











Th1-poised naive CD4 T cell subpopulation reflects anti-tumor immunity and autoimmune disease

Received: 5 August 2024

Accepted: 13 February 2025

Published online: 25 February 2025

 Check for updates

Jae-Won Yoon ^{1,10}, Kyung Min Kim^{2,10}, Sookyung Cho^{1,10}, Min-Ji Cho ¹, Seonjun Park ³, Daehee Hwang ², Hye Ryun Kim ⁴, Sung Ho Park³, Jae-Ho Cho ⁵, Hyobin Jeong ⁶  & Je-Min Choi ^{1,7,8,9} 

Naïve CD4 T cells are traditionally viewed as a quiescent, homogeneous, resting population, but emerging evidence reveals their heterogeneity, which can be crucial for understanding disease contexts and therapeutic outcomes. In this study, we identify distinct subpopulations within both murine and human naïve CD4 T cells by single cell-RNA-sequencing (scRNA-seq), particularly focusing on a subpopulation that expresses super-high levels of interleukin-7 receptor (IL-7R^{sup-hi}), along with CD97, IL-18R, and Ly6C. This subpopulation, absent in the thymus and peripherally induced, exhibits type 1 helper T cell (Th1)-poised characteristics and contributes to the inhibition of cancer progression in B16F10 tumor-bearing mice. In humans, this IL-7R^{sup-hi} subpopulation expressing CD97 correlates with the responsiveness to anti-PD-1 therapy in cancer patients and the disease state of multiple sclerosis. By elucidating the heterogeneity of naïve CD4 T cells and identifying a Th1-poised subpopulation capable of robust type 1 responses, we highlight the importance of this heterogeneity in inflammatory conditions for defining the disease states and predicting drug responsiveness.

Effector T cell subsets and the array of cytokines they secrete generally serve as critical indicators of inflammation and have been utilized as valuable biomarkers across various diseases, constituting targets for therapeutic interventions^{1–4}. Particularly in the context of immune checkpoint blockade therapy, the activation or exhaustion state of CD8 T cells, along with the proportion of regulatory T cells, are considered crucial biomarkers for predicting drug responsiveness in cancer patients^{5–7}. However, the associations between the disease and the naïve CD4 T cells generally deemed a

homogeneous and resting population^{8–10}, have not yet been largely explored.

Naïve T cells mature in the thymus upon self-antigen stimulation, undergoing positive and negative selection based on the strength of affinity towards self-antigens¹¹, potentially leading to fundamental intrinsic heterogeneity through varying T cell receptor (TCR) signaling strength. Although tonic TCR signals do not change the naïve T cell state, they can induce molecules reflecting their self-reactivity including CD5, Ly6C, and Nur77^{12–14}. These phenotypic diversities

¹Department of Life Science, College of Natural Sciences, Hanyang University, Seoul 04763, Republic of Korea. ²Department of Biological Sciences, Seoul National University, Seoul, Korea. ³Department of Biological Sciences, Ulsan National Institute of Science & Technology (UNIST), Ulsan, Republic of Korea. ⁴Division of Medical Oncology, Department of Internal Medicine, Yonsei Cancer Center Yonsei University College of Medicine, Seoul 03722, Republic of Korea. ⁵Medical Research Center for Combinatorial Tumor Immunotherapy, Department of Microbiology and Immunology, Chonnam National University Medical School, Hwasun 58128, Korea. ⁶Department of Systems Biology, College of Life Science and Biotechnology, Yonsei University, Seoul 03722, South Korea. ⁷Hanyang Institute of Bioscience and Biotechnology, Hanyang University, Seoul, Republic of Korea. ⁸Research Institute for Natural Sciences, Hanyang University, Seoul, Republic of Korea. ⁹Research Institute for Convergence of Basic Sciences, Hanyang University, Seoul 04763, Republic of Korea. ¹⁰These authors contributed equally: Jae-Won Yoon, Kyung Min Kim, and Sookyung Cho. ✉ e-mail: hyobinjeong@yonsei.ac.kr; jeminchoi@hanyang.ac.kr

also lead to functional heterogeneity in further activation or differentiation of naïve T cells^{15–20}.

The cytokine milieu resulting from infection²¹, autoimmune disease²², and cancer² triggers an inflammatory response, leading to T cell activation and differentiation. Naïve CD4 T cells express various cytokine receptors such as IL-27²³, IL-4R²⁴, and the type I interferon receptor²⁵ in the steady state. One previous report demonstrated that IL-27 inhibits Th17 differentiation and EAE development by rapidly inducing PD-L1 expression on naïve CD4 T cells²³. Moreover, naïve CD4 T cells from poly (I:C)-treated mice showed a propensity towards Treg differentiation, dependent on type I interferon signaling²⁶. Therefore, naïve CD4 T cells can be affected by the cytokines and might exhibit a predisposition for further activation and differentiation rather than remaining unaffected. Recent single-cell analysis revealed heterogeneity among naïve CD4 T cells in both mice^{27–29} and humans^{29,30}. However, clear evidence is still lacking regarding a subpopulation of naïve CD4 T cells that are already Th1-poised, as well as the clinical relevance of this heterogeneity.

We hypothesized that fate-biased cells may exist within the naïve CD4 T cell population and sought to explore how this heterogeneity arises in the periphery. Additionally, we aim to investigate the extent to which these naïve CD4 T cell subpopulations are associated with cancer and autoimmune diseases. In this study, we aim to elucidate the heterogeneity of naïve CD4 T cells in both mice and humans and, through comprehensive bioinformatic analysis, we identify a distinct subpopulation of Th1-poised naïve CD4 T cells (T_{N1}). Moreover, we present predictive modeling based on the composition of naïve CD4 T cells, which correlates with responsiveness to anti-PD-1 therapies and autoimmune disease states.

Results

Identification of *IL7*^{sup-hi} naïve CD4 T cells conserved in both murine and human

To elucidate the transcriptomic heterogeneity of naïve CD4 T cells, we performed single-cell RNA sequencing (scRNA-seq) of FACS-sorted mouse splenic naïve CD4 T cells (NK1.1, $\gamma\delta$ TCR, TCR- β^+ , CD4⁺, CD25⁺, CD62L^{hi}, CD44^{low}) from enriched cells obtained by MACS (removing CD8, CD11b, CD11c, CD19, B220, CD49b, CD105, anti-MHC class 2, Ter-119, $\gamma\delta$ TCR by negative selection) (Supplementary Fig. 1a). Comprehensive bioinformatic analysis was then conducted using optimized workflows (Supplementary Fig. 1b). Unsupervised clustering analysis revealed four distinct clusters within the naïve CD4 T cells (Fig. 1a), each defined by specific marker genes that highlight unique characteristics (Supplementary Fig. 2a). Cluster 0 (C0) was identified as the conventional resting naïve population, primarily expressing marker genes encoding cytoskeleton proteins. On the other hand, cluster 2 (C2, self-reactive) represented a population showing high self-reactivity, with the expression of *Nr4a1*^{12,20,31}, *Egr1*³², and *Cd5*^{13,15,17–20,33}, while cluster 3 (C3, ISG-high) exhibited increased expression of interferon-induced genes such as *Ift1L3*³⁴, and *Isg15*³⁵. Notably, Cluster 1 (C1) displayed a decrease in the expression of resting cluster marker genes but had highly elevated expression of *Il7r* and *Adgre5* and was annotated as an “unknown cluster (*IL7*^{sup-hi})” (Fig. 1b). Using ASURAT algorithm³⁶, we calculated single-cell pathway activities and projected naïve CD4 T cells into a low-dimensional space (Supplementary Fig. 2b). Cells clustered tightly according to their cluster identity (C0 – C4), indicating strong functional coordination within each naïve CD4 T cell clusters (Supplementary Fig. 2c). Further functional annotation revealed that C3 was enriched for type I interferon response gene sets, C2 was associated with inflammatory response gene sets, and C1 showed high expression of gene sets involved in growth and T cell differentiation (Fig. 1c). In addition to scRNA sequencing analysis, bulk-RNA sequencing further confirmed that *Il7r*, *Ly6c1* and *Il18r1* are highly expressed in *IL7*^{sup-hi} naïve CD4 T cells (Supplementary Fig. 3a–d). These differential marker genes were validated at the protein level

using flowcytometric analysis, followed by clustering analysis with FlowJo plugins (UMAP v4.0.4, FlowSOM v4.1.0) (Fig. 1d). We identified two main subpopulations within naïve CD4 T cells (Supplementary Fig. 1c), one of which (C1) is characterized by *IL-7R*^{high}, *Ly6C*⁺, *IL-18R*^{high}, *CD5*^{low}, and *Adgre5* (*CD97*)^{high} (Fig. 1e, f). These cells constitute ~20–40% of the naïve CD4 T cell population (Fig. 1g). Importantly, there are no differences in the molecules involved in T cell activation or naïve phenotype markers (Fig. 1h). To investigate whether similar heterogeneity or shared subpopulations also exist in human naïve CD4 T cells, we analyzed scRNA sequencing datasets of peripheral blood mononuclear cells (PBMCs) from four healthy human donors³⁷. Naïve CD4 T cells were annotated using the SingleR algorithm³⁸, with DICE as the reference profile based on the high-depth bulk RNA-seq of FACS-sorted human naïve CD4 T cells³⁹. Unsupervised clustering of human naïve CD4 T cells revealed six distinct clusters (Fig. 1i), with the human C0 cluster showing high similarity to the mouse C1 cluster (Fig. 1j). Marker gene analysis indicated significant overlap between the mouse C1 and human C0 subpopulations (adjusted *p*-value < 1.89e-7 from the hypergeometric test), revealing four common markers; *Il7r*, *Pim1*, *Adgre5*, *Stk38* (Fig. 1k). The differential expression of these marker genes was further validated by real-time PCR (Supplementary Fig. 4). Through comprehensive transcriptomic and proteomic bioinformatic analysis, we uncovered significant heterogeneity in both steady-state mice and healthy humans, identifying a shared distinct subpopulation (*IL-7R*^{sup-hi}) within naïve CD4 T cells.

IL-7R^{sup-hi} naïve CD4 T cells are peripherally induced

As *Ly6C* and *IL-18R* are distinct surface markers of *IL-7R*^{sup-hi} cells, we confirmed that *Ly6C/IL-18R* double-positive (DP) cells exhibit high *IL-7R* expression (Fig. 2a). These DP cells also express *CD97* at higher levels than *Ly6C/IL-18R* double-negative (DN) cells. Interestingly, *CD5* expression is significantly lower in DP cells, suggesting less self-reactivity. Notably, *Ly6C* expression is rare in *CD44*^{hi} memory-phenotype (MP) CD4 T cells (Fig. 2b), likely due to antigen receptor stimulation. Moreover, *Ly6C/IL-18R* DP cells are nearly absent in the thymus, likely due to the requirement for TCR stimulation during positive selection, but they are abundantly present in peripheral lymphoid and non-lymphoid tissues (Fig. 2c, d). Additionally, marker molecules associated with *IL-7R*^{sup-hi} naïve CD4 T cells, such as *Ly6C*, *IL-7R*, and *IL-18R*, are up-regulated in splenic naïve CD4 T cells compared to their thymic counterparts, whereas *CD5* expression is higher in thymic naïve CD4 T cells (Fig. 2e). These observations were further supported by protein-marker level clustering from FACS analysis (Supplementary Fig. 5a–c), indicating that these phenotypic differences are acquired in the periphery. To confirm this hypothesis, we performed scRNA-seq on naïve CD4 T cells from spleen and thymic CD4 single-positive (SP) cells of mice (Fig. 2f). A significant increase in the proportion of C1 cluster (*IL7*^{sup-hi}) in the spleen compared to the thymus (odds ratio = 6.0, *p*-value < 2.2e-16 from the fisher’s exact test) (Fig. 2g). Monocle pseudotime analysis indicated that, while the pseudotime of cells is dispersed in C0 (resting), the overall pseudotime of cells were shifted towards late stages in other clusters, including C1 (Fig. 2h). Furthermore, TCR clones from thymic C0 cluster largely matched those in the spleen C1 cluster (Fig. 2i), suggesting that *IL-7R*^{sup-hi} naïve CD4 T cells are generated in the periphery. There was no significant difference in TCR diversity among naïve CD4 T cell clusters, implying no significant clonal expansion due to TCR stimulation (Fig. 2j). Additionally, protein marker-based clustering revealed distinct characteristics in murine splenic (peripheral) naïve CD4 T cell subclusters across different ages, with cluster 2 (C2, *IL-7R*^{sup-hi}, *Ly6C*⁺, *IL-18R*^{hi}, *CD5*^{lo}) absent in 4 day-old mice but gradually emerging between 8 weeks and 6 months. (Supplementary Fig. 6a–c). Next, we aimed to determine which signals drive the phenotypic change in naïve CD4 T cells generated in the thymus as they transition to the periphery. When thymic CD4 naïve T cells were cultured without TCR stimulation,

