25. https://doi.org/10.1186/1479-5876-6-25.

72. Huang, D., Chen, X., Zeng, X., Lao, L., Li, J., Xing, Y., Lu, Y., Ouyang, Q., Chen, J., Yang, L., et al. (2021). Targeting regulator of G protein signaling 1 in tumor-specific T cells enhances their trafficking to breast cancer. Nat. Immunol. *22*, 865–879. https://doi.org/10.1038/s41590-021-00939-9.

73. Djenidi, F., Adam, J., Goubar, A., Durgeau, A., Meurice, G., de Montpréville, V., Validire, P., Besse, B., and Mami-Chouaib, F. (2015). CD8+CD103+ tumor–infiltrating lymphocytes are tumor-specific tissue-resident memory T cells and a prognostic factor for survival in lung cancer patients. J. Immunol. *194*, 3475–3486. https://doi.org/10.4049/jimmunol.1402711.

74. Zheng, C., Fass, J.N., Shih, Y.P., Gunderson, A.J., Sanjuan Silva, N., Huang, H., Bernard, B.M., Rajamanickam, V., Slagel, J., Bifulco, C.B., et al. (2022). Transcriptomic profiles of neoantigen-reactive T cells in human gastrointestinal cancers. Cancer Cell *40*, 410–423.e7. https://doi.org/10.1016/j.ccell.2022.03.005.

75. Judge, S.J., Murphy, W.J., and Canter, R.J. (2020). Characterizing the dysfunctional NK cell: assessing the clinical relevance of exhaustion, anergy, and senescence. Front. Cell. Infect. Microbiol. *10*, 49.

76. Sun, C., Xu, J., Huang, Q., Huang, M., Wen, H., Zhang, C., Wang, J., Song, J., Zheng, M., Sun, H., et al. (2017). High NKG2A expression contributes to NK cell exhaustion and predicts a poor prognosis of patients with liver cancer. OncoImmunology *6*, e1264562. https://doi.org/10.1080/2162402X.2016.1264562.

77. Sarhan, D., Hippen, K.L., Lemire, A., Hying, S., Luo, X., Lenvik, T., Curtsinger, J., Davis, Z., Zhang, B., Cooley, S., et al. (2018). Adaptive NK cells resist regulatory T-cell suppression driven by IL37. Cancer Immunol. Res. *6*, 766–775. https://doi.org/10.1158/2326-6066.CIR-17-0498.

78. Luoma, A.M., Suo, S., Wang, Y., Gunasti, L., Porter, C.B.M., Nabilsi, N., Tadros, J., Ferretti, A.P., Liao, S., Gurer, C., et al. (2022). Tissue-resident memory and circulating T cells are early responders to pre-surgical cancer immunotherapy. Cell *185*, 2918–2935.e29. https://doi.org/10.1016/j.cell.2022.06.018.

79. Polański, K., Young, M.D., Miao, Z., Meyer, K.B., Teichmann, S.A., and Park, J.E. (2020). BBKNN: fast batch alignment of single cell transcriptomes. Bioinformatics *36*, 964–965. https://doi.org/10.1093/bioinformatics/btz625.

80. Wolf, F.A., Angerer, P., and Theis, F.J. (2018). SCANPY: large-scale single-cell gene expression data analysis. Genome Biol. *19*, 15. https://doi.org/10.1186/s13059-017-1382-0.

81. Wolock, S.L., Lopez, R., and Klein, A.M. (2019). Scrublet: computational identification of cell doublets in single-cell transcriptomic data. Cell Syst. *8*, 281–291.e9. https://doi.org/10.1016/j.cels.2018.11.005.

82. Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A.H., Tanaseichuk, O., Benner, C., and Chanda, S.K. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat. Commun. *10*, 1523. https://doi.org/10.1038/s41467-019-09234-6.

83. R Core Team (2021). R: A language and environment for statistical computing (Vienna, Austria: R Foundation for Statistical Computing). https://www.R-project.org/.

84. Van Rossum, G., and Drake, F.L. (2009). Python 3 Reference Manual (CreateSpace).

85. Yang, C., Siebert, J.R., Burns, R., Gerbec, Z.J., Bonacci, B., Rymaszewski, A., Rau, M., Riese, M.J., Rao, S., Carlson, K.S., et al. (2019). Heterogeneity of human bone marrow and blood natural killer cells defined by single-cell transcriptome. Nat. Commun. *10*, 3931. https://doi.org/10.1038/s41467-019-11947-7.

86. Guo, C., Wu, M., Huang, B., Zhao, R., Jin, L., Fu, B., Wang, P., Wang, D., Zheng, M., Fang, J., et al. (2022). Single-cell transcriptomics reveal a unique memory-like NK cell subset that accumulates with ageing and correlates with disease severity in COVID-19. Genome Med. *14*, 46. https://doi.org/10.1186/s13073-022-01049-3.

87. Lanier, L.L., Yu, G., and Phillips, J.H. (1989). Co-association of CD3ζ with a receptor (CD16) for IgG Fc on human natural killer cells. Nature *342*, 803–805. https://doi.org/10.1038/342803a0.

88. Li, H., van der Leun, A.M., Yofe, I., Lubling, Y., Gelbard-Solodkin, D., van Akkooi, A.C.J., van den Braber, M., Rozeman, E.A., Haanen, J.B.A.G., Blank, C.U., et al. (2019). Dysfunctional CD8 T cells form a proliferative, dynamically regulated compartment within human melanoma. Cell *176*, 775–789.e18. https://doi.org/10.1016/j.cell.2018.11.043.

89. Prager, I., and Watzl, C. (2019). Mechanisms of natural killer cell-mediated cellular cytotoxicity. J. Leukoc. Biol. *105*, 1319–1329. https://doi.org/10.1002/JLB.MR0718-269R.

90. Huntington, N.D., Cursons, J., and Rautela, J. (2020). The cancer–natural killer cell immunity cycle. Nat. Rev. Cancer *20*, 437–454. https://doi.org/10.1038/s41568-020-0272-z.

91. Sangro, B., Melero, I., Wadhawan, S., Finn, R.S., Abou-Alfa, G.K., Cheng, A.L., Yau, T., Furuse, J., Park, J.W., Boyd, Z., et al. (2020). Association of inflammatory biomarkers with clinical outcomes in nivolumab-treated patients with advanced hepatocellular carcinoma. J. Hepatol. *73*, 1460–1469. https://doi.org/10.1016/j.jhep.2020.07.026.

92. Shimasaki, N., Jain, A., and Campana, D. (2020). NK cells for cancer immunotherapy. Nat. Rev. Drug Discov. *19*, 200–218. https://doi.org/10.1038/s41573-019-0052-1.

93. Guo, X., Zhang, Y., Zheng, L., Zheng, C., Song, J., Zhang, Q., Kang, B., Liu, Z., Jin, L., Xing, R., et al. (2018). Global characterization of T cells in non-small-cell lung cancer by single-cell sequencing. Nat. Med. *24*, 978–985. https://doi.org/10.1038/s41591-018-0045-3.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
CoraLite®488-conjugated Rabbit IgG control Polyclonal antibody	Proteintech	Cat# CL488-30000; RRID: AB_2919216
CoraLite®488-conjugated DNAJB1 Polyclonal antibody	Proteintech	Cat# CL488-13174; RRID: AB_2919069
Brilliant Violet 421™ anti-human CD3 Antibody	Biolegend	Cat# 317344; RRID: AB_2565849
APC anti-human Perforin Antibody	Biolegend	Cat# 353312; RRID: AB_2571969
CD3 Polyclonal antibody	Proteintech	Cat# 17617-1-AP; RRID: AB_1939430
Anti-rat IgG (H+L), (Alexa Fluor® 488 Conjugate)	Cell Signaling Technology (CST)	Cat# 4416S; RRID: AB_10693769
DC-LAMP Antibody (1010E1.01)	Novus Biologicals	Cat# DDX0191P-100; RRID: AB_2827532
Recombinant Anti-NCAM1 antibody [EP2567Y] (ab75813),100uL(CD56)	Abcam	Cat# ab75813; RRID: AB_2632384
PE anti-human CD158a (KIR2DL1) Antibody	Biolegend	Cat# 374903; RRID: AB_2832735
APC Mouse Anti-Human CD158e1 (NKB1)	BD Biosciences	Cat# 564103; RRID: AB_2738594
PE-CF594 Mouse Anti-Human Granzyme B	BD Biosciences	Cat# 562462; RRID: AB_2737618
Brilliant Violet 510™ anti-human CD45 Antibody	Biolegend	Cat# 304036; RRID: AB_2561940
Alexa Fluor® 700 anti-human CD16 Antibody	Biolegend	Cat# 302026; RRID: AB_2278418
PE/Cyanine7 anti-human CD56 (NCAM) Antibody	Biolegend	Cat# 362510; RRID: AB_2563927
anti-DNAJB1 polyclonal Antibody	Solarbio	Cat# K106727P
Annexin A1 Polyclonal antibody	Proteintech	Cat# 21990-1-AP; RRID: AB_11182596
CD68 Monoclonal antibody	Proteintech	Cat# 66231-2-Ig; RRID: AB_2881622
CD86 (E2G8P) Rabbit mAb	Cell Signaling Technology	Cat# 91882S; RRID: AB_2797422
Recombinant Anti-TGF beta 1 antibody [EPR21143]	Abcam	Cat# ab215715-40ul; RRID: AB_2893156
CD3 Monoclonal Antibody (SK7), APC-eFluor™ 780	Thermo Fisher Scientific	Cat# 47-0036-42; RRID: AB_10717514
CD19 Monoclonal Antibody (SJ25C1), APC-eFluor™ 780	Thermo Fisher Scientific	Cat# 47-0198-42; RRID: AB_10719114
CD56 (NCAM) Monoclonal Antibody (CMSSB), APC-eFluor™ 780	Thermo Fisher Scientific	Cat# 47-0567-42; RRID: AB_10854573
PerCP/Cyanine5.5 anti-human CD45	Biolegend	Cat# 368504; RRID: AB_2566352
Brilliant Violet 650(TM) anti-human CD197 (CCR7)	Biolegend	Cat# 353234; RRID: AB_2563867
Mouse Anti-Human CD208 Monoclonal Antibody, PE Conjugated	BD Biosciences	Cat# 558126; RRID: AB_647194
IL-15 Monoclonal Antibody (34559), APC	Thermo Fisher Scientific	Cat# MA5-23627; RRID: AB_2608838
Mouse Anti-PD-L1 Monoclonal Antibody, PE-Cy7 Conjugated	BD Biosciences	Cat# 558017; RRID: AB_396986
BV510 Mouse Anti-Human CD83	BD Biosciences	Cat# 563223; RRID: AB_2738080
BD OptiBuild™ BUV395 Mouse Anti-Human HLA-DR, DP, DQ	BD Biosciences	Cat# 740302; RRID: AB_2740041
PE/Cyanine7 anti-human Perforin	Biolegend	Cat# 353316; RRID: AB_2571973
PerCP/Cyanine5.5 anti-human CD56 (NCAM)	BioLegend	Cat# 362506; RRID: AB_2563915
Critical commercial assays		
LIVE/DEAD™ Fixable Near-IR Dead Cell Stain Kit, for 633 or 635 nm excitation	Invitrogen	Cat# L34975
Tumor Dissociation Kit, human	Miltenyi Biotec	Cat# 130-095-929
Zombie Violet Fixable Viability Kit	Thermo Fisher Scientific	Cat# 423114
PanoPANEL Kits	Panovue	Cat# 10234100100
eBioscience™ Foxp3 / Transcription Factor Staining Buffer Set	Invitrogen	Cat# 00-5523-00
Chemicals, peptides, and recombinant proteins		
100 μm Smart Straine	Miltenyi Biotec	Cat# 130-098-463
Red blood cell lysis buffer	TIANDZ	Cat# 90309
Human BD Fc Block	BD	Cat# 564219

