# Global mRNA and chromatin accessibility profiling elucidate how HIV-1 perturbs ILC and NK cell subsets

Yetao Wang[1], Lawrence Lifshitz[1], Noah J. Silverstein[1], Esther Mintzer[2], Kevin Luk[2], Pam St. Louis[3], Michael A Brehm[3], Scot A. Wolfe[2], Steven Deeks[4], and Jeremy Luban[1,5,6,7]*

[1] Program in Molecular Medicine, University of Massachusetts Medical School, Worcester, MA, USA

[2]Department of Molecular, Cell and Cancer Biology, University of Massachusetts Medical School, Worcester, MA, USA

[3]Diabetes Center of Excellence, University of Massachusetts Medical School, Worcester, MA, USA

[4]Department of Medicine, University of California, San Francisco, CA, USA

[5]Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, MA, USA

[6]Broad Institute of Harvard and MIT, Cambridge, MA, USA

[7]Massachusetts Consortium on Pathogen Readiness, Boston, MA, USA

*Correspondence to Jeremy Luban: jeremy.luban@umassmed.edu

**HIV-1 infection is associated with depletion of gut ILC3s, blood ILC2s and conventional CD56$^{dim}$ natural killer (NK) cells, as well as with expansion of CD56$^{-}$ NK cells. These cell types derive from a common precursor (ILCP) yet their interrelationship remains unclear. Bulk and single cell RNA-Seq and ATAC-Seq showed that blood ILCs cluster into ILC2s, ILCPs, a mixed cluster of both CD56$^{dim}$NK cells and CD56$^{-}$NK cells, and a distinct cluster of CD56$^{hi}$NK cells which shares features of both ILCs and CD56$^{dim}$NK cells. In surprising contrast to mice, tissue repair factor AREG was produced by human NK cells, with even higher levels in CD56$^{hi}$NK cells than in ILCs. AREG production was promoted by TCF7/WNT signaling and inhibited by TGFB1, a cytokine elevated in people living with HIV-1. Knockout of RUNX3, a WNT antagonist acting downstream of TGFB1, increased AREG production. Transcriptional profiling revealed that, when CD4$^{+}$T cells were depleted, or when MTOR was inhibited, CD56$^{dim}$NK cells shifted to become metabolically-defective CD56$^{-}$NK cells; in either case, exogenous IL-2 prevented this transition. These findings establish the interrelatedness of ILC populations disrupted by HIV-1 infection and provide insight into how disruption of these cells contributes to chronic inflammation in people living with HIV-1.**

## Introduction

Innate lymphoid cells (ILCs) are a diverse population of cells which contribute to a broad range of biological functions, including tissue homeostasis and repair, inflammation, and protection from infection[1–3]. All ILCs bear the leukocyte common antigen CD45, but lack markers of well-characterized cellular lineages, including T cells, B cells, hematopoietic stem cells, mast cells, and myeloid cells[4–6].

CD127, the cell surface protein IL-7Rα encoded by IL7R, defines a heterogeneous subset of cytokine-producing ILCs, with developmental and functional parallels to $CD4^+$ helper T cells[7,8]. Specifically, ILC1s are analogs of $T_H1$ cells, which are defined by the transcription factor TBX21, and produce IFN-γ. Similar to $T_H2$ cells, ILC2s are defined by $GATA3^+$, bear cell surface marker $CRTH2^+$, and produce IL-4, IL-5, and IL-13. ILC3s are counterparts of $T_H17$ cells, which express RORγT and produce IL-17 and IL-22.

Natural killer (NK) cells are a subset of ILCs, more akin to $CD8^+$ T cells in that they have granzyme and perforin-mediated cytolytic activity against tumor cells, and virus-infected cells[7,8]. NK cells bear proteins that distinguish them from the other ILCs, including CD56, CD16, and the transcription factor EOMES[6–9].

Despite diversity in gene expression, protein production, and phenotype, all human ILC subsets can be generated experimentally from a common innate lymphoid cell precursor (ILCP) found in the blood[10]. Consistent with shared derivation from a common precursor, the canonical markers described above do not always discriminate between the ILC subsets. For example, unlike the vast majority of $CD56^{dim}$NK cells, $CD56^{hi}$NK cells bear CD127, CD117, and TCF7[6], proteins typical of ILCs. Additionally, cytokines found in specific environments can reveal phenotypic and functional plasticity in ILCs. CD127 on ILC2s is downregulated by IL-33 in murine lung, mediastinal lymph node, and spleen[11], or when human blood ILCs are incubated *in vitro* with common γ chain cytokines[6]. NK cells are commonly distinguished from ILCs based on expression of EOMES, but this is not absolute[12,13]; TGF-β in the salivary gland suppresses EOMES and promotes the ILC-associated genes TNFSF10 and CD73[14,15]. Incubation of $Lin^-CD34^-CD56^+CD117^+$ cells from human tonsil with IL-7 and FLT3L, in the presence of OP9-DL1 feeder cells, generates a mixture of $CD94^+$NK cells and $NKp44^+$ILC3s[16].

Perturbation of ILC and NK cell subsets has been reported in association with Crohn's disease, psoriasis, chronic obstructive pulmonary disease, non-small cell lung cancer, or infection with any of several viruses, including human immunodeficiency virus type 1 (HIV-1), human cytomegalovirus (HCMV), hepatitis C virus (HCV), and influenza A virus[6,12,17]. People living with HIV-1 have permanent depletion of ILC2s in the blood, and of ILC3s in the intestinal lamina propria, even after viremia has been suppressed by antiviral therapy[6,18,19].

$CD56^{hi}$NK cells are increased in the blood of people living with HIV-1 or HCV, and in people with systemic lupus erythematosus[6,20], presumably as a result of chronic inflammation. The numbers of a functionally-defective $Lin^{-}CD56^{-}CD16^{+}$NK cell subset are also expanded by viral infection (HIV-1, HCV, HCMV, and hantavirus), or by autoimmune disease[17,21,22]. These $CD56^{-}$NK cells appear to be derived from $CD56^{dim}$NK cells in that the expansion of $CD56^{-}$NK cells is accompanied by a decrease in $CD56^{dim}$NK cells[17,21,22], and when $CD56^{-}$NK cells from people with untreated HIV-1 infection are incubated *ex vivo* with exogenous IL-2, CD56 becomes detectable, though cytolytic function is not restored to the level of $CD56^{dim}$ cells[22].

Taken together, these observations concerning human blood ILCs and NK cells, under conditions of normal homeostasis and in the context of pathogenic inflammation, indicate that much remains to be clarified regarding the relationship between these cell types. Here, global transcriptional and chromatin profiling was used to investigate blood ILCs and NK cell subsets from uninfected people, and from people living with HIV-1 infection. These studies provide new insight into the relationship between these innate lymphoid cell types and revealed previously unappreciated functions of NK cells.

## Results

**CD127 and CD56 identify four discrete cell populations among Lin$^-$ human PBMCs**

Human blood ILCs have been defined operationally as Lin$^-$CD127$^+$CD56$^-$, and NK cells have been defined as Lin$^-$CD127$^-$CD56$^+$ [5,6,12]. These definitions, though, do not accommodate the significant plasticity that these cell populations display, a property made more apparent by the inflammation that accompanies certain environmental conditions such as HIV-1 infection[6]. Additionally, even in the absence of inflammation, human blood contains an obscure population of Lin$^-$CD127$^-$CD56$^-$ cells that are poorly-characterized.

To clarify the identity of human blood ILCs at the transcriptional level, and the relatedness of these cells to each other, Lin$^-$CD45$^+$CD56$^-$ cells from five blood donors were sorted into CD127$^-$ and CD127$^+$ subpopulations, and each was subjected to bulk RNA-Seq (Fig. 1a and Extended Data Fig. 1a). 401 and 241 differentially expressed genes (DEgenes) were enriched among the CD127$^-$ and CD127$^+$ cells, respectively (Fig. 1b and Supplementary Table 1).

The Lin$^-$CD45$^+$CD56$^-$CD127$^+$ blood cells were enriched for genes specific to ILC2s, or shared with ILCPs, including IL7R (CD127), IL4R, PTGDR2 (CRTH2), CCR6, CCR7, GATA3 and TCF7, and gene ontology (GO) analysis highlighted expressed genes associated with cytokine production (Fig. 1c,d and Supplementary Table 1; adjusted p value cut-off <0.01).

The Lin$^-$CD45$^+$CD56$^-$CD127$^-$ population was found to be an NK cell subset, based on its expression of TBX21, EOMES, IFNG, KLRD1 (CD94), 2B4 (CD244), GZMB, GZMH, GNLY, KIR2DL3, KIR2DL4, KIR2DS4, KIR3DL1, KIR3DL2, and KIR3DL3 (Fig. 1c and Supplementary Table 1). SEMA4D regulates NK cell killing activity and IFN-γ production through interaction with CD72 on target cells[23–26], and the SEMA4D signaling components, ARHGEF12, ERBB2, RHOC, and MYL9, were well expressed on these cells (Fig. 1e and Supplementary Table 1), as were NK cell-specific receptors, killer Ig-like receptors (KIRs), CD94-NKG2C, and NKp44, that use DAP12 as a signal transduction element[27,28] (Fig. 1e, Supplementary Table 1). TBX21 and CD16 proteins were detected by flow cytometry on the majority of Lin$^-$CD45$^+$CD56$^-$CD127$^-$ cells, confirming that,

despite undetectable CD56 protein, this population consists of *bona fide* NK cells (Fig. 1f,g).

Taking together the transcriptional profiling and flow cytometry analysis, Lin$^{-}$ PBMCs can be divided into four main ILC subsets, CD56$^{hi}$NK cells, CD56$^{dim}$NK cells, CD56$^{-}$NK cells, and CD127$^{+}$ILCs, the latter including CD127$^{+}$CRTH2$^{+}$ILC2s and CD127$^{+}$CRTH2$^{-}$CD117$^{+}$ILCPs (Fig. 1f,g and Extended Data Fig. 1b).

**Global transcription and epigenetic features of the four ILC subsets in human blood**

To better characterize the four ILC subpopulations in human blood, PBMCs from four blood donors were sorted into ILCs (Lin$^{-}$CD45$^{+}$CD56$^{-}$CD127$^{+}$), CD56$^{-}$NK cells (Lin$^{-}$CD45$^{+}$CD56$^{-}$CD127$^{-}$CD16$^{+}$), CD56$^{dim}$NK cells (Lin$^{-}$CD45$^{+}$CD56$^{dim}$), and CD56$^{hi}$NK cells (Lin$^{-}$CD45$^{+}$CD56$^{hi}$). Each population was then separately subjected to bulk RNA-Seq (Extended Data Fig. 1c).

As compared with ILCs, CD56$^{-}$, CD56$^{dim}$, and CD56$^{hi}$ NK cells exhibited 1,128, 1,236, and 910 DEgenes, respectively (Fig. 2a and Supplementary Table 2). In contrast with CD56$^{-}$ and CD56$^{dim}$ NK cells, CD56$^{hi}$NK cells shared many transcripts with ILCs, including IL7R (CD127), KIT, IL4R, IL1RL1 (IL33R), CCR7, GPR183, TCF7, and MYC (Extended Data Fig. 2a and Supplementary Table 2)[29]. Surprisingly, amphiregulin (AREG), an epidermal growth factor (EGF)-like protein produced by ILC2s which promotes tissue repair and homeostasis[30], was expressed in all NK cell subsets, with highest levels in CD56$^{hi}$ NK cells (Fig. 2b, Extended Data Fig. 2a, and Supplementary Table 2). In contrast, multiple mouse datasets showed that AREG was expressed by ILC2s, but not by other ILC subsets nor by NK cells (Extended Data Fig. 2b-e and Supplementary Table 3)[31–33], indicating a major difference in AREG regulation between mouse and human.

Principal component analysis confirmed that the human blood ILC transcriptome was distinct from that of all three NK cell subsets, that CD56$^{dim}$ and CD56$^{-}$ NK cell transcriptomes were nearly indistinguishable from each other, and that the CD56$^{hi}$NK cell transcriptome was distinct from that of the other NK subsets (Fig. 2c).

Reactome analysis showed that, as compared with ILCs, 86 classified pathways were enriched in CD56$^{-}$ and CD56$^{dim}$ NK cells, whereas only 8 pathways were enriched in CD56$^{hi}$NK cells (Fig. 2d and Supplementary Table 2). All NK cell subsets, though, were distinguished from ILCs by five pathways, which highlight the fundamental role of IFN-γ and DAP12 signaling in the three NK cell subsets (Fig. 2e, Extended Data Fig. 2f and Supplementary Table 2). Typical of cells with a primary role in secretion of cytokines, the ILCs had high-level expression of genes encoding ribosomal proteins, and proteins involved in translation initiation and elongation (Fig. 2f, Extended Data Fig. 2g and Supplementary Table 2). The ILCs were further distinguished from CD56$^{hi}$NK cells by higher level expression of IL-4, IL-13, and SCF-KIT pathways, consistent with the established functions of ILC2s and ILCPs (Fig. 2g, Extended Data Fig. 2h, and Supplementary Table 2).

The global transcriptional profiles of the ILCs and NK cell subsets were mirrored by the chromatin accessibility of genes, as determined by ATAC-Seq. DAP12, GNLY, EOMES, TBX21, NKG7, ILR2B, and CST7 loci were more accessible to Tn5 transposase in NK cell subsets, whereas IL13, IL17RB, PTGDR2, TNFRSF25, SOCS3, IL23A, and IL32 loci were more accessible in ILCs (Fig. 2h and Extended Data Fig. 3a). The IL7R promoter was open in both ILCs and CD56$^{hi}$NK cells, but not in CD56$^{dim/-}$ NK cells, and the AREG promoter was accessible in all NK subsets, as well as in ILCs (Fig. 2h). In contrast to these results with human cells, the mouse Areg promoter was only accessible in mouse ILC2s, but not in other ILC subsets or in NK cells (Extended Data Fig. 3b). Interestingly, despite undetectable RORγT protein in blood ILCs (Fig. 1g), which are primarily ILC2s in phenotype, the chromatin at the RORC promoter was open (Extended Data Fig. 3a), consistent with the phenotypic plasticity of these cells and their potential to acquire an ILC3 phenotype[10].

**Single cell transcriptional analysis of human blood ILCs and NK cells**

To better understand the relationship between the different subsets of human blood ILCs, PBMCs from three blood donors were sorted into ILCs (Lin$^{-}$CD45$^{+}$CD56$^{-}$CD127$^{+}$), CD56$^{-}$ NK cells (Lin$^{-}$CD45$^{+}$CD56$^{-}$CD127$^{-}$CD16$^{+}$), CD56$^{dim}$NK cells (Lin$^{-}$CD45$^{+}$CD56$^{dim}$), and CD56$^{hi}$NK cells (Lin$^{-}$CD45$^{+}$CD56$^{hi}$). Equal numbers of each subset were then pooled and

subjected to single cell RNA-Seq (Extended Data Fig. 1c). Transcriptomes from 5,210 individual cells were analyzed, and 96% of the cells fit within one of four main clusters (Fig. 3a), though surprisingly, not the same four subsets that had been sorted.

Lin$^{-}$CD45$^{+}$CD56$^{-}$CD127$^{+}$ cells formed two distinct clusters. The first cluster was enriched for expression of genes that define ILC2s, including IL7R (CD127), TCF7, PTGDR2 (CRTH2), GATA3, IL2RA, TNFRSF25, LTB, IL17RB (Fig. 3b-c). The first cluster expressed GATA3-AS1, which increases transcription of IL-5, IL-13, and GATA3 [34], and also HPGD and HPGDS, genes encoding prostaglandin D2 biosynthetic enzymes required for ILC2 cytokine production[35].

The second cluster lacked critical ILC2-associated transcripts, and had higher expression of TNFRSF4 and TNFRSF18, defining features of ILCPs[10]. FOS, JUN, and JUNB, genes required for cell survival, proliferation, and development[36], were also enriched in this ILCP cluster (Fig. 3c-e and Supplementary Table 4).

The third cluster was defined as CD56$^{hi}$NK cells[6], based on expression of GZMK [6], and on enrichment for transcripts from both NK cells and ILCs, including TBX21, KLRD1, IL7R, LTB, GATA3, and TCF7 (Fig. 3b,c and Supplementary Table 4). Consistent with the PCA analysis of bulk RNA-Seq data (Fig. 2c), the fourth cluster consisted of both CD56$^{dim}$ and CD56$^{-}$ NK cells, despite the fact that the two subsets had been separated by flow cytometry. The fourth CD56$^{dim/-}$ cluster was distinguished from CD56$^{hi}$NK cells by exclusively expressing GZMH, and higher levels of CCL3, CCL4, and CCL5. The NK cell signature genes CMC1, NKG7, and CST7 were enriched by the two NK cell clusters (Fig. 3b,c and Supplementary Table 4)[37]. Consistent with bulk RNA-Seq data (Extended Data Fig. 2 and Supplementary Table 2), AREG transcript was detected in all four clusters (Fig. 3c).