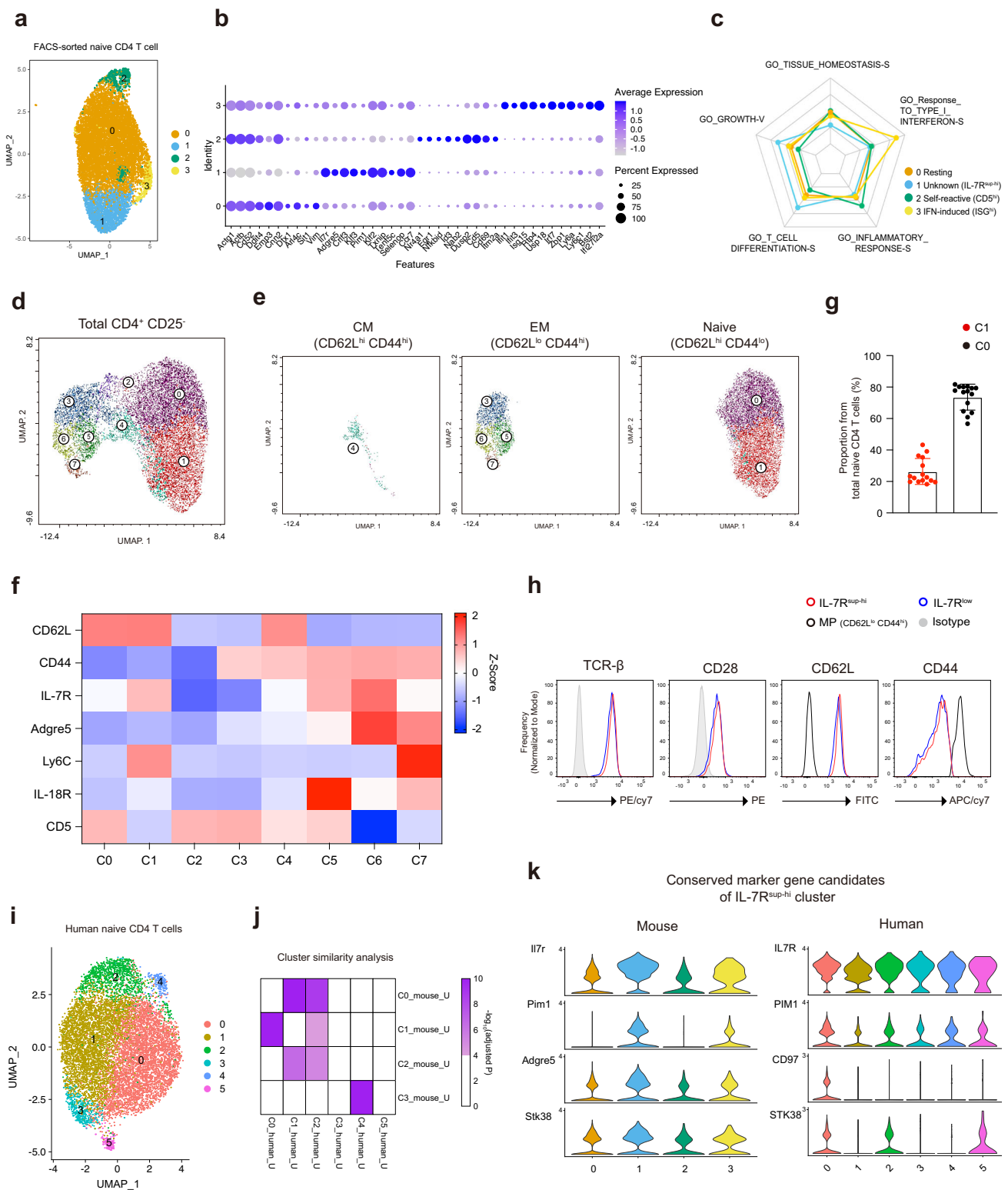
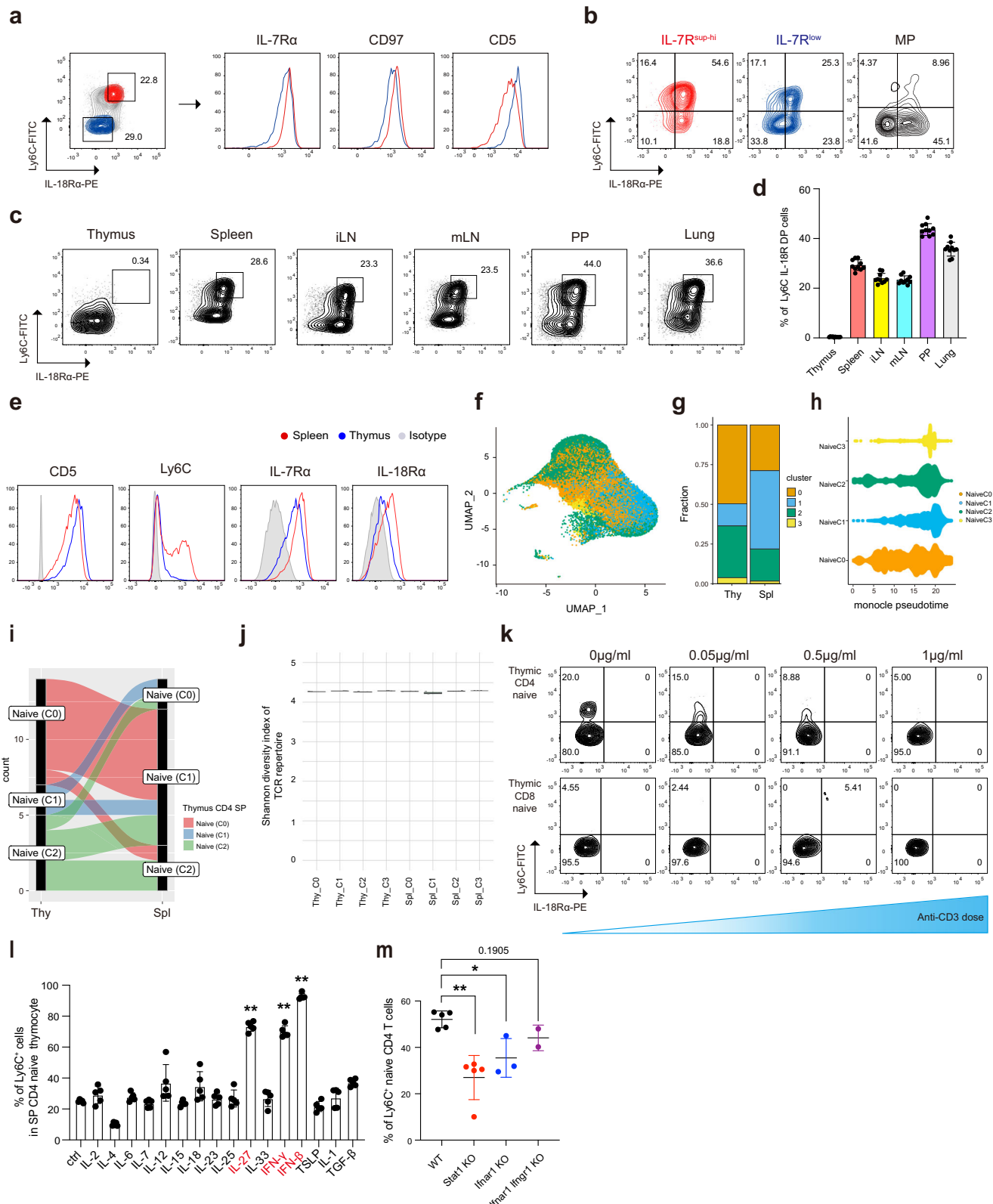


Fig. 1 | IL-7R^{sup-hi} naive CD4 T cell subpopulation, identified by both RNA and protein markers, is commonly present in both murine and human. **a** UMAP of scRNA-seq data from murine splenic naive CD4 T cells. Cells were colored by the unsupervised clustering result. **b** Dot plot showing the representative marker genes for each cluster. **c** Radar chart showing the characteristic pathway activities of each cluster obtained by ASURAT. Y-axis in the chart indicate sign score from 0 to 1 divided by 4 sectors which represent pathway activities. **d** UMAP of FACS data from murine splenic CD4 T cells. **e** UMAP from each indicated population (CM: CD62L^{high} CD44^{high}, EM: CD62L^{low} CD44^{high}, Naive: CD62L^{high} CD44^{low}). **f** Heatmap of indicated target protein expression from each cluster (Z-score). **g** Proportion of IL-7R^{sup-hi}

naive T cells from total naive CD4 T cells ($n = 15$). Data presented as the mean \pm S.D. **h** Representative histogram of flow cytometry data from IL-7R^{sup-hi} and IL-7R^{low} naive CD4 T cells. **i** UMAP of scRNA-seq data from human naive CD4 T cells. The color of the cells indicates the cluster identity. **j** Heatmap of similarity scores between murine and human naive CD4 T cell clusters. **k** Violin plots shows the significance of overlap between up-regulated genes in each cluster. Cluster similarity was evaluated using one-sided fisher's exact test. **k** Violin plots for the common marker genes between murine and human IL-7R^{sup-hi} naive T cell cluster. Source data are provided as a Source Data file.



Ly6C expression increased over their days, but was significantly suppressed upon anti-CD3 antibody stimulation (Fig. 2k), suggesting that antigen recognition may not be required for this phenotype in the periphery. In contrast, Ly6C expression significantly increased upon stimulation with STAT1-activating cytokines such as IFN- γ , IL-27, and type I interferon (Fig. 2l), and was reduced in STAT1 knockout (KO) mice (Fig. 2m). Collectively, these findings demonstrate that IL-7R^{sup-hi},

Ly6C/IL-18R DP naive CD4 T cells exhibit lower self-reactivity, are generated in the periphery via STAT1-activating cytokines.

Th1-poised characteristics of IL-7R^{sup-hi} naive CD4 T cells

Since IL-7R^{sup-hi} naive CD4 T cells exhibit high expression of IL-18R, we hypothesized that these cells could be particularly responsive to IL-18 or IL-18-related Th1 responses. Functional enrichment analysis of the

Fig. 2 | IL-7R^{sup-hi} naïve CD4 T cells are peripherally induced by STAT-1 activating cytokines. **a** Representative histogram of IL-7R, Adgre5, CD5 expression in IL-18R/Ly6C double-positive (DP) and double-negative (DN) naïve CD4 T cells. **b** Representative dot plot of IL-18R and Ly6C expression in IL-7R^{sup-hi} and IL-7R^{low} naïve CD4 T cells and memory-phenotype CD4 T cells (MP). **c** Representative dot plot of IL-18R/Ly6C DP population in naïve CD4 T cells from each indicated tissue. **d** Tissue distribution of IL-18R/Ly6C DP population in naïve CD4 T cells ($n = 10$). **e** Representative histogram of naïve CD4 T cells from spleen and thymus. **f** UMAP of murine naïve CD4 T cells from two different tissues (thymic CD4 SP and spleen). Cells were colored by cluster identity using label transfer of naïve CD4 T cell clusters from Fig. 1A. **g** Stacked bar graph showing the proportion of naïve CD4 T cell clusters from thymus (Thy) and spleen (Spl). **h** Dot plots showing the distribution of pseudotime for the single cells in each cluster. **i** Alluvial plot to track the

clonal population in thymus (Thy) and spleen (Spl) using TCR-seq data. **j** Box plots for the Shannon diversity index of TCR repertoire for each cluster in each tissue. **k** Expression of Ly6C and IL-18R in naïve CD4 and CD8 single-positive (SP) cells in thymocytes stimulated with anti-CD3 for 3 days. **l** Proportion of Ly6C⁺ cells in SP CD4 naïve thymocytes stimulated by the indicated cytokines for 3 days ($n = 5$). Statistical significance was confirmed by two-sided Mann-Whitney *U*-test with control group and each experimental group ($p = 0.0079$). **m** Proportion of Ly6C⁺ naïve CD4 T cells in the splenocyte of WT, Stat1^{-/-}, Ifnar1^{-/-}, and Ifnar1^{-/-} Ifngr1^{-/-} mice (WT $n = 5$, Stat1 KO $n = 5$, Ifnar1 KO $n = 3$, Ifnar1 Ifngr1 KO $n = 2$). Data presented as the mean \pm S.D. Statistical significance was confirmed by the two-sided Mann-Whitney *U*-test (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). Source data and exact statistics value are provided as a Source Data file.

up-regulated genes for the IL-7R^{sup-hi} naïve CD4 T cell cluster (C1), obtained from previous scRNA-seq data, revealed enrichment in the IL-12, IL-18, MyD88 pathways (Fig. 3a). Both mRNA and protein levels of IL-18R, as well as mRNA level of IL-12Rb1, were elevated in IL-7R^{sup-hi} naïve CD4 T cells (Fig. 3b-c). We found no significant difference in antigen-specific T cell proliferation between IL-7R^{sup-hi} and IL-7R^{low} naïve CD4 T cells (Fig. 3d), indicating comparable TCR responsiveness. However, IL-7R^{sup-hi} naïve CD4 T cells responded more sensitively to IL-12 and IL-18, upregulating T-bet and IFN- γ expression, whereas IL-7R^{low} naïve CD4 T cells showed weaker Th1 polarization (Fig. 3e-g). Additionally, we observed that naïve CD4 T cells from 4 day-old mice did not respond to IL-12 and IL-18, while their Th1 response gradually increased with age, implying Th1-poised characteristic of the subpopulation (Supplementary Fig. 6d). Upon adoptive transfer into Rag-deficient mice, IL-7R^{sup-hi} and IL-7R^{low} naïve CD4 T cells caused comparable weight loss (Fig. 3h). Notably, IL-7R^{low} naïve CD4 T cells predominantly secreted IL-17A, while IL-7R^{sup-hi} naïve CD4 T cells primarily secreted IFN- γ in the spleen and mLN, indicating an intrinsic Th1 bias in IL-7R^{sup-hi} naïve CD4 T cells (Fig. 3i-j). While maintaining the characteristics of naïve cells, we anticipated underlying epigenetic changes related to cytokine stimulation. Unsupervised clustering of publicly available scATAC-seq data²⁷ from FACS-sorted splenic naïve CD4 T cells (Supplementary Fig. 7a-d) revealed that the potential enhancer region of the *Ifngr1* gene was significantly more accessible in the IL-7R^{sup-hi} cluster compared to the IL-7R^{low} cluster (Fig. 3k). Additionally, transcription motif analysis identified that the motifs for T-bet, STAT1, STAT4, and Eomes were located at the center of the potential enhancer region of *Ifngr1*, which was highly accessible in IL-7R^{sup-hi} cluster (Fig. 3l). Collectively, these findings suggest that IL-7R^{sup-hi} naïve CD4 T cells possess a Th1-poised phenotype, characterized by heightened sensitivity to IL-12 and IL-18.