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Data files for human scRNA-seq dataset	This study	GEO: GSE212890
Raw data for human scRNA-seq dataset	This study	GSA for Human: HRA000321
Public human cancer scRNA-seq datasets	Other public studies	See Table S1 for details
Human breast cancer treatment-related scRNA-seq dataset	Zhang et al.[59]	GEO: GSE169246
Human melanoma treatment-related scRNA-seq dataset	de Andrade et al.[15]	GEO: GSE139249
Human lung cancer treatment-related RNA-seq dataset	Jung et al.[61]	GEO: GSE135222
Human melanoma treatment-related RNA-seq dataset	Riaz et al.[60]	GEO: GSE91061
Human metastatic urothelial carcinoma treatment-related RNA-seq dataset	Balar et al.[62]	IMvigor210
Software and algorithms		
Prism Software 9.1.1	GraphPad Software	https://www.graphpad.com/scientific software/prism/
HALO 3.3	Indica Labs	https://www.indicalab.com/halo
Flowjo Software 10.4	Flowjo, LLC	https://www.flowjo.com/solutions/flowjo
OlyVIA 3.8	OLYMPUS	https://olyvia.software.informer.com/
BBKNN 1.5.1	Polański et al.[79]	https://github.com/Teichlab/bbknn
Cellranger 3.0.0	10x Genomics	https://10xgenomics.com/
CellPhoneDB 2.1.7	Efremova et al.[63]	https://www.cellphonedb.org/
Scanpy 1.8.2	Wolf et al.[80]	https://scanpy.readthedocs.io/en/latest/
SCENIC 1.1.2-2	Aibar et al.[49]	https://github.com/aertslab/SCENIC
Scrublet 0.2.3	Wolock et al., 2018[81]	https://github.com/AllonKleinLab/scrublet
scVelo 0.2.3	Bergen et al.[54]	https://scvelo.readthedocs.io/
ROGUE	Liu et al.[34]	https://github.com/PaulingLiu/ROGUE
Metascape	Zhou et al.[82]	https://metascape.org/gp/index.html#/main/step1
Scaden	Menden et al.[58]	https://github.com/KevinMenden/scaden
R 4.1.2	R Core Team, 2021[83]	https://www.r-project.org/
Python 3.7.6	Van Rossum and Drake, 2009[84]	https://www.python.org/
Other		
Interactive explorer of human NK cells	This study	http://pan-nk.cancer-pku.cn/
Code	This study	https://github.com/TangFei965/pan-NK

RESOURCE AVAILABILITY

Lead contact

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

Materials availability

This study did not generate new unique reagents.

Data and code availability

- The accession numbers for the raw sequencing in this paper are GSA for Human (Genome Sequence Archive for Human in BIG Data Center, Beijing Institute of Genomics, Chinese Academy of Sciences): HRA000321. The gene expression data can be obtained from the NCBI GEO database with accession number GEO: GSE212890. Visualization of the scRNA-seq datasets in this study can be found at http://pan-nk.cancer-pku.cn/.

- Our analysis code has been uploaded to the github (https://github.com/TangFei965/pan-NK). GEO accession numbers for public treatment-associated datasets are listed in the key resources table. The detailed information of all public datasets has been provided in Tables S1 and S7.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Human participants

Ten patients enrolled in this study were those pathologically diagnosed with liver cancer, including nine males and one female, through the Beijing Shijitan Hospital, Beijing, China. The written informed consents were obtained from all participants before tissue and blood sample collection for research according to regular principles. Ethical approvals were gained from the Ethics Committee of Beijing Shijitan Hospital, Capital Medical University. Patients were pathologically diagnosed with hepatocellular carcinoma (HCC) except P20220722 and P20220520 who were diagnosed with intrahepatic cholangiocarcinoma (ICC). Patient P20220706 was treated with interventional therapy, while others had no prior chemotherapy or targeted systemic therapy. The ages of these patients ranged from 49 to 72, with a mean age of 60. The clinical stages of these patients, which were classified according to the guidance of AJCC, covered four stage I, three stage II, and three stage III. Detailed clinical metadata of these patients in this study are provided in Table S4. The patient specimens collected within this study were from East Asians. Paired samples of tumors and adjacent liver tissues were collected for NK cell isolation. Tumor tissues were dissected approximately 2 cm from tumor edge, and adjacent non-tumor tissues were at least 2 cm from the paired tumor.

The samples of forty-seven cancer patients were collected and analyzed in our previous study.[64] This study was also approved by the Research and Ethical Committee of Peking University Cancer Hospital and complied with all relevant ethical regulations. The detailed clinical information of the patients was shown in Table S1.

METHOD DETAILS

Sample collection and flow cytometry

Patients' or healthy donor's PBMCs were isolated through density gradient centrifugation using HISTOPAQUE®-1077 from anticoagulated whole blood. Fresh tumor and adjacent liver tissue from patients were dissociated with Tumor Dissociation Kit, human (Miltenyi Biotec, 130-095-929), according to the manufacturer's instructions. After being subsequently filtered by 100 μm Smart Straine (Miltenyi Biotec, 130-098-463), the dissociated cells were treated with red blood cell lysis buffer (TIANDZ, 90309) on ice for 1∼2 minutes to lyse red blood cells. Single-cell suspensions were centrifuged at 400 g for 5 min and resuspended with PBS, and then were stained with LIVE/DEAD™ Fixable Near-IR Dead Cell Stain Kit (Invitrogen, L34975) or Zombie Violet Fixable Viability Kit (Invitrogen, 423114) to distinguish live cells from labeled dead cells. After washing with PBS with 2% FBS, fluorochrome-conjugated antibodies for cell surface antigens were added and stained for 30 min at 4°C. For intracellular staining, cells were fixed and permeabilized using eBioscience™ Foxp3 / Transcription Factor Staining Buffer Set (invitrogen, 00-5523-00). In brief, cells were thoroughly resuspended using Fixation/permeabilization solution and fixed for 30min at 4°C, then washed with 1X Perm/Wash Buffer. After that, cells were blocked for 10min at 4°C using Human Fc Block (BD, 564219) and antibodies for intracellular antigens were added and then incubated for 1 hour at 4°C. During IL-15 detection, cells were not re-stimulated. Data were collected by a FACS Aria Sorp flow cytometer and analyzed with FlowJo software.

Multiplex immunofluorescent staining

Tumor samples from patients were placed in 4% Paraformaldehyde for 24 h, and then dehydrated and embedded in paraffin. Sections were stained using PanoPANEL Kits (panovue, 10234100100) to perform multiplex immunofluorescence according to the manufacturer's instructions. Briefly, slides were deparaffinized with xylene and a graded series of ethanol dilutions (100%, 95% and 70%), and were then fixed with 10% neutral buffered formalin for 10 minutes, followed by microwave-based antigen retrieval using the antigen restoration solution and antibody blocking for 30 minutes. Primary antibodies were incubated for 1h at room temperature, and HRP-labeled secondary antibodies were incubated at room temperature for 30 minutes, followed by TSA fluorescent dye working solution incubation for 30 minutes. Finally, after multi-antigen staining, nuclei were stained with DAPI for 20 minutes. Slides were enclosed using Nail Polish, scanned using the SLIDEVIEW VS200 (Olympus), and analyzed with OlyVIA software.

Single cell RNA library preparation and sequencing

Tumor and adjacent non-tumor tissues were collected from matched patients. Single cells were digested from tissues as described previously.[64] The tissues were cut into approximately 1-2 mm^3 small pieces in the RPMI-1640 medium (GIBCO) with 10% fetal bovine serum (FBS, GIBCO), and enzymatically digested by Tumor Dissociation Kit (Miltenyi Biotec) for 60 min on a 37 °C rotor according to the manufacturer's protocol.

