nature immunology

Article

https://doi.org/10.1038/s41590-022-01327-7

Clonal expansion and epigenetic inheritance of long-lasting NK cell memory

Received: 28 July 2022

Accepted: 9 September 2022

Published online: 26 October 2022

Check for updates

Timo Rückert[1] ✉, Caleb A. Lareau[2], Mir-Farzin Mashreghi[3], Leif S. Ludwig[4,5] and Chiara Romagnani[1,6,7] ✉

Clonal expansion of cells with somatically diversified receptors and their long-term maintenance as memory cells is a hallmark of adaptive immunity. Here, we studied pathogen-specific adaptation within the innate immune system, tracking natural killer (NK) cell memory to human cytomegalovirus (HCMV) infection. Leveraging single-cell multiomic maps of ex vivo NK cells and somatic mitochondrial DNA mutations as endogenous barcodes, we reveal substantial clonal expansion of adaptive NK cells in HCMV$^+$ individuals. NK cell clonotypes were characterized by a convergent inflammatory memory signature enriched for AP1 motifs superimposed on a private set of clone-specific accessible chromatin regions. NK cell clones were stably maintained in specific epigenetic states over time, revealing that clonal inheritance of chromatin accessibility shapes the epigenetic memory repertoire. Together, we identify clonal expansion and persistence within the human innate immune system, suggesting that these mechanisms have evolved independent of antigen-receptor diversification.

Vertebrates have evolved several mechanisms of immune adaptation, enabling optimized secondary responses to the same challenge. In a traditional paradigm, persistent immune memory has exclusively been associated with T and B cell clones selected from a pool expressing enormously diversified antigen receptors. Competitive clonal expansion after antigen encounter is accompanied by pronounced epigenetic remodeling, forming a repertoire of memory cells with enhanced and accelerated recall responses[1]. The dogma that adaptation mechanisms are exclusively confined to B and T cells has been challenged by observations of adjusted secondary responses within the innate immune system, with the most prominent example being 'trained immunity' of myeloid cells[2]. As for adaptive immune cells, these functional differences are epigenetically imprinted but are generally considered to be more transient and shorter lived than classical adaptive immunity. However, long-term maintenance of trained immunity has been observed in different cell types[3–6]. Further, in sharp contrast with the massive expansion of individual, epitope-specific T and B cell clones, trained immunity is mostly described at the population level, as innate immune cell activation occurs via widely expressed pattern recognition receptors.

More specific innate recognition mechanisms for several pathogens are used by natural killer (NK) cells, with the most prominent example being cytomegalovirus (CMV)[7]. CMV closely interacts with the host immune system to establish lifelong latent infection[8]. Consequently, CMV infection strongly imprints both the innate and adaptive immune repertoires, resulting in expansion of T cell clones and the emergence of 'adaptive' NK cells. While in C57BL/6 mice, these cells are expanded in response to the engagement of the activating receptor Ly49H by the murine CMV ligand m157 (ref. [9]), human adaptive NK cells specifically recognize peptides derived from the human CMV (HCMV) protein gpUL40 through the receptor CD94/NKG2C, facilitating affinity-dependent expansion of NKG2C$^+$ NK cells[10]. Recognition via conserved innate receptors expressed by cell populations is an

[1]Innate Immunity, Deutsches Rheuma-Forschungszentrum Berlin (DRFZ), ein Leibniz Institut, Berlin, Germany. [2]Department of Pathology, Stanford University, Stanford, CA, USA. [3]Therapeutic Gene Regulation, Deutsches Rheuma-Forschungszentrum Berlin (DRFZ), ein Leibniz Institut, Berlin, Germany. [4]Berlin Institute of Health at Charité Universitätsmedizin Berlin, Berlin, Germany. [5]Max-Delbrück-Center for Molecular Medicine in the Helmholtz Association (MDC), Institute for Medical Systems Biology (BIMSB), Berlin, Germany. [6]Charité—Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt Universität zu Berlin, Medizinische Klinik für Gastroenterologie, Infektiologie und Rheumatologie, Berlin, Germany. [7]Leibniz-Science Campus Chronic Inflammation, Berlin, Germany. ✉e-mail: timo.rueckert@drfz.de; romagnani@drfz.de

important difference to unique antigen receptors, likely affecting the size of the pool recruited into the response. In this context, the narrow, self-major histocompatibility complex (MHC)-specific killer cell immunoglobulin-like receptor (KIR) profile of human adaptive NK cells has sparked discussions on an oligoclonal origin[11]. Along this line, animal transfer studies of barcoded NK cells or progenitors have demonstrated the potential of NK cells to clonally expand in immunodeficient or depleted hosts[12–14], raising conceptual parallels to adaptive immune cells[15]. However, how these findings apply to natural infection in healthy individuals, where much larger populations of cells compete with each other and clonal diversity is not biased by transfer efficiency, is still unclear. Moreover, whether adaptive NK cell expansions induced by HCMV emerge from oligoclonal 'founder' populations and persist to form long-lasting memory remains elusive.

Here, we generated a single-cell multiomic map of human NK cells from HCMV$^+$ and HCMV$^-$ individuals, integrating measurements of gene expression and chromatin accessibility with lineage tracing using somatic mitochondrial DNA (mtDNA) mutations as endogenous barcodes[16,17]. We detected substantial clonal expansion of adaptive NK cells with a convergent inflammatory memory signature accompanied by private, clonotype-associated accessible chromatin sites, demonstrating a human innate immune population clonally persisting in a memory state.

Results

Mapping NK cell subsets onto transcriptional and epigenetic landscapes

To resolve the relationship of human NK cell subsets and the impact of HCMV on these populations, we performed single-cell assay of transposase-accessible chromatin (ATAC) with select antigen profiling by sequencing (ASAP-seq)[18] and cellular indexing of transcriptomes and epitopes (CITE-seq)[19] of NK cells from peripheral blood of HCMV$^+$ donors with NKG2C$^+$ NK cell expansions and HCMV$^-$ individuals (Fig. 1a and Supplementary Table 1). NKG2C$^+$/NKG2C$^-$ NK cells (sorted as CD3$^-$CD14$^-$CD19$^-$CD7$^+$; Methods) from all donors were equally enriched to a 1:1 ratio to correct for different frequencies of NKG2C$^+$ cells between HCMV$^+$ and HCMV$^-$ individuals (Extended Data Fig. 1a). After filtering for high-quality cells (Methods), 39,106 cells and 49,530 cells were analyzed for their epigenetic and transcriptional profiles, respectively. We detected expression of 19,975 genes and accessibility of 147,299 chromatin regions with consistent enrichment of transcription start sites (TSSs) across donors (Extended Data Fig. 1b). Using cell surface protein information, we annotated clusters defined by their transcriptomes and chromatin accessibility with long-standing NK cell subset definitions, while at the same time testing the validity of these population definitions in an unbiased manner (Fig. 1b–e). Namely, we observed four main clusters in both modalities that could be designated as CD56^{bright}, early CD56^{dim}, mature CD56^{dim} and adaptive NK cells (Fig. 1b,d) with clearly distinct genomic signatures (Extended Data Fig. 1d–f), in line with previous data[20–24]. An additional cluster of proliferating cells was only resolved via transcriptional profiling (Fig. 1d and Extended Data Fig. 1c). CD56^{bright} (CD16$^{-/lo}$CD57$^-$CD62L$^+$NKG2A$^+$) cells were characterized by high accessibility around loci of characteristic transcription factors (TFs) *TCF7*, *RUNX2*, *BACH2*, *ZEB1* and *MAML3*, and expression of signature transcripts, including *GZMK*, *XCL1* and *IL7R*. By contrast, CD56^{dim} (CD16$^+$CD57$^+$/CD57$^-$CD62L$^+$/CD62L$^-$NKG2A$^+$/NKG2A$^-$) NK cells were characterized by chromatin accessibility in the proximity of *ZBTB16*, *BCL11B*, *GLI3* and *ZEB2* and expression of *GZMB*, *PRF1* and *CX3CR1*. Between CD56^{bright} and CD56^{dim} NK cells, a population of seemingly intermediate, early CD56^{dim} (CD16$^+$CD57$^-$CD62L$^+$NKG2A$^+$) NK cells shared signatures of both subsets, such as high gene scores for *TCF7*, *ZEB1* and *ZBTB16* (Fig. 1b–g and Extended Data Fig. 1d–f). Label transfer between modalities further allowed for directly linking gene expression and chromatin accessibility, generating an integrated resource that enables correlation-based prediction of subset-specific gene regulatory elements for core regulators of NK cell identity, such as *TCF7*, *RUNX2* and *ZEB2* (Extended Data Fig. 2a–d). Overall, we provide a multimodal map of human NK cells, combining single-cell measurements of chromatin accessibility, gene expression and surface phenotypes.

Distinct signatures for naive and adaptive NKG2C$^+$ NK cells

Our multiomic analysis of human NK cells clearly revealed the presence of a separate and genomically distinct adaptive NK cell cluster (Fig. 1b–g), characterized by high transcript levels of *KLRC2*, *CD3E* and *PATL2* together with marked downregulation of *CD7*, *KLRB1* and *FCER1G* (Fig. 1g) and distinct surface marker expression of CD57 and CD2 coinciding with low NKp30, CD161 and NKG2A (Fig. 1c,e). While adaptive NK cells shared large parts of their genomic effector signature with conventional CD56^{dim} NK cells, they displayed a marked reduction of *ZBTB16* and *ZNF516* and further increased gene scores for *ZEB2* and the metabolic regulator *ARID5B*[25] compared to CD56^{dim} NK cells (Fig. 1f). Moreover, adaptive NK cells were characterized by expression and accessibility of *ZBTB38* and *JAKMIP1* (Fig. 1f,g and Extended Data Fig. 1e,f). Importantly, the integrated analysis demonstrated remodeling at chromatin regions around the *KLRCX* locus, with several of those being positively correlated with increased *KLRC2* expression and one region having a negative correlation with *KLRC1* expression (Extended Data Fig. 2e). Their functional role was supported by the presence of several TF binding motifs, with the most strongly correlating putative *KLRC1* enhancer containing a TCF7 motif, consistent with expression and accessibility in CD56^{bright} NK cells, while AP1 motifs were notable in the putative *KLRC2* enhancers specifically accessible in adaptive NK cells, which together may regulate the core adaptive NK cell phenotype characterized by high surface expression of NKG2C combined with lack of NKG2A[26].

While expansion of adaptive NKG2C$^+$ NK cells is associated with HCMV infection, NKG2C$^+$ NK cells are present also in HCMV$^-$ individuals, albeit at lower frequencies. To directly compare the transcriptional and epigenetic landscape of NKG2C$^+$/NKG2C$^-$ NK cells and their distribution within NK cell clusters in HCMV$^+$ and HCMV$^-$ individuals, we divided the datasets by HCMV serostatus (Fig. 2a–h). As expected, the adaptive cluster was selectively present in HCMV$^+$ individuals and was mostly comprised of NKG2C$^+$ NK cells (Fig. 2a–d). Comparative analysis of chromatin accessibility and gene expression after downsampling to equal cell numbers for both donor groups highlighted the specific impact of HCMV on NKG2C$^+$ cells. A total of 232 genes were differentially accessible between the NKG2C$^+$ and NKG2C$^-$ populations in HCMV$^+$ individuals, including key adaptive genes, such as *JAKMIP1*, *ZBTB38*, *ZBTB16* or *ZNF516* (Fig. 2i), along with 269 differentially expressed genes, including adaptive NK cell markers *CD3E*, *IL32*, *FCER1G* and *KLRB1* (Fig. 2k). Conversely, NKG2C$^+$ NK cells from HCMV$^-$ donors were randomly dispersed among NKG2C$^-$ NK cells (Fig. 2e–h), and differences in chromatin accessibility and gene expression, apart from *KLRC2* (encoding NKG2C), were minor (Fig. 2m,o), highlighting their largely similar epigenomes and transcriptomes. Genes differentially accessible or expressed between NKG2C$^+$ and NKG2C$^-$ cells in HCMV$^+$ individuals, including *ZBTB16* and *FCER1G*, followed a coordinated pattern separating conventional from adaptive NK cells, whereas this was not observed in HCMV$^-$ individuals (Fig. 2j,n,l,p), underlining that these differences reflect the adaptive remodeling induced in NKG2C$^+$ cells by HCMV.