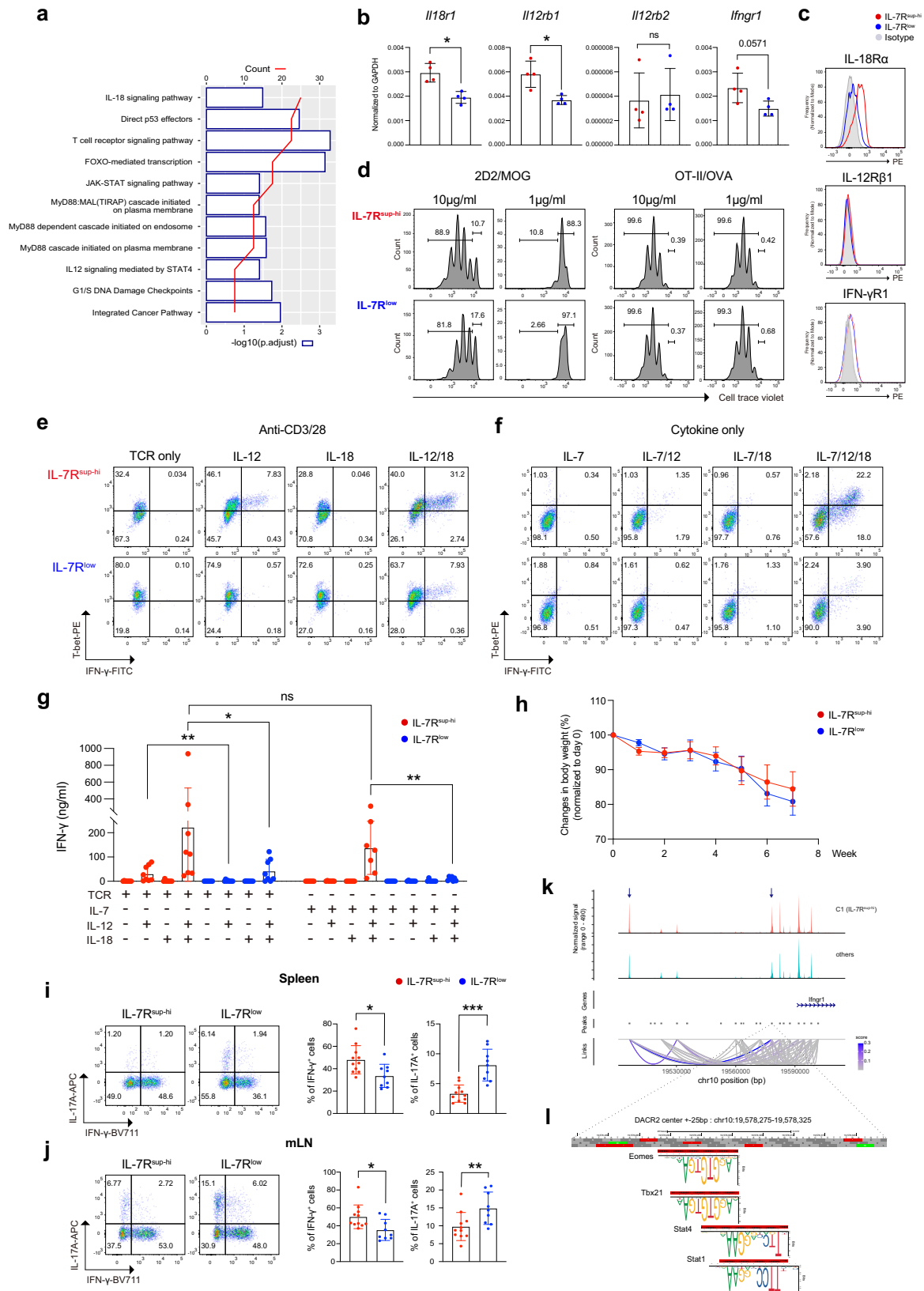
Tumoricidal effect of IL-7R^{sup-hi} naïve CD4 T cells via IFN- γ signaling

Given the heightened responsiveness of IL-7R^{sup-hi} naïve CD4 T cells to IL-12 and IL-18, we sought to investigate the changes of naïve CD4 T cells upon stimulation. To achieve this, we performed scRNA-seq after stimulating naïve CD4 T cells with IL-12 and IL-18 for integrative pseudotime analysis, comparing naïve CD4 T cells before and after stimulation (Fig. 4a-b). We observed that naïve CD4 T cells stimulated by IL-12 and IL-18 exhibited upregulation of *Tbx21*, *Ifng*, *Gzmb*, *Mki67* during their transition to late stages, along with a verified reduction in naïve score and an elevation in Th1 score (Fig. 4c). We applied the K-means clustering algorithm to group genes that changed with pseudotime, resulting in four groups of genes (G1-G4) (Fig. 4d). Functional enrichment analysis of each group (G1-G4) highlighted terms related to interferon gamma response, cytotoxicity, and pathways in cancer, predominantly enriched in G4, which are genes with increased expression along the pseudotime trajectory (Fig. 4e). Notably, the expression of granzyme B, a critical component in cytolytic Th1 effector function, was induced in IL-7R^{sup-hi} naïve CD4 T cells by IL-12

and IL-18 stimulation (Fig. 4f-g). IL-12 and IL-18-stimulated naïve CD4 T cells exhibited tumoricidal activity when co-cultured with B16F10 mouse melanoma cells, suggesting a potential role in tumor immunity (Fig. 4h-i). Furthermore, transferring IL-12/IL-18-stimulated OT-II IL-7R^{sup-hi} naïve CD4 T cells into B16F10 tumor-bearing mice demonstrated effective tumor rejection, while IL-7R^{low} naïve CD4 T cells did not result in significant tumor reduction (Fig. 4j-k). As IL-12 and IL-18 stimulation led to dominant IFN- γ expression by naïve CD4 T cells, we hypothesized that IFN- γ might play an important role in mediating tumoricidal effect. After injecting B16F10 cells into C57BL/6 mice, transferring IL-12/18-activated naïve CD4 T cells significantly inhibited tumor infiltration and growth in lung tissue. Conversely, IFN- γ knockout (KO) naïve CD4 T cells stimulated by IL-12 and IL-18 failed to exert any tumoricidal effect (Fig. 4l-m), and there was no elevated expression of T-bet and granzyme B upon IL-12 and IL-18 stimulation (Fig. 4n). Additionally, the MHC-II blocking antibody did not significantly reverse the anti-tumor effects of IL-12/18-activated naïve CD4 T cells (Supplementary Fig. 8a-b), and OT-II naïve T cells without IL-12 and IL-18 stimulation were unable to reject B16-OVA cells in the lungs (Supplementary Fig. 8c-d). These results suggest that IL-12 and IL-18 efficiently stimulate IL-7R^{sup-hi} naïve CD4 T cells, enabling them to contribute to the tumoricidal effect and mount a robust cytotoxic immune response via IFN- γ signaling rather than through direct tumor recognition.

Human relevance and clinical significance of IL-7R^{sup-hi} naïve CD4 T cells

Expanding upon our findings in humans, we confirmed that IL-7R^{sup-hi} naïve CD4 T cells within the healthy human PBMCs exhibited elevated levels of CD97 expression (Fig. 5a-c, Supplementary Fig. 1d), with a more evident difference compared to that observed in mouse naïve CD4 T cells. We also confirmed that human IL-7R^{sup-hi} naïve CD4 T cells are highly responsive to IL-12 and IL-18, producing IFN- γ (Fig. 5d). Therefore, human IL-7R^{sup-hi} naïve CD4 T cells are phenotypically and functionally comparable to their mouse counterparts. Then, we sought to explore the clinical significance of the IL-7R^{sup-hi} naïve CD4 T cell population and its potential correlation with human diseases. Firstly, we analyzed scRNA-seq datasets from PBMCs of two independent cohorts of gastric intestinal (GI) cancer patients and biliary tract (BT) cancer patients who had received anti-PD-1 treatment ($n = 21$) (Fig. 5e, Supplementary Fig. 9a-b). The proportion of conventional T cell subtypes did not differ significantly between responders and non-responders (Fig. 5f). However, the proportion of C1 (IL-7R^{sup-hi}) naïve CD4 T cells was significantly higher in the responder group compared to the non-responder group (p -value = 1.73×10^{-4} , fisher's exact test) (Fig. 5g). Motivated by these results, we constructed a classification model of anti-PD-1 therapy response using the proportion of various T cell types including naïve CD4 T cell subclusters. Projection of patients into the PLS-DA based latent space revealed that responders and non-responders were not distinguishable when only the proportion of conventional T cell types were included as feature sets without naïve T



cell subclusters (Fig. 5h). However, incorporating the proportion of naïve CD4 T cell subclusters allowed the groups to be clearly distinguished (Fig. 5i). During cross-validation, the model accuracy was increased from around 70% to above 90% by incorporating naïve CD4 T cell subcluster proportions (Fig. 5j), with significant increase of C1 proportion in responders (Fig. 5k). To further validate our data, we

made use of CITE-seq data from PBMCs of BT cancer patients and annotated T cell subsets using protein markers to extract protein marker-based naïve CD4 T cells. This time, we performed unsupervised clustering without relying on label transfer to define naïve T cell subclusters and observed consistent reproducibility in distinguishing responders and non-responders (Supplementary Fig. 10a-m).

Fig. 3 | IL-7R^{sup-hi} naïve CD4 T cells are poised for a Th1 response. **a** Graph showing significantly over-represented pathways for up-regulated genes in murine naïve CD4 T cell C1 cluster. The blue bar represents pathway significance (-log₁₀ p.adjust), while the red line shows the count of up-regulated C1 genes involved in each pathway. Statistical significance was calculated using a one-sided hypergeometric test in ConsensusPathDB. **b–c** Cytokine receptor expression in IL-7R^{sup-hi} and IL-7R^{low} naïve CD4 T cells analyzed by **(b)** qPCR ($n = 4$) and **(c)** flow cytometry. **d** Representative CTV histogram of MOG- or OVA-activated 2D2 or OT-II IL-7R^{sup-hi} and IL-7R^{low} naïve CD4 T cells. **e–f** Representative dot plots of IL-7R^{sup-hi} and IL-7R^{low} naïve CD4 T cells stimulated with IL-18 and/or IL-12 **(e)** with anti-CD3/CD28 for 4 days or **(f)** with IL-7 for 5 days. **g** IFN- γ concentration in cultured media measured by ELISA ($n = 8$ for TCR, $n = 7$ for without TCR, with independent experiments conducted for each n). **h–j** Adoptive naïve CD4 cell transfer to Rag-deficient mice. **h** Weekly weight normalized to day 0 (IL-7R^{sup-hi} $n = 11$; IL-7R^{low} $n = 9$; individual mice). **i, j** Cytokine production of transferred naïve CD4 T cells in **(i)** spleen and **(j)**

mLN (IL-7R^{sup-hi} $n = 11$, IL-7R^{low} $n = 9$). **k** Two differentially accessible chromatin regions (DACR) potentially regulating *Ifngr1* expression. The first two tracks show normalized accessibility signals for C1 and others from murine naïve CD4 T cell scATAC-seq. Two DACRs are highlighted with arrows. Tracks 3–4 display gene locus positions and peaks from the scATAC-seq, while the fifth track shows inferred cis-regulatory interactions linking DACRs to *Ifngr1*. scATAC-seq data were reanalyzed from a previous study²⁷. **l** DACR2 center +25bp region was zoomed in to show the TF motifs located in the middle of this region. The locations and the sequence logos of the Tbx21, Stat1, Stat4, and Eomes motifs were highlighted that are matched with the reference genome sequence in the reverse complementary manner. Motifs were annotated based on JASPAR database⁷⁴. Data presented as the mean \pm S.D. Statistical significance of in 3 **(b–j)** was confirmed by the two-sided Mann-Whitney *U*-test ($*p < 0.05$, $**p < 0.01$) and two-way-ANOVA done in 3 h. Source data and exact statistics value are provided as a Source Data file.