Dissociated cells were then passed through a 100μm SmartStrainer and centrifuged at 400g for 5 min. To lyse red blood cells, the pelleted cells were treated with red blood cell lysis buffer (Miltenyi Biotec) on ice for 1-2 min after removing the supernatant. After

washing twice with 1 x PBS (GIBCO), the cell pellets were re-suspended in sorting buffer (PBS supplemented with 1% FBS) and stained with antibodies against CD45 and 7AAD for FACS with a BD FACSAria™ III sorter (BD Biosciences) instrument. 10,000∼20,000 cells were loaded for library construction using the 10x Chromium Single cell 5' Library (10x Genomics, V3), according to the manufacturer's instructions. Purified libraries were subject to an Hiseq X Ten sequencer (Illumina) for sequencing with 150-bp paired-end reads.

Single-cell RNA-seq data preprocessing

Newly generated single-cell sequencing data were aligned with the GRCh38 human reference genome and quantified using Cell Ranger (version 3.0, 10x Genomics Inc). The preliminary filtered data generated from Cell Ranger were used for downstream filtering and analyses. The quality of cells was then assessed based on two metrics: (1) The number of detected genes per cell; (2) The proportion of mitochondrial gene counts per cell. Specifically, cells with detected genes fewer than 200 and mitochondrial unique molecular identifier (UMI) count percentage larger than 10% were filtered out. The genes detected in more than 3 cells were retained. To remove the potential doublets, Scrublet[81] was used for each sequencing library with the expected doublet rate set to be 0.05, and cells with the predicted doubletScore larger than 0.3 were further filtered out.

For scRNA-seq data from other publications, the same filtering steps were applied for the datasets with raw count matrix to obtain high-quality cells. As for the specific datasets with no available count data, the CPM or TPM matrix were downloaded and used directly. The detailed meta data (including tissue location, sample and patient identifier) were retrieved from the original studies (Table S1).

We next normalized the count data generated from the droplet-based platform (Such as 10x Genomics Chromium and IndropSeq) using the *normalize_total* function with parameter '*target_sum=1e4*' in Scanpy.[80] Count data generated from Smart-seq2 were normalized by length of each gene before using the Scanpy pipeline. All the normalized data were logarithmically transformed for downstream analyses. Although unsupervised clustering is widely used to identify cell populations in the analysis of scRNA-seq data, we found that it inevitably included other cell types during NK cell identification (data not shown). Thus, we firstly isolated the potential NK cells based on the positive expression (above zero) of either *NCAM1* or *KLRF1* as well as the negative expression of the CD3 genes, including *CD3D*, *CD3G* and *CD3E*.[16,17,85,86] We further excluded the cells expressing marker genes of other cell types in subsequent unsupervised clustering (details in the next section). The detailed reasons for the sorting strategy are as follows: 1) Both *NCAM1* and *KLRF1* were used for the *in silico* FACS of NK cells to address the potential bias introduced by dropout events. 2) NKp80 (*KLRF1*) is one of markers of human NK cells, which is not expressed on ILC1s,[20,45] and this gene is also extensively used as a NK cell marker in scRNA-seq data.[85,86] By contrast, we found that another NK cell marker used in mouse models, *EOMES*, could be detected in human T cells and ILC1s as well,[20] thereby not included in our strategy. 3) Since *CD3D*, *CD3G* and *CD3E* are highly expressed in effector T cells, NK-T cells and human ILC1s, we used these genes as negative selection markers to ensure the purity of NK cells. Other CD3-related genes associated with NK cell functions such as CD3ζ[87] were completely preserved in our scRNA-seq atlas.

Batch effect correction and unsupervised clustering

Highly variable genes (HVGs) were selected within each cancer type to ensure the selection of genes that are conserved at the pan-cancer level using the *scanpy.pp.highly_variable_genes* function with the parameter setting '*batch_key=cancer types*'. The unexpected effects of the total counts and the percentage of mitochondrial gene counts per cell were regressed out from the normalized expression matrix using the *scanpy.pp.regress_out* function on the matrix of HVGs following the common workflow of Scanpy. Next, principal component analysis (PCA) was performed on the matrix of HVGs to reduce noise and reveal the main axes of variation using the *scanpy.tl.pca* function, and the top 40 components were retained for downstream analyses.

The batch effects were corrected by the BBKNN algorithm, which detected the top nearest neighbors of each cell from each batch respectively instead of the entire cell pool.[79] The parameters of the *scanpy.external.pp.bbknn* function were set to ''*batch_key='patient', n_pcs=40*''. Uniform Manifold Approximation and Projection (UMAP) was employed for visualization via the *scanpy.tl.umap* function with default parameters.

We performed two rounds of unsupervised clustering to unveil the structure of NK cell population using the *scanpy.tl.leiden* function. After the first-round of unsupervised clustering (the resolution parameter of clustering was set to 1.0), each cell cluster was annotated as $CD56^{bright}CD16^{lo}$ or $CD56^{dim}CD16^{hi}$ NK cells based on the expression levels of CD56 (*NCAM1*) and CD16 (*FCGR3A*). The obtained $CD56^{bright}CD16^{lo}$ and $CD56^{dim}CD16^{hi}$ NK cell populations were analyzed separately to obtain the high-resolution map of NK cell populations in the second-round of unsupervised clustering with the resolution parameter set to 0.6. Notably, we removed clusters showing signature of B cells (*CD19*, *MS4A1*), myeloid cells (*LYZ*, *C1QA*, *CD68*) or epithelial cells (*EPCAM, KRT18, KRT19*). The cluster-specific marker genes were identified using the *scanpy.tl.rank_genes_groups* function with the parameter of '*method=wilcoxon*' (Table S2). We identified NK cell clusters by their whole signature genes, and we named them by selection of one specific signature gene after simultaneously considering the observed expression pattern and previous conventions.

To confirm that our clustering results sufficiently depicted the transcriptional heterogeneity of NK cells and no further sub-clustering was necessary, an entropy-based universal metric ROGUE was applied for the defined subsets.[34] The ROGUE index approaching one represents a relatively purer subset with fewer significant intra-cluster variations.

The proportion and tissue distribution of NK cell subsets

To quantify the distribution of NK cell population, we calculated the proportion of each NK cell subset for all samples from blood, tumor and adjacent non-tumor tissues. Samples with the number of cells less than 20 were excluded to avoid the bias caused by inaccurate estimations. Subsets from $CD56^{bright}CD16^{lo}$ and $CD56^{dim}CD16^{hi}$ NK cells were processed respectively. Due to insufficient sample numbers, all metastatic samples were incorporated into ''other tissue'' and not included in downstream analysis.

Calculation of signature score

We used the R package AUCell[49] to calculate the signature score of a specific gene set. AUCell is ranking-based, and is independent of the gene expression units and the normalization procedure. We first built the ranked expression matrix using the *AUCell_buildRankings* function, and then calculated the AUC value using the *AUCell_calcAUC* function. Detailed gene sets used in this study were defined in the following sections.

Definition of cytotoxicity, inflammatory and stress gene sets

To evaluate the functional variations between $CD56^{bright}CD16^{lo}$ and $CD56^{dim}CD16^{hi}$ NK cell subsets, we defined cytotoxicity, inflammatory and stress-related gene sets after comprehensive compilation of previous studies. The cytotoxicity gene sets were defined as *GZMA*, *GZMB*, *GZMH*, *GZMM*, *GZMK*, *GNLY*, *PRF1* and *CTSW*.[88,89] The inflammatory gene set was defined as *CCL2*, *CCL3*, *CCL4*, *CCL5*, *CXCL10*, *CXCL9*, *IL1B*, *IL6*, *IL7*, *IL15* and *IL18*.[90,91] The general stress gene set was defined as *BAG3, CALU, DNAJB1, DUSP1, EGR1, FOS, FOSB, HIF1A, HSP90AA1, HSP90AB1, HSP90B1, HSPA1A, HSPA1B, HSPA6, HSPB1, HSPH1, IER2, JUN, JUNB, NFKBIA, NFKBIZ, RGS2, SLC2A3, SOCS3, UBC, ZFAND2A, ZFP36* and *ZFP36L1*.[88] Each score was calculated by AUCell as we mentioned above.

Definition of HLA-dependent and -independent receptors

To evaluate the expression levels of inhibitory and activating receptors on NK cell surface, we defined four gene sets based on their inhibitory or activating properties and HLA-dependence, including HLA-dependent inhibitory receptors (*KIR2DL1*, *KIR2DL3*, *KIR3DL1*, *KIR3DL2*, *LILRB1* and *LAG3*), HLA-independent inhibitory receptors (*PDCD1*, *SIGLEC7*, *CD300A*, *CD96*, *IL1RAPL1*, *TIGIT* and *HAVCR2*), HLA-dependent activating receptors (*KIR2DL4*, *CD160* and *KLRC2*) and HLA-independent activating receptors (*NCR3*, *NCR1*, *KLRK1*, *CRTAM* and *FCGR3A*).[92] Each score was calculated by AUCell as we mentioned above.

Hierarchical clustering of cancer types

In order to compare the composition of NK cells among different cancer types, we performed unsupervised hierarchical clustering for all analyzed cancer types in this study based on the proportions of $CD56^{bright}CD16^{lo}$ and $CD56^{dim}CD16^{hi}$ NK cell subsets. The frequencies of all subsets were used as input for the R function *tree* to hierarchically cluster the cancer types. Only the tumor-derived NK cells were used for analysis.

Specificity and sensitivity assessment

We evaluated the sensitivity for selected differentially expressed genes between NK cells with different tissue origins using the permutation test. We randomly shuffled the tissue label of each NK cell using the *sklearn.model_selection.permutation_test_score* function with the parameter of '*n_permutation=1000*'. The selected gene could be a good predictor, if the score on original data is much greater than those obtained from the permuted data and the *P* value is less than 0.05.