Interestingly, even HCMV$^+$ individuals displayed a small fraction of NKG2C$^+$ NK cells that were present in all three subsets of the conventional NK cell compartment (Fig. 2a–d). Indeed, separate analysis of NKG2C$^+$ NK cells alone from HCMV$^+$ individuals revealed the coexistence of adaptive NKG2C$^+$ NK cells with a minority of 'naive' NKG2C$^+$ NK cells (Extended Data Fig. 3a,d) that lacked adaptive remodeling (Extended Data Fig. 3b,e) and spanned the whole spectrum of conventional NK cells, recapitulating the signatures identified from the full integrated analysis (Extended Data Fig. 3c,f). We confirmed by flow cytometry that the frequencies of naive NKG2C$^+$ NK cells were comparable between HCMV$^+$ and HCMV$^-$ individuals and that they were

https://doi.org/10.1038/s41590-022-01327-7

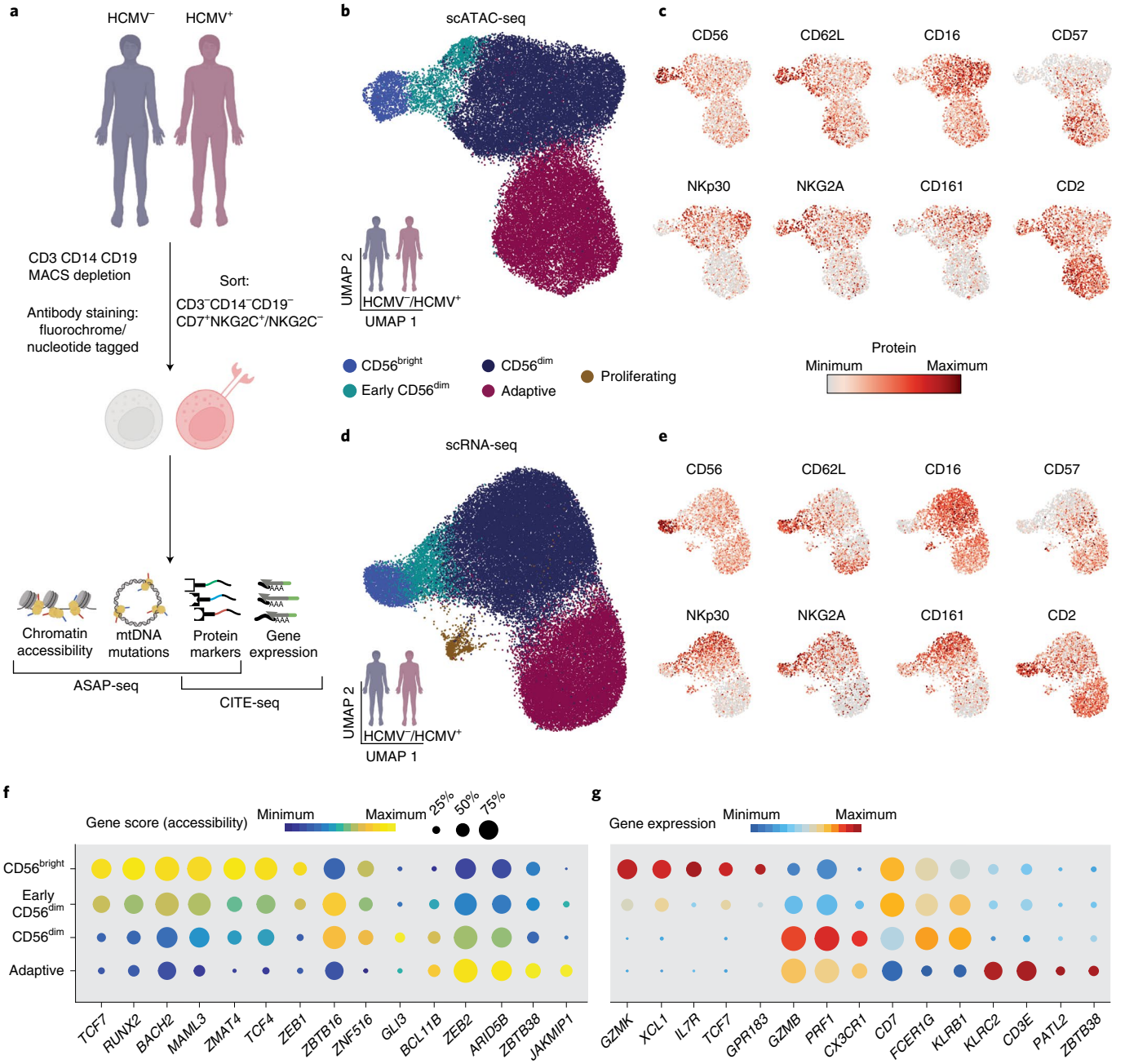


Fig. 1 | Mapping NK cell subsets onto transcriptional and epigenetic landscapes. a, NKG2C⁺ and NKG2C⁻ NK cells were isolated from four to five HCMV⁺ individuals and two HCMV⁻ healthy blood donors, stained with nucleotide barcode-labeled antibodies, mixed at a 1:1 ratio and analyzed by ASAP-seq ($n = 6$) and CITE-seq ($n = 7$). **b**,**d**, Integrated UMAP embedding of NK cells from donors analyzed by scATAC-seq (**b**; $n = 6$) and scRNA-seq (**d**; $n = 7$). **c**,**e**, Surface protein expression of the indicated markers as measured by ASAP-seq (**c**) or CITE-seq (**e**). **f**,**g**, Column-scaled accessibility scores (**f**) and expression (**g**) per cluster and detection frequencies for the indicated genes. Created with BioRender.com.

similarly distributed between the CD56^{bright} and CD56^{dim} compartment (Extended Data Fig. 3g–i).

Together, this comparative analysis of HCMV⁺ and HCMV⁻ donors highlighted the pronounced remodeling that HCMV imposes on the NKG2C⁺ NK cell pool that coexists with a minority of naive NKG2C⁺ NK cells during the latent phase of HCMV infection.

HCMV leaves an inflammatory memory footprint enriched in AP1 motifs

While gene expression profiles and accessibility of regulatory elements in the proximity of *trans*-acting factors provide a view on the current state of a cell, analysis of TF motif enrichment in *cis*-regulatory elements also captures molecular events that have left footprints in the generation of this state. Global analysis using chromVAR[27] identified a strong signature in CD56^{bright} NK cells with enhanced activity of motifs for TFs such as TCF7, RUNX2 or ZEB1 (Fig. 3a and Extended Data Fig. 4a,b,e), consistent with an important role of these TFs in shaping the epigenetic and transcriptional identity[21,24] of this subset (Fig. 1f,g) and increased motif activity for NF-κB-related TFs (REL) and STAT4 (Fig. 3a and Extended Data Fig. 4a,b,e), reflecting the enhanced responsiveness of CD56^{bright} NK cells to activation by proinflammatory cytokines[28]. Accordingly, an NF-κB- and STAT4-motif-containing peak in the proximity of the *IFNG*

https://doi.org/10.1038/s41590-022-01327-7

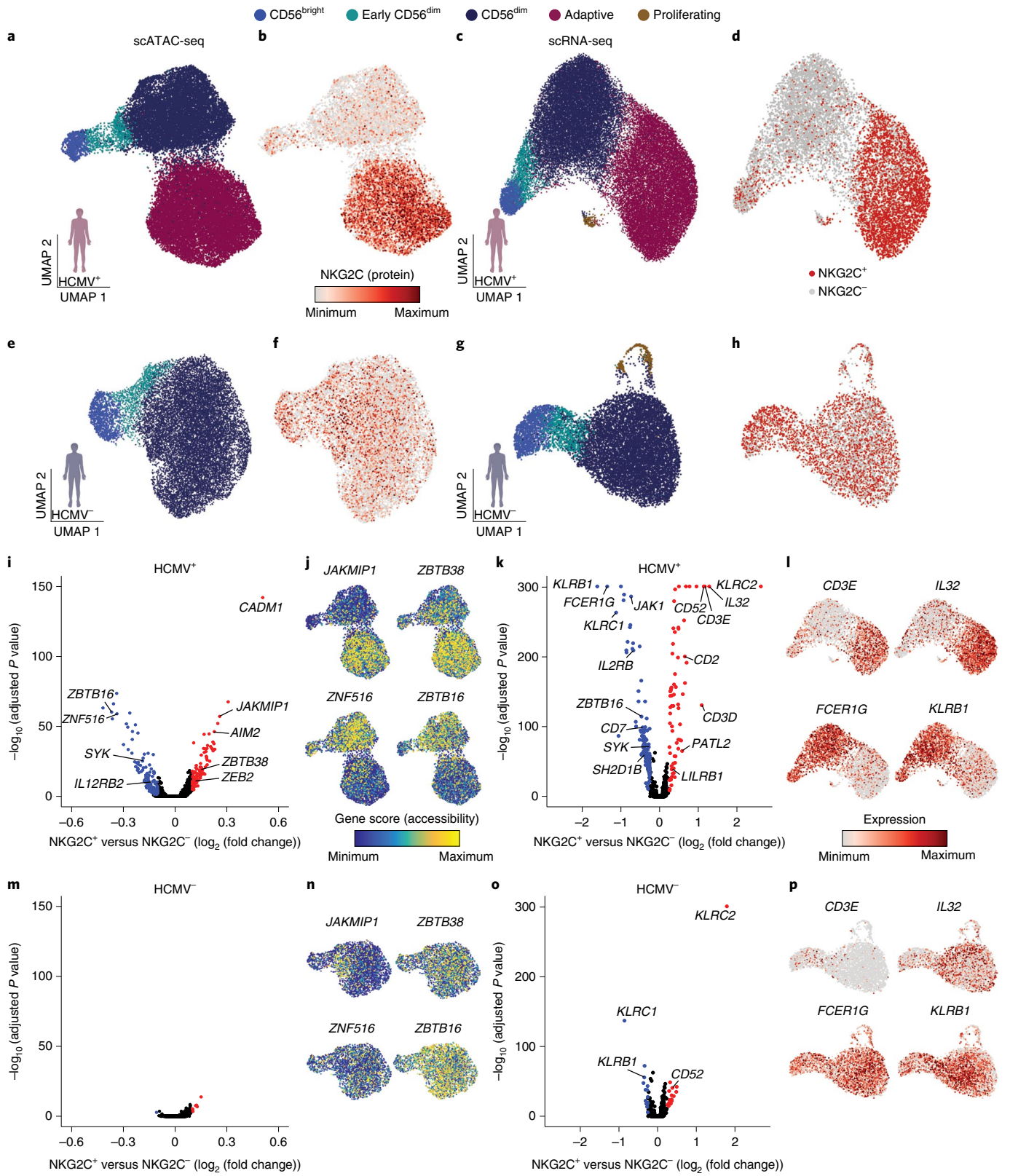
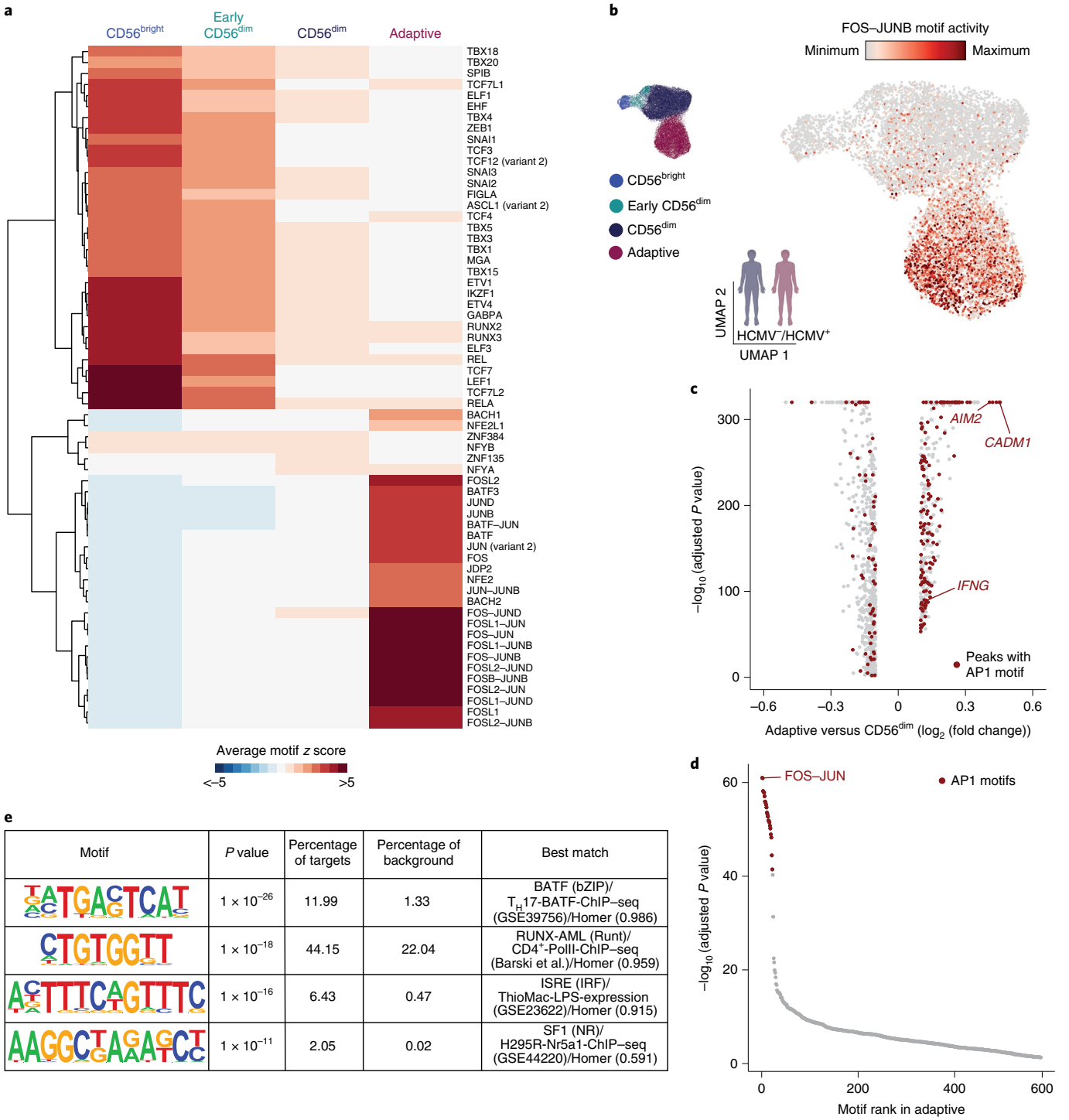