Additionally, analyzing scRNA-seq data of PBMCs from multiple sclerosis (MS) patients and healthy controls revealed the presence of various effector and naïve CD4 T cell populations in both groups (Fig. 5l). Interestingly, the proportion of the C1 population was significantly higher in MS patients compared to healthy donors (p -value = 1.1×10^{-3} , t -test) (Fig. 5m). Similar to our approach for classifying immunotherapy response, we developed a PLS-DA model to classify MS patients and healthy donors using cell-type proportions as feature sets. While known T cell populations like Th1, Th17, Tfh, Treg could distinguish between MS patients and healthy donors with an accuracy of around 80% (Fig. 5n–o), incorporating naïve CD4 T cell subcluster information increased the accuracy to over 90% (Fig. 5p–q). Naïve CD4 T cell subpopulations such as C1 and C0 were the most significant contributors to distinguishing between healthy individuals and MS patients, even more so than Th1 or Th17 cells (Fig. 5r). Finally, analysis of PBMCs from COVID-19 patients showed a positive correlation between disease severity and an increased abundance of the IL7R^{sup-hi} naïve CD4 T cell subcluster (H3-mouse C1) (Supplementary Fig. 11a–e). Interestingly, we observed a comparable increase in IFN-induced naïve CD4 subclusters (H2-mouse C3), which correlates with a population reported in a recent study²⁹. Thus, the IL-7R^{sup-hi} population in naïve CD4 T cells has significant human relevance and in silico PLS-DA modeling using the proportion of naïve T cell subclusters derived from scRNA-seq as key feature sets could be a valuable tool for estimating long-term drug responsiveness and distinguishing between inflammatory disease patients and healthy individuals.

Discussion

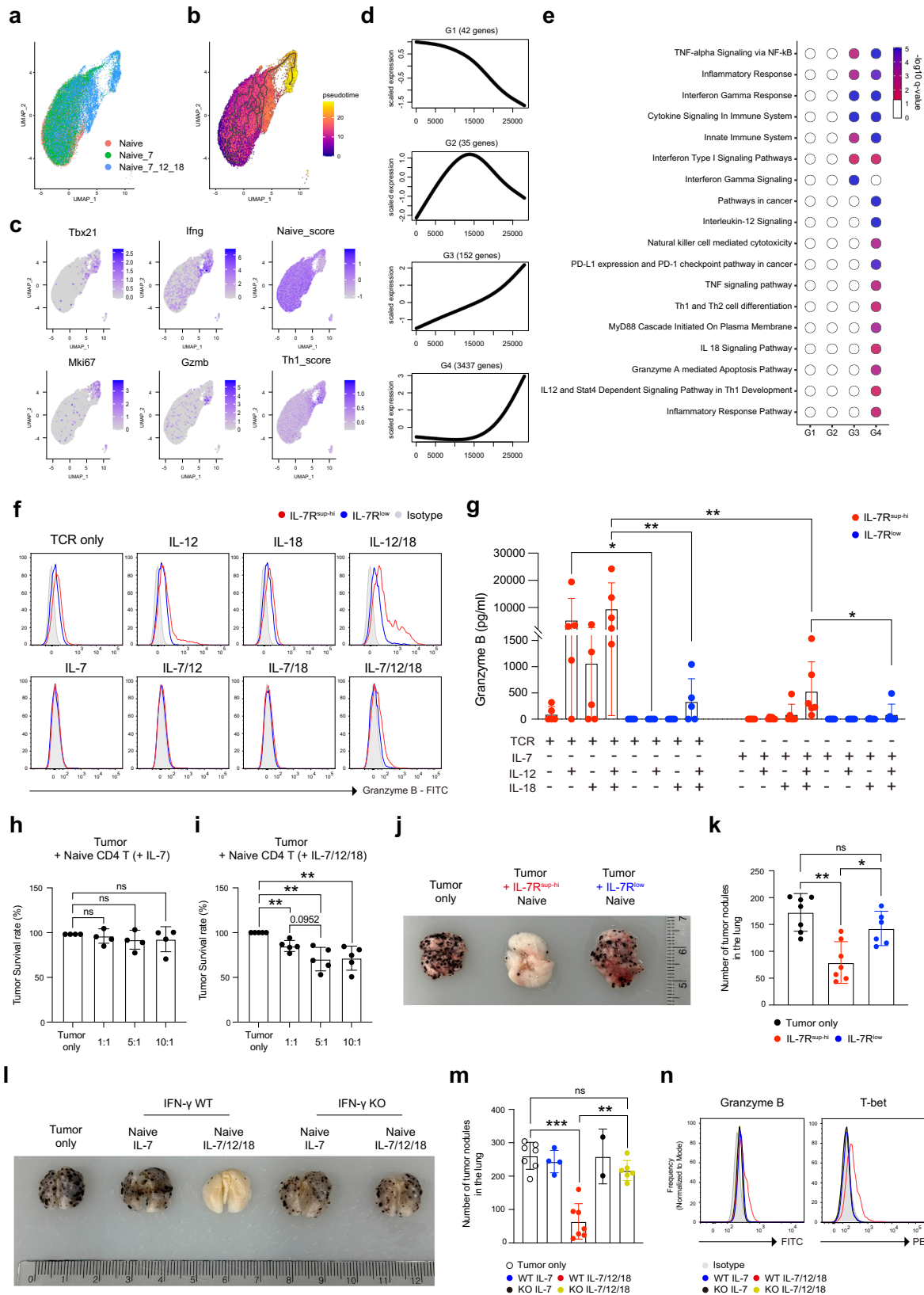
In this study, we reveal a multitude of distinct subpopulations within both murine and human peripheral naïve CD4 T cells, prominently spotlighting the previously unknown IL-7R^{sup-hi} subpopulation. This population is characterized by its expression of IL-18R and Ly6C, with higher expression of CD97, and lower levels of CD5, indicating less self-reactivity. These specific markers denote a subset notably involved in sensitive Th1 responses (T_{N1}), which shows a positive correlation in various inflammatory disease contexts. Noteworthy is the correlation found between this IL-7R^{sup-hi} subpopulation and the responsiveness to anti-PD-1 therapy in cancer patients, as well as its association with the disease state observed in multiple sclerosis and COVID-19 patients. This finding underscores that naïve CD4 T cells have not just remained naïve but instead reflects various inflammatory traces.

Naïve T cells have traditionally been defined as a resting, quiescent, and homogeneous population that has not yet encountered its antigen^{8–10}. The fundamental question of whether naïve T cells are truly a homogeneous population is the main initial inquiry of this study, and it has been partly explored in previous research by other groups. Early studies noted that CD5 expression exhibited stage-specific dynamics during thymic T cell maturation and aligns with changes in self-

reactivity of naïve T cells³³. This led to further investigation into how CD5 expression levels affect naïve T cell function. It was found that naïve CD8 T cells with high CD5 expression, which indicates higher MHC-I-mediated self-reactivity, demonstrate increased sensitivity to common gamma chain cytokines, which support their survival and proliferation¹⁵, as well as heightened responsiveness to antigen stimulation⁴⁰. Similarly, CD4 T cells with high CD5 expression show an intrinsically elevated TCR response^{41,42}, and are more prone to Tfh differentiation due to a chromatin accessibility landscape formed during thymic development¹⁸. Also, CD5^{hi} naïve CD4 T cells are inclined towards Treg differentiation due to their high self-affinity⁴³. Additionally, naïve CD4 T cells experiencing strong basal TCR signaling defined by *Nur77* expression exhibit an anergy-like responsiveness to further TCR stimulation²⁰. Consistently, we could also identify the *Cd5*^{hi} cluster (C2) in naïve CD4 T cells with high expression of *Nr4a1* (*Nur77*) and lower expression of *Il7r* (IL-7R), which constitutes only 4.88% of total naïve CD4 T cells. By comparison, *Il7r*^{sup-hi} cells (C1) express a low level of *Cd5*, suggesting less self-reactivity.

Ly6C expression, another marker of self-reactivity, decreases in naïve CD4 T cells due to MHC-II mediated self-recognition, and Ly6C⁺ cells are more likely to differentiate into regulatory T cells or Th17 cells in a Ca²⁺ signaling-dependent manner^{16,44}. IL-7R^{sup-hi} naïve CD4 T cell subpopulation, which expresses high levels of Ly6C, indicates less self-reactive character (Ly6C⁺ and CD5^{lo}). Notably, in TCR transgenic mice (OT-II, 2d2), all naïve CD4 T cells express high levels of Ly6C (Supplementary Fig. 12a, b), suggesting the absence of tonic TCR signaling, whereas the IL-7R levels are comparable to those in wild-type naïve CD4 T cells (Supplementary Fig. 12c).

However, neither CD5 nor Ly6C alone adequately characterizes Th1-poised naïve CD4 T cells, particularly regarding IL-18 sensitivity. First, IL-7R expression levels showed no significant differences between subpopulations divided by the top 20% versus low 20% of CD5 expression, or by Ly6C expression levels into positive and negative groups (Supplementary Fig. 13a). While the C2 (*Cd5*^{hi}) cluster in scRNA-seq showed low IL-7R levels, its small size (4.88% of total naïve CD4 T cells) likely limited its influence on IL-7R levels within the top 20% CD5-expressing cells. IFN- γ production was modestly increased in CD5^{low} or Ly6C⁺ cells when stimulated with anti-CD3 and anti-CD28 antibodies in addition to IL-12 and IL-18, aligning findings from a previous study under Th1 polarization conditions with IL-12⁴⁵. However, no difference was observed with cytokine stimulation alone in CD5^{high} vs CD5^{low} cells (Supplementary Fig. 13b). Notably, increased IFN- γ production induced by IL-12 and IL-18 is specifically observed in the IL-7R^{sup-hi} population within the Ly6C⁺ naïve CD4 T cells (Supplementary Fig. 13c). Even in TCR transgenic mice, where all naïve CD4 T cells express Ly6C, only IL-7R^{sup-hi} subpopulation responds sensitively to IL-12 and IL-18 (Supplementary Fig. 13d). This suggests that Th1-poised naïve CD4 T cells (IL-7R^{sup-hi}) constitute a unique population with steady-state IL-18R expression, similar to NK cells⁴⁶, ILC1⁴⁷, CD8



T cells⁴⁸, making them particularly responsive to IL-18 stimulation and important contributors to type 1 immunity.

We and others have recently reported heterogenous subpopulations within steady-state CD44^{hi} memory-phenotype CD4 T cells, as analyzed by single-cell RNA-sequencing^{49,50}, revealing distinct effector-like memory-phenotype cells that can be activated by cytokines such

as IL-1/IL-23, IL12/IL-18 to produce effector cytokines^{49,51-54}. Notably, a single-cell RNA sequencing study of murine naive CD4 T cells revealed distinct subpopulations, including a “memory-like naive” subset that expresses *Tcf7*, *Bcl2*, and *Il7r* as marker genes, which was observed to be increased in VISTA-deficient mice²⁷. Since VISTA is a crucial regulator of naive CD4 T cell quiescence, the memory-like naive

Fig. 4 | IL-7R^{sup-hi} naïve CD4 T cells exhibit cytolytic Th1 character, contributing to tumor rejection in response to IL-12/IL-18. **a** UMAP of scRNA-seq data from murine naïve CD4 T cells under different conditions including control (Naïve) and IL-7 treatment (Naïve_7) and IL-7 + 12 + 18 treatment (Naïve_7_12_18). Cells were colored by conditions. **b** UMAP of scRNA-seq colored by pseudotime inference using Monocle3. **c** UMAP of naïve CD4 T cells colored by the expression level of various marker genes (Tbs2l, Ifng, Mki67, Gzmb) or gene sets (Naïve score, Th1 score). **d** Group of pseudotime dependently changed genes clustered by k-means clustering algorithm. X-axis and y-axis correspond to rank of cells ordered by pseudotime and z-score of expression level, respectively. **e** Dot plot showing the significance of pathway over-representation by the marker genes of each clusters. **f** Representative histogram of granzyme B expression. **g** The concentration of granzyme B in cultured media was measured by ELISA ($n = 5$ for TCR, $n = 6$ for without TCR, with independent experiments conducted for each n). **h, i** Tumor

survival rate measured by CCK-8 in co-culture with B16F10 and naïve CD4 T cells cultured by **(h)** IL-7 ($n = 4$) or **(i)** IL-7/12/18 ($n = 5$) at indicated E:T ratio (E: effector, T: tumor). **j** Representative lung image and **k** the number of pleural tumor nodules in IL-12/18-activated OT-II IL-7R^{sup-hi} or IL-7R^{low} naïve CD4 T cell injected B16F10 lung metastasis model (tumor only $n = 7$, IL-7R^{sup-hi} $n = 7$, IL-7R^{low} $n = 6$). **l** Representative lung image and **m** the number of pleural tumor nodules in IL-12/18-activated WT or IFN- γ KO naïve CD4 T cell injected B16F10 lung metastasis model (tumor only $n = 7$, WT IL-7 $n = 4$, WT IL-7/12/18 $n = 7$, KO IL-7 $n = 2$, KO IL-7/12/18 $n = 6$). **n** Representative histogram of granzyme B and T-bet in WT or IFN- γ KO naïve CD4 T cells stimulated by cytokines indicated, analyzed by flow cytometry. DATA are presented as the mean \pm S.D. Statistical significance of in **4 (g-m)** was confirmed by the two-sided Mann-Whitney U -test ($*p < 0.05$, $**p < 0.01$, $***p < 0.001$). Source data and exact statistics value are provided as a Source Data file.