To compare the expression specificity for each selected gene, we calculated the relative fold change between tissue-infiltrating NK cells and circulating NK cells according to the following formula:

$$specificity = \frac{Nt_i}{Nt} \Big/ \frac{Nb_i}{Nb}$$

where *Nt*referred to the number of all tissue-infiltrating NK cells, Nt_i the number of NK cells expressing gene *i* in the tumor and adjacent non-tumor tissue, *Nb* the number of all blood-derived NK cells, and Nb_i the number of NK cells expressing gene *i* in blood. The above analyses were performed for $CD56^{bright}CD16^{lo}$ and $CD56^{dim}CD16^{hi}$ NK cells, respectively.

RNA velocity analysis

The newly generated 10x scRNA-seq data was used for the RNA velocity analysis.[53] First of all, the spliced and unspliced UMIs were recounted by the python package velocyto.[54] Then we normalized the count matrices by the library size and kept the genes detected in over 20 cells in both spliced and unspliced matrices. Next, HVGs were selected by the *scanpy.pp.highly_variable_genes* function to be used for PCA. The moments of normalized spliced and unspliced counts were calculated using the *scvelo.pp.moments* function for each cell with default parameters. The RNA velocity was estimated using the *scvelo.tl.velocity* function with the dynamical mode selected and the velocity graph was built with the cosine similarity using the *scvelo.tl.velocity_graph* function. The RNA velocities were projected into the UMAP coordinates with the *scv.pl.velocity_embedding_stream* function for visualization.

Pseudotime trajectory inference

The diffusion map algorithm was used to infer the differentiation trajectory of all tumor-infiltrating CD56^{dim}CD16^{hi} NK cells following the workflow in Scanpy.[55] Based on the above BBKNN-derived batch-balanced k nearest-neighbor graph, the diffusion map was built using the *scanpy.tl.diffmap* function with default parameters. We selected the CD56^{bright}CD16^{hi} NK cells as the root because they located at the start of the directed streamline inferred by RNA velocity, and calculated the diffusion pseudotime for CD56^{dim}CD16^{hi} NK cells using the *scanpy.tl.dpt* function. We then fit the expression profile of each gene to the pseudotime using a polynomial regression model.

SCENIC analysis

Activated regulons in tissue-enriched CD56^{bright}CD16^{lo} and CD56^{dim}CD16^{hi} NK cell subsets were analyzed by SCENIC, respectively.[49] The newly generated scRNA data in this study was used as input for the python implementation of the SCENIC algorithm (pyscenic). Briefly, the gene-gene co-expression relationships between transcription factors (TF) and their potential targets were inferred using the *grn* function with the gene regulatory network reconstruction algorithm '*grnboost2*' selected. Next, the command *aucell* was used to calculate the regulon activity for each cell. Finally, the differentially activated regulons were recognized in a specific NK cell subset by wilcoxon test with adjusted P value < 0.05.

Cell-cell interaction analysis by CellPhoneDB

The Python package CellPhoneDB[63] was applied to infer the cell-cell interactions between NK cells and other immune cells in our newly generated data. We first collected scRNA-seq data of myeloid cells and T cells from our previous studies and kept their original annotations for downstream analyses.[13,64] Of note, only the immune cells in tumors were used. For the cell-cell interactions between each CD56^{dim}CD16^{hi} NK cell subset and other immune cell types, the predicted ligand-receptor pairs with P values < 0.05 and average expression > 1, were extracted for counting and presentation in the figures (Table S6).

Analysis of immunotherapy datasets

To address whether TaNK cells were associated with ICB treatment response, we collected the published scRNA-seq data from patients subjected to the ICB therapy with sufficient NK cells. The bulk RNA-seq datasets of ICB therapy from large cohorts were also included in our analysis. The samples from primary tumors were used for downstream analyses.

For the scRNA-seq data, we compared the proportion of TaNK cells in all CD56^{dim}CD16^{hi} NK cells for each patient. Then we used wilcoxon test to compare the proportion of TaNK cells between responders and non-responders. As for the bulk RNA-seq data, all count data were normalized (TPM) and $\log_2$-transformed. To address the discordance in clinical information, all patients were renamed as responder (including CR/PR patients) or non-responder (including SD/PD patients) based on their clinical responses after treatment. To test the correlation of TaNK cells and immunotherapy responses, we applied the five marker genes (*KLRF1*, *DNAJB1*, *BAG3, SERPINE1* and *NR4A1*) as the signature of TaNK cells, and further compared the signature activity of TaNK cells between responders and non-responders.

TCGA RNA-seq data analysis

The expression data of TCGA and the clinical information of patients were downloaded using the TCGAbiolinks package. We calculated the signature scores of CD56^{dim}CD16^{hi} NK cells and LAMP3^{+} DCs to verify their connections. Based on previous studies, *LYZ*, *CST3*, *LAMP3*, *FSCN1* and *CCL19* were used as the LAMP3^{+} DC signature, and *NCR1*, *NCR3*, *KLRB1*, *CD160* and *PRF1* were used as the NK cell signature.[14,37,64] Patients with the *NCAM1* expression lower than the median level were regarded as the CD56^{dim}CD16^{hi} NK cell enrichment group. The correlation between LAMP3^{+} DCs and CD56^{dim}CD16^{hi} NK cells at the pan-cancer level was calculated by Pearson correlation.

Survival analysis

The prognostic performance of TaNK cells among tumors was evaluated by TCGA data of each cancer type. The signature score of TaNK cells was defined as above. The TCGA patients of each cancer type were classified into high and low groups based on the optimal cutpoint returned by the *surv_cutpoint* function. To explore the effect of TaNK cells on overall survival, we performed survival analyses by the Cox proportional hazards model implemented in the R package survival after correcting the effects of other clinical factors including tumor stage, gender and patient age. Survival curves were fit using the Kaplan-Meier formula in the R package survival, and visualized using the *ggsurvplot* function of the R package survminer. We performed deep learning–based cell composition analysis to further validate the association between the TaNK cell frequency and prognosis based on Scaden.[58] We trained the neural network model on 100,000 simulated samples which were generated by randomly combining different numbers of TaNK cells and other cell types, and resulted in a 96% accuracy on an independently simulated validation dataset. We then applied the optimized model on TCGA data and patients of each cancer type were classified into high and low groups based on the deconvoluted TaNK cell frequencies. The detailed code and models used to perform this analysis in this study are available at the following github repository: https://github.com/TangFei965/pan-NK.

Comparison with human peripheral blood for NK cells

Here, we used all blood-derived scRNA-seq data for downstream analyses with the exception of GES158055, because it was generated using a unique alignment and quantification algorithm (Table S7). In order to define the level of change in circulating NK cells, the similarities of each major circulating NK cell population between healthy donors and patients were measured based on the cosine distance of their respective transcriptome. First, we combined the cell-by-gene expression matrix from healthy donors and patients from each of the examined cancer types separately. The HVGs were then selected from each joint matrix to serve as the input for PCA projection. Finally, the low-dimensional space spanned by the top 40 PCs was used to calculate the cosine distance between healthy donors and patients.

We further evaluated the changes for NK cells in peripheral blood in terms of the compositional changes of NK cell subset and gene expression fluctuations. For each cancer type, we calculated the fold change of the average proportion for each NK cell subset between tumor patients and healthy donors.

We also compared the ratio of the observed and expected number of cells (*Ro/e*) for c8−KLRC2 NK cell subset in each cancer types to examine their tissue preference, where the expected cell numbers for each subset were obtained from the chi-square test.[93] One subset was assumed as an enriched population in a specific tissue if the corresponding *Ro/e* index is greater than 1.