Fig. 2 | Distinct signatures of naive and adaptive NKG2C⁺ NK cells. **a**,**e**, UMAP embedding of NK cells from HCMV⁺ (**a**; $n = 4$) and HCMV⁻ (**e**; $n = 4$) donors analyzed by scATAC-seq. **b**,**f**, NKG2C surface expression as measured by ASAP-seq. **c**,**g**, UMAP embedding of NK cells from HCMV⁺ (**c**; $n = 5$) and HCMV⁻ (**g**; $n = 2$) donors analyzed by scRNA-seq. **d**,**h**, Distribution of barcoded NKG2C⁺ and NKG2C⁻ populations. **i**,**m**, Differentially accessible genes between NKG2C⁺ and NKG2C⁻ NK cells for HCMV⁺ (**i**) and HCMV⁻ (**m**) donors as determined by logistic regression. **j**,**n**, Representative accessibility scores of genes defining adaptive NK cells for HCMV⁺ (**j**) and HCMV⁻ (**n**) donors. **k**,**o**, Differentially expressed genes between NKG2C⁺ and NKG2C⁻ NK cells for HCMV⁺ (**k**) and HCMV⁻ (**o**) donors determined by two-sided Wilcoxon rank-sum test with Bonferroni adjustment. **l**,**p**, Representative expression of genes defining adaptive NK cells for HCMV⁺ (**l**) and HCMV⁻ (**p**) donors. Created with BioRender.com.

https://doi.org/10.1038/s41590-022-01327-7



e

Motif	*P* value	Percentage of targets	Percentage of background	Best match
	1×10^{-26}	11.99	1.33	BATF (bZIP)/ T_H17-BATF-ChIP–seq (GSE39756)/Homer (0.986)
	1×10^{-18}	44.15	22.04	RUNX-AML (Runt)/ CD4+-PolII-ChIP–seq (Barski et al.)/Homer (0.959)
	1×10^{-16}	6.43	0.47	ISRE (IRF)/ ThioMac-LPS-expression (GSE23622)/Homer (0.915)
	1×10^{-11}	2.05	0.02	SF1 (NR)/ H295R-Nr5a1-ChIP–seq (GSE44220)/Homer (0.591)

Fig. 3 | HCMV infection leaves an inflammatory memory footprint enriched in AP1 motifs. a, Heat map of differentially active motifs represented as average chromVAR deviation scores per cluster. **b**, FOS–JUNB motif activity projected on UMAP embedding. **c**, DARs between adaptive and CD56^{dim} NK cells determined by two-sided Wilcoxon rank-sum test with Bonferroni adjustment. AP1-motif-containing regions are marked in red, and selected regions are annotated by gene proximity. **d**, Motif enrichment in chromatin regions specifically accessible in adaptive NK cells as determined by one-sided hypergeometric test with Benjamini–Hochberg adjustment. **e**, De novo motif analysis results on chromatin regions specifically accessible in adaptive NK cells. *P* values were calculated by comparing enrichment to the cumulative binomial distribution. Created with BioRender.com.

gene was specifically accessible in CD56^{bright} cells, while this was closed in the more mature subsets (Extended Data Fig. 4f). While CD56^{dim} NK cells were mostly characterized by a lack of the CD56^{bright}-specific TF motif activity, both CD56^{dim} and adaptive NK cells were further distinguished from CD56^{bright} NK cells by enhanced CTCF motif activity (Extended Data Fig. 4a,e), as previously described[24]. Importantly, adaptive NK cells showed a strikingly separated signature compared to both conventional NK cell subsets, characterized by enhanced activity of AP1 motifs, with the strongest difference observed for FOS–JUNB (Fig. 3a,b and Extended Data Fig. 4d), and AP1 motifs were the most significantly enriched in adaptive NK-cell-specific accessible regions (Fig. 3c,d). De novo motif prediction on these peaks also yielded a canonical AP1 motif as the most

significant hit, further corroborating these results (Fig. 3e). Recently, AP1 motifs have been described to be at the core of a unifying inflammatory memory signature in different immune cell types and murine epithelial stem cells, the latter characterized by specific accessibility in the proximity of the *Aim2* and *Cadm1* loci[4]. Notably, their human orthologs also showed highly specific opening of AP1-motif-containing regions in adaptive NK cells (Fig. 3c and Extended Data Fig. 4g) and were among the genes with the most pronounced difference in accessibility between NKG2C$^+$ and NKG2C$^-$ cells in HCMV$^+$ individuals (Fig. 2i). Finally, among the more accessible peaks with AP1 motifs were several enhancers of the *IFNG* gene (Extended Data Fig. 4f), providing a possible mechanism for the increased interferon-γ (IFNγ) production by adaptive NK cells in response to activating receptor stimulation[29,30]. Overall, these findings demonstrate how differential TF activity shapes NK cell epigenetic remodeling during differentiation and identify AP1 TFs as potential drivers shaping adaptive NK cell chromatin accessibility, an intriguing parallel to the inflammatory imprinting used to establish immune memory in different cell types[4,31,32].

Synergistic imprinting by HCMV peptides and proinflammatory cytokines

Activation by interleukin-12 (IL-12) and IL-18 in combination with engagement of NKG2C by HCMV-derived peptides induces the specific expansion of NKG2C$^+$ NK cells with some of the transcriptional characteristics also observed ex vivo[10]. To assess whether these signals drive epigenetic reprogramming, including AP1 TF activity, and mimic HCMV-induced activation of naive NKG2C$^+$ NK cells, we cocultured NK cells from HCMV$^-$ donors for 12 h with IL-15 and RMA-S/HLA-E target cells pulsed with the NKG2C-activating VMAPRTLFL (LFL) or non-activating VMAPQSLLL peptide in the presence or absence of IL-12 and IL-18. We marked the different conditions with individual nucleotide-barcoded hashtags, sorted for NKG2C$^+$ cells and performed ASAP-seq (Fig. 4a).

Analysis of high-quality cells clustering by condition and irrespective of donors (Methods) enabled us to identify the epigenetic signatures acquired by activated NKG2C$^+$ NK cells in a stimulus-dependent fashion, defining four main clusters clearly enriched for cells cultured under the respective conditions, that is, non-activated ('control'), activated by LFL peptide ('LFL'), proinflammatory cytokines ('IL-12 + IL-18') or a combination thereof ('LFL + IL-12 + IL-18'; Fig. 4b, Extended Data Fig. 5a,b and Methods). The resulting chromatin accessibility profiles had a similar TSS enrichment (Extended Data Fig. 5c) as observed ex vivo (Extended Data Fig. 1b), irrespective of activation. Activation status of cells recruited in stimuli-associated clusters was exemplified by surface upregulation of CD137 (4-1BB), with maximal expression observed on cells receiving both peptide and cytokine stimulation (Fig. 4c). Differences in expression levels of CD137 across conditions were confirmed by flow cytometry (Extended Data Fig. 5d,e), validating the proteogenomic-based cluster annotations.

While IL-12 + IL-18 stimulation alone promoted the activation of NKG2A$^+$NKG2C$^+$ naive NK cells, hardly any NKG2A$^+$NKG2C$^+$ cells were recruited into response patterns induced by peptide stimulation (Fig. 4d), in line with the dominant inhibitory function of NKG2A engagement[33]. As, in contrast to adaptive NKG2C$^+$ NK cells[26], a large fraction of naive NKG2C$^+$ NK cells coexpress NKG2A (Extended Data Fig. 5g,h), these data suggest that peptide recognition might be decisive in biasing the pool of naive NKG2C$^+$ cells recruited during HCMV infection toward NKG2A$^-$NKG2C$^+$ cells.

Activation via proinflammatory cytokines or by peptide recognition via NKG2C induced pronounced chromatin remodeling, manifesting in 1,232 and 936 differentially accessible regions (DARs), respectively, compared to the control condition (Fig. 4e). Most regions underwent remodeling only after synergistic activation by the two combined stimuli (2,693 DARs; Fig. 4e), with different TF families contributing to these effects (Fig. 4f). Besides IRFs, EGR, MYC and NF-κB, LFL peptide activation induced remarkable activity of AP1-family TFs. Consistent with reports in mice[34], proinflammatory cytokines IL-12 and IL-18 were highly specific in their induction of POU TF-family activity, along with NF-κB. However, clear synergism between cytokines and peptide became apparent also on the TF motif activity level, in particular for AP1 (Fig. 4f and Extended Data Fig. 5f). Importantly, the two stimuli induced several features associated with adaptive NK cells ex vivo. This included opening of chromatin in the *ZEB2* and *ARID5B* loci, which was mainly dependent on proinflammatory cytokines or similarly enabled by each individual stimulus, respectively, as well as closing of the *ZBTB16* locus (Fig. 4g–i and Extended Data Fig. 5i). However, most of these changes were more pronounced after synergistic activation, as was also the case for chromatin accessibility in the *AIM2* region (Fig. 4g–i and Extended Data Fig. 5i). To globally assess which of the activation-induced signatures best reconciles the adaptive signature, we performed anchor-based integration of the in vitro stimulated cells with our ex vivo dataset (Methods)[35], thereby enabling a classification of the in vitro induced signatures based on chromatin profiles observed ex vivo. The highest adaptive prediction score was observed for cells activated by both stimuli, in concert with a reduction of the prediction score for conventional CD56^{dim} NK cells (Fig. 4j,k). The close relatedness of the adaptive ex vivo signature to the chromatin remodeling induced by synergistic in vitro activation was further supported by a strong correlation of overlapping DARs when comparing CD56^{dim} to adaptive NK cells and when comparing control to LFL + IL-12 + IL-18 activated cells (Extended Data Fig. 5j). Among these conserved DARs, we identified core features of adaptive NK cells, such as reduced accessibility at the *ZBTB16* locus and increased accessibility of *AIM2* or *IFNG* loci (Extended Data Fig. 5j). Importantly, these modulated chromatin regions showed a strong enrichment of AP1 motifs, whereas NF-κB, NFAT and STAT4 motifs were much less enriched (Extended Data Fig. 5k). Together, these findings directly connect proinflammatory cytokines and NKG2C-activating peptide, both provided during HCMV infection, to the characteristic phenotype and epigenetic remodeling of adaptive NK cells as observed ex vivo. Their synergism manifested on several layers from surface upregulation of CD137, over TF motif activity, to global chromatin remodeling, thereby recapitulating features of adaptive NK cell signatures including open chromatin enriched for AP1 motifs.