Pseudotime analysis[38,39] was used to determine how the four clusters of blood ILCs are related to each other. ILCs and ILCPs formed one branch, and CD56$^{-}$ and CD56$^{dim}$ NK cells formed a second distinct branch (Fig. 3f). ETS1, ID2, IKZF1, and IL2RG, genes essential for ILC and NK cell development[40–42], were expressed along the full trajectory from ILCs/ILCPs to CD56$^{dim/-}$NK cells (Fig. 3g). Consistent with CD56$^{hi}$NK cells exhibiting epigenetic and transcriptional features of both ILCs (IL7R, KIT, IL4R, IL1RL1, CCR7, GPR183, MYC and TCF7), and NK cells (EOMES, TBX21, NCAM1, KLRD1,

GZMB, GNLY, and KIRs) (Fig. 2 and Supplementary Table 2), pseudotime analysis placed CD56$^{hi}$NK cells along the trajectory at the junction between ILCs or CD56$^{dim/-}$ NK cells (Fig. 3g), demonstrating that CD56$^{hi}$NK cells occupy an intermediate state between ILCs/ILCPs and CD56$^{dim/-}$NK cells.

**Human blood NK cells are major producers of amphiregulin**

Given the unexpectedly high-level AREG expression in human blood NK cells (Fig. 2b,g and 3c), flow cytometry was used to assess cell-associated AREG-encoded protein. AREG was readily detected in ILCs and all NK subsets (Fig. 4a). Whether assessed for mean fluorescence intensity or percent positive cells, the AREG protein signal in CD56$^{hi}$NK cells was higher than in ILCs (Fig. 4a-c).

AREG-expressing ILCs maintain tissue homeostasis, and limit the tissue damage that results from inflammation[19,30,43]. People living with HIV-1 infection have reduced numbers of ILCs, and this reduction correlates inversely with systemic inflammation[6]. Compared with HIV-1$^-$ individuals, people living with HIV-1 who were not taking antiretroviral therapy (ART) had a decreased percentage of AREG$^+$ cells among all NK cell subsets (Fig. 4c). This was even the case among the controllers who spontaneously maintained the viral load below 2,000 copies of HIV-1 gRNA/ml (Fig. 4c). In HIV-1$^+$ individuals on antiretroviral therapy (ART), the percentage of AREG$^+$ was decreased, but only in CD56$^{hi}$NK cells and ILCs (Fig. 4c). Interestingly, the percentage of AREG$^+$NK cells correlated with the frequency of ILCs (Fig. 4d).

CD56$^{hi}$NK cells and ILCs were distinguished from the other NK cell subsets in that they are enriched for expression of TCF7, and other genes in the WNT signaling pathway (Extended Data Fig. 2a and Supplementary Table 1). The open chromatin region surrounding the AREG promoter (chr4:75,310,271-75,311,679) contains six TCF7-binding sites (Fig. 4e). Transcription factor RUNX3, an antagonist of TCF7 and WNT signaling[44,45], was expressed at a higher level in all NK cell subsets than in ILCs (Fig. 4f), and six RUNX3 binding sites were present in the open chromatin region of AREG (Fig. 4e). Consistent with the biological significance of these differences in TCF7 and RUNX3 expression, WNT agonist CHIR99021 upregulated AREG production in NK cells, but not in ILCs (Fig. 4g,h). Furthermore, TGF-β1, which activates RUNX3 signaling and is

upregulated during HIV-1 infection[46–49], attenuated the CHIR99021-induced AREG upregulation in NK cells (Fig. 4i). Finally, the effect of Cas12a/RNP-mediated knockout[50] of RUNX3 was assessed in primary blood NK cells. The editing rate for either of two separate crRNAs was greater than 90% (Extended Data Fig. 4a, left panel) and RUNX3 protein was undetectable in most of these cells by flow cytometry (Extended Data Fig. 4a, right panel). As compared to cells treated with Cas12a/RNP and a control crRNA targeting the AAVS1, gene editing by either of the two RUNX-specific crRNAs increased AREG production in NK cells (Extended Data Fig. 4b and Fig. 4j). These results demonstrate that AREG production in NK cells and ILCs is differentially regulated by TCF7/WNT and RUNX3.

**CD56$^{dim}$NK cells become CD56$^{-}$ in the absence of IL-2-producing CD4$^{+}$ T cells**

The majority of Lin$^{-}$CD45$^{+}$CD56$^{-}$PBMCs are ILCs or CD56$^{-}$NK cells (Fig. 1f). In people living with HIV-1, ILCs were permanently depleted and this reduction correlated inversely with expansion in the number of CD56$^{-}$NK cells (Fig. 5a,b, Extended Data Fig. 4c and Supplementary Table 5)[6,18], However, among people living with HIV-1, CD56$^{-}$NK cell expansion was greatest in HIV-1$^{+}$ people who were not on ART (Fig. 5c, Extended Data Fig. 4d and Supplementary Table 5)[22]. The close similarity between CD56$^{-}$ and CD56$^{dim}$ NK cells revealed by our cluster algorithm (Fig. 3), along with the inverse correlation in their numbers (Fig. 5d), suggested that CD56$^{dim}$ NK cells give rise to CD56$^{-}$NK cells in the context of HIV-1 infection (Extended Data Fig. 4e).

To identify an experimental condition in tissue culture, under which CD56$^{dim}$ NK cells give rise to CD56$^{-}$ NK cells, and to determine whether a specific cell type among PBMCs stabilizes CD56$^{dim}$NK cells, PBMCs from HIV-1$^{-}$ blood donors were maintained in culture for 5 days after selective depletion of either total T cells (anti-CD3), CD4$^{+}$T cells (anti-CD4), CD8$^{+}$ T cells (anti-CD8), B cells (anti-CD19 and -CD20), monocytes and macrophages (anti-CD14 and -CD11b), stem cells (anti-CD34), myeloid cells (anti-CD33), or dendritic cells (anti-DC-SIGN and -BDCA3). Depletion of total CD3$^{+}$T cells, or of CD4$^{+}$T cells, but not of any of the other cell types, increased the proportion of CD56$^{-}$NK cells among Lin$^{-}$TBX21$^{+}$ cells in the cultures (Extended Data Fig. 5a,b and Fig. 5e). Consistent with this *ex vivo* experiment, CD56$^{-}$NK cells were maximally increased in HIV-1$^{+}$

individuals who were not treated with ART, in whom $CD4^+$T cells were most severely depleted (Fig. 5f and Supplementary Table 5); $CD56^-$NK cells correlated inversely with $CD4^+$ T cells (Fig. 5g).

Since $CD4^+$T cells maintain NK cell physiology by secreting IL-2[51,52], the effect of IL-2 was tested next. Addition of exogenous IL-2 to the PBMC cultures that had been depleted of $CD4^+$ T cells maintained the stability of $CD56^{dim}$NK cells, and prevented the increase of $CD56^-$NK cells (Fig. 5h,i and Extended Data Fig. 5c). The effect of IL-2 was counteracted by anti-IL-2 antibody (Fig. 5i,j and Extended Data Fig. 5d), or by inhibition of IL-2 signaling by STAT3 or JAK3 inhibitors, but not by STAT6, AKT, or ERK1/2 inhibitors, without compromising cell survival (Extended Data Fig. 5e,f). Moreover, TGF-β1 which suppresses IL-2 production by $CD4^+$T cells during HIV-1 infection[46–49,53], increased the percentage of $CD56^-$NK cells and decreased IFN-γ production, and exogenous IL-2 counteracted this effect (Fig. 5k,l).

To test the importance of IL-2 in an *in vivo* model, NOD-scid $Il2rg^{null}$ (NSG) mice reconstituted with human $CD34^+$ hematopoietic stem cells were injected intravenously with PBS, or with recombinant adeno-associated virus expressing IL-2 (hIL-2-rAAV). 6 weeks later, the ratio of $CD56^{dim}$ to $CD56^-$ NK cells was greater in animals injected with hIL-2-rAAV than in controls (Fig. 5m and Extended Data Fig. 5g), and the NK cells from the hIL-2-rAAV-treated mice showed higher IFN-γ production than did cells from the control mice (Fig. 5n and Extended Data Fig. 5h). Thus, IL-2 is critical for maintaining the physiologically intact $CD56^{dim}$NK cells *in vivo*.

**Metabolic difference between $CD56^{dim}$ and $CD56^-$NK cells**

Comparison of bulk RNA-Seq data from sorted $CD56^{dim}$ and $CD56^-$NK cells revealed that genes associated with immune function (CD6, TRAF3, and IRAK2), and heightened glycolysis and oxidative phosphorylation (NCAM1, MRPL24, ACAT2, B3GAT1, DGKK), were enriched in $CD56^{dim}$NK cells (Fig. 6a and Supplementary Table 6)[54]. Genes regulating transcription, protein modification, and membrane trafficking (TXNDC5, PIGL, GORASP1, ARAP3 and AGAP1) were also enriched in $CD56^{dim}$NK cells, whereas genes encoding proteins that inhibit AKT, G-protein signaling, transcription, and NK cell

activation and survival (PRMT6, RGS1, ZBTB46, MED20, GNAQ) were enriched in CD56$^{-}$NK cells (Fig. 6a and Supplementary Table 6).

Upon stimulation with IL-12, IL-15 and IL-18, as compared with CD56$^{-}$NK cells, CD56$^{dim}$NK cells had higher level activation of genes involved in oxidative phosphorylation, translation, membrane trafficking, and immune activation, including DDX19B, RAB9B, RAB39B, TRMT13, GOLGA8B, NDUFAF4P1, COQ4, ACAD8, USP30, PPARGC1B, KLRF1, CXCR6, and STRA13 (Fig. 6b and Supplementary Table 6). Of note, the mitochondrial deubiquitinase USP30, which is important for protecting against depolarization-induced cell death[55], was also upregulated by stimulation to a greater extent in CD56$^{dim}$NK cells. In contrast, genes encoding proteins that inhibit AKT-mTOR activity and transcription (CITED2 and TSHZ2), and genes associated with fatty acid and cholesterol catabolism (HADH and FDX1), all features of functionally impaired NK cells, were enriched in CD56$^{-}$NK cells (Fig. 6b and Supplementary Table 6).

The findings above indicate that CD56$^{dim}$NK cells exhibit higher metabolic fitness upon stimulation than do CD56$^{-}$ cells. Consistent with this, multiple activators and targets of mTOR were upregulated, including MAPK6, SREF2, RPS6KA3, TFRC and SLC transporters. mTOR inhibitors, such as CITED2, FOXO3, and GSK3B, were downregulated by stimulation in CD56$^{dim}$NK cells (Fig. 6c and Supplementary Table 6). Stimulation of CD56$^{-}$NK cells downregulated mTOR activators or targets, such as MAPK3, MAPK13, MAPK14, SLC35A4 and COQ4 (Fig. 6c and Supplementary Table 6). Biological processes that are critical for NK cell effector function were only enriched in CD56$^{dim}$NK cells (Fig. 6d,e and Supplementary Table 6).

To better assess the effects of CD4$^{+}$T cells and IL-2 on cultured CD56$^{dim}$NK cells, RNA-Seq was performed on CD56$^{dim}$NK cells which had been sorted from 5 day cultures of PBMCs, PBMCs depleted of CD4$^{+}$T cells, or PBMCs depleted of CD4$^{+}$T cells, but supplemented with exogenous IL-2. Gene set enrichment analysis (GSEA) revealed that CD56$^{dim}$NK cells from the full PBMC culture, or from the CD4$^{-}$PBMC culture supplemented with IL-2, but not from the CD4$^{-}$PBMC culture, were enriched for transcripts that are highly expressed in sorted CD56$^{dim}$NK cells, as compared with sorted CD56$^{neg}$NK cells from PBMC culture (Fig. 6f,g and Supplementary Table 7). The convergence of RNA-Seq data from the freshly sorted cells, and the cells cultured under

different conditions, indicates that maintenance of metabolically healthy and functional CD56$^{dim}$NK cells is dependent upon IL-2-producing CD4$^{+}$T cells.

**IL-2 prevents NK cell functional defects caused by mTOR inhibition**

The above results indicate that CD56$^{dim}$ and CD56$^{-}$NK cells have distinct metabolic profiles, with the former having greater mTOR activity (Fig. 6). Glycolysis, oxidative phosphorylation, and RNA and protein synthesis, all processes regulated by mTOR signaling, are required for NK cell killing activity and cytokine production[54,56], and IL-2 induced AKT-mTOR activation is critical for NK cell effector function[57,58]. In fact, IL-2 treatment increased the activity of multiple mTOR components in NK cells, as evidenced by enhanced phosphorylation of mTOR itself on Ser2448, AKT on Ser473, 4EBP1 on Thr36 and Thr45, and S6 on Ser235 and Ser236, and increased surface levels of the transferrin receptor CD71 (Fig. 7a), the synthesis of which is known to be positively regulated by mTOR[59]. Depletion of CD4$^{+}$T cells from cultured PBMCs resulted in downregulation of these markers (Fig. 7b), consistent with the role of CD4$^{+}$T cells in maintaining metabolic fitness of NK cells (Fig. 5e-i and Fig. 6f).

When mTOR was inhibited by incubating PBMCs in rapamycin, the percentage of CD56$^{-}$NK cells was increased (Fig. 7c and Extended Data Fig. 5i) and IFN-γ production of NK cells cultured without CD4$^{+}$T cells was decreased. Moreover, IL-2 treatment rescued IFN-γ production in the presence of rapamycin (Fig. 7d). More potent inhibition by Torin 1 (mTOR kinase inhibitor) led to impaired IFN-γ production by NK cells, even in the presence of CD4$^{+}$T cells; however, this inhibition was overcome by addition to the culture of exogenous IL-2 (Fig. 7e). Thus, NK cell functionality is dependent on mTOR, the activity of which is maintained by IL-2-producing CD4$^{+}$ T cells.

## Discussion

The bulk RNA-Seq profiling reported here identified four discrete subpopulations of lineage negative cells among human PBMCs: ILC2s, ILCPs, CD56$^{hi}$NK cells, and a distinct cluster which included both CD56$^{dim}$ and CD56$^{-}$NK cells (Fig. 2c). Though ILCs and all NK cell subsets were well-separated, CD56$^{hi}$NK cells shared many features with both ILCs and conventional NK cells (Extended Data Fig. 2a,b). Reactome analysis showed that CD56$^{hi}$NK cells share DAP12 signaling with the other NK subsets, as well as translation regulators typical of ILCs (Fig. 2e,f). Confirming the expression data, ATAC-Seq showed that CD56$^{hi}$NK cells have open chromatin at loci found in both CD56$^{dim/-}$NK cells and ILCs (Fig. 2h). Pseudotime analysis of single cell transcriptome data placed CD56$^{hi}$NK cells at a position between ILCs/ILCPs and CD56$^{dim/-}$NK cells (Fig. 3f,g). These observations provide new insight into the relationship between human ILCs and NK cells.

An important discovery here was that NK cells express AREG (Fig. 4a,b), an epidermal growth factor family member which plays important roles in tissue repair and immune tolerance in both mouse models and human disease[30]. Multiple datasets demonstrate that mouse ILC2s express Areg, but that in mouse NK cells, the Areg locus is closed, and Areg is not expressed (Extended Data Fig. 2b, 3b). Furthermore, unlike in humans, the Areg open chromatin region in mouse ILC2s (chr5: 91,286,257-91,290,763) does not have TCF7 binding sites (http://jaspar.genereg.net/). The fact that the regulation of this gene is so different between the two species (Fig. 4e) may explain why production of this homeostatic protein by NK cells has not been reported before. In fact, AREG expression was even higher in CD56$^{hi}$NK cells than in ILC2s (Fig. 4a,b). AREG expression by NK cells suggests that, in addition to their well-characterized antiviral function, human NK cells may actively limit tissue damage and inflammation that accompanies immune activation. Such non-immune host defense strategies that promote host survival without decreasing pathogen burden have been called disease tolerance[60,61].