population in the VISTA-deficient mice was suggested as their increased sensitivity to antigen stimulation²⁷. We suspect that the annotated cluster might overlap with IL-7R^{sup-hi} cells, at least in part; however, the study lacked functional analysis of those cells and their relevance or clinical implications in humans. Along with the comprehensive scRNA-sequencing analysis, we demonstrate the protein marker-based clustering of naïve CD4 T cells confirms the existence of the IL-7R^{sup-hi} cluster expressing IL-18R within steady-state naïve CD4 T cells. Furthermore, in the Rag KO adoptive transfer model, where all naïve T cells transition into a memory-phenotype (CD44^{hi}), the IL-7R^{sup-hi} naïve CD4 T cell subpopulation preferentially promotes a Th1 response. Additionally, the scATAC-sequencing analysis revealed that these populations have elevated chromatin accessibility at the *Ifngr1*, with accessible regions predicted to be binding sites for Th1-related transcription factors such as Eomes and T-bet, even though these TFs are not expressed at the protein level in the steady state. Collectively, these findings suggest that IL-7R^{sup-hi} naïve CD4 T cells exhibit Th1-poised characteristics, which we designate as Th1-poised naïve (T_{NI}) cells. Thus, from an integrated perspective that considers both our findings and previously reported knowledge, naïve CD4 T cells exhibit cumulative history of peripheral stimulation as well as an intrinsic property, which influences subpopulations of naïve CD4 T cells and their predestined differentiation into effector or memory T cells.

The characteristics of IL-7R^{hi}, Tcf7^{hi} with Bcl-2 expression is somewhat similar to more differentiated cells, such as stem-like memory CD8⁵⁵ or CD4^{56,57} T cells with a naïve phenotype as previously reported. However, it is generally improbable that specific pathogen-free (SPF) housed intact mice would harbor these cells under steady-state conditions. Moreover, virtual memory T cells⁵⁸ are memory-phenotype T cells expressing CD122 and CD44, and stem-like memory cells with a naïve phenotype would likely constitute a very small population, if present at all. Approximately 15–30% of naïve CD4 T cell subpopulations we identified in this study exhibit high clonal diversity comparable to the other naïve T cell populations suggesting an absence of clonal expansion. In addition, IL-7R^{sup-hi} naïve CD4 T cells exhibit lower CD5 and higher Ly6C expression, suggesting less-self reactivity. Moreover, CD44^{hi} memory-phenotype CD4 T cells do not express Ly6C, collectively indicating that IL-7R^{sup-hi} naïve CD4 T cells have not undergone antigen stimulation.

Regarding the tumoricidal response of IL-7R^{sup-hi} naïve CD4 T cells stimulated by IL-12 and IL-18, we observed increased production of IFN- γ and granzyme B, similar to cytolytic CD4 T cells identified in cancer studies^{59,60}. In addition, IFN- γ -deficient IL-7R^{sup-hi} naïve CD4 T cells were unable to reduce tumor growth or secrete granzyme B in both in vitro and in vivo experiments, and together with findings from the MHC-II blocking antibody and B16-OVA experiments, this suggests that the anti-tumor effects of IL-12/18-activated naïve CD4 T cells did not rely solely on direct tumor recognition. Upon administration of anti-CD8 and anti-NK1.1 antibodies, the tumor-reducing effect of IL-12/18-activated naïve CD4 T cells was almost completely abolished

(Supplementary Fig. 14), indicating that these cells enhance the activity of other cytolytic cells, including CD8 T cells and NK cells, in vivo. Nonetheless, it remains clear that T_{NI} cells possess the potential to mount a robust cytotoxic immune response, either directly or indirectly.

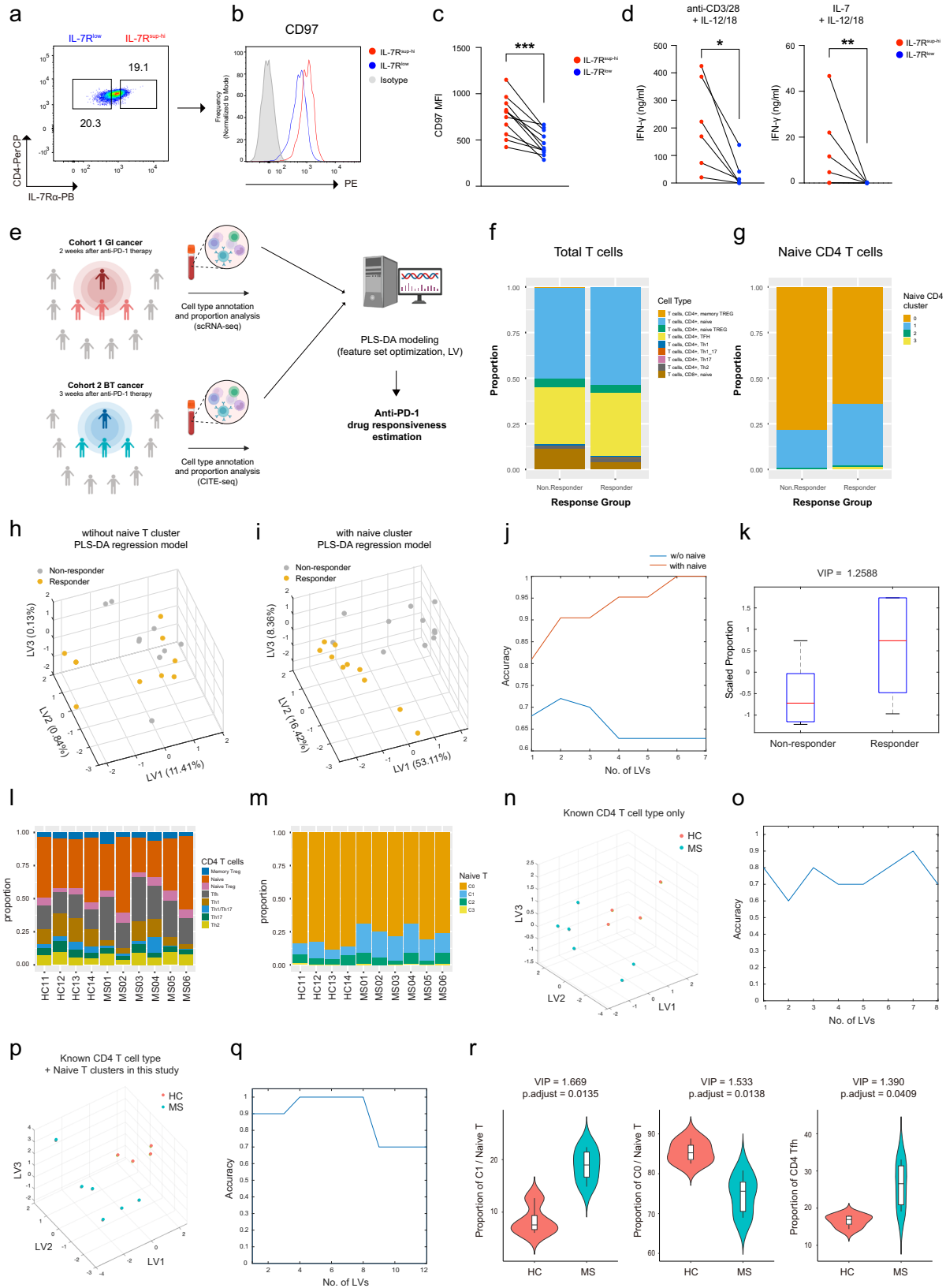
The lack of a clinical biomarker for predicting responsiveness to anti-PD-1 therapy complicates the identification of effective drugs for individual patients, posing a substantial obstacle to the widespread adoption of costly cancer immunotherapy treatments. Therefore, previous single-cell-based studies have revealed correlated biomarkers, such as CD8 T cells^{61,62} and Treg cells⁶³ in tumor micro-environment, which could potentially serve as predictive markers for anti-PD-1 therapy responsiveness. Subsequently, recent studies suggest the use of non-invasive PBMC-based biomarkers, revealing a correlation between effector T cells, such as CD8 T cells⁶, and the subset of monocytes (CD14⁺ CD16⁺ HLA-DR^{hi})⁷, and their association with the responsiveness to anti-PD-1 therapy. Here, we reveal that the IL-7R^{sup-hi} naïve CD4 T cell subpopulation in human PBMC, identified as a common subpopulation between mice and humans highly expressing CD97 and IL-7R, could serve as a marker in distinguishing responders from non-responders to anti-PD-1 therapy in two independent cohorts of gastrointestinal (GI) cancer patients and biliary tract (BT) cancer patients. Using this property, we successfully developed a supervised model (PLS-DA model) to estimate anti-PD-1 therapy responsiveness in cancer based on the proportion of naïve CD4 T cell subpopulation as key feature sets. Furthermore, this modeling approach is effectively applicable to distinguish autoimmune disease and viral infection patients from healthy groups. These results demonstrate that the proportion of IL-7R^{sup-hi} naïve CD4 T cells reflects the inflammatory context, facilitating the estimation of disease states and drug responsiveness. Further studies are necessary to analyze naïve CD4 T cell subpopulations in other cancer types and inflammatory diseases, as this holds significant value for non-invasively collected clinical information for developing personalized treatment strategies along with other clinical information.

In summary, we identified a unique subpopulation of IL-7R^{sup-hi} naïve CD4 T cells, present in both humans and mice, that is peripherally induced and constitutes a significant portion of naïve CD4 T cell population and exhibit Th1-poised characteristics, which we refer to as T_{NI} cells. Our findings underscore the importance of the proportion of these naïve CD4 T cells in the blood as a potential biomarker for estimating therapeutic outcomes and distinguishing disease states. This highlights the role of shaped naïve CD4 T cell heterogeneity, including fate-biased characteristics, in the context of inflammation.

Methods

Mice

All mouse experimental procedures used in this study were approved by the Institutional Animal Care and Use Committee of Hanyang University (2021-0045 A, 2022-0041 A, 2022-0132 A, 2023-0132 A).



C57BL/6J mice were purchased from DBL (Chungcheong-buk-do, Korea). Rag^{-/-}, 2D2 TCR-transgenic mice, and OT-II TCR-transgenic mice were purchased from The Jackson Laboratory (Bar Harbor, Me, USA). IFN-γ^{-/-} mice were provided by Prof. Jongsoo Lee (Soonchunhyang University) and Stat1^{-/-}, Ifnar^{-/-}, Ifnar.lfng^{-/-} mice were provided by Prof. Jae-Ho Cho (Chonnam National University) Male

and female mice aged 7–12 weeks were used, and all mice within the same experiment were matched for both age and sex. The study design and analysis did not account for sex as a variable. No sex-specific data were collected in this study, as investigating sex-based effects was outside the scope of our research focus. Mice were housed and bred in a specific pathogen-free animal facility at

Fig. 5 | Human relevance and clinical impact of IL-7R^{sup-hi} naïve CD4⁺ T cell population in PBMC. **a–b** Representative dot plot of IL-7R gating strategy (**a**) and histogram of CD97 expression (**b**) on human naïve CD4 T cells in PBMC. **c** MFI value of CD97 ($n = 11$, individual human PBMCs). **d** IFN- γ concentration in supernatants of IL-12/18-cultured FACS-sorted IL-7R^{sup-hi} or IL-7R^{low} naïve CD4 T cells measured by ELISA ($n = 6$, pooled with age/sex-matched $n = 16$). **e** Schematic of experimental design. Created in BioRender. Choi, J. (2025) <https://BioRender.com/v86d887> **f–g** Proportion of conventional T cells (**f**) and naïve CD4 subpopulations (**g**) among non-responders or responders to anti-PD-1 immunotherapy. **h–i** 3D PLS-DA plots distinguishing non-responders (gray-colored dots, $n = 11$) and responders (yellow-colored dots, $n = 10$), based on conventional T cells (**h**) or including naïve CD4 T cell subpopulations (**i**). **j** Accuracy plot after leave-one-out cross validation (blue/orange lines: models without/with naïve CD4 subpopulations) **k** Boxplots of C1 (IL-7R^{sup-hi}), a significant PLS-DA feature distinguishing non-responders and responders. Boxplots were defined by minima = 25th percentile - 1.5X interquartile range

(IQR), maxima = 75th percentile + 1.5X IQR, center = median, and bounds of box = 25th and 75th percentile. **l–m** Proportion of effector T cells in healthy controls (HC) vs. multiple sclerosis (MS) patients (**l**) and naïve CD4 clusters (**m**). **n–o** 3D PLS-DA plots (**n**) and accuracy plots (**o**) distinguishing HC and MS using conventional T cells without naïve CD4 subpopulations. **p, q** 3D PLS-DA plot (**p**) and accuracy plot (**q**) for models incorporating naïve CD4 subpopulations. **r** Violin plots of the top three PLS-DA features distinguishing HC ($n = 4$) and MS ($n = 6$). Y-axis indicates cluster proportions; differences are evaluated by two-sided t -test followed by Benjamini Hochberg multiple correction. Boxplots were defined by minima = 25th percentile - 1.5X interquartile range (IQR), maxima = 75th percentile + 1.5X IQR, center = median, and bounds of box = 25th and 75th percentile. DATA are presented as the mean \pm S.D. Statistical significance of in 5c and 5d was confirmed by the two-sided Mann-Whitney U -test ($*p < 0.05$, $**p < 0.01$, $***p < 0.001$). Source data and exact statistics value are provided as a Source Data file.