Convergent and divergent epigenetic features of adaptive NK cells

The integrated analysis of different HCMV$^+$ donors revealed consistent epigenetic, transcriptional and phenotypic features of the adaptive NK cell signature ex vivo. However, integration of donors masked donor-specific heterogeneity within the adaptive NK cell compartment. Indeed, separate analysis of the epigenetic landscape of individual donors revealed clearly defined adaptive subclusters (Fig. 5a and Extended Data Fig. 6a). These populations were characterized by an adaptive phenotype, largely lacking surface expression of NKp30 and NGK2A, and were mostly positive for NKG2C and self-MHC-specific KIRs (Extended Data Fig. 6b and Supplementary Table 2). Interestingly, subcluster 4 in donor P1 likely represents an NKG2C$^-$ adaptive NK cell expansion (Extended Data Fig. 6b), as previously described by others[11,36,37]. To assess the degree of convergence of the adaptive NK cell signature, we analyzed the overlap of DARs between the total adaptive and the CD56^{dim} compartment across HCMV$^+$ donors (P1, P2, P3 and P4), revealing that around half of all DARs were shared between at least three donors, similar to the signatures distinguishing CD56^{bright} from CD56^{dim} NK cells (Fig. 5b,c). Consistent with integration into a donor-overarching adaptive cluster (Fig. 1b), this adaptive signature was shared by all adaptive subclusters, separating them from the conventional CD56^{dim} populations (Extended Data Fig. 6c). Many of the features identified in the integrated analysis were highly penetrant across donors and subclusters, such as the peak in the proximity of *AIM2* (Extended Data Fig. 6d) and the strongly enhanced AP1 motif activity (Extended Data Fig. 6e), further supporting the idea that this

https://doi.org/10.1038/s41590-022-01327-7

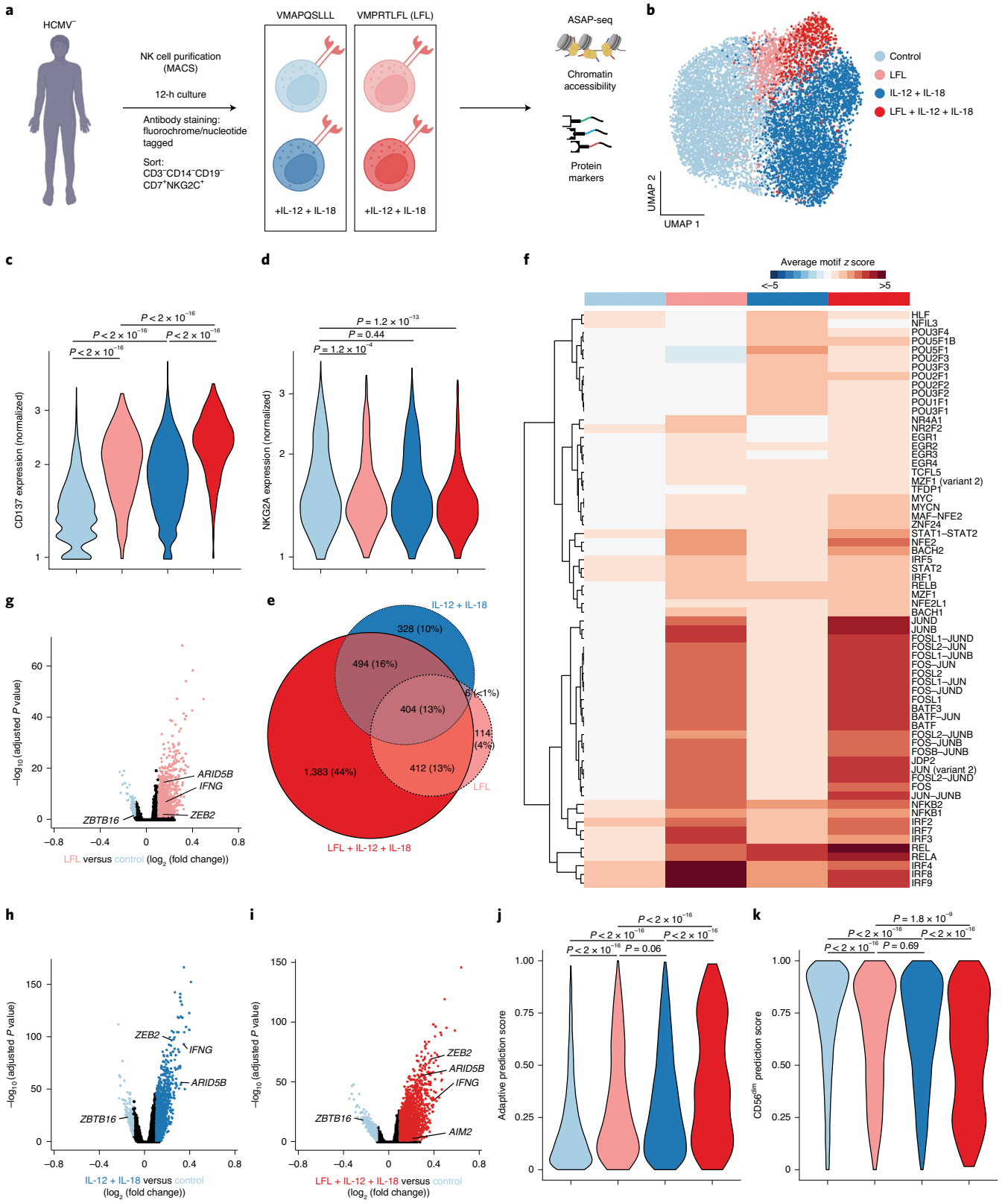


Fig. 4 | Synergistic imprinting by HCMV peptides and proinflammatory cytokines. a, NK cells from two HCMV⁻ individuals were cocultured for 12 h with RMA-S/HLA-E target cells pulsed with the indicated peptides and 10 ng ml^{-1} IL-15 in the presence (dark shades) or absence (light shades) of IL-12 and IL-18. Different conditions were marked with nucleotide-labeled hashtags and were analyzed by ASAP-seq. **b**, UMAP embedding, clusters were annotated based on clear enrichment of cells from the indicated conditions. **c**,**d**, Surface expression of CD137 (**c**) and NKG2A (**d**) per cluster; data were analyzed by two-sided Wilcoxon rank-sum test with a Benjamini–Hochberg adjustment. **e**, Euler diagram illustrating overlap of DARs between clusters. **f**, Motif activity of differentially active motifs represented as average chromVAR deviation scores per cluster. **g**–**i**, DARs between LFL and control (**g**), IL-12 + IL-18 and control (**h**) and LFL + IL-12 + IL-18 and control (**i**) by logistic regression. **j**,**k**, Prediction scores for classification as adaptive (**j**) and CD56^{dim} (**k**) NK cells based on integration with the ex vivo dataset as reference; data were analyzed by two-sided Wilcoxon rank-sum test with a Benjamini–Hochberg adjustment. Created with BioRender.com.

https://doi.org/10.1038/s41590-022-01327-7

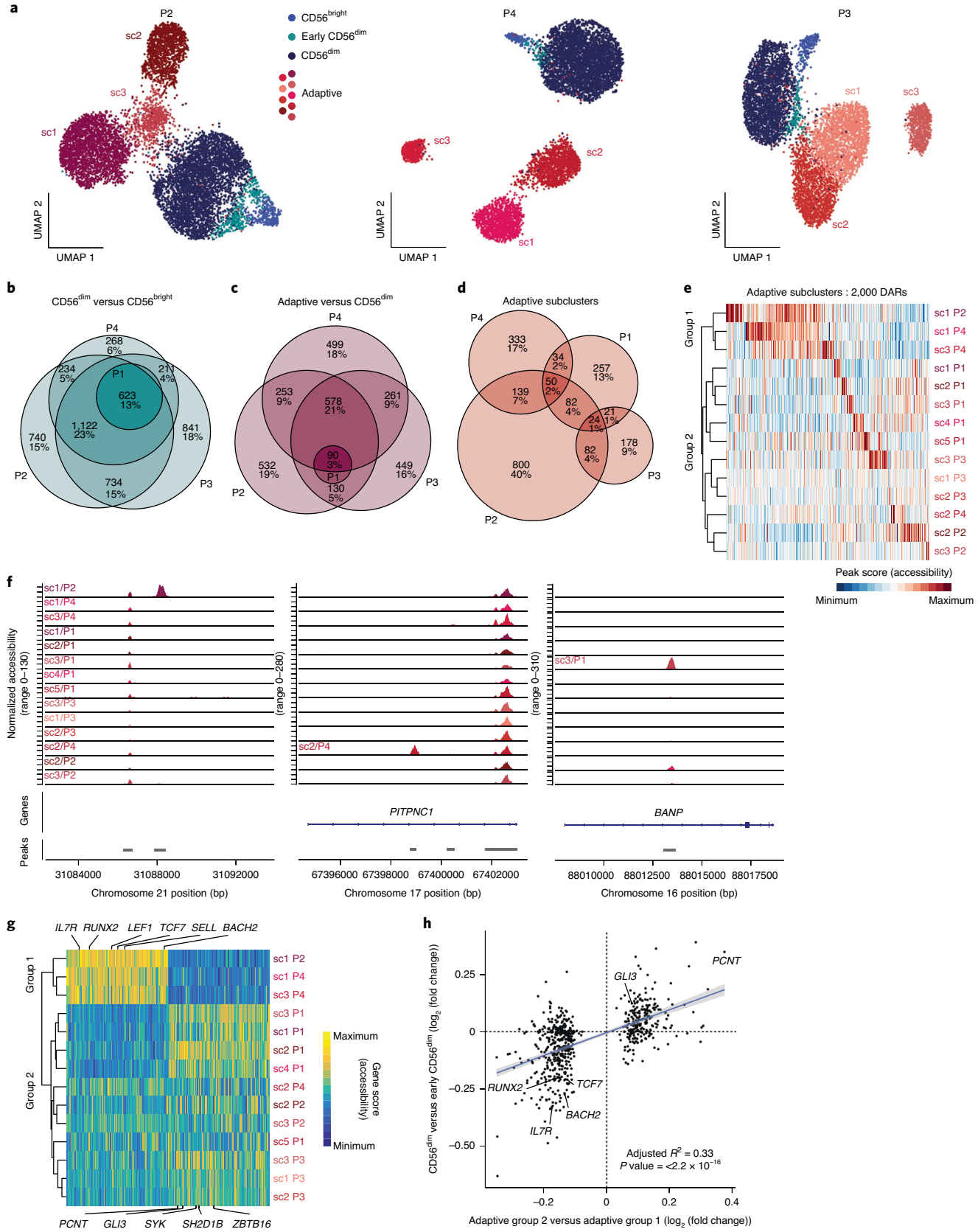


Fig. 5 | Convergent and divergent epigenetic features of adaptive NK cells. **a**, UMAP embedding of NK cells from three HCMV$^+$ individuals analyzed by scATAC-seq; sc, subcluster. **b**–**d**, Euler diagrams illustrating the overlap of DARs across donors comparing CD56^{dim} and CD56^{bright} (**b**), adaptive and CD56^{dim} (**c**) or adaptive subclusters (**d**). **e**, Column-scaled accessibility of subcluster-defining DARs and hierarchical clustering of adaptive subclusters for all individuals. **f**, Accessibility of representative subcluster-specific chromatin regions for each donor; bp, base pairs. **g**, Top 100 differential column-scaled gene scores for each adaptive subcluster group by Wilcoxon rank-sum test. **h**, Linear correlation between DARs comparing the two adaptive subcluster groups and CD56^{dim} to early CD56^{dim} NK cells; error bands show 95% confidence interval. The *P* value was determined by two-sided *F*-test.

convergent epigenetic signature reflects a coordinated program induced in response to signals received during HCMV infection. In contrast to these shared aspects, DARs between adaptive subclusters were much less conserved (Fig. 5d). Each adaptive subcluster showcased highly unique open chromatin regions (Fig. 5e,f), underlining their distinctive epigenetic makeup. Nevertheless, hierarchical clustering divided the adaptive subclusters into two groups, mainly driven by overlapping signatures of three subclusters (Fig. 5e). Importantly, when directly comparing the gene scores of the two groups, two clear signatures emerged, which had an intriguing overlap with those defining CD56^{bright}/early CD56^{dim} NK cells, such as *RUNX2*, *BACH2*, *TCF7*, *IL7R* and *SELL* (encoding CD62L), on the one end and CD56^{dim}-associated genes, such as *GLI3* or *ZBTB16*, on the other end (Fig. 5g). Indeed, differential accessibility of the respective chromatin regions between the adaptive subcluster groups correlated significantly with the differences between CD56^{bright}/early CD56^{dim} and CD56^{dim} NK cells (Fig. 5h and Extended Data Fig. 7a) and followed consistent patterns as exemplified for chromatin regions in the proximity of *TCF7* or *PCNT* (Extended Data Fig. 7b), suggesting that conventional NK cell maturation signatures are preserved after adaptive differentiation. This was also supported by differential activity of CD56^{bright}/early CD56^{dim}-associated motifs, such as TCF7L2, on the one hand and the CD56^{dim}-associated CTCF motif on the other hand (Extended Data Fig. 7c). Strikingly, the group resembling CD56^{bright}/early CD56^{dim} NK cells was also marked by surface expression of CD62L (Extended Data Fig. 7e), consistent with its expression by CD56^{bright} and early CD56^{dim} NK cells within the conventional compartment[38] (Fig. 1c,e). Additionally, the clusters that shared part of the CD56^{bright}/early CD56^{dim} signature had relatively higher AP1 motif activity (Extended Data Fig. 7d), suggesting that the two groups might also reflect a different extent of adaptive chromatin remodeling. Overall, this donor-centric analysis revealed an unappreciated heterogeneity within the adaptive NK cell compartment defined by a highly diverse set of unique, subcluster-specific peaks and signatures resembling gradients in conventional NK cell maturation, possibly reflecting the cell of origin.