The AREG-deficiency described here in NK cells isolated from people living with HIV-1 would be expected to contribute to systemic inflammation and increase the risk of cardiovascular disease and cancer in these individuals[62]. Elevated TGF-β1 has been associated with non-AIDS inflammatory disorders[48] and the locus-specific chromatin

assays performed here showed that TGF-β1-RUNX3 signaling negatively regulates AREG expression in NK cells (Fig. 4i,j). Perhaps these effects on AREG expression explain in part the persistence of inflammation in this TGF-β1-induced immunosuppressive environment.

In people living with HIV-1, AREG$^{+}$NK cells were not only decreased in viremic individuals, but also in spontaneous controllers (Fig. 4c). In light of this observation, it is interesting that elite controllers experience more frequent hospitalizations for non-AIDS inflammatory complications than do HIV-1$^{+}$ people on ART[63]. Though many people in the spontaneous controller group studied here do not satisfy the strict criteria required to be labeled as elite controllers, the reduction in AREG$^{+}$ NK cells suggests that spontaneous control of viremia comes at the cost of inflammation. Drugs such as CHIR990212, which boost WNT signaling and AREG production by NK cells (Fig. 4g) might provide benefit to people living with HIV-1 by countering HIV-1-associated inflammation.

Previous studies suggest that the numbers of CD56$^{-}$NK cells increase in people with active HIV-1 replication at the expense of CD56$^{dim}$NK cells[21,22]. However, the mechanism behind this shift towards a functionally defective NK cell population required further qualification. To mimic the CD4$^{+}$T cell depletion caused by HIV-1 *in vivo*, CD4$^{+}$T cells were depleted from PBMCs cultured *ex vivo*. This resulted in decrease in CD56$^{dim}$NK cells and concomitant increase in CD56$^{-}$NK cells (Fig. 5e,h and Extended Data Fig. 5b). In this experimental system, the basal levels of IL-2 secretion by the CD4$^{+}$T cells maintained stable numbers of CD56$^{dim}$NK cells, since exogenous IL-2 was sufficient to replace CD4$^{+}$T cells, and IL-2 neutralization rendered CD4$^{+}$T cells ineffective (Fig. 5h-j). Humanized mice treated with hIL-2-rAAV exhibit higher CD56$^{+}$NK cells in the blood, spleen and liver, and elevated IFN-γ production than their counterparts, indicating that IL-2 plays a critical role in maintaining NK cell homeostasis (Fig. 5m,n).

As compared to CD56$^{dim}$NK cells, CD56$^{-}$NK cells were defective in multiple metabolic pathways, as evidenced by reduced expression, under both resting and stimulated conditions, of genes involved in glycolysis and oxidative phosphorylation (Fig 6a-e). In response to stimulation, many mTOR related genes that required for the upregulation of RNA and protein synthesis (Supplementary Table 6), were upregulated in CD56$^{dim}$NK cells, but not in CD56$^{-}$NK cells (Fig. 6c-e). Given the effect on NK cell subsets

of culturing PBMCs in the absence of $CD4^+$T cells or in the presence of exogenous IL-2 (Fig. 6f,g), it was particularly interesting that exogenous IL-2 promoted mTOR activity whereas $CD4^+$T cell depletion inhibited it (Fig. 7 a,b). Further, NK cell functional defects caused by pharmacologic inhibition of mTOR were rescued by exogenous IL-2 (Fig. 7d,e). Gene set enrichment analysis demonstrated that $CD4^+$T cells and IL-2 are similarly important for maintaining the transcriptional profile of $CD56^{dim}$NK cells (Fig. 6e,f), indicating that mTOR maintains normal NK cell physiology. It has been suggested that mTOR inhibition may serve as an effective intervention to control HIV-1 infection[64–67]. However, the data obtained here suggests that mTOR inhibition would compromise NK cell effector functions important for viral clearance.

## References


1. Barrow, A. D. & Colonna, M. Innate lymphoid cell sensing of tissue vitality. *Curr. Opin. Immunol.* **56**, 82–93 (2019).
2. Klose, C. S. N. & Artis, D. Innate lymphoid cells as regulators of immunity, inflammation and tissue homeostasis. *Nat. Immunol.* **17**, 765–774 (2016).
3. Ebbo, M., Crinier, A., Vély, F. & Vivier, E. Innate lymphoid cells: major players in inflammatory diseases. *Nat. Rev. Immunol.* **17**, 665–678 (2017).
4. Yudanin, N. A. *et al.* Spatial and Temporal Mapping of Human Innate Lymphoid Cells Reveals Elements of Tissue Specificity. *Immunity* **50**, 505–519 (2019).
5. Hazenberg, M. D. & Spits, H. Human innate lymphoid cells. *Blood* **124**, 700–709 (2014).
6. Wang, Y. *et al.* HIV-1-induced cytokines deplete homeostatic innate lymphoid cells and expand TCF7-dependent memory NK cells. *Nat. Immunol.* **21**, 274–286 (2020).
7. Artis, D. & Spits, H. The biology of innate lymphoid cells. *Nature* **517**, 293–301 (2015).
8. Cherrier, D. E., Serafini, N. & Di Santo, J. P. Innate Lymphoid Cell Development: A T Cell Perspective. *Immunity* **48**, 1091–1103 (2018).
9. Spits, H., Bernink, J. H. & Lanier, L. NK cells and type 1 innate lymphoid cells: partners in host defense. *Nat. Immunol.* **17**, 758–764 (2016).
10. Lim, A. I. *et al.* Systemic Human ILC Precursors Provide a Substrate for Tissue ILC Differentiation. *Cell* **168**, 1086–1100 (2017).
11. Li, B. W. S. *et al.* Group 2 Innate Lymphoid Cells Exhibit a Dynamic Phenotype in Allergic Airway Inflammation. *Front. Immunol.* **8**, 1684 (2017).
12. Bal, S. M., Golebski, K. & Spits, H. Plasticity of innate lymphoid cell subsets. *Nat. Rev. Immunol.* **20**, 552–565 (2020).
13. Colonna, M. Innate Lymphoid Cells: Diversity, Plasticity, and Unique Functions in Immunity. *Immunity* **48**, 1104–1117 (2018).
14. Cortez, V. S. *et al.* Transforming Growth Factor-β Signaling Guides the Differentiation of Innate Lymphoid Cells in Salivary Glands. *Immunity* **44**, 1127–1139 (2016).
15. Cortez, V. S. *et al.* SMAD4 impedes the conversion of NK cells into ILC1-like cells by curtailing non-canonical TGF-β signaling. *Nat. Immunol.* **18**, 995–1003 (2017).
16. Chen, L. *et al.* CD56 Expression Marks Human Group 2 Innate Lymphoid Cell Divergence from a Shared NK Cell and Group 3 Innate Lymphoid Cell Developmental Pathway. *Immunity* **49**, 464–476 (2018).
17. Lugli, E., Marcenaro, E. & Mavilio, D. NK Cell Subset Redistribution during the Course of Viral Infections. *Front. Immunol.* **5**, 390 (2014).
18. Kløverpris, H. N. *et al.* Innate Lymphoid Cells Are Depleted Irreversibly during Acute HIV-1 Infection in the Absence of Viral Suppression. *Immunity* **44**, 391–405 (2016).
19. Monticelli, L. A. *et al.* IL-33 promotes an innate immune pathway of intestinal tissue protection dependent on amphiregulin–EGFR interactions. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 10762–10767 (2015).
20. Poli, A. *et al.* CD56brightnatural killer (NK) cells: an important NK cell subset. *Immunology* **126**, 458–465 (2009).
21. Björkström, N. K., Ljunggren, H.-G. & Sandberg, J. K. CD56 negative NK cells: origin, function, and role in chronic viral disease. *Trends Immunol.* **31**, 401–406 (2010).

22. Mavilio, D. *et al.* Characterization of CD56–/CD16+ natural killer (NK) cells: A highly dysfunctional NK subset expanded in HIV-infected viremic individuals. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 2886–2891 (2005).
23. Mizrahi, S. 'ar, Markel, G., Porgador, A., Bushkin, Y. & Mandelboim, O. CD100 on NK Cells Enhance IFNγ Secretion and Killing of Target Cells Expressing CD72. *PLoS One* **2**, e818 (2007).
24. Eriksson, E. M. *et al.* Expansion of CD8+ T cells lacking Sema4D/CD100 during HIV-1 infection identifies a subset of T cells with decreased functional capacity. *Blood* **119**, 745–755 (2012).
25. He, Y. *et al.* Interferon-α-Enhanced CD100/Plexin-B1/B2 Interactions Promote Natural Killer Cell Functions in Patients with Chronic Hepatitis C Virus Infection. *Front. Immunol.* **8**, 1435 (2017).
26. Kumanogoh, A. & Kikutani, H. Biological functions and signaling of a transmembrane semaphorin, CD100/Sema4D. *Cell. Mol. Life Sci.* **61**, 292–300 (2004).
27. Turnbull, I. R. & Colonna, M. Activating and inhibitory functions of DAP12. *Nat. Rev. Immunol.* **7**, 155–161 (2007).
28. Lanier, L. L. DAP10- and DAP12-associated receptors in innate immunity. *Immunol. Rev.* **227**, 150–160 (2009).
29. Vivier, E. *et al.* Innate Lymphoid Cells: 10 Years On. *Cell* **174**, 1054–1066 (2018).
30. Zaiss, D. M. W., Gause, W. C., Osborne, L. C. & Artis, D. Emerging functions of amphiregulin in orchestrating immunity, inflammation, and tissue repair. *Immunity* **42**, 216–226 (2015).
31. Shih, H.-Y. *et al.* Developmental Acquisition of Regulomes Underlies Innate Lymphoid Cell Functionality. *Cell* **165**, 1120–1133 (2016).
32. Pokrovskii, M. *et al.* Characterization of Transcriptional Regulatory Networks that Promote and Restrict Identities and Functions of Intestinal Innate Lymphoid Cells. *Immunity* **51**, 185–197 (2019).
33. Yoshida, H. *et al.* The cis-Regulatory Atlas of the Mouse Immune System. *Cell* **176**, 897–912 (2019).
34. Gibbons, H. R. *et al.* Divergent lncRNA GATA3-AS1 Regulates GATA3 Transcription in T-Helper 2 Cells. *Front. Immunol.* **9**, 2512 (2018).
35. Maric, J. *et al.* Cytokine-induced endogenous production of prostaglandin D2 is essential for human group 2 innate lymphoid cell activation. *J. Allergy Clin. Immunol.* **143**, 2202–2214 (2019).
36. Shaulian, E. & Karin, M. AP-1 in cell proliferation and survival. *Oncogene* **20**, 2390 (2001).
37. Crinier, A. *et al.* High-Dimensional Single-Cell Analysis Identifies Organ-Specific Signatures and Conserved NK Cell Subsets in Humans and Mice. *Immunity* **49**, 971–986 (2018).
38. 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).
39. Qiu, X. *et al.* Single-cell mRNA quantification and differential analysis with Census. *Nat. Methods* **14**, 309–315 (2017).
40. Zook, E. C. *et al.* The ETS1 transcription factor is required for the development and cytokine-induced expansion of ILC2. *J. Exp. Med.* **213**, 687–696 (2016).
41. Barton, K. *et al.* The Ets-1 transcription factor is required for the development of

natural killer cells in mice. *Immunity* **9**, 555–563 (1998).
42. Walker, J. A., Barlow, J. L. & McKenzie, A. N. J. Innate lymphoid cells — how did we miss them? *Nat. Rev. Immunol.* **13**, 75–87 (2013).
43. Monticelli, L. A. *et al.* Innate lymphoid cells promote lung-tissue homeostasis after infection with influenza virus. *Nat. Immunol.* **12**, 1045–1054 (2011).
44. Shan, Q. *et al.* The transcription factor Runx3 guards cytotoxic CD8+ effector T cells against deviation towards follicular helper T cell lineage. *Nat. Immunol.* **18**, 931–939 (2017).
45. Ito, K. *et al.* RUNX3 attenuates beta-catenin/T cell factors in intestinal tumorigenesis. *Cancer Cell* **14**, 226–237 (2008).
46. Ito, Y. & Miyazono, K. RUNX transcription factors as key targets of TGF-beta superfamily signaling. *Curr. Opin. Genet. Dev.* **13**, 43–47 (2003).
47. Ikushima, H. & Miyazono, K. TGFβ signalling: a complex web in cancer progression. *Nat. Rev. Cancer* **10**, 415–424 (2010).
48. Theron, A. J., Anderson, R., Rossouw, T. M. & Steel, H. C. The Role of Transforming Growth Factor Beta-1 in the Progression of HIV/AIDS and Development of Non-AIDS-Defining Fibrotic Disorders. *Front. Immunol.* **8**, 1461 (2017).
49. Wiercińska-Drapalo, A., Flisiak, R., Jaroszewicz, J. & Prokopowicz, D. Increased Plasma Transforming Growth Factor-β1 Is Associated with Disease Progression in HIV-1-Infected Patients. *Viral Immunol.* **17**, 109–113 (2004).
50. Liu, P. *et al.* Enhanced Cas12a editing in mammalian cells and zebrafish. *Nucleic Acids Res.* **47**, 4169–4180 (2019).
51. Jost, S. *et al.* CD4+ T-cell help enhances NK cell function following therapeutic HIV-1 vaccination. *J. Virol.* **88**, 8349–8354 (2014).
52. Wu, Z., Frascaroli, G., Bayer, C., Schmal, T. & Mertens, T. Interleukin-2 from Adaptive T Cells Enhances Natural Killer Cell Activity against Human Cytomegalovirus-Infected Macrophages. *J. Virol.* **89**, 6435–6441 (2015).
53. Das, L. & Levine, A. D. TGF-β inhibits IL-2 production and promotes cell cycle arrest in TCR-activated effector/memory T cells in the presence of sustained TCR signal transduction. *J Immunol.* **180**, 1490-1498 (2008).
54. Poznanski, S. M. & Ashkar, A. A. What Defines NK Cell Functional Fate: Phenotype or Metabolism? *Front. Immunol.* **10**, 1414 (2019).
55. Liang, J.-R. *et al.* USP 30 deubiquitylates mitochondrial P arkin substrates and restricts apoptotic cell death. *EMBO Rep.* **16**, 618–627 (2015).
56. Ganeshan, K. & Chawla, A. Metabolic regulation of immune responses. *Annu. Rev. Immunol.* **32**, 609–634 (2014).
57. Ray, J. P. *et al.* The Interleukin-2-mTORc1 Kinase Axis Defines the Signaling, Differentiation, and Metabolism of T Helper 1 and Follicular B Helper T Cells. *Immunity* **43**, 690–702 (2015).
58. Wu, Y., Tian, Z. & Wei, H. Developmental and Functional Control of Natural Killer Cells by Cytokines. *Front. Immunol.* **8**, 930 (2017).
59. Zheng, Y. *et al.* A role for mammalian target of rapamycin in regulating T cell activation versus anergy. *J. Immunol.* **178**, 2163–2170 (2007).
60. Medzhitov, R., Schneider, D. S. & Soares, M. P. Disease tolerance as a defense strategy. *Science* **335**, 936–941 (2012).
61. Ayres, J. S. The Biology of Physiological Health. *Cell* **181**, 250–269 (2020).