QUANTIFICATION AND STATISTICAL ANALYSIS

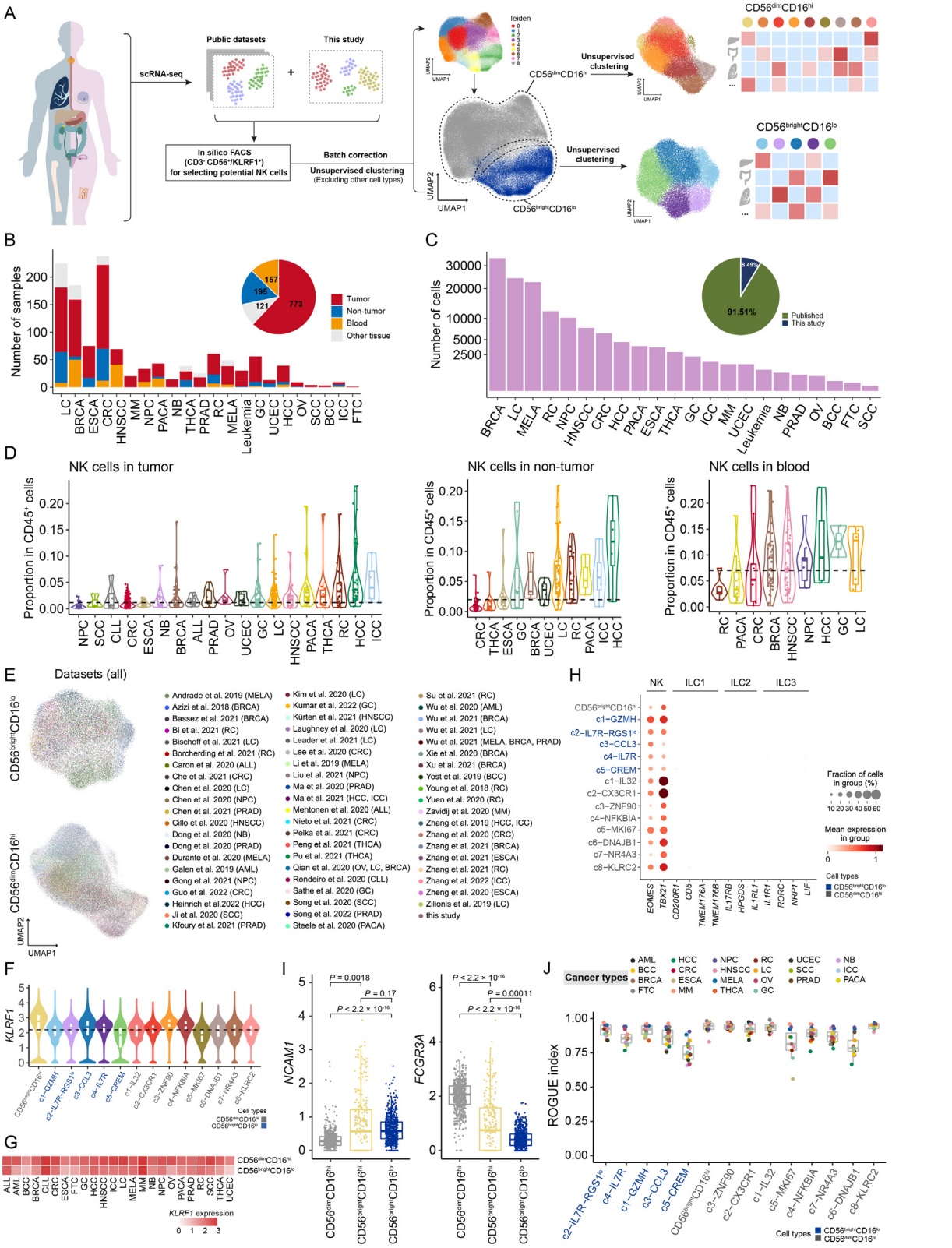
Statistical analysis

Statistical analyses used in this study included Kruskal-Wallis test, Wilcoxon test and t-test as described in the Figure legends. The Kaplan-Meier method was applied in survival analyses. Correlations among immune cell subsets were estimated by Pearson correlation.

Cell
Resource

Supplemental figures

CellPress
OPEN ACCESS



(legend on next page)

Figure S1. Basic information of the data and the integration of NK cells, related to Figure 1

(A) Schematic diagram of construction for human pan-cancer NK cell atlas.

(B) The sample number of NK cells derived from blood (n = 157), tumor (n = 773) and adjacent non-tumor (n = 195) tissues.

(C) The number of NK cells across cancer types. The y axis is scaled by using a square root transformation.

(D) The proportion of NK cells in all $CD45^{+}$ cells for each sample across cancer types. Dashed line indicates the median value.

(E) UMAP plot showing integrated $CD56^{bright}CD16^{lo}$ and $CD56^{dim}CD16^{hi}$ NK cell atlas with information of datasets.

(F) Violin plots showing the expression level of *KLRF1* across all NK cell subsets.

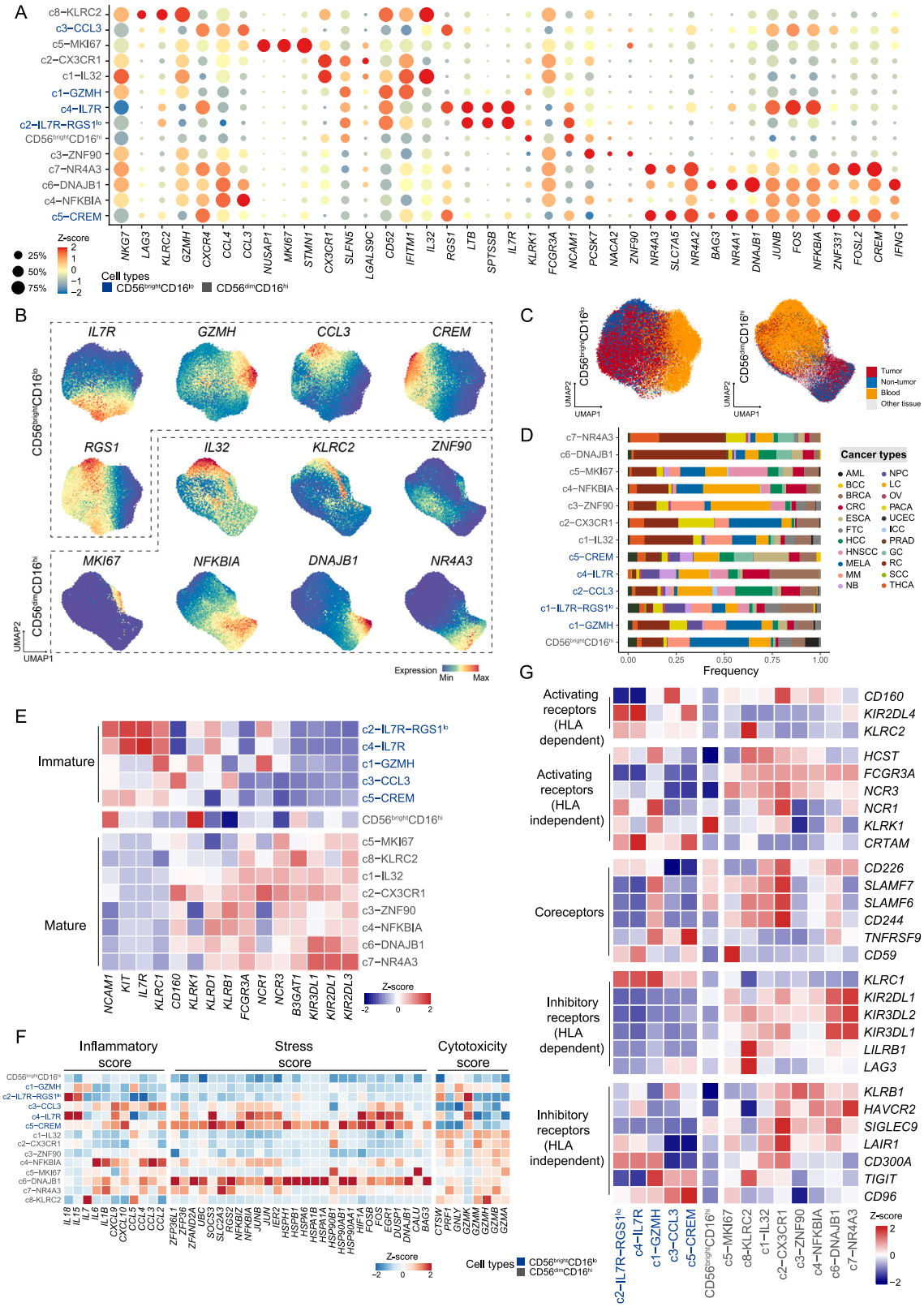
(G) The comparison of *KLRF1* expression level across cancer types in both $CD56^{bright}CD16^{lo}$ and $CD56^{dim}CD16^{hi}$ NK cells.

(H) The expression pattern of ILC markers among NK cell subsets. The color represents the mean expression within each of the NK subsets and the dot size indicates the fraction of cells in the subsets expressing the corresponding gene.

(I) Boxplots comparing the expression of *NCAM1* and *FCGR3A* in NK cell major populations. Each dot corresponds to one cell. Two-sided unpaired Wilcoxon test.

(J) Boxplots showing cell purity for each NK cell subset across tumors by ROGUE index (STAR Methods).

See also Table S1.



(legend on next page)

Figure S2. Characteristics of NK cell subsets, related to Figure 1

(A) Bubble heatmap showing the expression pattern of selected signature genes of the NK cell subsets. Dot size represents the fraction of expressing cells. Color indicates the *Z* score scaled gene expression levels.

(B) UMAP plots showing expression levels of signature genes used to name every $CD56^{bright}CD16^{lo}$ or $CD56^{dim}CD16^{hi}$ NK cell subset.

(C) UMAP plots showing the tissue distribution of $CD56^{bright}CD16^{lo}$ and $CD56^{dim}CD16^{hi}$ NK cell.

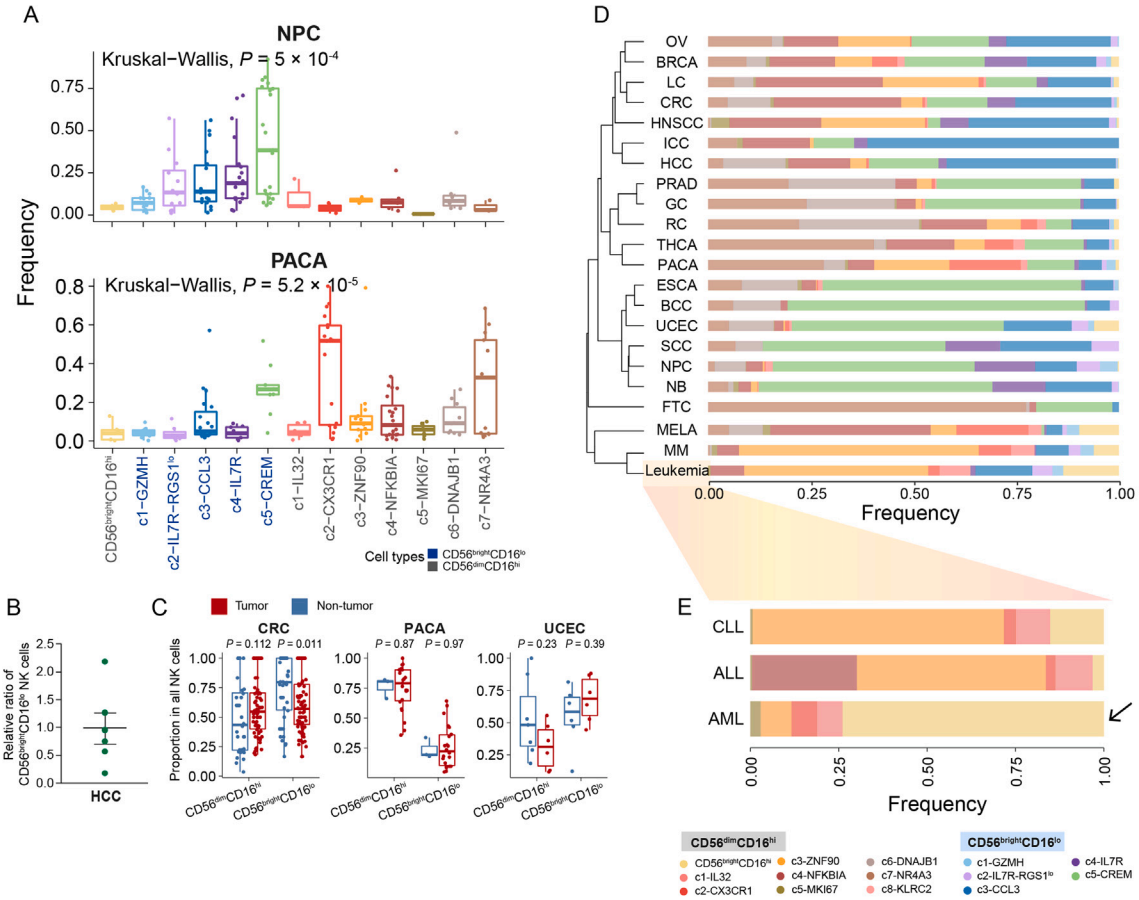
(D) Barplots showing the contribution of cancer types to each NK cell subset.