Clonal expansion underlies divergent signatures of adaptive NK cells

In contrast to the convergent remodeling at key adaptive genes and a pronounced enrichment of AP1 motifs, the large number of unique peaks detected only in individual adaptive subclusters represents a diversification that is difficult to reconcile with a coordinated differentiation program. We hypothesized that this apparent diversification might result from the expansion of individual naive NKG2C^{+} cells with unique accessible chromatin regions. Expansion of rare founder cells would amplify these regions beyond the detection threshold while at the same time reduce the epigenetic heterogeneity of their progeny. Hence, as a measure for cluster heterogeneity, we analyzed the average distance between the *k* nearest neighbors within the latent semantic indexing (LSI) space representing their chromatin accessibility landscape (Methods). Notably, all adaptive subclusters were less heterogeneous than their corresponding conventional CD56^{dim} compartment, as demonstrated by a uniformly lower average distance (Fig. 6a and Extended Data Fig. 8a). This heterogeneity measure was robust over a wide range of values for *k* (Extended Data Fig. 8b), indicating a focused epigenetic profile of individual adaptive NK cell clusters, potentially resulting from a selection bottleneck during recruitment, followed by clonal expansion. To test the hypothesis that divergent epigenetic profiles of adaptive subclusters are associated with clonal expansions, we applied a recently published method that exploits mtDNA mutations as endogenous barcodes to reconstruct clonal relationships and concomitantly links them to epigenetic and cell surface phenotypes[16–18]. Applying this method to four HCMV^{+} and three HCMV^{-} donors, we identified informative mtDNA mutations by high strand concordance and variance-to-mean ratio (Extended Data Fig. 8c). In the HCMV^{+} donors, many of these mutations were specifically enriched in the adaptive NK cell compartment (Fig. 6b and Extended Data Fig. 8d). Importantly, clonotypes defined by individual mutations (Extended Data Fig. 9a and Methods) were significantly associated not only with adaptive NK cells as a whole but also with specific subclusters, as demonstrated by the χ^2 statistic of the observed compared to randomly permuted clonotype–cluster relationship (Fig. 6c,d). This striking concordance between mitochondrial mutations and epigenetic identities reveals the inheritance of epigenetic states as a defining clonal mark. Conversely, such an association was absent in the conventional NK cell compartment (Fig. 6e and Extended Data Fig. 9b), suggesting that mitochondrial mutations of conventional NK cells do not flag clonal expansions but might instead be present already on the progenitor level. Similarly, there was no significant association between clonotypes and epigenetic identities in HCMV^{-} donors (Extended Data Fig. 9c), underlining the specific clonal expansion of NKG2C^{+} NK cells in response to HCMV. Donor-specific patterns further highlighted the degree of clonal expansion of the adaptive subclusters. For example, a large fraction of cells in subcluster 3 of donor P2 carried the 5590G>A mutation at near homoplasmic frequencies (Fig. 6b), coinciding with expression of a second self-MHC-specific KIR3DL1, which was absent from the other adaptive subclusters in this individual (Extended Data Fig. 6b).

Finally, association of clonotypes to individual subclusters defined by chromatin accessibility highlighted the epigenetic similarity of cells belonging to one clonotype. This was further supported by the significant association between individual open chromatin regions and clonotypes (Fig. 6f). Importantly, we found regions that were uniquely accessible in individual donors and clonotypes (Extended Data Fig. 9d), while a part of this diversity was again driven by the opposing accessibility in the proximity of genes associated with conventional NK cell maturation, such as *IL7R* and *GLI3* (Extended Data Fig. 9e), suggesting that these signatures are indeed clonally inherited. Together, we demonstrate that subclusters of adaptive NK cells contain individual clonotypes with private chromatin accessibility profiles, strongly supporting clonal expansion as a driving force of adaptive NK cell generation and epigenetic diversification.

Adaptive NK cell clonotypes are stably maintained over time

The specific association of non-overlapping clonotypes to adaptive subclusters suggests a degree of clonal and epigenetic stability. To test this hypothesis, we performed a longitudinal follow-up analysis of three HCMV^{+} donors after 11, 7 and 19 months, respectively. Clusters defined by chromatin accessibility were unchanged between the two time points, supporting stable maintenance of the frequency and composition of the adaptive NK cell pool down to the subcluster level (Fig. 7a and Extended Data Fig. 10a,b). Further, the subcluster-specific open chromatin regions remained stable in this time frame (Fig. 7b,c and Extended Data Fig. 10c–f). Importantly, the same clonotype-defining mutations were detected at both time points and remained associated with the original subclusters (Fig. 7d and Extended Data Fig. 10g,h), highlighting their persistence and demonstrating that adaptive subclusters represent independent and stably maintained clonal expansions. Along these lines, even frequencies of clonotypes significantly associated with the adaptive compartment were stably maintained over time (Fig. 7e and Extended Data Fig. 10i,k), which was further supported by an unskewed total clonotype distribution across time points (Fig. 7f and Extended Data Fig. 10j,l). Overall, the longitudinal follow-up demonstrates the long-term clonal maintenance of adaptive NK cells with stable chromatin accessibility profiles in progeny arising from individual clonal founders, which are key features of 'epigenetic memory', as previously ascribed to the adaptive immune system.

Discussion

Previous studies have revealed conserved signatures of adaptive NK cells driven by coordinated changes in their gene regulatory networks, such as downregulation of the TF PLZF[36] or increased activity of the T cell

https://doi.org/10.1038/s41590-022-01327-7

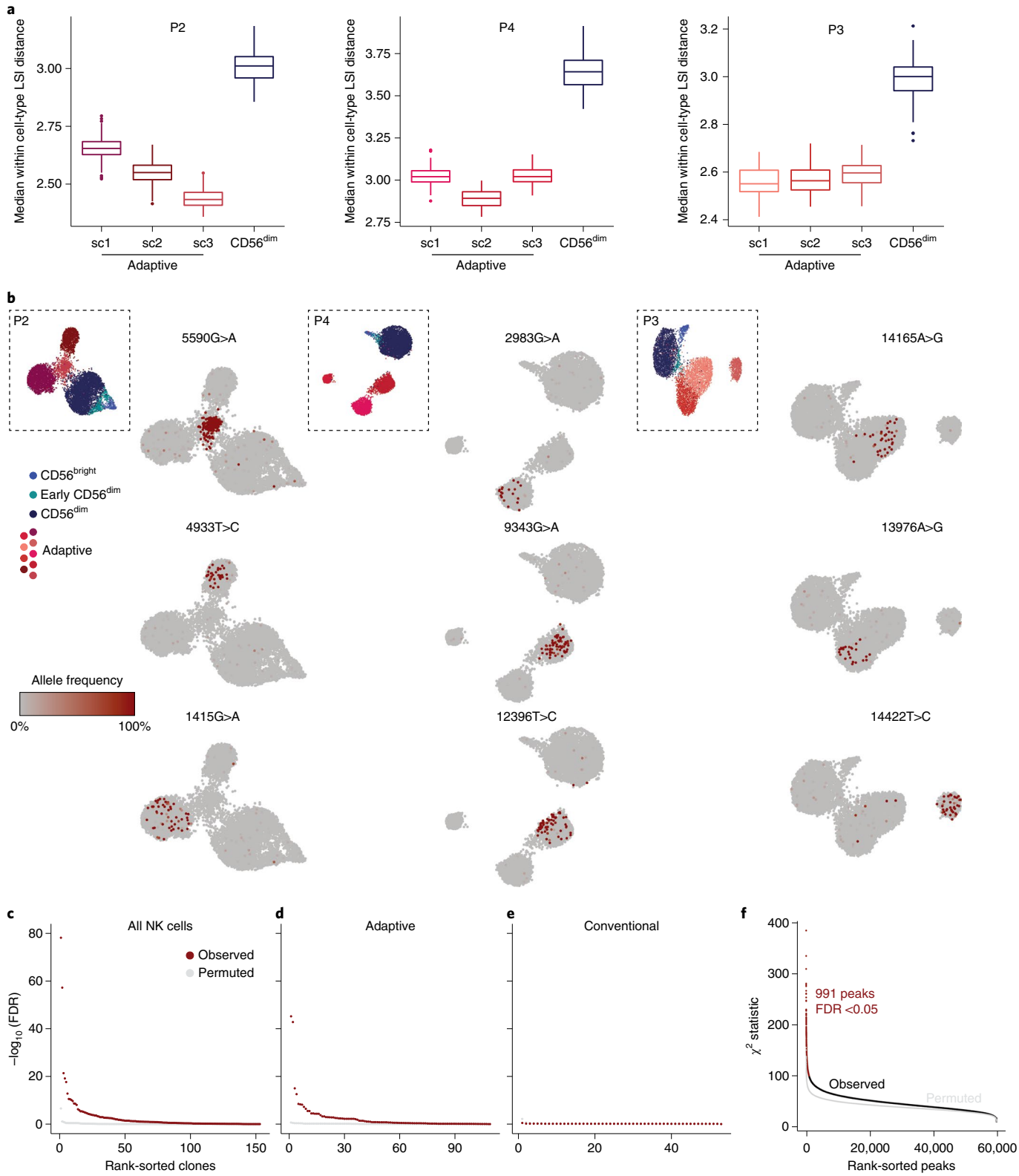


Fig. 6 | Clonal expansion underlies divergent epigenetic signatures of adaptive NK cells. a, Cluster heterogeneity as assessed by measuring the median distance of 200 randomly sampled cells to their 10 nearest neighbors and repeating this process 100 times. The upper and lower hinges of the box plots correspond to the first and third quartiles, respectively. The upper and lower whiskers extend to the largest/smallest value no further than 1.5× the interquartile range from the hinges. Outliers beyond the whiskers are displayed as individual points. **b**, Allele frequency of representative somatic mtDNA mutations projected onto UMAP embedding for each HCMV$^+$ donor. **c**–**e**, Association of clonotypes to clusters defined by chromatin accessibility for all NK cells (**c**), adaptive NK cells (**d**) and conventional NK cells (**e**); false discovery rate (FDR) from χ^2 tests for the observed and randomly permuted clonotype–cluster relationships for all donors. **f**, Association of clonotypes to open chromatin regions as assessed by χ^2 test for the observed and randomly permuted clonotype–peak relationships.

https://doi.org/10.1038/s41590-022-01327-7

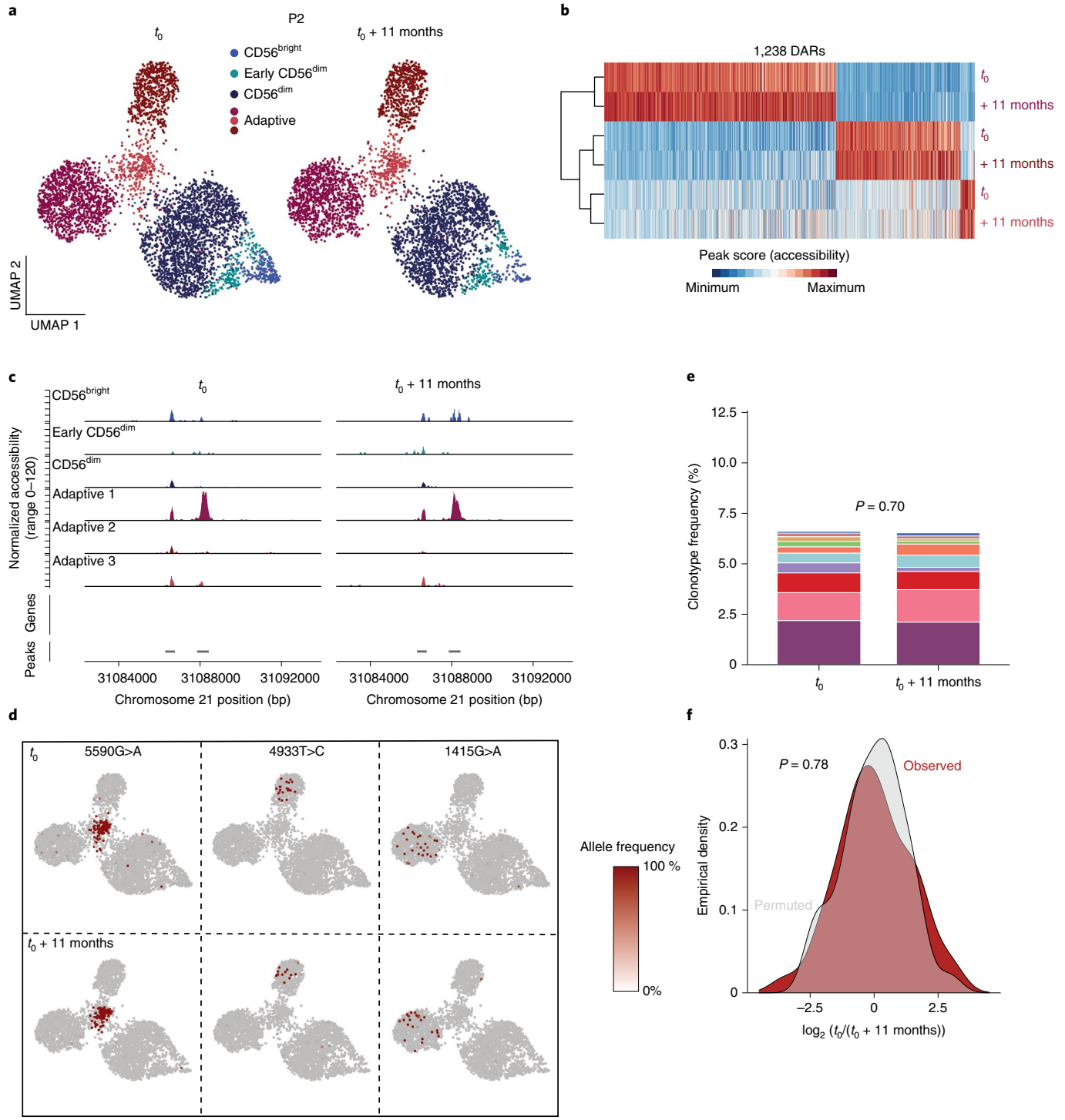


Fig. 7 | Adaptive NK cell clonotypes are stably maintained over time. a, UMAP embedding of NK cells from HCMV$^+$ donor P2 analyzed by scATAC-seq at two different time points. **b**,**c**, Stability of overall (**b**) and representative (**c**) subcluster-defining DARs over time; data are column scaled. **d**, Representative clonotype-defining mutations projected onto UMAP embeddings at the two time points. **e**, Clonotype frequency of adaptive NK cell clonotypes within the total adaptive NK cell compartment over time; data were analyzed by Fisher's exact test with Monte Carlo simulation. **f**, Observed and permuted distribution of clonotype $\log_2$ (fold change) values between time points; data were analyzed by two-sided Kolmogorov–Smirnov test.

lineage TF BCL11B[24]. Conversely, insights into their heterogeneity have remained limited to the diversification of signaling adaptor expression between and within individuals[36]. Here, we provide an integrated, multiomic single-cell analysis of human NK cells, enabling us to study both convergent and divergent aspects of adaptive NK cell biology.