Hanyang University under controlled conditions with a constant temperature ($21 \pm 1^\circ\text{C}$) and humidity ($50 \pm 5\%$) and a 12 h light/dark cycle with regular chow and autoclaved water.

Mouse naïve CD4 T cell isolation and in vitro stimulation

Mouse splenocytes were harvested and enriched by CD4 T cell isolation kit (#130-104-454, Miltenyi Biotec) in negative selection. Then using the enriched CD4⁺ T cells, naïve CD4⁺ T cells (TCR- β^+ , CD4⁺, CD25⁻, CD62L^{hi}, CD44^{low}, IL-7R^{sup-hi} or ^{low}) were sorted using FACS (ARIA Fusion, BD Bioscience). Purified naïve CD4⁺ T cells were seeded in 96-well-plate in 2×10^5 cells/well and stimulated by plate-bound anti-mouse CD3/CD28 (5 $\mu\text{g}/\text{ml}$, BD bioscience) and/or mouse IL-7 (10 ng/ml, R&D systems), IL-12 (20 ng/ml, peprotech), IL-18 (20 ng/ml, R&D systems) as indicated in Fig. 2. For Fig. 4, thymocytes were stimulated with the following cytokine conditions: mouse IFN- γ , mouse IL-12 (20 ng/ml, peprotech), mouse IL-27, mouse IFN- β , mouse IL-7, mouse IL-18, mouse IL-4, mouse IL-25, mouse IL-33, mouse TSLP, mouse IL-6, mouse IL-1 β , mouse IL-23, mouse IL-2 (20 ng/ml, R&D systems), TGF- β (10 ng/ml, R&D systems). For further intracellular cytokine staining, cultured cells were re-stimulated by PMA (P-1585; Sigma, 40 ng/ml) and ionomycin (I-0634; Sigma, 0.8 $\mu\text{g}/\text{ml}$) in the presence of monensin (51-2092kz; BD Bioscience).

In vitro T cell proliferation

Purified 2d2 (V β 11⁺, CD4⁺, CD25⁻, CD62L^{hi}, CD44^{low}, IL-7R^{sup-hi} or ^{low}), OT-II (V α 2⁺, CD4⁺, CD25⁻, CD62L^{hi}, CD44^{low}, IL-7R^{sup-hi} or ^{low}), or WT naïve CD4⁺ T cells were labeled with CellTrace Violet (CTV; #C34557, ThermoFisher) using the provided protocol. The CTV-labeled 2d2 or OT-II naïve CD4⁺ cells were then stimulated by MOG₃₅₋₅₅ (10 $\mu\text{g}/\text{ml}$, 1 $\mu\text{g}/\text{ml}$) or OVA (10 $\mu\text{g}/\text{ml}$, 1 $\mu\text{g}/\text{ml}$) for 3 days, in the presence of irradiated (30 Gy) non-CD4 cells obtained from the splenocytes of the corresponding mice. The CTV-labeled WT naïve CD4⁺ cells were stimulated by anti-CD3/28 (5 $\mu\text{g}/\text{ml}$, BD bioscience) for 3 days.

Flow cytometry

Cell staining was performed at 4°C for 30 min using the following monoclonal antibodies:

anti-mouse CD45 Pacific Blue (30-F11; Biolegend, dilution 1:500) #103125 lot B431357
 anti-mouse TCR gamma/delta PE (GL-3, Invitrogen, dilution 1:500) #118107 lot B404051
 anti-mouse CD25 FITC (PC61.5; Biolegend, dilution 1:500) #102005 lot B432520
 anti-mouse CD25 PE (PC61.5; Invitrogen, dilution 1:500) #12-0251-82 lot 2506997
 anti-mouse CD19 PE (6D5, Biolegend, dilution 1:500) #115507 lot B433650
 anti-mouse NK1.1 PE (PK136; Biolegend, dilution 1:200) #108707 lot B399975

anti-mouse TCR- β chain PE/cy7 (H57-597; Biolegend, dilution 1:200) #109222 lot B394527
 anti-mouse TCR V β 11 APC (RR3-15; Invitrogen, dilution 1:500) #17-5827-82 lot 2372924
 anti-mouse TCR Va2 APC (B20.1; Invitrogen, dilution 1:500) #17-5812-82 lot B2367480
 anti-mouse CD4 PerCP (GK1.5; Biolegend, dilution 1:500) #100432 lot B415616
 anti-mouse CD4 BV605 (GK1.5; Biolegend, dilution 1:500) #100451 lot B386408
 anti-mouse CD8 PE (53-6.7; Biolegend, dilution 1:500) #100707 lot B409666
 anti-mouse CD8 BV785 (53-6.7; BD Bioscience, dilution 1:500) #563332 lot 3037490
 anti-mouse CD62L PE/cy7 (MEL-14; Invitrogen, dilution 1:500) #25-0621-81 lot2405782
 anti-mouse CD62L PE/Dazzle (MEL-14; Biolegend, USA, dilution 1:500) #104448 lot B436257
 anti-mouse/human CD44 APC/cy7 (IM7; Biolegend, dilution 1:500) #103028 lot B399687
 anti-mouse CD127 FITC (A7R34; Biolegend, dilution 1:200) #135008 lot B424331
 anti-mouse CD127 BV785 (A7R34; Biolegend, dilution 1:200) #135037 lot B429495
 anti-mouse Ly6C FITC (HK1.4; Biolegend, dilution 1:200) #128005 lot B419426
 anti-mouse CD5 Pacific Blue (53-7.3, Invitrogen, dilution 1:1000) #100641 lot B390906
 anti-mouse CD218 PE (A17071D, Biolegend, dilution 1:200) #157904 lot B397549
 anti-mouse CD97v2 APC (587702; R&D systems, dilution 2 μl per test) #FAB3734A lot ABOO022204
 anti-mouse IL-12Rb1 PE (FAB1998P; R&D systems, dilution 2 μl per test) #FAB1998P lot ABJC0421092
 anti-mouse IL-12Rb2 APC (FAB1959A; R&D systems, dilution 2 μl per test) #FAB1959A lot ADTU0419091
 anti-mouse IFN-gR1 PE (2E2; Invitrogen, dilution 1:100) #12-1191-80 lot 2489323
 After staining of the surface markers, the cells were fixed and permeabilized using eBioscience™ FOXP3/Transcription Factor Staining Buffer Set (00-5523-00; Invitrogen) for 30 min at RT. Intracellular staining was performed for 30 min at RT using the following monoclonal antibodies:
 anti-mouse IFN-g APC (XMG1.2; Invitrogen, dilution 1:500) #17-7311-81 lot 2410272
 anti-mouse IFN-g BV711 (XMG1.2; Biolegend, dilution 1:500) #17-7311-81 lot B427828
 anti-T-bet PE (4B10; Biolegend, dilution 1:50) #644809 lot B385153

anti-mouse granzyme B FITC (NGZB; Invitrogen, dilution 1:200) #11-8898-80 lot 2547936

anti-mouse IL-17A APC (eBio17B7; Invitrogen, dilution 1:200) #17-1777-81 lot 2702171.

Stained cells were acquired by flow cytometry (FACS Symphony A3, BD Bioscience), and data were analyzed using FlowJo software (version 10.10). For FACS data clustering analysis, we utilized FlowJo plugins (UMAP v4.0.4, FlowSOM v4.1.0) in the FlowJo software (version 10.10). All antibodies used in this study were commercially available and were validated by the manufacturers.

Cytokine measurement (ELISA)

Secreted cytokines in the cell-cultured supernatant were measured using a mouse IFN- γ ELISA kit (430801; Biolegend), mouse granzyme B ELISA kit (88-8022-88; Invitrogen), human IFN- γ ELISA kit (430104; Biolegend) according to the manufacturer's instructions. Well-plates (90181; Corning Costar) were overnight coated with capture antibodies at 4 °C and blocked with assay diluent 1X for 1 h at RT. Diluted samples and standards were incubated at RT for 2 h, and then detection antibodies and streptavidin-HRP were sequentially added. Then 1X substrate solution (TMB) was added and incubated for 15 min to 30 min; then the enzymatic reaction was stopped by 2NH₂SO₄. The optical density was detected at 450 nm. Through all steps, plates were washed 4 times with PBS-T buffer (1X PBS, 0.05% Tween-20).

Quantitative RT-PCR

Total RNAs were isolated with RNeasy Mini Kit (74106; Qiagen) following the manufacturer's protocol. Extracted RNAs were reverse transcribed using ReverTra Ace qPCR RT Master Mix (FSQ-201; Toyobo). RT-PCR was performed on QuantStudio 1 (Applied Biosystems) using iQ SYBR Green Supermix (1708882; Bio-Rad). The target gene expression levels were normalized with GAPDH. Specific primer sequences are provided in the Supplementary Data 1.

Tissue processing and preparation

After mice were sacrificed and perfused with PBS, lung tissues were obtained and digested using the Lung Dissociation Kit (130-095-927; Miltenyi Biotec), GentleMACS C Tubes (130-093-237; Miltenyi Biotec), and the GentleMACS Octo Dissociator with Heaters (130-096-427; Miltenyi Biotec) at the Biospecimen-Multiomics Digital Bioanalysis Core Facility of Hanyang University, following the protocol provided by the manufacturers. Subsequently, lung lymphocytes were enriched using Percoll gradients (30%, 70%). Lymphoid tissues, including the thymus, spleen, lymph nodes (iLN, mLN), and Peyer's patch, were collected and gently dissociated into single-cell suspensions for further analysis.

In vitro cytotoxicity assay (CCK-8)

B16F10 melanoma cells were seeded in 96-well plates with 100 μ L of complete DMEM (10% FBS, 1% penicillin-streptomycin) at 37 °C, 5% CO₂. Following activation of naive CD4⁺ T cells with IL-7 (10 ng/ml) and/or IL-12 and IL-18 (20 ng/ml) for 5 days, these cells were added to wells containing B16F10 cells at indicated effector-to-target (E:T) ratios. Remaining live cells were stained with CCK-8 (CK04; DOJINDO) and the tumor survival rate was calculated relative to the OD value of the B16F10-only control.

Adoptive cell transfer in pulmonary melanoma (B16F10) metastasis model

The B16F10 mouse melanoma cell line, derived from C57BL6/J mouse melanoma, was purchased from ATCC (CRL-6475). B16-OVA cell line was provided by Prof. Seung-Woo Lee, POSTECH, Pohang, Republic of Korea. B16F10 and B16-OVA cells are cultured in complete DMEM (10% FBS, 1% penicillin/streptomycin). Upon reaching 80–90% confluence in complete DMEM, cells were harvested, adjusted to a concentration

of 10⁶ cells/mL in pre-warmed PBS, and 5 \times 10⁵ cells were intravenously injected into the mice (day 0). After 2 days, IL-12 and IL-18-activated WT or IFN- γ KO naive CD4⁺ T cells were intravenously injected into the mice. On day 14, mice were euthanized, and the number of melanoma colonies, visualized as black dots on the lung surface, were counted. For antibody blocking or depletion experiment MHC-II blockade was performed by weekly i.p injection of 200 μ g of anti-MHC-II antibody (clone Y-3P, BioXcell, #BE0178, lot 796422M2). CD8 and NK depletion was performed by a single i.p injection of 200 μ g of anti-CD8 antibody (clone 2.43, BioXcell, #BE0061, lot 811521N1) and anti-NK1.1 antibody (clone PK136, BioXcell, #BE0036 lot 828622A2). All antibodies were diluted in the manufacturer-recommended pH Dilution Buffer (BioXcell). All antibodies used in this study were commercially available and were validated by the manufacturers.

Adoptive cell transfer colitis model

FACS-sorted naive CD4⁺ T cells were resuspended into pre-warmed PBS and 5 \times 10⁵ cells were intraperitoneally injected into the Rag^{-/-} mice. The weight of recipient mice measured one or two times per week. After 10 weeks mice were sacrificed, and CD4⁺ T cells from spleen, mLN, and colon were analyzed by flow cytometry.