62. Deeks, S. G., Tracy, R. & Douek, D. C. Systemic effects of inflammation on health during chronic HIV infection. *Immunity* **39**, 633–645 (2013).
63. Crowell, T. A. *et al.* Hospitalization Rates and Reasons Among HIV Elite Controllers and Persons With Medically Controlled HIV Infection. *J. Infect. Dis.* **211**, 1692–1702 (2015).
64. Heredia, A. *et al.* Targeting of mTOR catalytic site inhibits multiple steps of the HIV-1 lifecycle and suppresses HIV-1 viremia in humanized mice. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 9412–9417 (2015).
65. Besnard, E. *et al.* The mTOR Complex Controls HIV Latency. *Cell Host Microbe* **20**, 785–797 (2016).
66. Martin, A. R. *et al.* Rapamycin-mediated mTOR inhibition uncouples HIV-1 latency reversal from cytokine-associated toxicity. *J. Clin. Invest.* **127**, 651–656 (2017).
67. Donia, M., McCubrey, J. A., Bendtzen, K. & Nicoletti, F. Potential use of rapamycin in HIV infection. *Br. J. Clin. Pharmacol.* **70**, 784–793 (2010).
68. Hashimshony, T. *et al.* CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq. *Genome Biol.* **17**, 77 (2016).
69. Kim, D. *et al.* TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36 (2013).
70. Derr, A. *et al.* End Sequence Analysis Toolkit (ESAT) expands the extractable information from single-cell RNA-seq data. *Genome Res.* **26**, 1397–1410 (2016).
71. Love, M., Anders, S. & Huber, W. Differential analysis of count data--the DESeq2 package. *Genome Biol.* **15**, 550 (2014).
72. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
73. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).
74. Qiu, X. *et al.* Reversed graph embedding resolves complex single-cell trajectories. *Nat. Methods* **14**, 979–982 (2017).
75. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
76. Robinson, J. T. *et al.* Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).
77. Kleinstiver, B. P. *et al.* Engineered CRISPR-Cas12a variants with increased activities and improved targeting ranges for gene, epigenetic and base editing. *Nat. Biotechnol.* **37**, 276–282 (2019).
78. Rautela, J., Surgenor, E. & Huntington, N. D. Drug target validation in primary human natural killer cells using CRISPR RNP. *J. Leukoc. Biol.* **108**, 1397–1408 (2020).
79. Brehm, M. A. *et al.* Parameters for establishing humanized mouse models to study human immunity: analysis of human hematopoietic stem cell engraftment in three immunodeficient strains of mice bearing the IL2rgamma(null) mutation. *Clin. Immunol.* **135**, 84–98 (2010).
80. Hasgur, S., Aryee, K. E., Shultz, L. D., Greiner, D. L. & Brehm, M. A. Generation of Immunodeficient Mice Bearing Human Immune Systems by the Engraftment of Hematopoietic Stem Cells. *Methods Mol. Biol.* **1438**, 67–78 (2016).
81. He, Y. *et al.* Kinetics of adeno-associated virus serotype 2 (AAV2) and AAV8 capsid antigen presentation in vivo are identical. *Hum. Gene Ther.* **24**, 545–553 (2013).
82. McCarty, D. M., Monahan, P. E. & Samulski, R. J. Self-complementary recombinant

adeno-associated virus (scAAV) vectors promote efficient transduction independently of DNA synthesis. *Gene Ther.* **8**, 1248–1254 (2001).

## Acknowledgements

We thank the study participants who provided blood samples and their caretakers. M. Krone (UCSF) provided Institutional Review Board regulatory assistance, sample preparation, and record keeping. S. Maitland provided advice on the purification of programmable nuclease proteins. Ruijia Wang assisted with data analysis. This research was supported by NIH grants R37AI147868 to J.L., R01HL150669 to S.A.W., R24OD026440, R01AI132963, and UC4DK104218 to M.A.B., F30HD100110 to N.J.S., and F31HL147482 to K.L. The UCSFbased SCOPE cohort was supported by the UCSF/Gladstone Institute of Virology and Immunology CFAR (P30 AI027763), and the CFAR Network of Integrated Systems (R24 AI067039). The content of this publication does not necessarily reflect the views or policies of DHHS, nor does the mention of trade names, commercial products, or organizations imply endorsement by the US Government.

## Declaration of Interests

M.A.B. is a consultant for The Jackson Laboratory. S.A.W. is a consultant for Chroma Medicine. K.L. and S.A.W. have filed a patent application related to genome editing reagents described in this work.

## Methods

**Clinical samples.** PBMCs of HIV-1$^{-}$, HIV-1$^{+}$ viremic, HIV-1$^{+}$ on ART, and HIV-1$^{+}$, controllers used in this study, were from the University of California San Francisco SCOPE Cohort. The clinical characteristics of each individual in the SCOPE cohort was provided in Supplementary Table 5. All participants provided written informed consent for protocols that were included in study of cellular immunity in HIV-1 infection, in accordance with procedures approved by the University of Massachusetts Medical School (UMMS), and the University of California, San Francisco (UCSF) Institutional Review Boards. All participants were consented the day of the procedure. All PBMCs from the SCOPE cohort were tested blindly with the code broken after analysis, and no samples were excluded from the analysis.

**Human mononuclear cell isolation and NK cell enrichment.** Human peripheral blood was diluted in equal volume of RPMI-1640 (Gibco), overlaid on Lymphoprep (STEMSELL, 07851), and centrifuged at 500 x g at room temperature for 30 minutes. Mononuclear cells were washed 3 times with MACS buffer (0.5% BSA and 2 mM EDTA in PBS), and either used immediately, or frozen in FBS containing 10% DMSO. NK cells were negatively enriched using EasySep™ Human NK Cell Isolation Kit (STEMCELL, 17955), according to the manufacturer's instructions.

**Flow cytometry.** Cells were first stained with Live and Dead violet viability kit (Invitrogen, L-34963). To detect surface molecules, cells were stained in MACS buffer with antibodies for 30 min at 4°C in the dark. To detect IFN-γ or AREG, cells were stimulated with the IL-12 (10 ng/ml), IL-15 (50 ng/ml) and IL-18 (50 ng/ml) for 16 hrs, or with PMA and ionomycin (eBioscience, 00-4970-03) for 3 hrs. In both cases, protein transport inhibitors (eBioscience, 00-4980-03) were present during the stimulation. To detect transcription factors or cytokines, cells were fixed and permeabilized using Foxp3 staining buffer kit (eBioscience, 00-5523-00), then intracellular molecules were stained in permeabilization buffer with antibodies.

**NK cell and ILC sorting.** PBMCs were stained with a panel of lineage markers, CD56, CD16 and CD127. The CD56$^{hi}$, CD56$^{dim}$, CD56$^{-}$ and ILCs were sorted as indicated in Extended Data Fig. 1a and c using a BD FACSAria IIu. Enrichment of the sorted cell populations was confirmed by flow cytometry before initiation of downstream experiments.

**Bulk RNA-Seq Library preparation.** The sequencing library was prepared using CEL-Seq2[68]. RNA from sorted cells was extracted using TRIzol reagent (ThermoFisher, 15596018). 10 ng RNA was used for first strand cDNA synthesis using barcoded primers (the specific primers for each sample were listed in Supplementary Table 8). The second strand was synthesized by NEBNext Second Strand Synthesis Module (NEB, E6111L). The pooled dsDNA was purified with AMPure XP beads (Beckman Coulter, A63880), and subjected to in vitro transcription (IVT) using HiScribe T7 High Yield RNA Synthesis Kit (NEB, E2040S), then treated with ExoSAP-IT (Affymetrix, 78200). IVT RNA was fragmented using RNA fragmentation reagents (Ambion), and underwent another reverse transcription step using random hexamer RT primer-5'-GCC TTG GCA CCC GAG AAT TCC ANN NNN N-3' to incorporate the second adapter. The final library was amplified with indexed primers: RP1 and RPI1 or RPI2 (as indicated in Supplementary Table 8), and the beads purified library was quantified with 4200 TapeStation (Agilent Technologies), and paired-end sequenced on Nextseq 500 V2 (Illumina), Read 1: 15 cycles; index 1: 6 cycles; Read 2: 60 cycles.

**Library preparation for single cell RNA-Seq.** The sequencing library was prepared by Single Cell 3' Reagent Kits v2 (10xgenomics, 120234). Isolated cells were 2x washed and resuspended in 1xPBS containing 0.05% BSA. Cell number and viability were measured by Bio-Rad TC 20 cell counter, cell concentration was adjusted around 1000-1500 cells/ul (viability>90%). Single cell suspension was loaded onto Chromium Controller (10xGenomics) to participate 3000-6000 single cells into gel beads in emulsions (GEMs). Libraries were constructed according to the instruction of single cell 3' reagent kits v2. The yield and the quality of amplified cDNA were checked using High Sensitivity D5000 ScreenTape on TapeStation (Agilent Technologies). The final library

was amplified using PCR cycles determined by cDNA quantification, and the quality of the library was checked again by TapeStation. The sequencing depth was controlled at more than 50,000 reads per cell. Sorted $CD56^{-}$, $CD56^{dim}$, $CD56^{hi}$NK cells and ILCs from three donors were pooled and were sequenced as one sample, indexed using primers from the B2 position of the index plate.

**ATAC-Seq.** Nuclei were precipitated after sorted $CD56^{hi}$, $CD56^{dim}$, $CD56^{-}$ NK cells and ILCs were suspended in lysis buffer (10 mM Tris·Cl, pH 7.4; 10 mM NaCl; 3 mM $MgCl_2$, 0.1% NP-40). 2.5 ul of Tn5 transposase (Nextera DNA Library Prep Kit, Illumina, FC-121-1030) was added per 50 ul reaction, for 30 min at 37°C. Released DNA fragments were purified by PCR product purification kit (Promega, A9282) and used to generate libraries. The barcode primers and common primers used for each sample from 2 donors were listed in Supplementary Table 8. Libraries were paired-end sequenced on Nextseq 500 V2 (Illumina) using Read 1: 42 cycles; Index 1: 8 cycles, and Read 2: 42 cycles.

**Bulk RNA-Seq Processing and Analysis.** The pooled reads were separated by CEL-Seq2 barcodes and mapped to the hg19 genome using Tophat[69] (version 2.0.14, default parameters). ESAT[70] was used to quantify aligned reads using a transcript annotation file containing all RefSeq genes filtered to select only 'NM' transcripts, and extending the transcripts up to 1,000 bases past the end of the annotated 3' end (*-wExt 1000, -task score3p*), multi-mapped reads (*-multimap ignore*) were discarded. DEBrowser was used to analyze the most varied genes, DESeq2 was used to perform differential expression analysis[71]. Data were transformed using rlog within DEseq2, and prcomp was used to calculate the PCs for principal component analysis (PCA). Mouse RNA-seq data were downloaded from GSE77695, GSE109125, and GSE116092, and aligned to the mouse reference genome (mm10) using STAR (version 2.1.6)[72]. Counts of reads aligned to RefSeq genes were quantified using RSEM (version 1.3.1)[73] and normalized using DEseq2.

**Single cell RNA-Seq processing and analysis**. The Cell Ranger software package (10X Genomics) was used to perform sample demultiplexing, barcode processing, and

single cell 3' gene counting. Default settings were used. The filtered gene matrices generated by Cell Ranger were used as input into the open-source R package Seurat 3.0 (http://satijalab.org/seurat/) (Satija et al., 2015). FindNeighbors and FindClusters (using the default resolution of 0.6) were then run on the PCA reduction. FindMarkers was used to identify differentially expressed genes among clusters. Pseudotime analysis was done using monocle (version 2)[74]. Cluster information (for color coding the cells for display) was imported from the Seurat analysis. The raw count data from the Seurat object was read into monocle and the standard workflow was followed: estimateSizeFactors, estimateDispersions, and detectGenes(min_expr=1) were called. Then, the dispersion table was calculated and subsetted to generate ordering genes with mean_expression >= 0.1 and dispersion_empirical >= dispersion_fit. The data was reduced to two dimensions and the cells were ordered (using the orderCells function). plot_cell_trajectory was used to create pseudotime trajectories and plot_genes_in_pseudotime (with its default settings, which rescales expression to relative values) used to show individual gene expression as a function of pseudotime.

**ATAC-Seq analysis.** Paired-end reads were filtered with trimmomatic (version 0.32), aligned with Bowtie2 (version 2.2.3) to a reference genome hg19. The duplicates were removed using Picard's MarkDuplicates (version 0.32). To be able to accurately call the peaks, each aligned read was first trimmed to the 9-bases at the 5' end, the region where the Tn5 transposase cuts the DNA. To smooth the peaks, the start site of the trimmed reads were extended up and down stream for 10 bases. Peaks were called using these adjusted aligned reads with MACS2 [75]. The adjusted aligned reads were converted to tdf files using IGVTools (IGVtools count –w5) (version 2.3.31)[76] for visualization. The Tn5 transposase accessible region was used as input for JASPAR motif analysis (http://jaspar.genereg.net/). Mouse ATAC-Seq data was downloaded from GSE77695, GSE100738, and GSE116091, and was analyzed as above using mm10 as the reference genome.

**Cas12a RNP-mediated knockout in NK cells.** NK cells from PBMCs were isolated with EasySep™ Human NK Cell Isolation Kit (STEMCELL, 17955), and after culture in NK

MACS Medium (MACS,130-114-429) for 7 days, the NK cells were ready for electroporation. Two Cas12a target sites (PAM underlined) within RUNX3 sequences were chosen for editing: 5'-TTTCACCCTGACCATCACTGTGTTCAC-3' and 5'-CTTACCTCGCCCACTGCGGCCCACGAA-3'. The following Cas12a target site in adeno-associated virus integration site 1 (AAVS1) was used as a control chromosomal DNA target: 5'-TTTATCTGTCCCCTCCACCCCACAGTG-3'. Chemically end-protected (Alt-R) AsCas12a crRNAs were synthesized for each target site by IDT. Enhanced AsCas12a (enAsCas12a)[77] coding sequence was engineered to include three nuclear localization signal sequences and cloned into a pet21a expression vector for expression in, and purification from *E. coli*, as previously described[50]. For each electroporation, 200 pmol crRNA and 100 pmol recombinant enAsCas12a protein were mixed at room temperature for 20 min. 1 million NK cells were resuspended in 20 ul 4D nucleofector master mix (82% P3 + 18% supplement 1; Lonza, V4XP-3032) and then mixed with Cas12a RNPs for electroporation using program CM137[78]. The electroporated NK cells were ready for downstream experiments after culture in MACS medium for 5 days.

***In vivo* treatment of IL-2.** NOD-*Prkdc*$^{scid}$*IL2rg*$^{tm1Wjl}$ (NSG) mice were purchased from The Jackson Laboratory (Bar Harbor, ME) and engrafted with human CD34+ hematopoietic stem cells derived from umbilical cord blood (UCB). UCB was obtained from donors that were consented under an approved IRB protocol at the UMass Memorial Medical Center, Department of General Obstetrics and Gynecology (Worcester, MA), and all samples used for engraftment were de-identified. UCB was processed as previously described and underwent CD3 T cell depletion[79]. For engraftment, NSG mice between 3 to 4 weeks of age received 100 cGy irradiation and were then injected IV with 100,000 CD34+ cells[80]. Human immune system development was evaluated by flow cytometry at 16 weeks post-HSC injection.

The double-stranded (ds) adeno-associated virus (AAV) vectors were engineered and packaged as previously described[81]. Briefly, full-length cDNA encoding human IL-2 were subcloned into a dsAAV plasmid[82] containing the murine preproinsulin II promoter. dsAAV vector packaging with serotype 8 capsid protein was produced by the Viral Vector Core at the University of Massachusetts Medical School Horae Gene Therapy Center

(Worcester, MA, USA). HSC-engrafted NSG mice were intraperitoneally injected with 2.5 × $10^{11}$ particles of the purified AAV8-huIL-2 (AAV-IL-2). 6 weeks later, the blood, spleen, and liver were harvested for detection of CD56 or IFN-γ production from NK cells.

**Statistical analysis**. Statistical test was performed using GraphPad Prism. The usage of paired or unpaired two-tailed student's *t*-test was indicated in the figure legends. $p<0.05$ was considered as significant. Variance was estimated by calculating the mean ± s.e.m. in each group. Variances among groups of samples were compared using the F-test function in GraphPad.

**Data availability.** The data that support the findings of this study are available within the manuscript and in its supplementary information files. Bulk and single-cell RNA-Seq, and ATAC-Seq datasets, can be found at: https://www.ncbi.nlm.nih.gov/gds/?term=GSE168212.

## Figure Legends

**Fig.1 | Characterization of Lin$^-$ lymphoid cells among human PBMCs.**

**a,** CD127$^+$ and CD127$^-$ cells from Lin$^-$CD56$^-$ population after gating on lymphoid, singlet, live, CD45$^+$ cells of PBMCs. Lineage (Lin) markers include antibodies against: CD3, CD4, TCRαβ, TCRγδ, CD19, CD20, CD22, CD34, FcεRIα, CD11c, CD303, CD123, CD1a, and CD14. **b**, Heatmap of differentially expressed genes by RNA-Seq, sorted Lin$^-$CD56$^-$CD16$^-$CD127$^+$ versus Lin$^-$CD56$^-$CD16$^-$CD127$^-$ cells, from PBMCs of five donors (log2 fold change >1, padj<0.01 determined by DESeq2). **c,** normalized counts of Lin$^-$CD56$^-$CD16$^-$CD127$^+$ (red) and Lin$^-$CD56$^-$CD16$^-$CD127$^-$ (blue) cells related genes from **b** (n=5). **d,e,** Reactome analysis based on enriched transcripts of Lin$^-$CD56$^-$CD16$^-$CD127$^+$ cells (**d**) or Lin$^-$CD56$^-$CD16$^-$CD127$^-$ cells (**e**). **f,** Gating strategy for CD56$^{hi}$NK, CD56$^{dim}$NK, CD56$^-$NK cells and ILCs. **g,** The indicated populations as in **f** were detected with TBX21, CRTH2 and RORγT. All data were generated using blood from HIV-1$^-$ donors.