Our finding of a strong enrichment of AP1 motifs in adaptive NK cell chromatin extends observations made in different cell types[4,31,39–42], further supporting the concept of a conserved inflammatory memory signature imprinted by AP1 TFs[4,32]. Here, we demonstrate that HCMV infection can induce this characteristic inflammatory memory in human NK cells, revealing a surprising conservation of this signature across stimuli and species. Consistent with reports in mice[43,44] and our own studies in humans[10] we demonstrate that synergistic activation of naive NGK2C$^+$ NK cells via CD94/NKG2C and proinflammatory

cytokines most closely recapitulated the adaptive NK cell signature observed ex vivo and induced the most pronounced activity of AP1 TFs.

Conversely, our donor-specific analysis revealed two different layers of adaptive NK cell heterogeneity. First, we identified two underlying signatures conserved across adaptive NK cells from HCMV$^+$ donors, which resembled maturation differences within the conventional NK cell compartment. Second, we unveiled unique, donor-specific, that is, private, open chromatin regions characterizing individual subclusters. Based on the high degree of clonality of adaptive NK cell expansions, we propose that both of these levels of heterogeneity can be explained by clonal inheritance of epigenetic traits, enabling an epigenetic founder effect. Acting as a clonal bottleneck, the recruitment of individual NK cells into the adaptive response and their massive expansion might skew the distribution of open chromatin regions. Hence, the two layers of adaptive NK cell heterogeneity likely reflect the degree of diversity and maturation stage within the original naive NKG2C$^+$ pool before acquisition of the superimposed adaptive differentiation program. Together with our observation that naive NKG2C$^+$ NK cells span the whole spectrum of conventional NK cell maturation, these data suggest that naive NKG2C$^+$ NK cells might be recruited into the adaptive response at these different stages and preserve a part of their initial epigenetic identity. This is consistent with findings in mice, where pre-existing NK cell subtypes were preserved after murine CMV-induced expansion[45]. Alternatively, these might represent different stages of adaptive NK cells, although the association of clonotypes with individual subclusters and their stability in time argue against this interpretation. Regarding the unique open chromatin regions associated with adaptive NK cell subclusters, we envisage that they become evident due to the drastic expansion of individual NK cell clones, which at the same time underlies the reduced heterogeneity of adaptive subpopulations. Consequently, we propose that the macroscopic picture of epigenetic diversification, manifesting, for example, in variable expression of signaling adaptors[36], is created by the coexistence of several clonal expansions that are, in fact, epigenetically focused.

Importantly, our findings enable us to generalize some of the mechanisms driving clonal expansion and their consequences within the immune system as a whole. One such unifying prerequisite for clonal expansion is the competitive recruitment of a limited number of cells to emerge as clones from a diverse population. Despite their reliance on a conserved receptor for HCMV recognition, other cell-intrinsic features of NK cells are strongly diversified, especially their combined cell surface receptor expression profiles[46], leading to selection of optimal receptor combinations. The low frequency of cells expressing NKG2C before HCMV infection is the first limitation of the original pool activated by HCMV peptides. Further, in response to peptide and cytokine stimulation, only NKG2A$^-$ NK cells underwent remodeling that resembled adaptive NK cells the closest, suggesting that NKG2A expression further limits the pool of naive NKG2C$^+$ cells optimally responding to HCMV. Expression levels of NKG2C might also contribute to selection, as adaptive NKG2C$^+$ NK cells express higher levels of NKG2C than their naive counterparts[47], and a similar mechanism occurs for Ly49H$^+$ NK cells in mice[12,48]. Clonal selection can further be influenced by complex biological traits that are heterogeneously distributed among cells before or after activation, such as proliferation and differentiation kinetics[49,50] or metabolic capacity[51]. More studies will be needed to distinguish the adaptive NK cell features modulated by external signals from those conferring a selection advantage within the naive repertoire and the ones that are stochastic passengers of the clonal expansion process. The potential of non-genetic inheritance mechanisms to impact cell fitness has recently been demonstrated in the context of tumor cell resistance against anticancer therapeutics[52]. As T cells can also clonally inherit transcriptional and epigenetic diversity[53], these analyses might provide general insights into the drivers of clonal success in different cell types independent of receptor specificity.

Finally, prolonged survival of adaptive NK cells has been postulated based on the observation of stable expression of KIR and signaling adaptor patterns[36], and evidence from individuals with mutations in *GATA2* or *PIGA* supports this concept[54,55]. Here, we demonstrated long-term persistence on the clonal level for at least 19 months, although the exact cellular mechanisms enabling clonal maintenance remain to be defined.

Here, we demonstrate the clonal expansion and long-term persistence of an innate immune population in a memory state in response to a naturally occurring infection and introduce clonal selection and epigenetic inheritance as mechanisms that shape human innate immune composition for optimized secondary responses. Understanding the mechanisms determining clonal fitness in adaptive and innate immune cells independent of diversified antigen receptors might open new avenues for fundamental research and therapeutic interventions.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41590-022-01327-7.

References

1. Pradeu, T. & Du Pasquier, L. Immunological memory: what's in a name? *Immunol. Rev.* **283**, 7–20 (2018).
2. Netea, M. G. et al. Defining trained immunity and its role in health and disease. *Nat. Rev. Immunol.* **20**, 375–388 (2020).
3. Kaufmann, E. et al. BCG educates hematopoietic stem cells to generate protective innate immunity against tuberculosis. *Cell* **172**, 176–190 (2018).
4. Larsen, S. B. et al. Establishment, maintenance, and recall of inflammatory memory. *Cell Stem Cell* **28**, 1758–1774.e8 (2021).
5. Kleinnijenhuis, J. et al. Bacille Calmette–Guerin induces NOD2-dependent nonspecific protection from reinfection via epigenetic reprogramming of monocytes. *Proc. Natl Acad. Sci. USA* **109**, 17537–17542 (2012).
6. Mitroulis, I. et al. Modulation of myelopoiesis progenitors is an integral component of trained immunity. *Cell* **172**, 147–161 (2018).
7. Hammer, Q., Rückert, T. & Romagnani, C. Natural killer cell specificity for viral infections. *Nat. Immunol.* **19**, 800–808 (2018).
8. Berry, R., Watson, G. M., Jonjic, S., Degli-Esposti, M. A. & Rossjohn, J. Modulation of innate and adaptive immunity by cytomegaloviruses. *Nat. Rev. Immunol.* **20**, 113–127 (2020).
9. Sun, J. C., Beilke, J. N. & Lanier, L. L. Adaptive immune features of natural killer cells. *Nature* **457**, 557–561 (2009).
10. Hammer, Q. et al. Peptide-specific recognition of human cytomegalovirus strains controls adaptive natural killer cells. *Nat. Immunol.* **19**, 453–463 (2018).
11. Béziat, V. et al. NK cell responses to cytomegalovirus infection lead to stable imprints in the human KIR repertoire and involve activating KIRs. *Blood* **121**, 2678–2688 (2013).
12. Grassmann, S. et al. Distinct surface expression of activating receptor Ly49H drives differential expansion of NK cell clones upon murine cytomegalovirus infection. *Immunity* **50**, 1391–1400 (2019).
13. Truitt, L. L. et al. Impact of CMV infection on natural killer cell clonal repertoire in CMV-naive rhesus macaques. *Front. Immunol.* **10**, 2381 (2019).
14. Wu, C. et al. Clonal expansion and compartmentalized maintenance of rhesus macaque NK cell subsets. *Sci. Immunol.* **3**, eaat9781 (2018).
15. Adams, N. M., Grassmann, S. & Sun, J. C. Clonal expansion of innate and adaptive lymphocytes. *Nat. Rev. Immunol.* **20**, 694–707 (2020).

16. Ludwig, L. S. et al. Lineage tracing in humans enabled by mitochondrial mutations and single-cell genomics. *Cell* **176**, 1325–1339 (2019).
17. Lareau, C. A. et al. Massively parallel single-cell mitochondrial DNA genotyping and chromatin profiling. *Nat. Biotechnol.* **39**, 451–461 (2021).
18. Mimitou, E. P. et al. Scalable, multimodal profiling of chromatin accessibility, gene expression and protein levels in single cells. *Nat. Biotechnol.* **39**, 1246–1258 (2021).
19. Stoeckius, M. et al. Simultaneous epitope and transcriptome measurement in single cells. *Nat. Methods* **14**, 865–868 (2017).
20. 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).
21. Collins, P. L. et al. Gene regulatory programs conferring phenotypic identities to human NK cells. *Cell* **176**, 348–360 (2019).
22. Yang, C. et al. Heterogeneity of human bone marrow and blood natural killer cells defined by single-cell transcriptome. *Nat. Commun.* **10**, 3931 (2019).
23. Smith, S. L. et al. Diversity of peripheral blood human NK cells identified by single-cell RNA sequencing. *Blood Adv.* **4**, 1388–1406 (2020).
24. Holmes, T. D. et al. The transcription factor Bcl11b promotes both canonical and adaptive NK cell differentiation. *Sci. Immunol.* **6**, eabc9801 (2021).
25. Cichocki, F. et al. ARID5B regulates metabolic programming in human adaptive NK cells. *J. Exp. Med.* **215**, 2379–2395 (2018).
26. Gumá, M. et al. Imprint of human cytomegalovirus infection on the NK cell receptor repertoire. *Blood* **104**, 3664–3671 (2004).
27. 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).
28. Cooper, M. A. et al. Human natural killer cells: a unique innate immunoregulatory role for the CD56^{bright} subset. *Blood* **97**, 3146–3151 (2001).
29. Foley, B. et al. Cytomegalovirus reactivation after allogeneic transplantation promotes a lasting increase in educated NKG2C^{+} natural killer cells with potent function. *Blood* **119**, 2665–2674 (2012).
30. Luetke-Eversloh, M. et al. Human cytomegalovirus drives epigenetic imprinting of the *IFNG* locus in NKG2C^{hi} natural killer cells. *PLoS Pathog.* **10**, e1004441 (2014).
31. Ostuni, R. et al. Latent enhancers activated by stimulation in differentiated cells. *Cell* **152**, 157–171 (2013).
32. Yukawa, M. et al. AP-1 activity induced by co-stimulation is required for chromatin opening during T cell activation. *J. Exp. Med.* **217**, e20182009 (2020).
33. Valés-Gómez, M., Reyburn, H. T., Erskine, R. A., López-Botet, M. & Strominger, J. L. Kinetics and peptide dependency of the binding of the inhibitory NK receptor CD94/NKG2-A and the activating receptor CD94/NKG2-C to HLA-E. *EMBO J.* **18**, 4250–4260 (1999).
34. Wiedemann, G. M. et al. Deconvoluting global cytokine signaling networks in natural killer cells. *Nat. Immunol.* **22**, 627–638 (2021).
35. Stuart, T. et al. Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902 (2019).
36. Schlums, H. et al. Cytomegalovirus infection drives adaptive epigenetic diversification of NK cells with altered signaling and effector function. *Immunity* **42**, 443–456 (2015).
37. Liu, L. L. et al. Critical role of CD2 co-stimulation in adaptive natural killer cell responses revealed in NKG2C-deficient humans. *Cell Rep.* **15**, 1088–1099 (2016).
38. Juelke, K. et al. CD62L expression identifies a unique subset of polyfunctional CD56^{dim} NK cells. *Blood* **116**, 1299–1307 (2010).
39. Moskowitz, D. M. et al. Epigenomics of human CD8 T cell differentiation and aging. *Sci. Immunol.* **2**, eaag0192 (2017).
40. Scharer, C. D., Barwick, B. G., Guo, M., Bally, A. P. R. & Boss, J. M. Plasma cell differentiation is controlled by multiple cell division-coupled epigenetic programs. *Nat. Commun.* **9**, 1698 (2018).
41. Lau, C. M. et al. Epigenetic control of innate and adaptive immune memory. *Nat. Immunol.* **19**, 963–972 (2018).
42. de Laval, B. et al. C/EBPβ-dependent epigenetic memory induces trained immunity in hematopoietic stem cells. *Cell Stem Cell* **26**, 657–674 (2020).
43. Nabekura, T. & Lanier, L. L. Tracking the fate of antigen-specific versus cytokine-activated natural killer cells after cytomegalovirus infection. *J. Exp. Med.* **213**, 2745–2758 (2016).
44. Sun, J. C. et al. Proinflammatory cytokine signaling required for the generation of natural killer cell memory. *J. Exp. Med.* **209**, 947–954 (2012).
45. Flommersfeld, S. et al. Fate mapping of single NK cells identifies a type 1 innate lymphoid-like lineage that bridges innate and adaptive recognition of viral infection. *Immunity* **54**, 2288–2304 (2021).
46. Horowitz, A. et al. Genetic and environmental determinants of human NK cell diversity revealed by mass cytometry. *Sci. Transl. Med.* **5**, 208ra145 (2013).
47. Lopez-Vergès, S. et al. Expansion of a unique CD57^{+}NKG2C^{hi} natural killer cell subset during acute human cytomegalovirus infection. *Proc. Natl Acad. Sci. USA* **108**, 14725–14732 (2011).
48. Adams, N. M. et al. Cytomegalovirus infection drives avidity selection of natural killer cells. *Immunity* **50**, 1381–1390 (2019).
49. Buchholz, V. R. et al. Disparate individual fates compose robust CD8^{+} T cell immunity. *Science* **340**, 630–635 (2013).
50. Grassmann, S. et al. Early emergence of T central memory precursors programs clonal dominance during chronic viral infection. *Nat. Immunol.* **21**, 1563–1573 (2020).
51. Hartmann, F. J. et al. Single-cell metabolic profiling of human cytotoxic T cells. *Nat. Biotechnol.* **39**, 186–197 (2021).
52. Shaffer, S. M. et al. Memory sequencing reveals heritable single-cell gene expression programs associated with distinct cellular behaviors. *Cell* **182**, 947–959 (2020).
53. Mold, J. E. et al. Clonally heritable gene expression imparts a layer of diversity within cell types. Preprint at *bioRxiv* https://doi.org/10.1101/2022.02.14.480352 (2022).
54. Corat, M. A. F. et al. Acquired somatic mutations in PNH reveal long-term maintenance of adaptive NK cells independent of HSPCs. *Blood* **129**, 1940–1946 (2017).
55. Schlums, H. et al. Adaptive NK cells can persist in patients with *GATA2* mutation depleted of stem and progenitor cells. *Blood* **129**, 1927–1939 (2017).