Human peripheral blood mononuclear cell (PBMC) analysis

Studies of human PBMCs were exempted by the institutional review committee of Hanyang University (HYUIRB-202405-030). Human PBMCs were purchased from STEMCELL Technologies (ST70025.1). PBMCs analyzed in study were obtained from 16 healthy donors aged 23–43 years of different sex (8 male and 8 female). This study has not involved in sex- or race- based analysis. Cells were carefully handled according to the manufacturer's instructions. Cell staining was performed at 4 °C for 30 min using the following monoclonal antibodies:

anti-human CD3 BV605 (SK7; Biolegend, dilution 1 μ l per test) #344835 lot B376001
 anti-human CD4 PerCP (RPA-T4; Invitrogen, dilution 1 μ l per test) #300527 lot B403061
 anti-human CD8 BV785 (SK1; Biolegend, dilution 1 μ l per test) #344739 lot B365765
 anti-human CD25 APC/cy7 (BC96; Biolegend, dilution 1 μ l per test) #302613 lot B405392
 anti-human CD45RA FITC (5H9; BD Bioscience, dilution 1 μ l per test) #556626 lot 0079912
 anti-human CCR7 PE/cy7 (3D12; BD Bioscience, dilution 1 μ l per test) #557648 lot 2265749
 anti-human CD95 Alexa Fluor 700 (DX2; Biolegend, dilution 5 μ l per test) #305647 lot B374252
 anti-human CD127 Pacific Blue (A019D5; Biolegend, dilution 1 μ l per test) #351305 lot B274562
 anti-human CD97 PE (VIM3b; Biolegend, dilution 5 μ l per test) #336307 lot B364659.

Stained cells were acquired by flow cytometry (FACS Symphony A3, BD Bioscience), and data were analyzed using FlowJo software (version 10.10). All antibodies used in this study were commercially available and were validated by the manufacturers.

Human naive CD4⁺ T cell isolation and in vitro stimulation

Human PBMCs were harvested and naive CD4 T cells (CD3⁺, CD4⁺, CD8⁻, CD25⁻, CD45RA⁺, CCR7^{high}, IL-7R^{sup-hi} or low) sorted using FACS (ARIA Fusion, BD Bioscience). Purified naive CD4⁺ T cells were seeded in 96-well-plate in 1–1.5 \times 10⁵ cells/well and stimulated by plate-bound anti-human CD3/CD28 (5 μ g/ml, Invitrogen) and/or human IL-7 (10 ng/ml, R&D systems), IL-12 (20 ng/ml, peprotech), IL-18 (20 ng/ml, R&D systems).

Processing and analysis of scRNA-seq data

For the scRNA-seq of mouse Naive T cells generated in this study, sequenced transcripts were aligned to mouse genome (mm10) and

quantified into count matrices using Cellranger count workflows (10X Genomics, v.6.1.1). The R package Seurat (v.4.3.0)⁶⁴ was used for quality checking of single cells and unsupervised clustering of the data (Supplementary Fig 1b). To obtain high-quality cells, only those cells with >750 genes, >1500 total UMI counts, and a percentage of mitochondrial reads <5% were accepted. The normalization and scaling were performed using the Seurat package. When different conditions or tissue types were available, we integrated those data sets using the canonical correlation analysis (CCA) integration method in Seurat. Dimensionality reduction and marker gene identification were performed as standard using Seurat (adjusted p -value < 0.05, log₂ fold change > 0.25). Additionally, we performed ASURAT³⁶ analysis using default parameters for the functional annotation of each clusters and to create dimensionality reduction map based on the pathway activities of each cells. Pseudotime trajectory analysis was performed using Monocle3 (v.1.3.1)⁶⁵. TCR clonality analysis was performed using scRepertoire (v.1.7.2) with default parameters⁶⁶.

Analysis of publicly available scRNA-seq data from healthy human

We collected scRNA-seq data of T cells from healthy humans profiled by 10X genomics platform (GSE193770) and performed batch correction between donors using Harmony (v.0.1.0)⁶⁷. The quality checking, normalization, and scaling were done using Seurat⁶⁴. Previously known T cell types were annotated using SingleR³⁸ and DICE reference³⁹. Afterwards, we subset naïve T cells and performed unsupervised clustering, and differential gene expression analysis using Seurat⁶⁴. Cluster similarity analysis of mouse and human Naïve T cell clusters was performed by overlapping up-regulated genes for each cluster (adjusted p -value < 0.05, log₂ fold change > 0.1) and evaluating the significance of overlap using Fisher's exact test. In parallel, label transfer analysis was done using CCA method in Seurat⁶⁴ to transfer cluster label of mouse naïve T clusters onto human naïve T cells. For this analysis, we performed ortholog mapping of mouse genes into human genes using the ortholog information provided by HomoloGene (<https://www.ncbi.nlm.nih.gov/homologene/>).

Analysis of scATAC-seq of naïve T cells

We collected publicly available scATAC-seq of mouse T cells from different tissues using GEO database which were profiled by 10X genomics platform (PRJNA587562, GSE156112). In case of the samples the fragment files are not available, we downloaded the fastq files and processed them using Cell Ranger ATAC workflows (10X Genomics, v.2.1.0). When the fragment files were available, we downloaded those files and processed them with Signac (v.1.6.0)⁶⁸. Quality checking, normalization, and scaling was done using Signac⁶⁸. To annotate previously known T cell types, we performed label transfer analysis using Signac and reference single-cell RNA-seq data set⁶⁸. Afterward, we subset naïve T cells and further performed unsupervised clustering using Signac. Differentially accessible chromatin regions (DACRs) and differentially accessible genes (DAGs) were identified using FindAllMarkers function of Signac. For the cluster similarity analysis, we evaluated the overlap between DAGs of scATAC-seq clusters and DEGs of scRNA-seq clusters using the same criteria (adjusted p -value < 0.05, log₂ fold change > 0.1). For the cluster of interest, we performed differential motif activity analysis using chromVAR (v.1.16.0)⁶⁹ and reconstructed 3D chromatin interaction map using Cicero (v.1.3.9)⁷⁰.

Functional enrichment analysis

For the marker genes of each cluster, we performed functional enrichment analysis using ConsensusPathDB⁷¹. Those pathways passing the criteria of adjusted p -value < 0.1 and gene counts ≥ 3 was considered as significantly enriched pathways.

Analysis of scRNA-seq of PBMC from GI and BT cancer patients

To evaluate clinical significance of IL-7R^{sup-hi} naïve CD4⁺ T cells, we collected publicly available PBMC scRNA-seq data of gastrointestinal (GI) cancer patients⁷² and biliary tract (BT) cancer patients⁷³. To obtain high quality cells, data preprocessing was performed using R package Seurat (version 4.3.0) as aforementioned. Author-annotated cell type information was reconstructed either using Database of Immune Cell Expression/eQTLs/Epigenomics (DICE) from celldex (version 1.12.0) R package or CITE-seq based annotation. Subclusters of naïve CD4⁺ T cells were further annotated via label transfer of mouse naïve T clusters.

PLS-DA analysis for the prediction of therapy response

To evaluate discriminatory power of naïve CD4 T cells in immune therapy response, we performed partial least squares discriminant analysis (PLS-DA) for GI cancer patients⁷² and BTC patients⁷³. Patients were grouped into two; either responder or non-responder to anti-PD-1 drug therapy. We generated feature sets from the cell type proportions of 8 conventional T cell types, as well as the proportion of C1 to C4 within the naïve T cell clusters. Features were quantile normalized then autoscaled. Using the 10-feature set (proportion of conventional T cell types) or 14 feature set (proportion of conventional T cell types + Naïve T cell clusters) for N donors, we performed partial least square discriminant analysis (PLS-DA) to classify responders and non-responders. An X matrix [N donors-by-features] and a Y matrix [N donors-by-MS status] were prepared for PLS-DA, in order to find latent variables which can explain the variability in Y using linear combinations of features in the X matrix. Variable importance in projection (VIP) were computed for each feature. Null distribution for VIP was estimated by random permutation of group information 1000 times. Based on the null distribution, we selected features with VIP higher than 90th percentile. To estimate accuracy of our discrimination model performance, we performed leave-one-out cross validation (LOOCV).

PLS-DA analysis for the classification of MS status

To assess the utility of Naïve T cell cluster proportions for the classification of MS and healthy status, we collected the scRNA-seq data of pan T cells from MS ($n = 6$) and healthy ($n = 4$) donors (GSE193770). We processed the data and annotated conventional T cell types using SingleR and Naïve T cell clusters using label transfer as described above. We generated feature sets from the cell type proportions of 8 conventional T cell types, as well as the proportion of C1 to C4 within the Naïve T cell clusters. We then performed PLS-DA and calculated VIP values for each feature as we performed analysis for anti-PD-1 therapy response as described above.

Statistics & Reproducibility

All experimental data were analyzed by nonparametric analysis using the Mann-Whitney U -test in Prism version 9.0 (GraphPad Software, San Diego, CA). Data were presented as the mean \pm S.D. or mean \pm S.E.M. For all data, significance was defined as $p < 0.05$. Sample size and statistical information are provided in each figure legend. For in vitro experiments, mice were assigned to different experimental conditions with identical age and sex to ensure consistency across groups. When age was explicitly noted in the dataset, these parameters were also controlled accordingly. For in vivo experiments, animals were matched for age, sex, and body weight prior to experimentation to establish baseline equivalency. Following tumor injection, additional randomization was conducted to minimize bias in tumor induction levels, thereby enhancing the robustness of the experimental design. For human PBMC experiments, control and experimental groups were matched for age and sex to ensure comparability and minimize potential confounding factors during data analysis. Investigators were not blinded to group allocation during in vitro experiments but were single-blinded during in vivo experiments. No statistical method was

used to predetermine sample sizes, and no data were excluded from the analyses.

Reporting summary

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

Data availability

The scRNA-seq data generated in this study (scRNA-seq of murine Naïve CD4 T cells) have been deposited in the ENA database under accession code [PRJEB72620](https://www.ebi.ac.uk/ena/record/PRJEB72620). The publicly available scRNA-seq and CITE-seq data downloaded and reanalyzed in this study are available in the GEO database under accession code [GSE130157](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE130157), [GSE193770](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE193770), and [GSE210065](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE210065). The publicly available scATAC-seq data downloaded and reanalyzed in this study are available in the [PRJNA587562](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=PRJNA587562) and [GSE156112](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE156112).

Details about the scRNA-seq data generated in this study and the publicly available datasets are provided in Supplementary Data 2, while information about individual subjects from the human public data is included in Supplementary Data 3.

Source data are provided as a Source Data file. Source data are provided with this paper.

Code availability

The code can be accessed at https://github.com/kyungmink123/Tcell_Project/. This code can be cited using the following <https://doi.org/10.5281/zenodo.14436974>