**Fig.2 | Transcriptional and chromatin accessibility analysis of human blood NK cells and ILCs.**

**a,** Heatmap of differentially expressed genes by RNA-Seq, sorted CD56$^{hi}$NK, CD56$^{dim}$NK, CD56$^-$NK cells and ILCs, from PBMCs of 4 donors (log2 fold change >1, p<0.01 determined by DESeq2). **b,** Normalized counts of AREG from **a**. **c,** PCA based on RNA-Seq data of indicated populations. **d,** Number of enriched pathways based on differentially expressed genes of indicated NK subsets versus ILCs by Go Enrichment Analysis. **e,** Enriched pathways shared by all NK subsets as compared with ILCs. **f,** Shared enriched pathways of ILCs compared with CD56$^{dim}$ or CD56$^-$NK cells. **g,** Enriched pathways of ILCs compared with CD56$^{hi}$NK cells. **h,** ATAC-Seq analysis of sorted CD56$^{hi}$NK, CD56$^{dim}$NK, CD56$^-$NK cells and ILCs at indicated gene loci (representative of two donors). All data were generated using blood from HIV-1$^-$ donors.

**Fig.3 | Single cell transcriptome analysis of blood NK cells and ILCs.**

**a,** Two-dimensional t-SNE plot of single cell RNA-Seq of sorted CD56$^{hi}$NK, CD56$^{dim}$NK, CD56$^-$NK cells and ILCs. **b,c,** Expression and density of the indicated genes within t-SNE

plots. **d,** Heatmap of uniquely higher or lower expressed genes of NK cell or ILC clusters. **e,** Representative genes of indicated clusters. **f,** Minimum spanning tree based on the transcriptome of individual cells from **a**, showing pseudotime trajectory (black line, cells are color coded by clusters). **g,** the expression of genes that are enriched by NK cells and/or ILCs along the pseudotime trajectory. All data were generated using blood from HIV-1$^-$ donors.

**Fig.4 | NK cells are major AREG producers among lineage negative PBMCs.**

**a,** PBMCs were stimulated with PMA and ionomycin for 3 hrs then AREG from different subsets of NK cells and ILCs were detected. **b,** The mean fluorescence intensity (MFI) of AREG detected in **a** (n=20). **c,** PBMCs from HIV-1$^-$ (n=20), HIV-1$^+$ viremic (n=20), HIV-1$^+$ ART suppressed (n=19), and HIV-1$^+$ spontaneous controllers (n=20) were stimulated as described in **a**, the AREG$^+$ cells from indicated populations were detected. **d,** Correlation of AREG$^+$NK cells with ILCs (n=79) Correlation coefficient (r) by Pearson, zero slope p value determined by the F-test. **e,** Schematic map of AREG open chromatin region in NK cells detected by ATAC-Seq, TCF7 and RUNX3 binding motif were predicted by JASPAR motif analysis (http://jaspar.genereg.net/). **f,** Normalized counts of RUNX3 by RNA-Seq (n=4). **g,h,** PBMCs were treated with CHIR99021 (10uM) for 48 hrs, then stimulated as described in **a**, the percentage of AREG$^+$NK cells (**g**) and AREG$^+$ILCs (**h**) were shown (n=15). **i,** As in **g**, PBMCs were additionally treated with CHIR99021 and TGF-β1 before AREG$^+$NK cells were detected (n=8). **j,** AREG$^+$NK cells were detected after IL-12+IL-15+IL-18 stimulation for 16 hrs from control or RUNX3 knockout groups (n=4). Data are mean ± s.e.m., **b,f-j,** two-tailed paired *t*-test. **c,** two-tailed unpaired *t*-test. ns, not significant, *p<0.05, **p<0.01, ***p<0.001. For **c,d,** cohort was described in Supplementary Table 5.

**Fig.5 | CD4$^+$T cell depletion destabilizes NK cells.**

**a,** Percentage of ILCs from Lin$^-$CD56$^-$ population. HIV-1$^-$ (n=40), HIV-1$^+$ viremic (n=40), HIV-1$^+$ ART suppressed (n=37), and HIV-1$^+$ spontaneous controllers (n=40). **b,** Correlation of ILCs and CD56$^-$NK cells in Lin$^-$CD56$^-$ population from **a**. **c,** Percentage of CD56$^-$ and CD56$^{dim}$ NK cells in total NK cells in indicated groups (n=40 for each group).

**d,** Correlation of CD56$^{-}$ and CD56$^{dim}$NK cells from **c**. **e,** Percentage of CD56$^{-}$NK cells in total NK cells were detected from PBMCs and CD4$^{-}$PBMCs after culture for 5 days (n=7). **f,** CD4$^{+}$T cell number (counts/mm$^{3}$) from indicated groups (n=40 for each group). **g,** Correlation of CD4 number and CD56$^{-}$NK cells (n=160). **h,** Detection of CD56$^{dim}$ and CD56$^{-}$ NK cells after culture for 5 days in indicated conditions. **i,** Percentage of CD56$^{-}$NK cells in total NK cells was detected from CD4$^{-}$PBMCs in the presence or absence of IL-2 (10ng/ml) combined with or without IL-2 neutralizing antibody (4ug/ml) (n=7). **j,** Percentage of CD56$^{-}$NK cells in total NK cells was detected from PBMCs after treatment with isotype or IL-2 neutralizing antibody (4ug/ml) for 5 days (n=10). **k,l,** Percentage of CD56$^{-}$NK (**k**) or IFN-γ$^{+}$ NK (**l**) cells in total NK cells was detected from PBMCs after treatment with TGF-β1 (50ng/ml) or TGF-β1 (50ng/ml)+IL-2 (20ng/ml) for 5 days (n=8). **m,n,** NSG mice reconstituted with human CD34$^{+}$ hematopoietic stem cells were injected intravenously with PBS (n=5) or with hIL-2-rAAV (n=4), 6 weeks later, the blood, spleen and liver were harvested for detection of CD56 (**k**) or IFN-γ production (**l**) from human NK cells. Data are mean ± s.e.m., **a,c,f,m,n,** two-tailed unpaired *t*-test. **e,i-l,** two-tailed paired *t*-test. ns, not significant, *p<0.05, **p<0.01, ***p<0.001. For **a-d,f-h,** cohort was described in Supplementary Table 5.

**Fig. 6 | CD56$^{dim}$NK cells are distinguished from CD56$^{-}$NK cells by metabolic gene expression.**

**a,b,** Heatmap of metabolism and immune related genes that differentially express between CD56$^{-}$ and CD56$^{dim}$ NK cells directly sorted from PBMCs before (a, n=4) or after IL-12+IL-15+1L-18 stimulation (b, n=3) (log2 fold change >1, p<0.01 determined by DESeq2). **c,** Log2 fold change of mTOR signaling related genes of IL-12+IL-15+IL-18 stimulated versus un-stimulated CD56$^{-}$ or CD56$^{dim}$ NK cells. ns, not significant, **p<0.01, ***p<0.001. **d,e,** Protein classes analysis of CD56$^{-}$ or CD56$^{dim}$NK cells enriched genes after IL-12+IL-15+IL-18 stimulation. **f,g,** GSEAs comparing the CD56$^{dim}$NK cell expression signature in PBMCs (determined by RNA-seq, CD56$^{dim}$ versus CD56$^{-}$NK cells, Supplementary Table 7) from sorted CD56$^{dim}$NK cells in PBMC versus CD56$^{dim}$NK cells in CD4$^{-}$PBMC (f) or CD56$^{dim}$NK in CD4$^{-}$PBMC with IL-2 versus CD56$^{dim}$NK in CD4$^{-}$PBMC (g) after culture for 5 days. All data were generated using blood from HIV-1$^{-}$

donors.

**Fig. 7 | IL-2 overcomes the effects of CD4$^+$T cell depletion or MTOR inhibition.**
**a,** PBMCs were cultured with or without IL-2 (50ng/ml) for 5 days, the phosphorylation of mTOR on Ser2448, AKT on Ser473, 4EBP1 on Thr36 and Thr45, and S6 on Ser235 and Ser236, and the surface CD71 from NK cells (Lin$^-$TBX21$^+$) were detected by flow cytometry (n=8). **b,** PBMCs or CD4$^-$PBMCs were cultured for 5 days, the indicated targets from NK cells were detected as in **a** (n=8). **c,** PBMCs were treated with or without rapamycin (10nM) for 5 days, the percentage of CD56$^-$NK cells in Lin$^-$TBX21$^+$ population was detected (n=7). **d,e,** PBMCs or CD4$^-$PBMCs were cultured in the presence or absence of rapamycin (10nM) (**d**) or Torin 1 (250nM) (**e**) combined with or without IL-2 (10ng/ml) for 5 days, then were stimulated with IL-12+IL-15+IL-18 for 16 hrs, the IFN-γ from Lin$^-$TBX21$^+$ population were detected. Data are mean ± s.e.m., two-tailed paired *t*-test. ns, not significant, *p<0.05, **p<0.01, ***p<0.001. All data were generated using blood from HIV-1$^-$ donors.

**Extended Data Fig. 1 | Analysis of lineage negative PBMCs.**

**a,** Sorting strategy for Lin$^{-}$CD45$^{+}$CD56$^{-}$CD127$^{-}$ and Lin$^{-}$CD45$^{+}$CD56$^{-}$CD127$^{+}$cells. **b,** Lin$^{-}$CD45$^{+}$CD56$^{-}$PBMCs were detected with CD127, CD117 and CRTH2. **c,** Sorting strategy for CD56$^{-}$, CD56$^{dim}$ and CD56$^{hi}$NK cells and ILCs. All data were generated using blood from HIV-1$^{-}$ donors.

**Extended Data Fig. 2 | Representative transcripts of NK cells and ILCs**.

**a,** Representative genes shared between CD56$^{hi}$NK cells and ILCs. **b-e,** DESeq2 normalized counts of Areg (**b**), Gata3 (**c**), Rorc (**d**), and Tbx21 (**e**), in mouse ILC1s (n=7), ILC2s (n=12), ILC2ps (n=2), ILC3s (n=19) and NK cells (n=19). Data are mean ± s.e.m., two-tailed unpaired *t*-test. *p<0.05, **p<0.01, ***p<0.001. **f-h,** Enriched genes in all human NK cell subsets compared with ILCs (**f**), enriched genes in ILCs and CD56$^{hi}$NK cells compared with CD56$^{-}$ and CD56$^{dim}$NK cells (**g**), enriched genes in ILCs compared with CD56$^{hi}$NK cells (**h**). For **a,f-h,** Data were generated using blood from HIV-1$^{-}$ donors. For **b-e,** Mouse data was derived from GSE77695, GSE109125, and GSE116092.

**Extended Data Fig. 3 | Chromatin accessibility in NK cells and ILCs.**

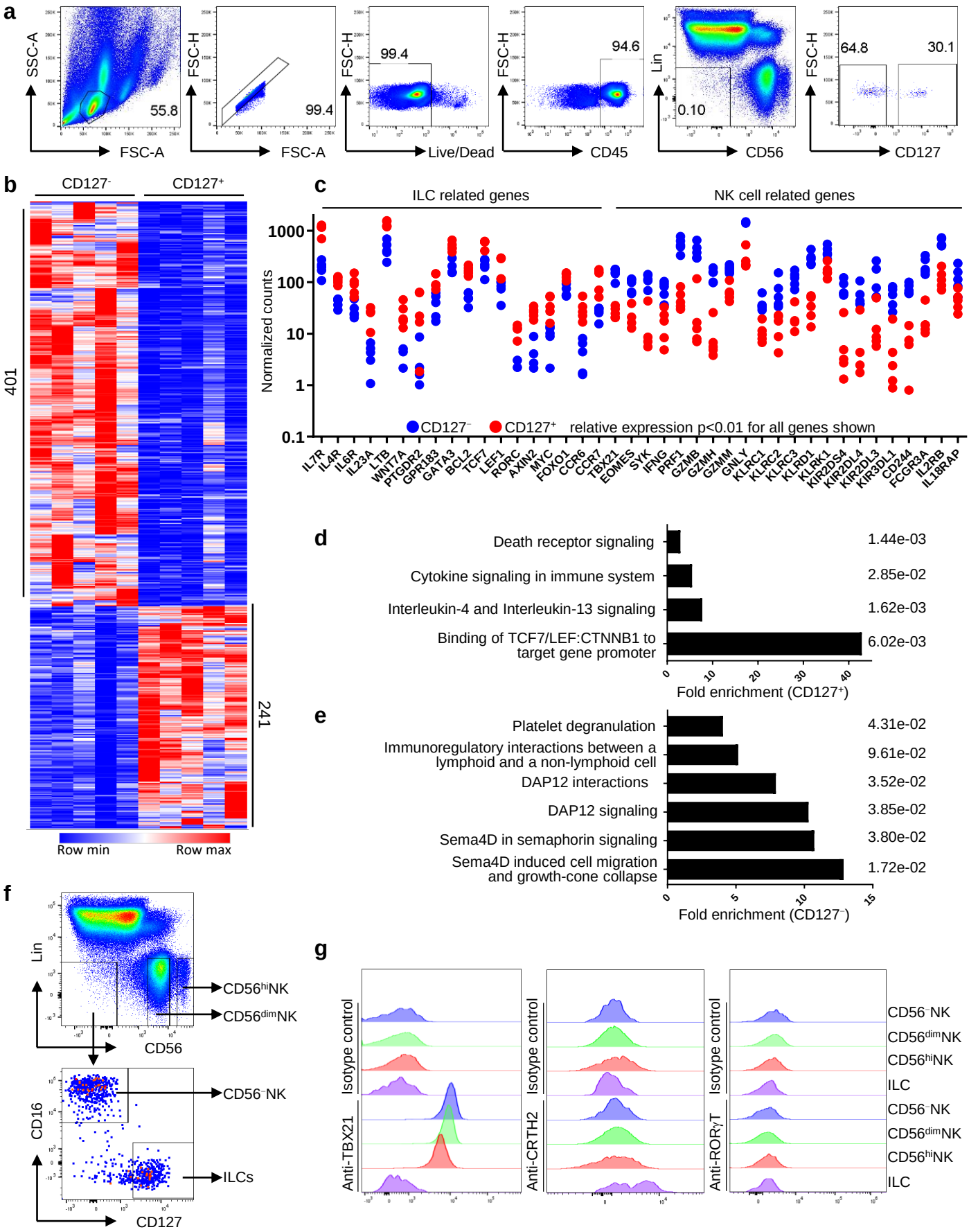
**a,** ATAC-Seq signal of gene loci associated with human NK cells and ILCs. **b,** ATAC-Seq signal from the Areg promoter region in mouse ILC1s, ILC2s, ILC2ps, ILC3s and NK cells, data was derived from GSE77695, GSE109125, and GSE116092.