Methods

Human studies

All analyses were performed in compliance with the relevant ethical regulations, and all donors gave informed consent. Primary NK cells were isolated from freshly drawn peripheral blood of healthy donors (Supplementary Table 1) or from buffy coats obtained from Deutsches Rotes Kreuz (DRK) Blutspendedienst Nord-Ost. The Charité Ethics Committee approved the study (EA4/196/18 and EA4/059/17).

Cell lines

RMA-S/HLA-E cells[56] were maintained in complete medium (RPMI-1640 containing glutamine and supplemented with 10 % (vol/vol) fetal bovine serum (FBS), 20 µM β-mercaptoethanol and 100 U ml^{-1} penicillin–streptomycin; all Thermo Fisher) in the presence of 400 µg ml^{-1} hygromycin B (InvivoGen).

Cell isolation

Peripheral blood mononuclear cells were isolated by density gradient centrifugation (Ficoll Paque Plus, GE Healthcare) and either processed immediately or cryopreserved in FBS containing 10% DMSO.

HCMV serology

For buffy coats, CMV serology was performed at DRK Dresden. Serological status of fresh blood donors was analyzed by HCMV IgG enzyme-linked immunosorbent assay (IBL International) following the manufacturer's instructions. All samples were measured in duplicate and were either clearly above or below the cutoff control.

Flow cytometry

Cell suspensions were stained with combinations of the following fluorochrome-conjugated antibodies following established guidelines[57]: CD56 PE/Dazzle 594 (1:200), CD137 PE/cyanine7 (1:50) and CD57 BV605 (1:25; all BioLegend); CD7 BV786 (1:25), CD56 BUV737 (1:50), CD3 BUV805 (1:50), streptavidin BUV395 (1:100), CD16 BUV496 (1:50) and CD337 (NKp30) BV421 (1:25; all BD Biosciences); CD159a (NKG2A) biotin/PE-Vio770 (both 1:50) and CD159c (NKG2C) PE (1:100; all REAfinity, Miltenyi Biotec); CD337 (NKp30) eFluor 450 (1:25), CD3 APC-eFluor 780 (1:50), CD14 APC-eFluor 780 (1:50) and CD19 APC-eFluor 780 (1:50; all Thermo Fisher) and anti-FcεRI, γ subunit-FITC (1:50; Merck). Dead cells were excluded using Fixable Viability Dye eFluor 780 (Thermo Fisher) or a Zombie Aqua Fixable Viability kit (BioLegend). Data were acquired on an LSR Fortessa (BD Biosciences). Cell sorting was performed on a FACSAria II (BD Biosciences). FlowJo v10 was used for analysis of flow cytometry data, and GraphPad Prism v8.4.3 was used for statistical analysis.

KIR ligand genotyping

DNA was isolated from 100 µl of peripheral blood using spin-column-based purification implemented with the DNeasy Blood and Tissue kit (Qiagen), and concentrations were measured using a Nanodrop 2000c spectrophotometer (Thermo Fisher). KIR ligand genotyping was performed using the Olerup SSP KIR HLA ligand typing kit (CareDx) following the manufacturer's instructions. Reactions were analyzed on a 2 % (wt/vol) agarose gel prestained with GelRed (Biotium) in 0.5× TBE buffer. Gels were documented on a UV transilluminator and interpreted using the manufacturer-supplied reference tables. All lanes displayed the required control bands.

Cell preparation for scATAC-seq

NK cells were enriched from peripheral blood mononuclear cells isolated from freshly drawn peripheral blood by magnetic depletion using microbeads against CD3, CD14 and CD19. The enriched fraction was stained with the following fluorochrome-conjugated antibodies: CD7 BV786 (1:25; BD Biosciences), CD3 APC-eFluor 780 (1:50), CD14 APC-eFluor 780 (1:50) and CD19 APC-eFluor 780 (1:50; all Thermo Fisher), CD159a (NKG2A) biotin (1:50) and CD159c (NKG2C) PE (1:100; both REAfinity, Miltenyi Biotec). Fc receptors were blocked using Human TruStain FcX (1:50; BioLegend) for 15 min at 4 °C. Dead cells were excluded using Fixable Viability Dye eFluor 780 (Thermo Fisher). In a second step, cells were stained with combinations of the following nucleotide barcode-labeled antibodies (BioLegend) for 30 min at 4 °C: TotalSeq-A0084 CD56 (1:200), A0083 CD16 (1:500), A0436 anti-biotin (1:100), A0147 CD62L (1:100), A0168 CD57 (1:100), A0367 CD2 (1:1,000), A0801 CD337 (NKp30; 1:100), A0149 CD161 (1:100), A0420 CD158 (1:100), A0592 CD158b (1:100), A0599 CD158e1 (1:100), A0390 CD127 (1:400), A0902 CD328 (Siglec-7; 1:1,000), A0867 CD94 (1:100), A0896 CD85j (ILT2; 1:100), A0911 anti-phycoerythrin (1:100), A0250 KLRG1 (1:100), A0061 CD117 (1:100), A1018 HLA-DR, HLA-DP and HLA-DQ (1:400), A0355 CD137 (1:100), A0152 CD223 (1:100), A0366 CD184 (1:100), A0251 hashtag 1, A0252 hashtag 2, A0253 hashtag 3, A0254 hashtag 4, A0255 hashtag 5, A0256 hashtag 6, A0257 hashtag 7, A0258 hashtag 8, A0259 hashtag 9, A0260 hashtag 10, A0262 hashtag 12, A0263 hashtag 13, A0264 hashtag 14 and A0265 hashtag 15 (all 1:200–1:400). As described in the ASAP-seq and mitochondrial scATAC-seq (mtscATAC-seq) protocols, cells were fixed with 1% paraformaldehyde for 10 min. Fixation was quenched by adding glycin to a final concentration of 0.125 M, and cells were washed twice with PBS/bovine serum albumin (BSA). NK cells were sorted on a FACSAria II (BD Biosciences) as viable, single cells being $CD3^-CD14^-CD19^-CD7^+$, separating $NKG2C^+$ and $NKG2C^-$ cells. We used CD7 to sort for NK cells, as it is widely expressed on all NK cell subsets[58] and enabled us to stain markers informative for NK cell differentiation, such as CD56 and CD16, with nucleotide barcode-labeled antibodies. After sorting, $NKG2C^+$/$NKG2C^-$ NK cells were pooled and lysed for 3 min on ice using modified lysis buffer (10 mM Tris-HCl (pH 7.5), 10 mM NaCl, 3 mM $MgCl_2$, 0.1% NP40 and 1% BSA). Afterward, lysed cells were washed with washing buffer (10 mM Tris-HCl (pH 7.5), 10 mM NaCl, 3 mM MgCl and 1% BSA) and resuspended in diluted nuclei buffer (10x Genomics). Lysis was confirmed by trypan blue staining, and cells were counted and further processed as described below.

scATAC-seq

scATAC-seq libraries were prepared using the Chromium Next GEM single-cell ATAC reagent kit v1.1 following the manufacturer's instructions with the modifications described in the ASAP-seq and mtscATAC-seq protocols to retrieve TotalSeq antibody-derived tags (ADTs), hashtag oligonucleotides (HTOs) and mtDNA[16–18]. The saved eluate from the silane bead elution and the supernatant of the first SPRIselect purification were combined for amplification of ADT/HTO libraries using the KAPA HiFi ready mix (Roche) with sample-specific index primers (Illumina small RNA RPIx/Truseq D7xx) and purified using SPRIselect reagent. Library size and quality were analyzed using a Fragment Analyzer (Advanced Analytical) before fragmentation and after final purification. Final library concentrations were measured on a Qubit 2.0 fluorometer (Thermo Fisher). Libraries were sequenced on a NextSeq500 or NovaSeq6000 sequencer (Illumina) using longer read1/2 configurations than suggested by 10x Genomics to improve mitochondrial genotyping (NextSeq500: R1 72 cycles, R2 72 cycles, I1 8 cycles, I2 16 cycles; NovaSeq6000: R1 88 cycles, R2 88 cycles, I1 8 cycles, I2 16 cycles).

Cell preparation for scRNA-seq

NK cells were enriched, stained and sorted as described above for scATAC-seq (without fixation). $NKG2C^+$ and $NKG2C^-$ cells were separately labeled with different nucleotide-labeled hashtag antibodies (BioLegend). $NKG2C^+$/$NKG2C^-$ populations were pooled at a 1:1 ratio, and the cell concentration was adjusted to 1,000 cells per µl for further processing.

scRNA-seq

scRNA-seq libraries were prepared using the Chromium Single Cell 3′ Reagent kit v2/v3.1 chemistry (10x Genomics) following

the manufacturer's instructions with the modifications required for recovery of ADTs and HTOs[19]. ADT and HTO libraries were amplified using the KAPA HiFi ready mix (Roche) with sample-specific index primers (Illumina small RNA RPIx/Truseq D7xx)[19] and purified using SPRIselect reagent. Library size and quality were analyzed using a Fragment Analyzer (Advanced Analytical) before fragmentation and after final purification. Final library concentrations were measured on a Qubit 2.0 fluorometer (Thermo Fisher). Libraries were sequenced on a NextSeq500 or NovaSeq6000 sequencer (Illumina) using the read configurations recommended by 10x Genomics (v2: R1 26 cycles, R2 98 cycles, I1 8 cycles; v3: R1 28 cycles, R2 91 cycles, I1 8 cycles).

Stimulation of NK cells from HCMV$^-$ donors

RMA-S/HLA-E target cells were prepared by seeding cells in 24- to 96-well plates at a cell concentration of 2×10^6 cells per ml in OptiMEM (Thermo Fisher) and irradiated with 3,000 cGy. Afterward, cells were pulsed either with the negative control peptide VMAPQSLLL or the activating peptide LFL (Peptides&Elephants) at a concentration of 300 µM for 6–12 h.