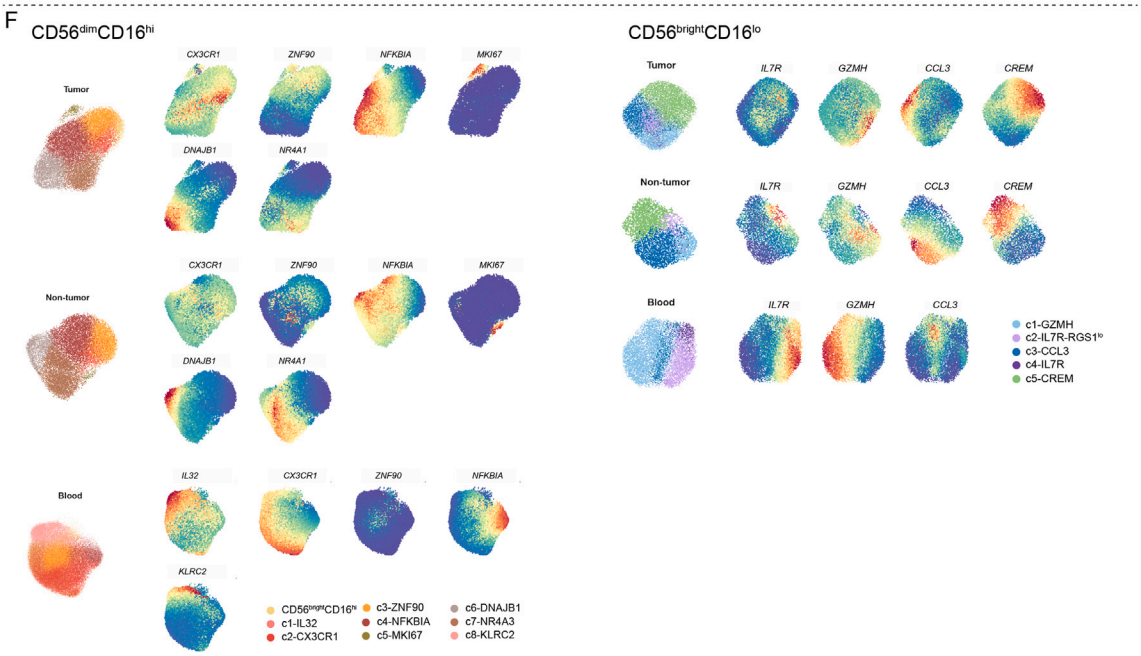
(E) Heatmap showing the differential expression of cell markers used for development identification among NK cell subsets. Color represents the *Z* score scaled gene expression levels.

(F) Heatmap showing the expression of corresponding genes used to define the functional scores in Figure 1G. Color indicates the *Z* score scaled gene expression levels.

(G) Heatmap showing the selected activating and inhibitory receptors differentially expressed in $CD56^{bright}CD16^{lo}$ and $CD56^{dim}CD16^{hi}$ NK cells. Color is coded by the *Z* score scaled gene expression.

See also Table S2.





(legend on next page)

Figure S3. Heterogeneity of the tumor-infiltrating NK cell populations across different cancer types, related to Figure 2

(A) Boxplots showing that $CD56^{bright}CD16^{lo}$ NK cell subsets were enriched in NPC and $CD56^{dim}CD16^{hi}$ NK cell subsets were enriched in PACA. The frequency of each NK cell subset is divided by the total NK cell number in tumors. Kruskal-Wallis test.

(B) Flow cytometry analysis showing the ratio of $CD56^{bright}CD16^{lo}$ NK cells to $CD56^{dim}CD16^{hi}$ NK cells in the tumor tissue of HCC patients (n = 6). Data are represented as mean ± SEM.

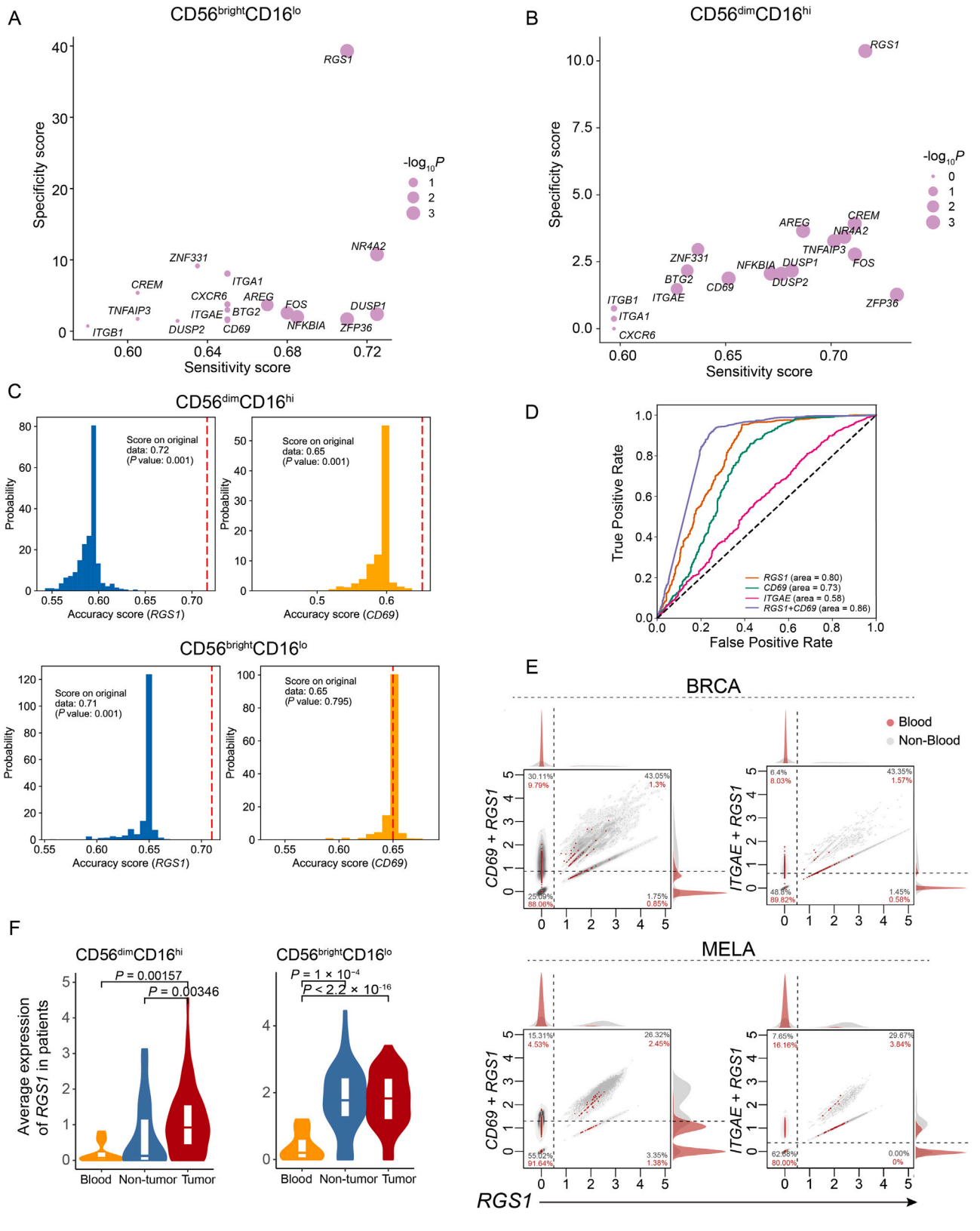
(C) Boxplots comparing the proportion of two NK cell major populations divided by the total NK cell number between tumor and adjacent non-tumor tissues, respectively. Two-sided unpaired Wilcoxon test.

(D) Unsupervised hierarchical clustering of cancer types based on the cellular composition of all NK cell subsets in each cancer type. The average proportion of each subset is shown.

(E) The cellular compositions of $CD56^{dim}CD16^{hi}$ NK cell subsets in three common types of leukemia including Acute Lymphocytic Leukemia (ALL), Chronic Lymphocytic Leukemia (CLL), and Acute Myeloid Leukemia (AML).

(F) UMAP plots showing the signature genes of NK cell subsets for each tissue type.

See also Table S4.