**Extended Data Fig. 4 | CD56$^{-}$NK cells are expanded in HIV-1 infection.**

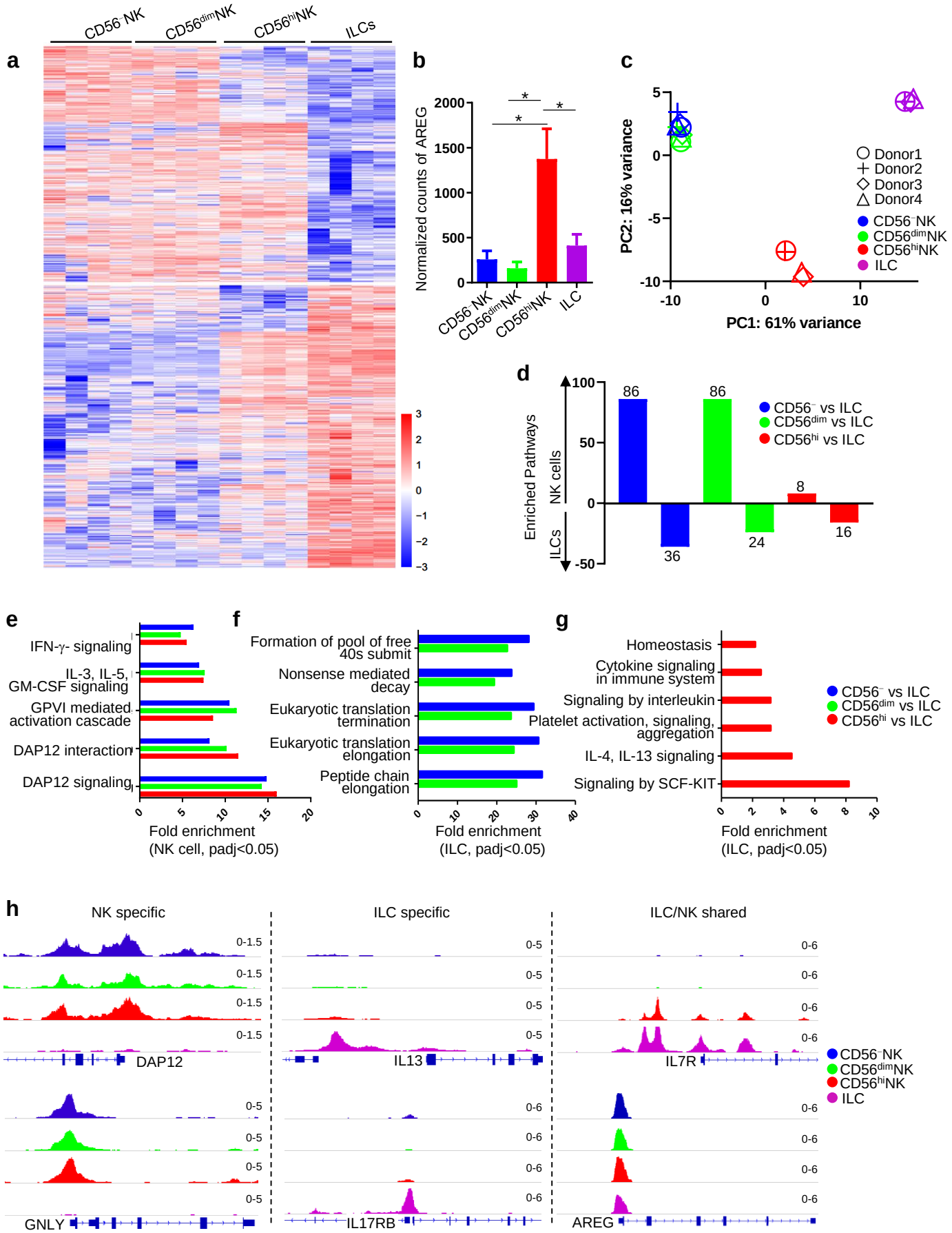
**a,** Detection of RUNX3 editing rate by ICE analysis (https://ice.synthego.com/) (left panel), and protein in NK cells by flow cytometry after control (AAVS1) or RUNX3 knockout (right panel). **b,** Detection of NK cells produced AREG after IL-12+IL-15+IL-18 stimulation for 16 hrs from control or RUNX3 knockout groups. **c,** Detection of ILCs and CD56$^{-}$NK cells from indicated groups. **d,** Detection of CD56$^{dim}$ and CD56$^{-}$ NK cells in indicated groups. **e,** Detection of CD56 MFI in CD56$^{dim}$NK cells in indicated groups (n=20 for each group). Data are mean ± s.e.m., *p<0.05, ***p<0.001. For **c-e,** cohort was described in Supplementary Table 5.

**Extended Data Fig. 5 | $CD4^+$T cells and IL-2 signaling are required for NK cell stability.**

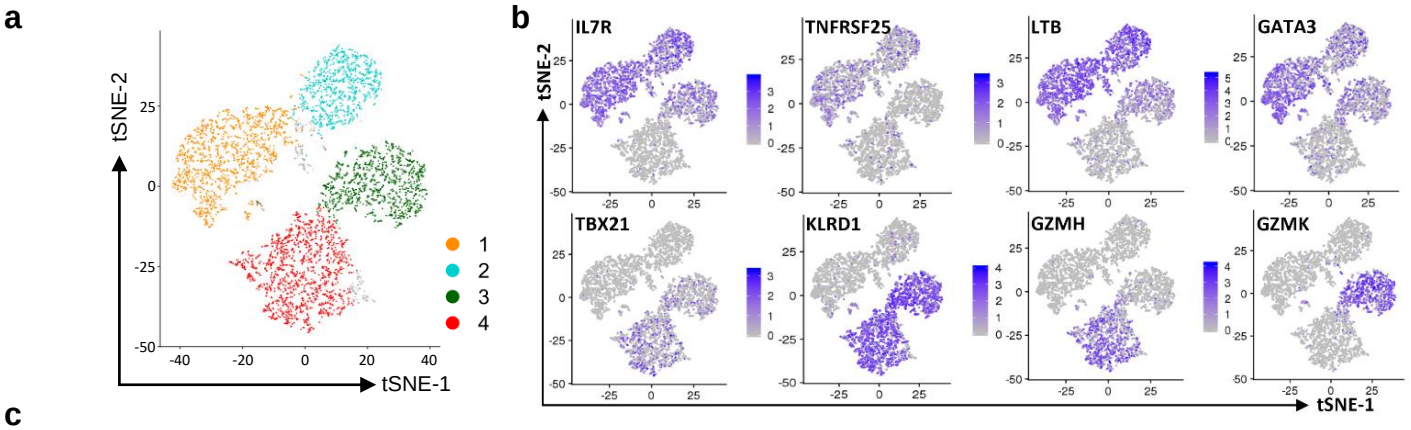
**a,** Specific cell types were depleted from PBMCs by magnetic beads based on surface markers as indicated, cultured for 5 days, then $Lin^-CD56^-$PBMCs were detected with CD127 and CD16 by flow cytometry. **b,** PBMC or $CD4^-$PBMCs were cultured for 5 days, CD56 MFI in $CD56^{dim}$NK cells were detected (n=7). **c,** CD56 MFI in $CD56^{dim}$NK cells was detected from $CD4^-$PBMCs in the presence or absence of IL-2 (10ng/ml) combined with or without IL-2 neutralizing antibody (4ug/ml) (n=7). **d,** CD56 MFI in $CD56^{dim}$NK cells was detected from PBMCs after treatment with isotype or IL-2 neutralizing antibody (4ug/ml) for 5 days (n=10). **e,f,** CD56 MFI in $CD56^{dim}$NK cells (**e**) and live cell percentage in $Lin^-$ $TBX21^+$ cells (**f**) were detected from PBMCs after treatment with c188-9 (STAT3i), AS-1517499 (STAT6i), MK-2206 (AKTi), Ravoxertinib (ERK1/2i) and CP690550 (JAK3i) for 5 days (n=4). **g,h,** CD56 (**g**) or IFN-γ production (**h**) from human NK cells in humanized NSG mice were detected as Fig. 5k,l. **i,** PBMCs were treated with or without rapamycin (10nM), the MFI of CD56 in $CD56^{dim}$NK cells was detected (n=7). Data are mean ± s.e.m., two-tailed paired *t*-test. ns, not significant, **p<0.01, ***p<0.001. All data were generated using blood from $HIV\text{-}1^-$ donors.