References

- Kopf, M., Bachmann, M. F. & Marsland, B. J. Averting inflammation by targeting the cytokine environment. *Nat. Rev. Drug Discov.* **9**, 703–718 (2010).
- Lippitz, B. E. Cytokine patterns in patients with cancer: a systematic review. *Lancet Oncol.* **14**, e218–e228 (2013).
- Neurath, M. F. Cytokines in inflammatory bowel disease. *Nat. Rev. Immunol.* **14**, 329–342 (2014).
- Liu, C. et al. Cytokines: From clinical significance to quantification. *Adv. Sci. (Weinh)* **8**, e2004433 (2021).
- Au, L. et al. Determinants of anti-PD-1 response and resistance in clear cell renal cell carcinoma. *Cancer Cell* **39**, 1497–1518 e1411 (2021).
- Kim, K. H., Kim, C. G. & Shin, E. C. Peripheral blood immune cell-based biomarkers in anti-PD-1/PD-L1 therapy. *Immune. Netw.* **20**, e8 (2020).
- Krieg, C. et al. High-dimensional single-cell analysis predicts response to anti-PD-1 immunotherapy. *Nat. Med.* **24**, 144–153 (2018).
- van den Broek, T., Borghans, J. A. M. & van Wijk, F. The full spectrum of human naive T cells. *Nat. Rev. Immunol.* **18**, 363–373 (2018).
- Chapman, N. M., Boothby, M. R. & Chi, H. Metabolic coordination of T cell quiescence and activation. *Nat. Rev. Immunol.* **20**, 55–70 (2020).
- Sprent, J. & Surh, C. D. Normal T cell homeostasis: the conversion of naive cells into memory-phenotype cells. *Nat. Immunol.* **12**, 478–484 (2011).
- Klein, L., Kyewski, B., Allen, P. M. & Hogquist, K. A. Positive and negative selection of the T cell repertoire: what thymocytes see (and don't see). *Nat. Rev. Immunol.* **14**, 377–391 (2014).
- This, S., Rogers, D., Mallet Gauthier, E., Mandl, J. N. & Melichar, H. J. What's self got to do with it: sources of heterogeneity among naive T cells. *Semin. Immunol.* **65**, 101702 (2023).
- Eggert, J. & Au-Yeung, B. B. Functional heterogeneity and adaptation of naive T cells in response to tonic TCR signals. *Curr. Opin. Immunol.* **73**, 43–49 (2021).
- Lee, S. W., Lee, G. W., Kim, H. O. & Cho, J. H. Shaping heterogeneity of naive CD8(+) T cell pools. *Immune. Netw.* **23**, e2 (2023).
- Cho, J. H., Kim, H. O., Surh, C. D. & Sprent, J. T cell receptor-dependent regulation of lipid rafts controls naive CD8+ T cell homeostasis. *Immunity* **32**, 214–226 (2010).
- Guichard, V. et al. Calcium-mediated shaping of naive CD4 T-cell phenotype and function. *Elife* **6**, e27215 (2017).
- Ju, Y. J. et al. Self-reactivity controls functional diversity of naive CD8(+) T cells by co-opting tonic type I interferon. *Nat. Commun.* **12**, 6059 (2021).
- Rogers, D. et al. Pre-existing chromatin accessibility and gene expression differences among naive CD4(+) T cells influence effector potential. *Cell Rep.* **37**, 110064 (2021).
- Sood, A. et al. CD5 levels define functionally heterogeneous populations of naive human CD4(+) T cells. *Eur. J. Immunol.* **51**, 1365–1376 (2021).
- Zinzow-Kramer, W. M., Weiss, A. & Au-Yeung, B. B. Adaptation by naive CD4(+) T cells to self-antigen-dependent TCR signaling induces functional heterogeneity and tolerance. *Proc. Natl Acad. Sci. USA* **116**, 15160–15169 (2019).
- Crouse, J., Kalinke, U. & Oxenius, A. Regulation of antiviral T cell responses by type I interferons. *Nat. Rev. Immunol.* **15**, 231–242 (2015).
- Wahren-Herlenius, M. & Dorner, T. Immunopathogenic mechanisms of systemic autoimmune disease. *Lancet* **382**, 819–831 (2013).
- Hirahara, K. et al. Interleukin-27 priming of T cells controls IL-17 production in trans via induction of the ligand PD-L1. *Immunity* **36**, 1017–1030 (2012).
- Deimel, L. P., Li, Z., Roy, S. & Ransinghe, C. STAT3 determines IL-4 signalling outcomes in naive T cells. *Sci. Rep.* **11**, 10495 (2021).
- Li, G., Ju, J., Weyand, C. M. & Goronzy, J. J. Age-associated failure to adjust type I IFN receptor signaling thresholds after T cell activation. *J. Immunol.* **195**, 865–874 (2015).
- Thompson, L. J. et al. Conditioning of naive CD4(+) T cells for enhanced peripheral Foxp3 induction by nonspecific bystander inflammation. *Nat. Immunol.* **17**, 297–303 (2016).
- ELTanbouly, M. A. et al. VISTA is a checkpoint regulator for naive T cell quiescence and peripheral tolerance. *Science* **367**, eaay0524 (2020).
- Even, Z. et al. The amalgam of naive CD4(+) T cell transcriptional states is reconfigured by helminth infection to dampen the amplitude of the immune response. *Immunity* **57**, 1893–1907 e1896 (2024).
- Deep, D. et al. Precursor central memory versus effector cell fate and naive CD4+ T cell heterogeneity. *J. Exp. Med.* **221**, e20231193 (2020).
- Zhang, H. et al. Aging-associated HELIOS deficiency in naive CD4(+) T cells alters chromatin remodeling and promotes effector cell responses. *Nat. Immunol.* **24**, 96–109 (2023).
- Winoto, A. & Littman, D. R. Nuclear hormone receptors in T lymphocytes. *Cell* **109**, S57–S66 (2002). Suppl.
- Shao, H., Kono, D. H., Chen, L. Y., Rubin, E. M. & Kaye, J. Induction of the early growth response (Egr) family of transcription factors during thymic selection. *J. Exp. Med.* **185**, 731–744 (1997).
- Azzam, H. S. et al. CD5 expression is developmentally regulated by T cell receptor (TCR) signals and TCR avidity. *J. Exp. Med.* **188**, 2301–2311 (1998).
- Schoggins, J. W. Interferon-stimulated genes: what do they all do? *Annu. Rev. Virol.* **6**, 567–584 (2019).
- Peng, Y. C. & Lenschow, D. J. ISG15 in antiviral immunity and beyond. *Nat. Rev. Microbiol.* **16**, 423–439 (2018).
- Iida, K., Kondo, J., Wibisana, J. N., Inoue, M. & Okada, M. ASURAT: functional annotation-driven unsupervised clustering of single-cell transcriptomes. *Bioinformatics* **38**, 4330–4336 (2022).
- Li, J. et al. KIR(+)CD8(+) T cells suppress pathogenic T cells and are active in autoimmune diseases and COVID-19. *Science* **376**, eabi9591 (2022).

38. Aran, D. et al. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat. Immunol.* **20**, 163–172 (2019).
39. Schmiedel, B. J. et al. Impact of genetic polymorphisms on human immune cell gene expression. *Cell* **175**, 1701–1715 e1716 (2018).
40. Fulton, R. B. et al. The TCR's sensitivity to self peptide-MHC dictates the ability of naive CD8(+) T cells to respond to foreign antigens. *Nat. Immunol.* **16**, 107–117 (2015).
41. Mandl, J. N., Monteiro, J. P., Vrsekooop, N. & Germain, R. N. T cell-positive selection uses self-ligand binding strength to optimize repertoire recognition of foreign antigens. *Immunity* **38**, 263–274 (2013).
42. Persaud, S. P., Parker, C. R., Lo, W. L., Weber, K. S. & Allen, P. M. Intrinsic CD4+ T cell sensitivity and response to a pathogen are set and sustained by avidity for thymic and peripheral complexes of self peptide and MHC. *Nat. Immunol.* **15**, 266–274 (2014).
43. Henderson, J. G., Opejin, A., Jones, A., Gross, C. & Hawiger, D. CD5 instructs extrathymic regulatory T cell development in response to self and tolerizing antigens. *Immunity* **42**, 471–483 (2015).
44. Martin, B. et al. Highly self-reactive naive CD4 T cells are prone to differentiate into regulatory T cells. *Nat. Commun.* **4**, 2209 (2013).
45. Sood, A. et al. Differential interferon-gamma production potential among naive CD4(+) T cells exists prior to antigen encounter. *Immunol. Cell Biol.* **97**, 931–940 (2019).
46. Fang, D. et al. Differential regulation of transcription factor T-bet induction during NK cell development and T helper-1 cell differentiation. *Immunity* **55**, 639–655 e637 (2022).
47. Corral, D. et al. ILC precursors differentiate into metabolically distinct ILC1-like cells during Mycobacterium tuberculosis infection. *Cell Rep.* **39**, 110715 (2022).
48. Tomura, M. et al. Differential capacities of CD4+, CD8+, and CD4-CD8- T cell subsets to express IL-18 receptor and produce IFN-gamma in response to IL-18. *J. Immunol.* **160**, 3759–3765 (1998).
49. Cho, M. J. et al. Steady-state memory-phenotype conventional CD4(+) T cells exacerbate autoimmune neuroinflammation in a bystander manner via the Bhlhe40/GM-CSF axis. *Exp. Mol. Med.* **55**, 1033–1045 (2023).
50. Kawabe, T. et al. Redefining the foreign antigen and self-divergent memory CD4(+) T-cell compartments via transcriptomic, phenotypic, and functional analyses. *Front Immunol.* **13**, 870542 (2022).
51. Kawabe, T. et al. Memory-phenotype CD4(+) T cells spontaneously generated under steady-state conditions exert innate T(H)1-like effector function. *Sci. Immunol.* **2**, eaam9304 (2017).
52. Kawabe, T. et al. Requirements for the differentiation of innate T-bet(high) memory-phenotype CD4(+) T lymphocytes under steady state. *Nat. Commun.* **11**, 3366 (2020).
53. Lee, H. G. et al. Pathogenic function of bystander-activated memory-like CD4(+) T cells in autoimmune encephalomyelitis. *Nat. Commun.* **10**, 709 (2019).
54. Shim, C. H., Cho, S., Shin, Y. M. & Choi, J. M. Emerging role of bystander T cell activation in autoimmune diseases. *BMB Rep.* **55**, 57–64 (2022).
55. Gattinoni, L., Speiser, D. E., Lichterfeld, M. & Bonini, C. T memory stem cells in health and disease. *Nat. Med.* **23**, 18–27 (2017).
56. Long, H. M. et al. MHC II tetramers visualize human CD4+ T cell responses to Epstein-Barr virus infection and demonstrate atypical kinetics of the nuclear antigen EBNA1 response. *J. Exp. Med.* **210**, 933–949 (2013).
57. Orlando, V. et al. Human CD4 T-cells with a naive phenotype produce multiple cytokines during mycobacterium tuberculosis infection and correlate with active disease. *Front Immunol.* **9**, 1119 (2018).
58. White, J. T., Cross, E. W. & Kedl, R. M. Antigen-inexperienced memory CD8(+) T cells: where they come from and why we need them. *Nat. Rev. Immunol.* **17**, 391–400 (2017).
59. Oh, D. Y. & Fong, L. Cytotoxic CD4(+) T cells in cancer: expanding the immune effector toolbox. *Immunity* **54**, 2701–2711 (2021).
60. Cachot, A. et al. Tumor-specific cytolytic CD4 T cells mediate immunity against human cancer. *Sci. Adv.* **7**, eabe3348 (2021).
61. Byrne, A. et al. Tissue-resident memory T cells in breast cancer control and immunotherapy responses. *Nat. Rev. Clin. Oncol.* **17**, 341–348 (2020).
62. Held, W., Siddiqui, I., Schaeuble, K. & Speiser, D. E. Intratumoral CD8(+) T cells with stem cell-like properties: implications for cancer immunotherapy. *Sci. Transl. Med.* **11**, eaay6863 (2019).
63. Tay, C., Tanaka, A. & Sakaguchi, S. Tumor-infiltrating regulatory T cells as targets of cancer immunotherapy. *Cancer Cell* **41**, 450–465 (2023).
64. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587 e3529 (2021).
65. Trapnell, C. et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381–386 (2014).
66. Borchering, N., Bormann, N. L. & Kraus, G. scRepertoire: An R-based toolkit for single-cell immune receptor analysis. *F1000Res.* **9**, 47 (2020).
67. Korsunsky, I. et al. Fast, sensitive and accurate integration of single-cell data with harmony. *Nat. Methods* **16**, 1289–1296 (2019).
68. Stuart, T., Srivastava, A., Madad, S., Lareau, C. A. & Satija, R. Single-cell chromatin state analysis with signac. *Nat. Methods* **18**, 1333–1341 (2021).
69. Schep, A. N., Wu, B., Buenrostro, J. D. & Greenleaf, W. J. chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. *Nat. Methods* **14**, 975–978 (2017).
70. Pliner, H. A. et al. Cicero predicts cis-regulatory DNA interactions from single-cell chromatin accessibility data. *Mol. Cell* **71**, 858–871 e858 (2018).
71. Kamburov, A. & Herwig, R. ConsensusPathDB 2022: molecular interactions update as a resource for network biology. *Nucleic Acids Res.* **50**, D587–D595 (2022).
72. Griffiths, J. I. et al. Circulating immune cell phenotype dynamics reflect the strength of tumor-immune cell interactions in patients during immunotherapy. *Proc. Natl Acad. Sci. USA* **117**, 16072–16082 (2020).
73. Keenan, B. P. et al. Circulating monocytes associated with anti-PD-1 resistance in human biliary cancer induce T cell paralysis. *Cell Rep.* **40**, 111384 (2022).
74. Castro-Mondragon, J. A. et al. JASPAR 2022: the 9th release of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* **50**, D165–D173 (2022).

Acknowledgements

We thank Ms. In Young Song for providing technical support in the FACS sorting conducted at the Hanyang LINC Analytical Equipment Center (Seoul). Computational analysis was performed using a high-performance computing system (HPC) at the Biospecimen-Multiomics Digital Bioanalysis Core Facility of Hanyang University. Illustrations in this study were generated using BioRender (<https://biorender.com>), with permission. This research was supported by the Basic Science Research Program of the National Research Foundation funded by the Korean government received by J.-M.C. (RS-2024-00334289, RS-2023-00217324), and H.J. (RS-2024-00336518, NRF-2018R1A6A1A03025607).

Author contributions

J.-W.Y., and J.-M.C. conceptualized and designed this study. K.M.K. performed bioinformatic analysis with public human data sets. J.-W.Y. and S.C. performed and analyzed most of the experiments. H.J. performed and executed the bioinformatic analyses. S.P., and S.-H.P. supported bulk RNA-seq analysis. M.-J.C. supported the experiments. D.H. provided advice on the bioinformatic analyses. J.-H.C. supported Stat1^{-/-},

Ifnar1^{-/-}, and Ifnar1/Ifngr1^{-/-} mice experiment. H.R.K supported human PBMC samples during the revision. J.-W.Y., K.M.K, H.J., and J.-M.C. wrote the draft manuscript, and all authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at

<https://doi.org/10.1038/s41467-025-57237-3>.

Correspondence and requests for materials should be addressed to Hyobin Jeong or Je-Min Choi.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025