(legend on next page)

Figure S4. *RGS1* as a better identifier of tissue-infiltrating NK cells, related to Figure 3

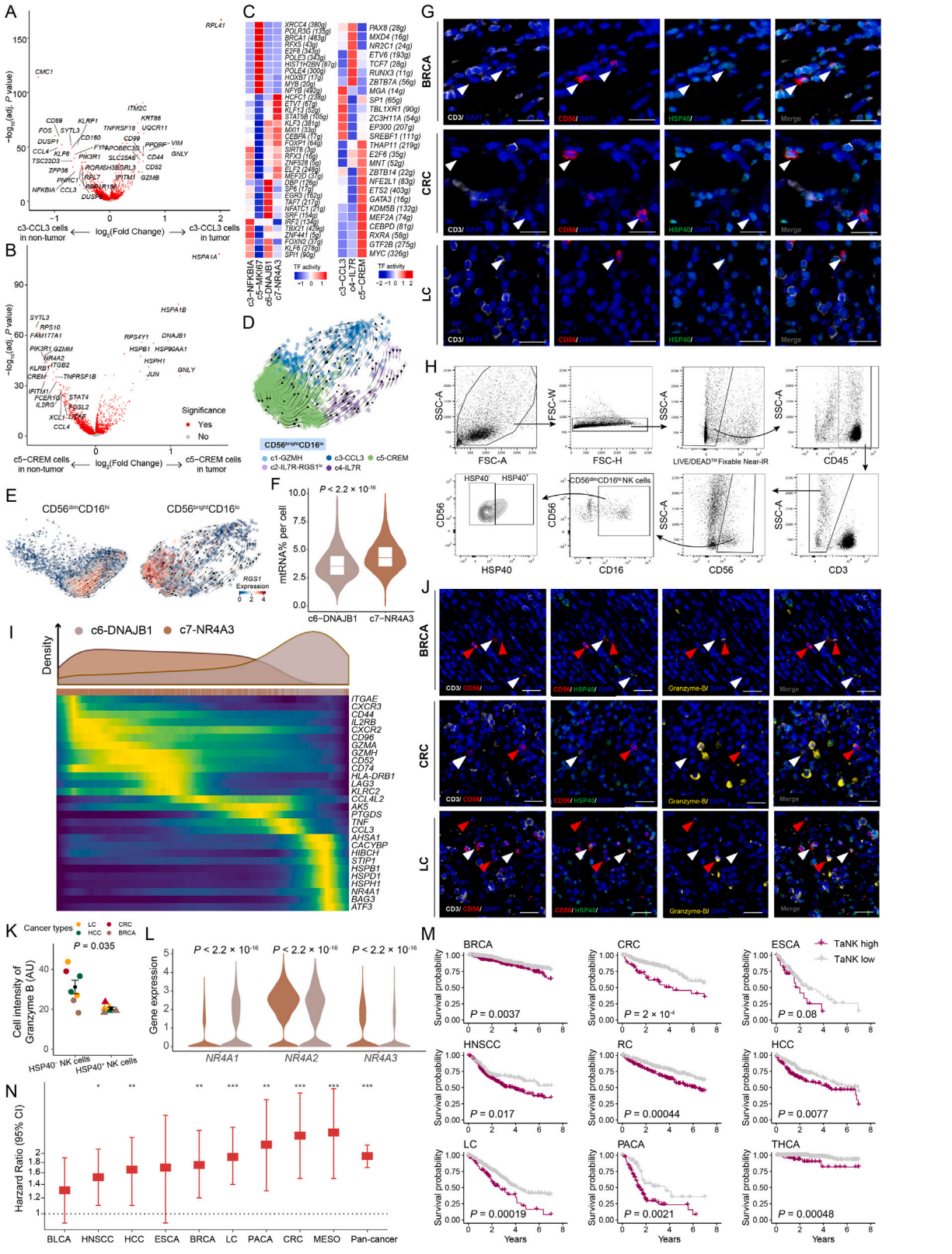
(A and B) The sensitivity and specificity of representative genes to distinguish the tissue origin of (A) $CD56^{bright}CD16^{lo}$ and (B) $CD56^{dim}CD16^{hi}$ NK cells (STAR Methods). Dot size represents the significance of sensitivity score evaluated by $-\log_{10}$ (p value). *P* values are generated by permutation test.

(C) The permutation test for *CD69* and *RGS1* as tissue-infiltrating marker of NK cells, respectively. Red line marks the score on original labeled data. *P* values are generated by permutation test (STAR Methods).

(D) Receiver Operator Characteristic (ROC) curves showing the performance in prediction of non-blood NK cells. The dash line represents the result of random selection.

(E) The performance of combinations of *RGS1* with either *CD69* or *ITGAE* in distinguishing tissue-derived NK cells based on *in silico* FACS. Red dots denote the NK cells derived from blood and gray dots denote tissue-derived NK cells. Dashed line indicates predicted boundary of blood and non-blood NK cells based on the expression level of the combination.

(F) Violin plots showing the significant alterations of *RGS1* expression among tissues in CRC. Two-sided unpaired Wilcoxon test.



(legend on next page)

Figure S5. Dynamic changes of CD56^{bright}CD16^{lo} NK cell infiltration and the existence of TaNK cells, related to Figures 4 and 5

(A and B) Volcano plots showing differentially expressed genes for (A) CD56^{bright}CD16^{lo} c3-CCL3 NK cells and (B) c5-CREM NK cells from tumor and adjacent non-tumor tissues. Each red point indicates significantly up- or down-regulated gene with adjusted p value < 0.05, two-sided t test.

(C) Heatmap showing TF activity for tissue-enriched CD56^{dim}CD16^{hi} and CD56^{bright}CD16^{lo} NK cell subsets (STAR Methods).

(D) RNA velocities overlaid on the UMAP of CD56^{bright}CD16^{lo} NK cells by using the newly generated data in this study. Arrows show the RNA velocity field. Dots are colored by CD56^{bright}CD16^{lo} NK cell subsets.

(E) RNA velocities overlaid on the UMAP of (left) CD56^{dim}CD16^{hi} and (right) CD56^{bright}CD16^{lo} NK cells by using the newly generated data in this study. Dots are colored by the expression of *RGS1*. Arrows show the RNA velocity field.

(F) Violin plots showing the percentage of counts in mitochondrial genes of CD56^{dim}CD16^{hi} c6-DNAJB1 and c7-NR4A3 NK cells. Two-sided unpaired Wilcoxon test.

(G) Multiplex immunofluorescence staining in other representative tumors to show the TaNK cells (arrows). The scale bar represents 20 μm.

(H) Gating strategy of flow cytometry for HSP40^{+} CD56^{dim}CD16^{hi} NK cells, exemplified by patient P20220722.

(I) Heatmap showing that genes changed along with the inferred pseudotime (STAR Methods). The density plot on top of the heatmap showing the distribution of CD56^{dim}CD16^{hi} c6-DNAJB1 and c7-NR4A3 NK cells along the pseudotime.

(J) Multiplex immunofluorescence staining in other representative tumors to show the expression of granzyme B in HSP40^{+} (white arrows) and HSP40^{-} NK cells (red arrows). The scale bar represents 20 μm.

(K) Quantification of the fluorescence intensity for granzyme B in HSP40^{+} NK cells (n = 6) and HSP40^{-} NK cells (n = 7) from (I) and Figure 4H using the Halo v3.3 image analysis platform. Data are represented as mean ± SEM. Two-sided unpaired Wilcoxon test.

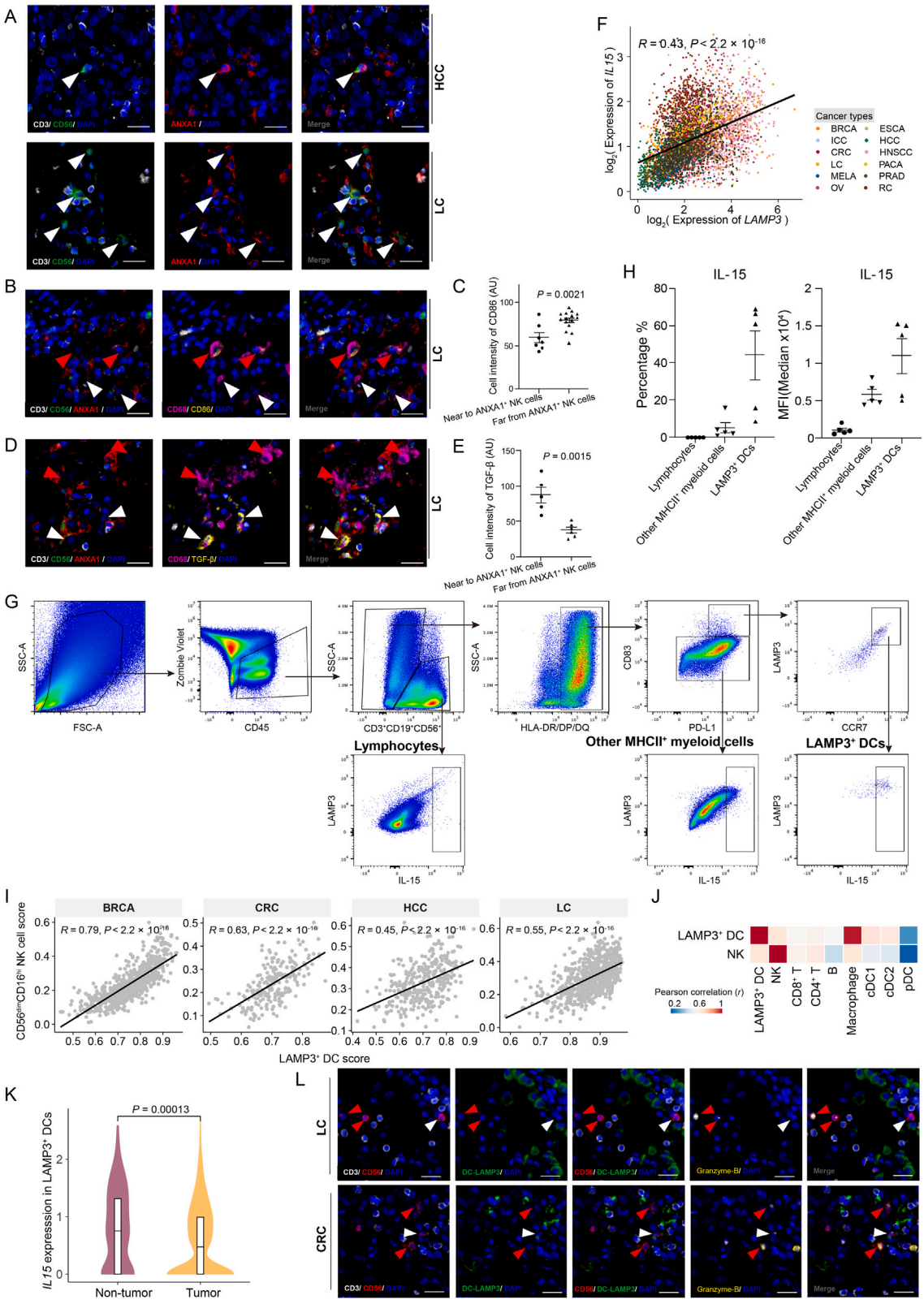
(L) Violin plots showing the expression levels of NR4As in CD56^{dim}CD16^{hi} c6-DNAJB1 and c7-NR4A3 NK cells. Two-sided unpaired Wilcoxon test.