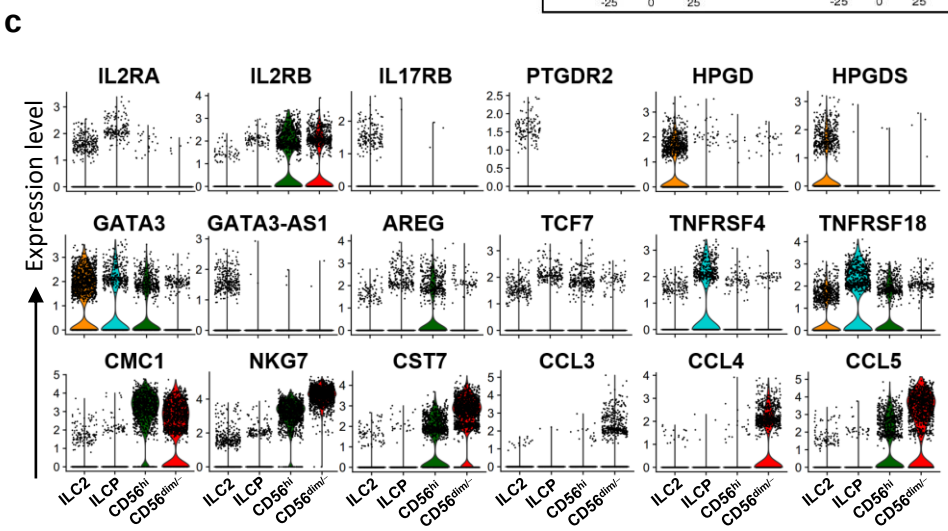


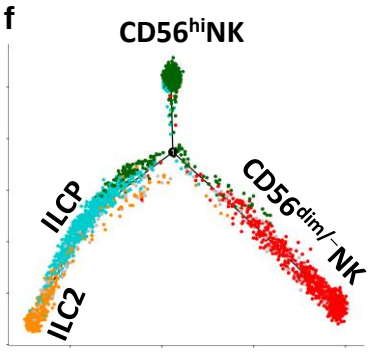

**Wang et al, Figure 1**

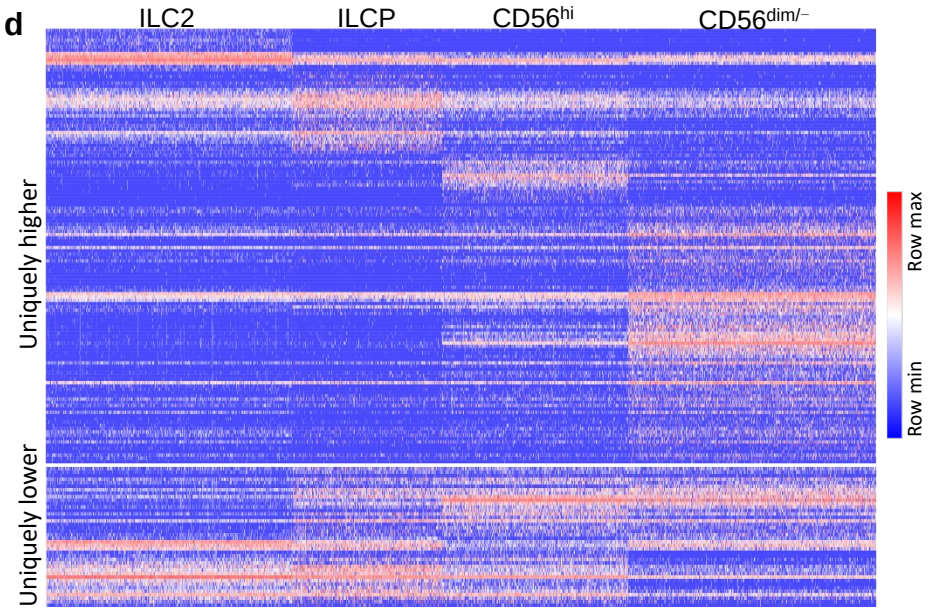


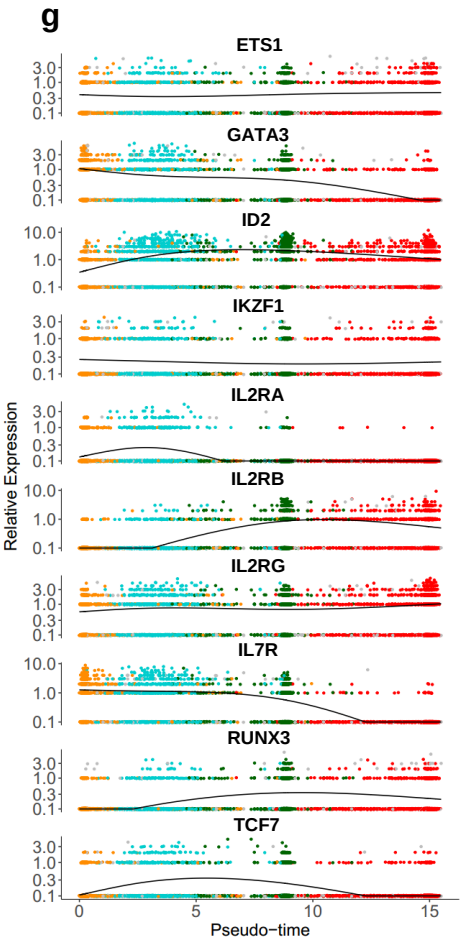

**Wang et al, Figure 2**



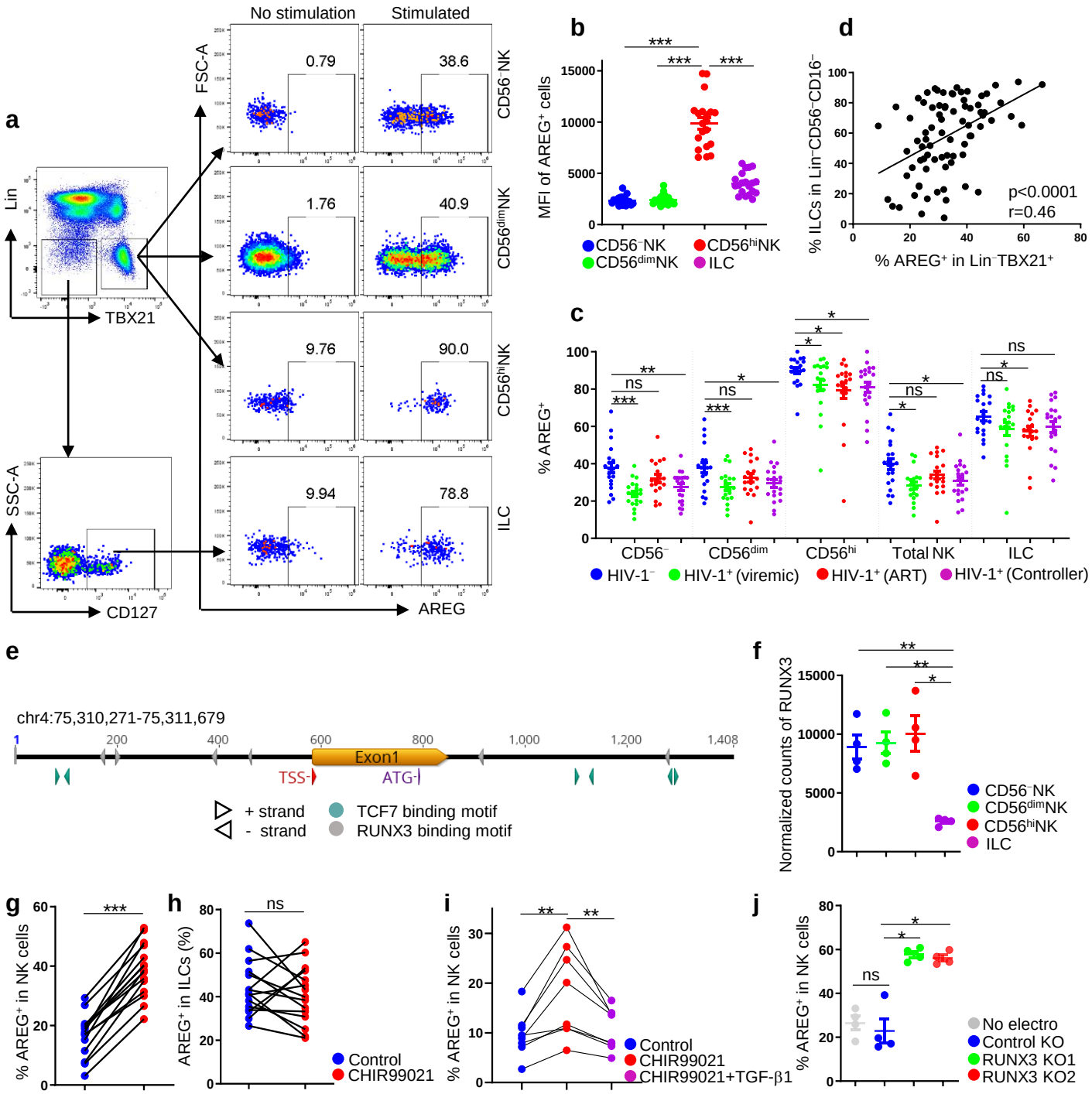



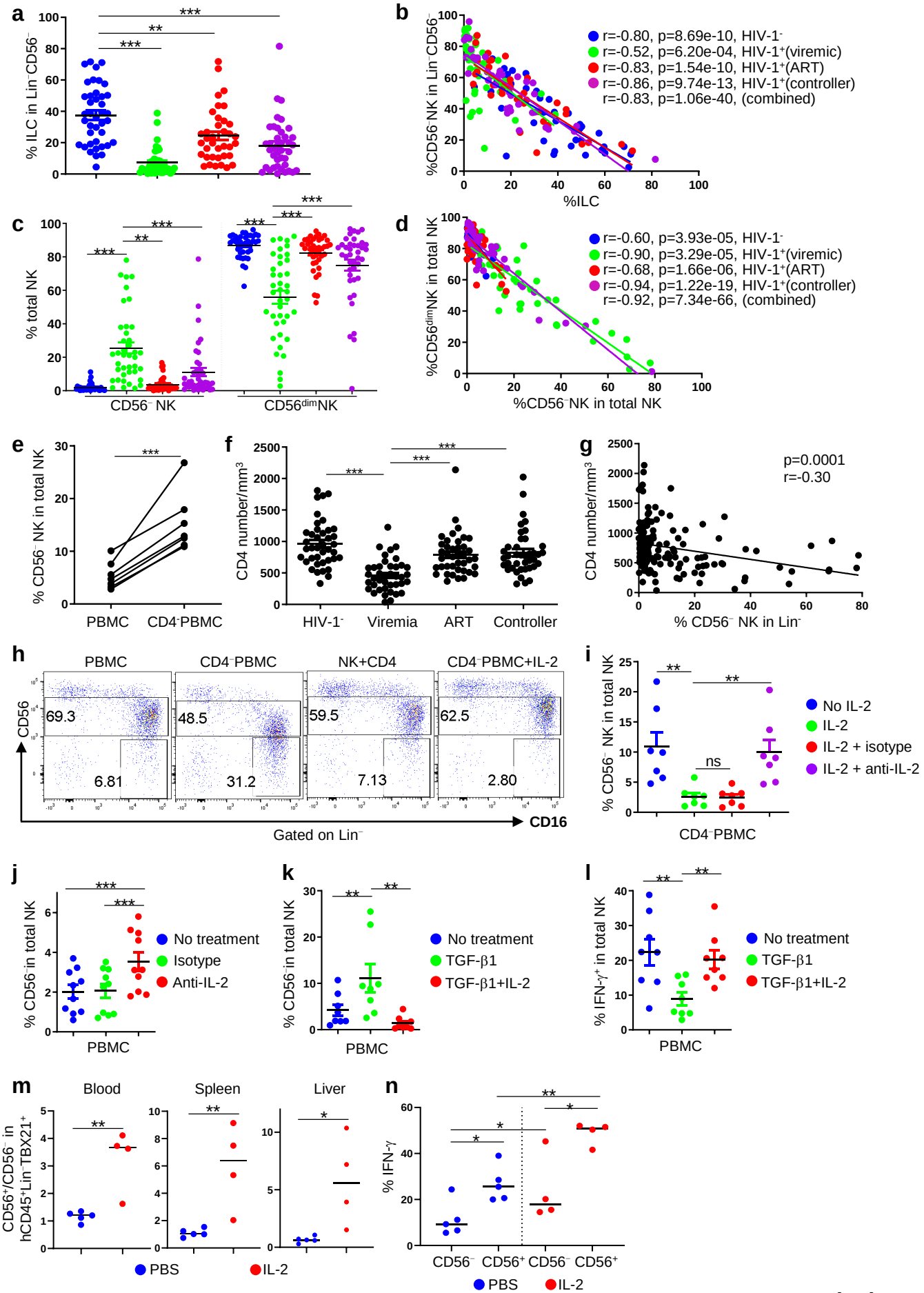

e

| Cluster | Representative genes that are highest in a given cluster |
|---|---|
| **ILC2** | IL-32, KLRB1, GATA3-AS1, HPGD, HPGDS, RPL18A, RPLP0 |
| **ILCP** | FOS, JUN, JUNB, FXYD5, TNFRSF4, TNFRSF18 |
| **CD56$^{hi}$NK** | CD2, CD74, KIR2DL1, DUSP4, XCL1, CMC1, KLRC1, KLRC2, GZMK, NCAM1 |
| **CD56$^{dim/-}$NK** | PRF1, GZMA, GZMB, GZMH, GZMM, NKG7, CST7, FGFBP2, CCL3, CCL4, CCL5 |

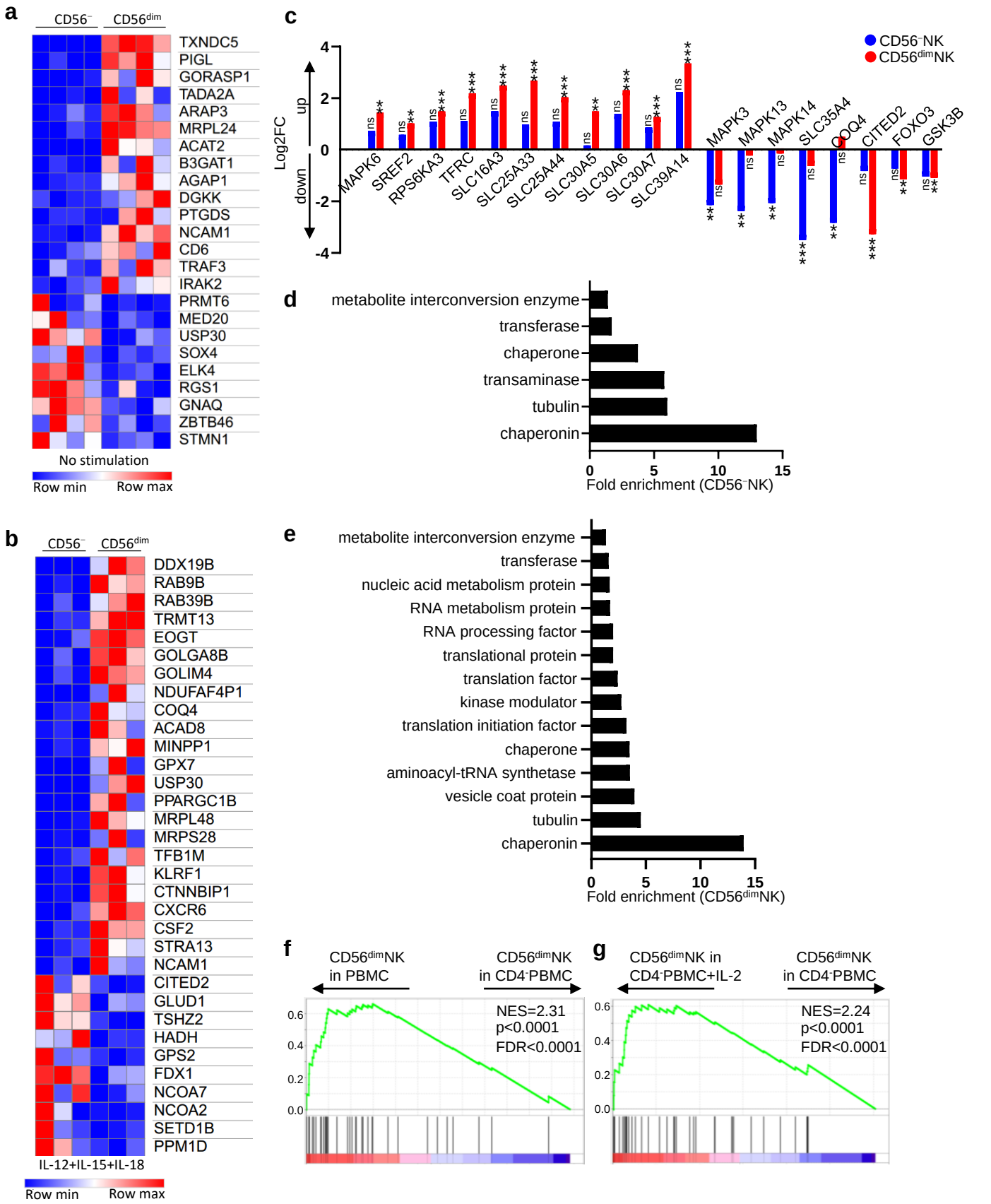
**Wang et al, Figure 3**



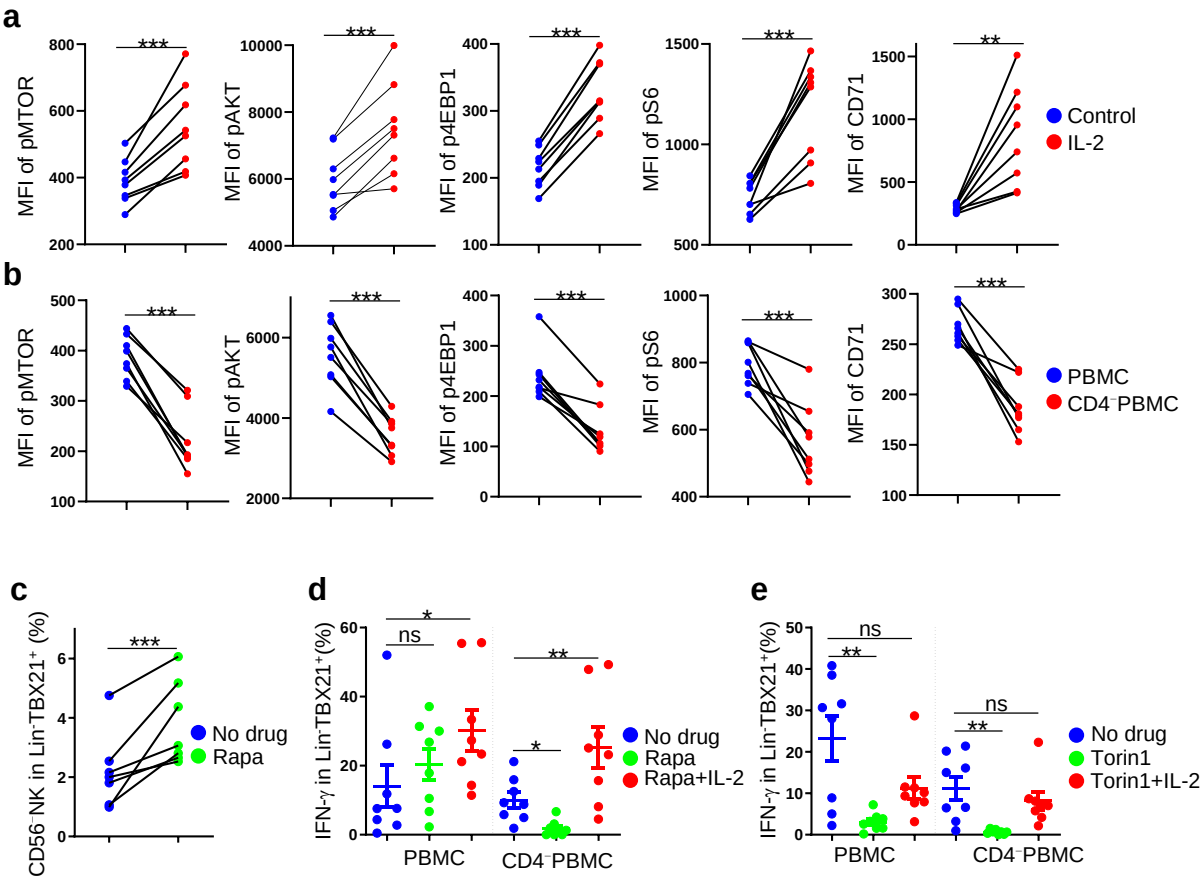

Wang et al, Figure 4



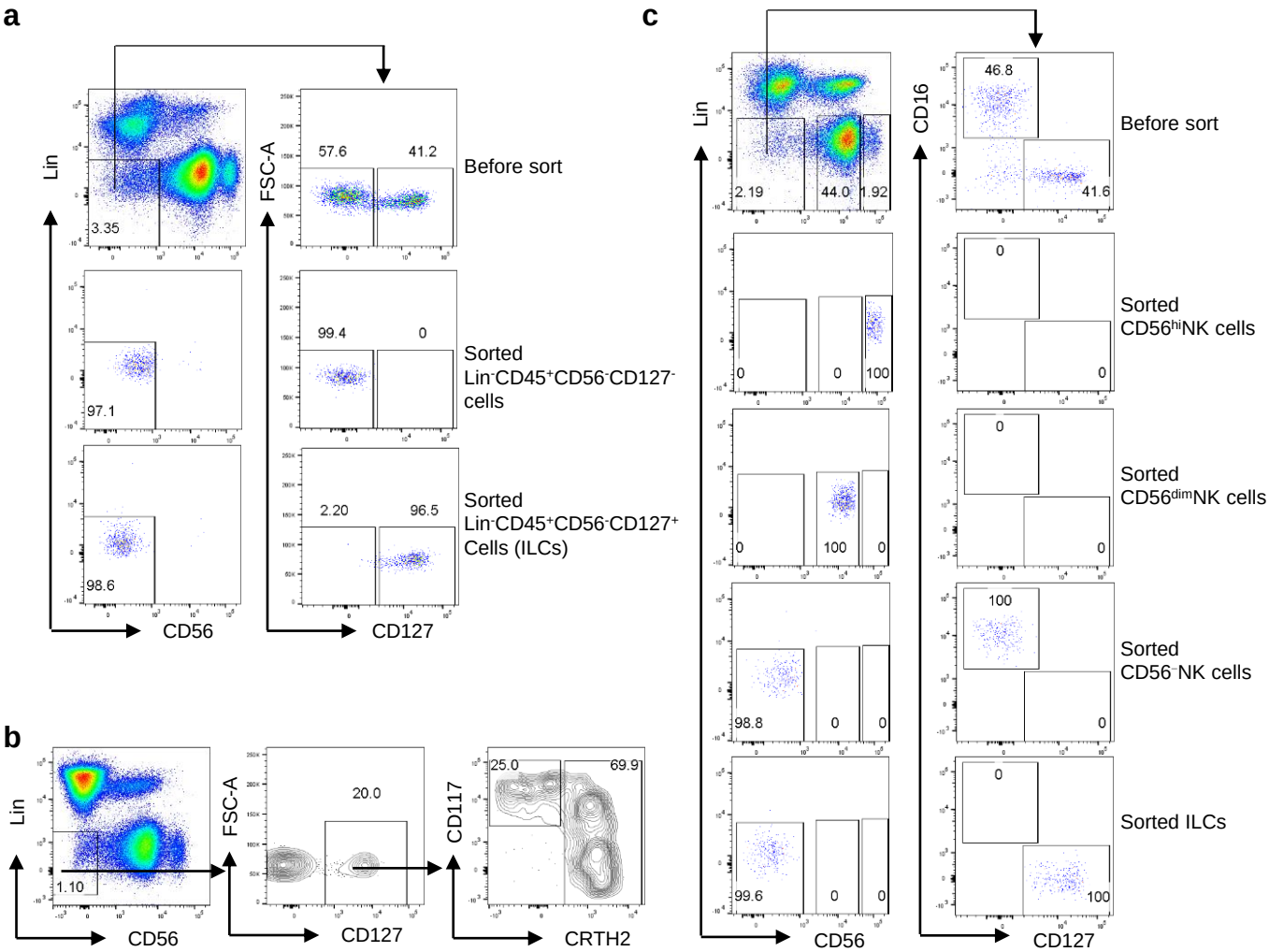

**Wang et al, Figure 5**



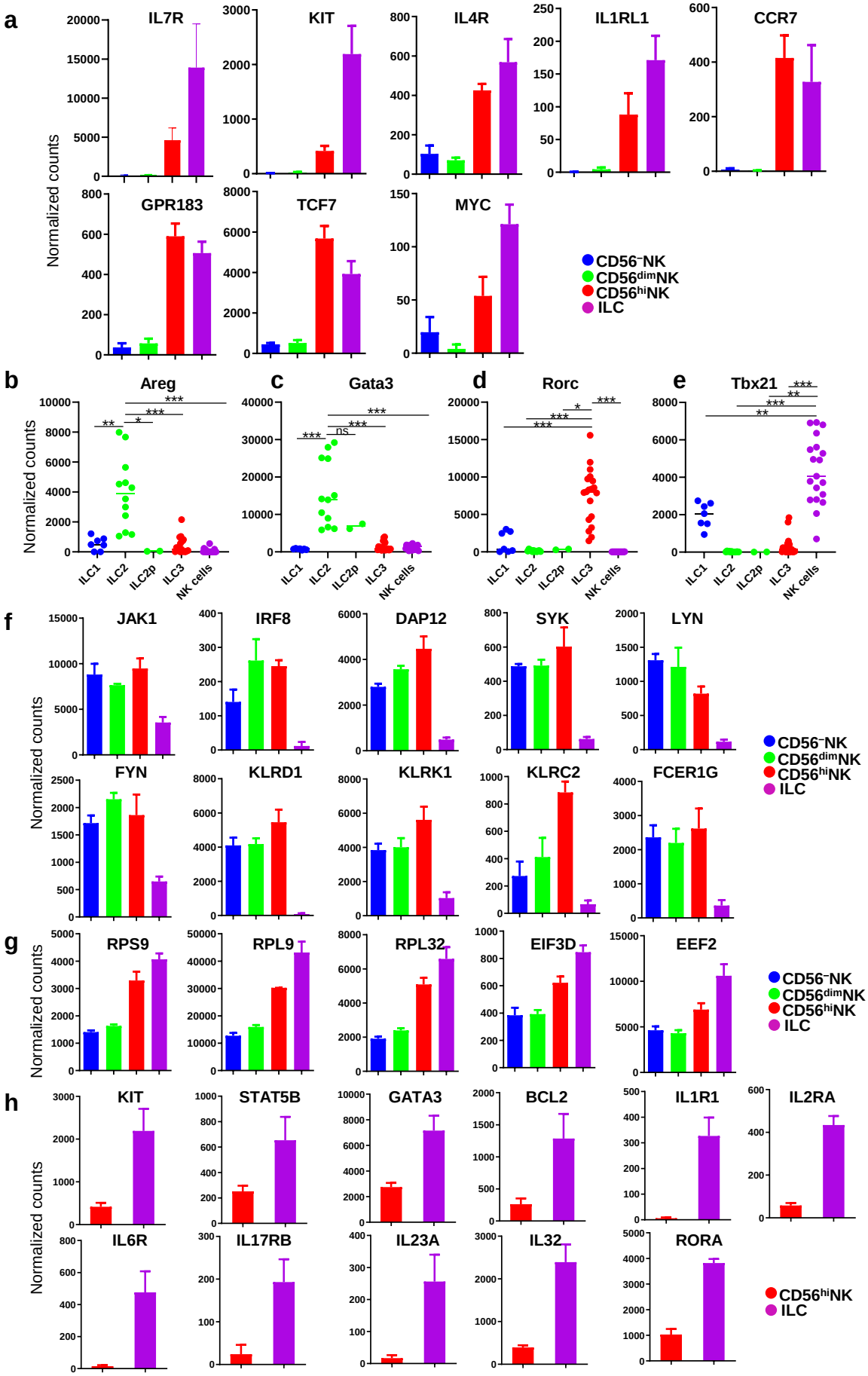

Wang et al, Figure 6



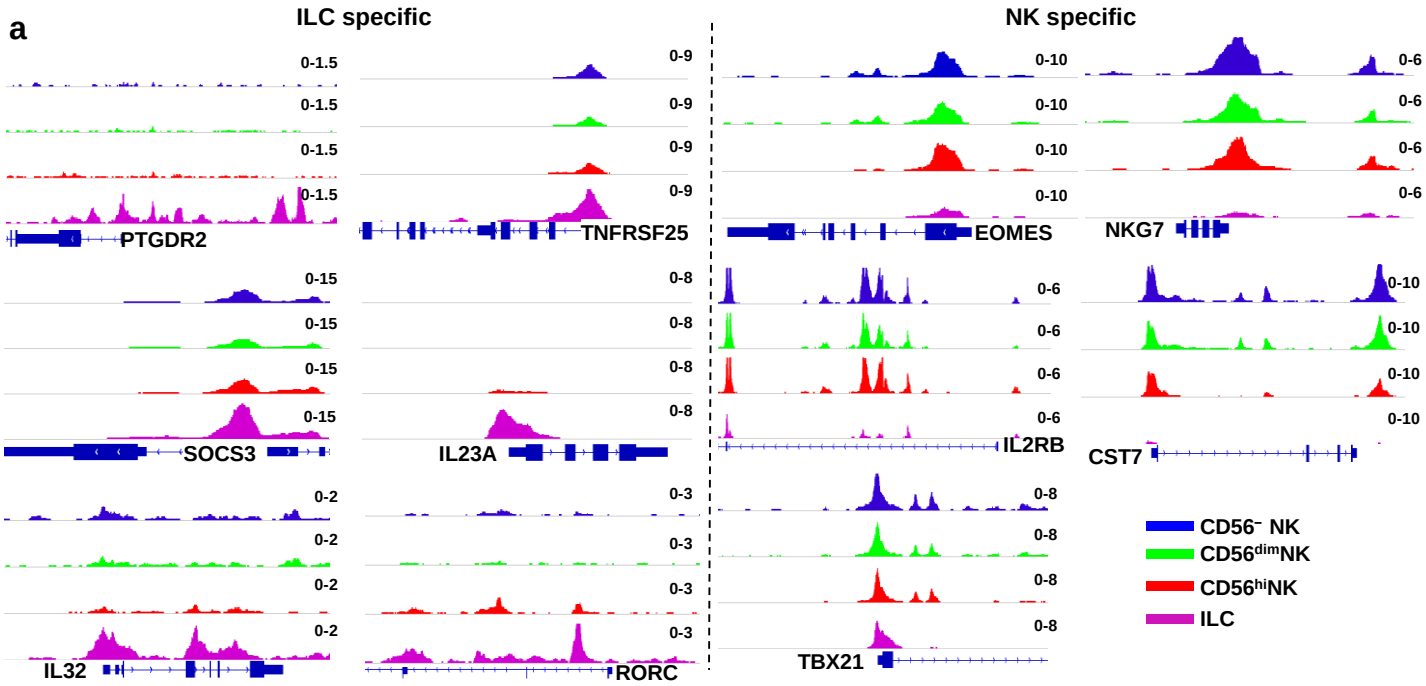

Wang et al, Figure 7

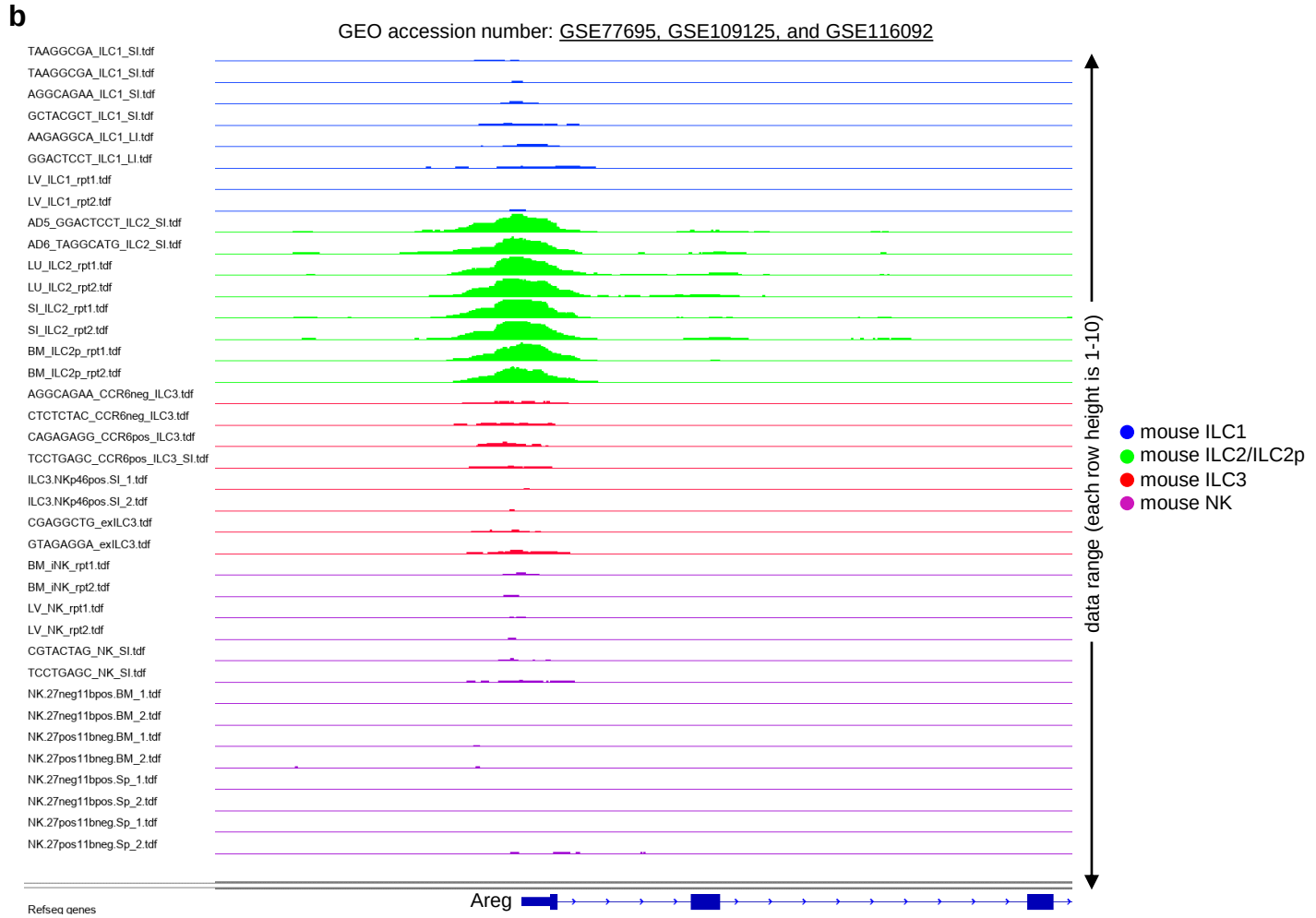



Wang et al, Extended Data Fig. 1

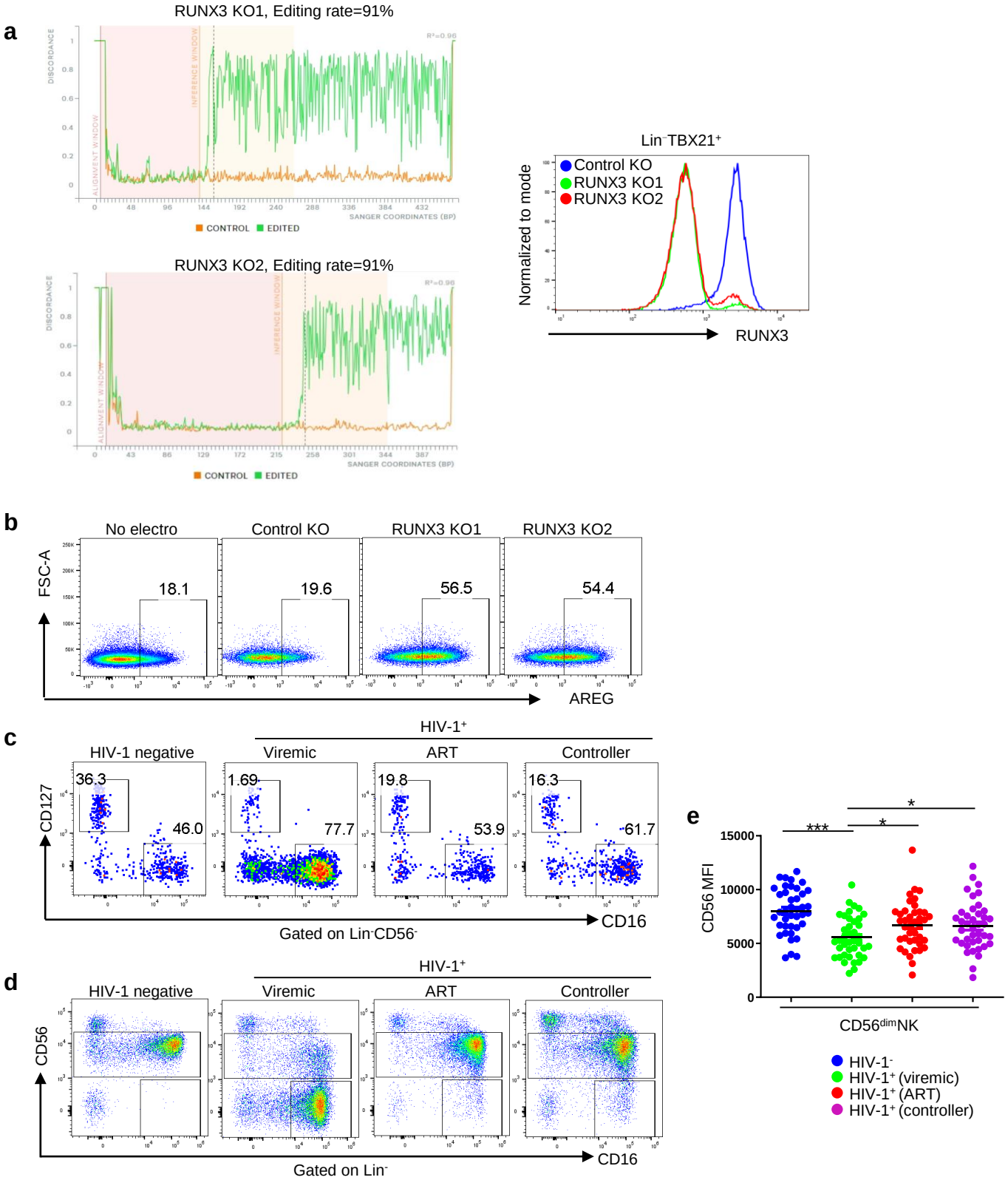



Wang et al, Extended Data Fig. 2

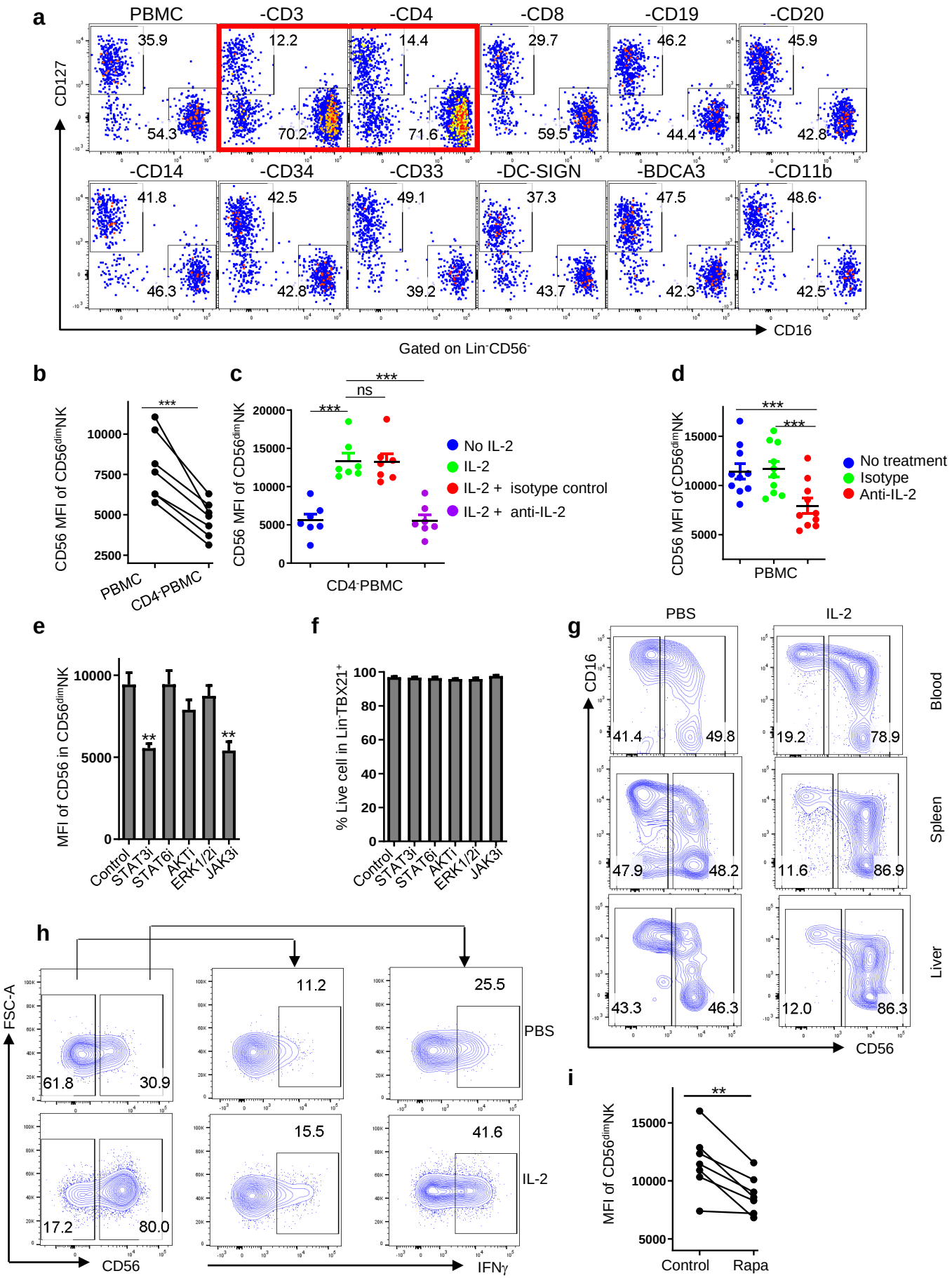



**Wang et al, Extended Data Fig. 3**




**Wang et al, Extended Data Fig. 4**




Wang et al, Extended Data Fig. 5