(M) Kaplan-Meier plots showing the signature activity of TaNK cells in different cancer types based on the deconvoluted TaNK cell proportions (STAR Methods). +, censored observations; log-rank test.

(N) Forest plot showing the effect of TaNK cells on overall survival based on the deconvoluted TaNK cell proportions (STAR Methods). The hazard ratios (HRs) are calculated by Cox regression models with the age, gender and tumor stage corrected. The y axis is scaled by using a $\log_{10}$ transformation. Dashed line indicates HR = 1. *p < 0.05, **p < 0.01, ***p < 0.001. *P* values are adjusted by Benjamini-Hochberg.

(D) and (E) are plotted using our newly generated data.

See also Tables S4 and S5.



(legend on next page)

Figure S6. The interaction of myeloid cells and NK cells in various cancer types, related to Figure 6

(A) Representative multiplex immunofluorescence images of HCC and lung cancer (LC) tumors. White arrows represent ANXA1^{+} NK cells. Scale bar represents 20 μm.

(B–E) Representative images and quantification of fluorescence intensity of (B and C) CD86 and (D and E) TGF-β expression on macrophages adjacent to or distant from ANXA1^{+} NK cells in tumor tissue of lung cancer. White arrows represent macrophages adjacent to ANXA1^{+} NK cells, whereas red arrows far from ANXA1^{+} NK cells. Scale bar represents 20 μm. Data are represented as mean ± SEM. Two-sided t test.

(F) Scatterplot showing the correlation of *LAMP3* and *IL15* in TCGA data. Pearson correlation test.

(G) The gating strategy of flow cytometry for tumor-infiltrating immune cells and IL-15 production.

(H) Flow cytometry analysis of the proportion and expression level of IL-15 among different tumor-infiltrating immune cells in HCC patients (n = 5). Data are represented as mean ± SEM.

(I) Scatterplots showing correlations of CD56^{dim}CD16^{hi} NK cells with LAMP3^{+} DCs in the TCGA dataset on each cancer type, based on their signature gene expressions. Pearson correlation test.

(J) Heatmap showing the Pearson correlation coefficient for gene signatures of distinct cell types in the TCGA dataset. *r*, Pearson correlation coefficient.

(K) Violin plot showing the comparison of *IL15* expression between tumor and adjacent non-tumor tissues in scRNA-seq data from GEO: GSE154763.

(L) Multiplex immunofluorescence staining to show the co-localization of LAMP3^{+} DCs (DC-LAMP3^{+}) and NK cells (CD3^{-}CD56^{+}) in LC and CRC. The scale bar represents 20 μm.

See also Table S4.

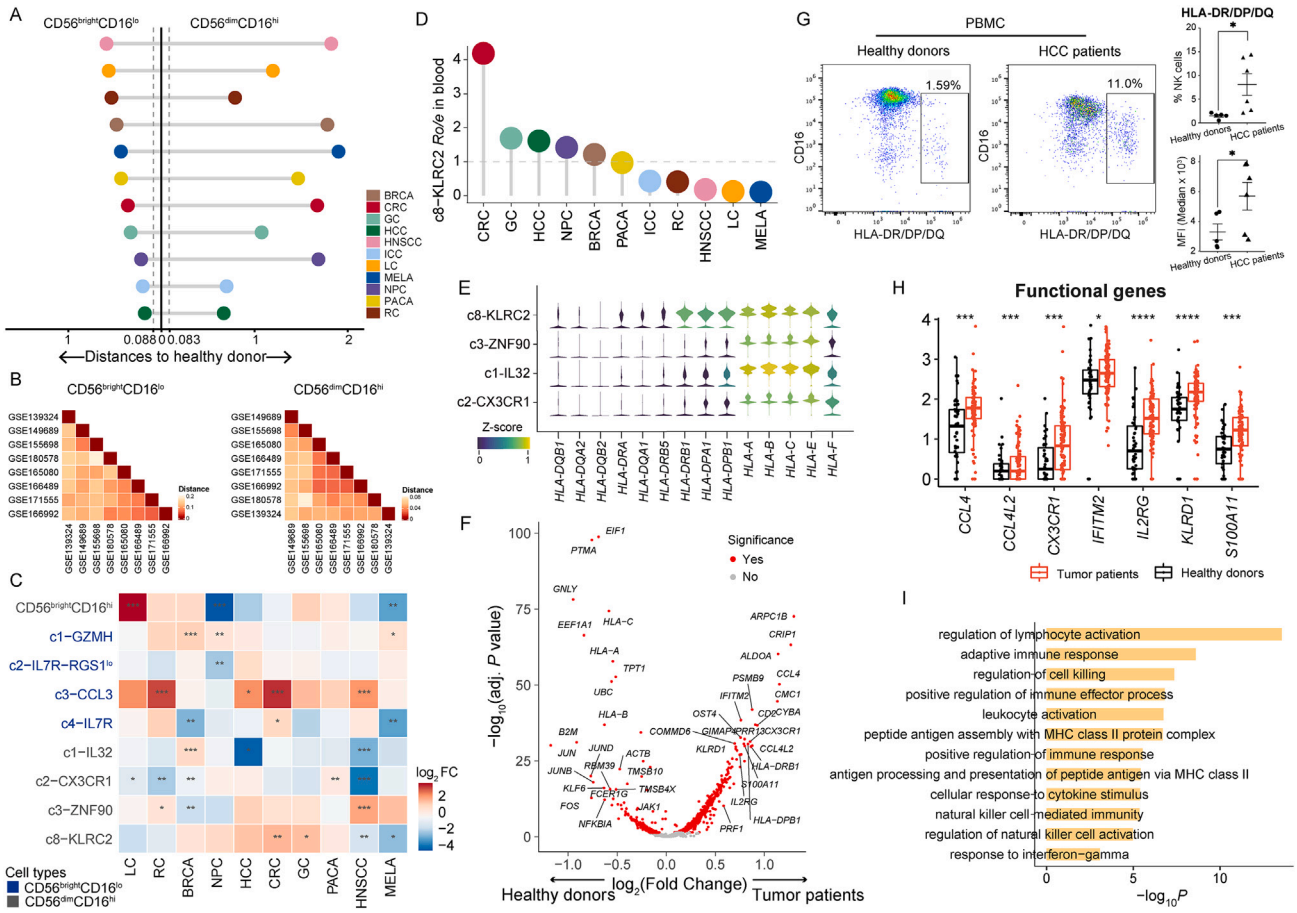


Figure S7. Alterations of circulating NK cells in patients, related to Figure 1

(A) Lollipop plot showing the similarities of NK cell major subsets (CD56^{dim}CD16^{hi} and CD56^{bright}CD16^{lo} NK cells) between healthy donors and patients, measuring by cosine distance (STAR Methods). The distance approaching zero indicates higher similarity. Dashed lines represent the average similarity of CD56^{dim}CD16^{hi} (right) and CD56^{bright}CD16^{lo} (left) NK cells among the healthy donors.

(B) Heatmap showing the similarities of NK cell major subsets (CD56^{dim}CD16^{hi} and CD56^{bright}CD16^{lo} NK cells) among healthy donors from varied datasets.

(C) Heatmap showing log fold changes (FCs) of mean proportion between patients and healthy donors for each subset. *p < 0.05, **p < 0.01, ***p < 0.001, two-sided unpaired Wilcoxon test.

(D) Lollipop plot showing the tissue distribution of c8-KLRC2 NK cells by *Ro/e* score. The cluster is characterized as being enriched in a specific cancer type with *Ro/e* > 1.

(E) Violin plots showing the expression levels of MHC class II genes in circulating CD56^{dim}CD16^{hi} NK cell subsets.

(F) Volcano plot showing differentially expressed genes for c8-KLRC2 NK cells from patients and healthy donors. Red dots denote gene with adjusted p value < 0.05, two-sided unpaired Wilcoxon test.

(G) The representative plots (left) and proportions of MHC class II molecule expression (right) on peripheral blood NK cells from healthy donors and HCC patients, analyzed by flow cytometry. Data are represented as mean ± SEM. *p < 0.05, **p < 0.01, ***p < 0.001, two-sided t test.

(H) Boxplots showing the average expression of functional genes in c8-KLRC2 NK cells from patients and healthy donors. Each dot corresponds to one patient. *p < 0.05, **p < 0.01, ***p < 0.001, two-sided t test.

(I) Pathways enriched in c8-KLRC2 NK cells in tumor patients by using Metascape.[82] Hypergeometric test. Benjamini-Hochberg adjusted p value < 0.01.

See also Tables S4 and S7.