For scATAC-seq, NK cells were isolated from freshly drawn blood by magnetic depletion of CD3$^+$ cells, followed by enrichment of CD56$^+$ cells. For read-out by flow cytometry, cells were isolated from buffy coats, and CD56$^+$ cells were enriched using CD56 microbeads and were cryopreserved in FBS containing 10% DMSO. Viable CD56$^+$CD3$^-$ cells were sorted from the cryopreserved CD56-enriched fractions. Purified NK cells were cocultured for 12 h with peptide-pulsed RMA-S/HLA-E cells in complete medium supplemented with 10 ng ml^{-1} IL-15 (Miltenyi Biotec) and the respective peptides at 300 µM in the presence or absence of 10 ng ml^{-1} IL-12 (Miltenyi Biotec) and 100 ng ml^{-1} IL-18 (MBL International). Cells were stained with fluorochrome-labeled antibodies and either analyzed by flow cytometry or additionally stained with nucleotide barcode-labeled antibodies, including hashtag antibodies, marking experimental conditions and donors as described above before sorting for viable CD7$^+$NKG2C$^+$ cells. Sorted cells were pooled at equal numbers per condition and processed for scATAC-seq as described above.

Removal of sensitive genetic variant information

Raw BAM files for all experiments were processed with BAMboozle[59] to remove potentially identifying donor-related single nucleotide polymorphisms and indels by replacing them with the sequences present in the corresponding reference genomes. To enable reproducibility of the clonotype analysis based on mitochondrial mutations, the mitochondrial genotyping results were deposited with the raw and processed data for each experiment.

Preprocessing of scATAC-seq data

Base call files were demultiplexed using cellranger-atac v1.2.0 mkfastq and bcl2fastq v2.20.0.422 into scATAC-seq and ADT/HTO libraries. scATAC-seq reads were mapped with cellranger-atac count to the GRCh38 reference genome, hardmasked for regions that would otherwise interfere with mapping to the mitochondrial genome[17], generating count tables and fragment files as output for further analysis. ADT and HTO reads were preprocessed using ASAP to kite[18], followed by pseudoalignment with kallisto[60] to a mismatch map generated with kite and counting with bustools[61].

Integrated analysis of ex vivo scATAC-seq data

Peak sets called from different experiments were reduced to a joint peak set as a basis for combined analysis of individual experiments using Signac[62]. Cells were filtered for outliers based on nucleosome signal (<1–1.2), TSS enrichment (>2.5–4) and the frequency of reads in peaks (>55–70 %) and blacklisted regions (<0.0001). ADT and HTO counts were imported, centered log ratio normalized and joined with the scATAC-seq data. Normalized hashtag reads were used for exclusion of doublets and to demultiplex donors. Chromatin accessibility counts were normalized by term frequency inverse document frequency, data from different experiments were merged, and dimensionality was reduced by singular value decomposition. The resultant LSI was integrated across experiments using harmony[63] and used as input into further dimensionality reduction by uniform manifold approximation (UMAP) and neighborhood graph-based clustering. A small population of non-NK cells, as judged by their lack of CD56 and CD16 expression concomitant with high levels of CD127 and CD117, was excluded. The merged and integrated dataset was used for peak calling using MACS2 (ref. [64]) to enable more sensitive calling of cluster-specific open chromatin regions. This new assay was again processed by LSI, followed by anchor-based integration of donors[35]. The UMAP embeddings and clustering for final analysis were based on this integrated LSI. Unbiased clustering initially yielded three major clusters that were annotated based on surface protein expression, with early CD56^{dim} NK cells separated from CD56^{bright} NK cells by further subclustering. Gene activities per cell were calculated by summing up counts within the gene body and up to 2,000 bases upstream of the TSS.

For analysis of the global signatures, differential accessibility analysis of individual chromatin regions was performed using logistic regression with the number of peaks as latent variable[65], while gene activities were assessed by Wilcoxon rank-sum test. For both, the $\log_2$ (fold change) threshold was set at 0.1, and significant hits from the direct comparison of CD56^{bright} to CD56^{dim} were added for visualization to include the CD56^{dim} signature that was mostly shared with adaptive NK cells. Averaged values were visualized in heat maps using tidyheatmap, an implementation of pheatmap.

To compare HCMV$^+$ and HCMV$^-$ donors, these were separated from the integrated dataset followed by UMAP embedding and clustering as described above. To directly compare the number of DARs between NKG2C$^+$ and NKG2C$^-$ cells, these populations were defined based on anti-PE ADT counts marking NKG2C surface expression and then downsampled to equal numbers (4,300 NKG2C$^+$/NKG2C$^-$ cells) for both donor groups. Further, NKG2C$^+$ NK cells from HCMV$^+$ donors were extracted, followed by UMAP embedding, and their cluster identities were assessed based on the signatures from the integrated analysis of the full dataset.

Motif analysis

Position frequency matrices for TF motif analysis were downloaded from the JASPAR2020 human core database and amended by Stat-TF motifs from the murine database[66]. Motif activities per cell were calculated using chromVAR[27]. Enrichment of motifs in differentially accessible peaks between clusters was assessed using the hypergeometric test implemented in Signac. De novo motif analysis on the adaptive NK-cell-specific peaks was performed with homer v4.11 findMotifsGenome.pl using the standard parameters[67].

Mitochondrial genotyping and clonotypes

Mitochondrial genotyping was performed with mgatk[17] in tenx mode using the barcodes identified as cells by cellranger. Only cells with at least 5× coverage of the mitochondrial genome were included in the analysis, achieving a median coverage of 11–20× across experiments. As sensitivity and positive predictive value are relatively independent of coverage for mutations with high heteroplasmy[17], those could be confidently detected at this coverage. Using a combination of highly abundant, homoplasmic mitochondrial mutations characteristic of each donor, we noted that donor demultiplexing was slightly incomplete in some cases, so we used these mutations as additional donor barcodes to further improve demultiplexing before the donor-specific analysis. Mitochondrial

mutations called by mgatk were filtered for high-confident variants detected in at least three cells, with strand concordance of >65% and a variance-to-mean ratio of >0.01. Clonotypes were identified by clustering on a neighborhood graph constructed on mitochondrial mutation frequency using the Euclidean distance metric, because we noted that the majority of adaptive clonotypes were defined by mutations with high allele frequencies and thereby achieved better clonotype resolution than with the cosine distance metric originally suggested for mtscATAC-seq[17]. Clustering parameters were empirically optimized per donor so that clonotypes were defined by approximately one mutation, and only those clonotypes were included that had at least one significantly enriched mutation. Clonotypes only defined by differential frequency of highly abundant mutations were excluded from the analysis. Association of clonotypes to clusters defined by their chromatin accessibility profiles was analyzed by χ^2 test and comparing the resultant false discovery rates (FDRs) to a randomly permuted matrix[17]. Analysis of clonotype association to open chromatin regions was performed using a χ^2 test on the binarized accessibility matrix and compared to a randomly permuted matrix.

Donor-specific and donor-unique signatures

Individual HCMV$^+$ donors were analyzed based on the peak set called for the integrated dataset. Both time points were included in this analysis and integrated using harmony[62]. Dimensionality reduction and clustering was performed for each individual donor as described above. To analyze overlaps and differences between the signatures defining clusters across donors, differential accessibility analysis was performed per donor as described for the integrated dataset, and overlaps were visualized using eulerr. To assess the actual signatures, adaptive subclusters defined in each donor were merged, and averaged accessibility per cluster was visualized with tidyheatmap. Differential gene scores and motif activities between the two emerging adaptive subcluster groups were analyzed by directly comparing them to each other within the merged object. To assess similarity of their signatures to those of CD56^{bright}/early CD56^{dim} and CD56^{dim} NK cells, fold changes of the adaptive subcluster group-defining DARs were plotted for both comparisons and analyzed for linear correlation.

Cluster heterogeneity

Cluster heterogeneity was assessed based on the median Euclidian distance of 200 randomly sampled cells per cluster to their 10 nearest neighbors in LSI space (components 2–30). Sampling was repeated 100 times, and results were plotted as box plots of the distribution of median distances of all simulations. The *k* parameter was varied from 5 to 30 to test for robustness.

Longitudinal analysis

The two time points for donors P2, P3 and P4 were separated from the merged, harmony-integrated objects, and the signatures defining adaptive subclusters were analyzed separately for each time point using hierarchical clustering and visualization with tidyheatmap. Allele frequency of clonotype-defining mutations was visualized for each time point. To analyze the stability of clonotypes over time, only clonotypes that had an association with chromatin accessibility-defined clusters with an FDR of <0.05 were included; these were further curated by assessment of their association with adaptive subclusters in UMAP embeddings. Stability of clonotypes significantly associated with adaptive subclusters was analyzed by testing association of clonotypes with time points by Fisher's exact test. The stability of the total clonotype distribution was further assessed by calculating their $\log_2$ (fold change) values between time points and comparing the distribution to a randomly permuted clonotype–time point relationship by Kolmogorov–Smirnov test.

Analysis of scATAC-seq data after activation

Quality control, joining of ADT, HTO and scATAC-seq counts, demultiplexing, peak calling on the full dataset with MACS2, normalization and dimensionality reduction by LSI were performed as described for the ex vivo analysis. The two analyzed donors were integrated with harmony, and the corrected LSI was used for UMAP embedding and clustering. After inspection of the distribution of cells from the different conditions and donors, the majority of cells were clearly separated based on whether they had received peptide or cytokine stimulation, but we also noted two smaller clusters, each enriched for cells from one donor, respectively, in which cells from different conditions were mixed, likely driven by underlying donor-specific signatures that were not corrected by harmony integration. To analyze stimulus-induced changes in chromatin accessibility, we focused the analysis on the majority of cells that clustered by conditions and irrespective of donor origin. The cell barcodes used for the final analysis were deposited together with the raw and processed data. After clustering these cells at higher resolution, we annotated clusters according to a clear enrichment of cells derived from different conditions (Extended Data Fig. 4b). To assess the similarity of the in vitro induced signatures to those defining NK cell subsets ex vivo, the MACS2-called peak set of the ex vivo analysis was quantified for the cells stimulated in vitro, and the two datasets were integrated using anchor-based integration, setting the ex vivo dataset as reference and the in vitro dataset as query[35]. Integration anchors were then used to transfer cluster labels from the ex vivo dataset to the in vitro dataset, and the per-cell prediction scores were plotted for each cluster. To identify which individual chromatin regions overlap between adaptive NK cells ex vivo and the cells stimulated with cytokines and peptide, fold changes for DARs overlapping between the comparisons of CD56^{dim} cells to adaptive NK cells ex vivo and control cells to LFL + IL-12+IL-18 cells in vitro were analyzed for linear correlation.

Preprocessing of scRNA-seq data

Base call files were demultiplexed using cellranger v3.0.2 mkfastq and bcl2fastq v2.20.0.422 into scRNA-seq and ADT/HTO libraries. scRNA-seq reads were mapped to the hg19 reference genome with cellranger count, generating gene-per-cell count tables as output for further analysis. ADT and HTO reads were either quantified with CITE-Seq-Count[68] or pseudoaligned with kallisto[60] to a mismatch map generated with kite and counted with bustools[61].

Downstream analysis of scRNA-seq data

scRNA-seq counts were imported into Seurat[69] and filtered for outliers regarding frequency of mitochondrial transcripts (v2: <5–6%; v3: <10%), total number of transcripts (v2: <6,000; v3: <10,000) and number of genes (v2: >500–800, <2,000–2,500; v3: >1,000, <5,000) per cell. Contaminating erythrocytes and B cells were excluded based on the expression of *HBA1*/*HBA2* and *IGJ*, respectively. ADT/HTO counts were imported, centered log ratio normalized and joined with the scRNA-seq data. Normalized hashtag reads were used for exclusion of doublets and to demultiplex NKG2C$^+$/NKG2C$^-$ populations and donors. Initially, donors were analyzed individually. Counts were normalized with scran[70] and used for principal-component analysis (PCA) on the 2,000 most variable genes. The number of principal components used for UMAP embedding and clustering was chosen based on elbow plots for each individual. After initial high-resolution clustering, a small population of non-NK cells, as judged by their lack of CD56 expression concomitant with high levels of *IL7R*, *GATA3*, *IL2RA* and *CD40LG*, were excluded. The PCA was recalculated on the remaining cells and used as input for UMAP embedding and clustering as before. HCMV$^+$ and HCMV$^-$ donors were then integrated separately based on the variable features overlapping between HCMV$^+$ donors by using anchor-based integration. The integrated data were scaled, regressing out the number of transcripts per cell. The HCMV$^+$/

HCMV$^-$ objects were merged, and the integrated data were rescaled, also regressing out the number of transcripts per cell. A PCA was performed on the shared variable features and was used as input for UMAP embedding and clustering. The normalized, non-integrated RNA assay was used for all differential expression analyses. For differential expression analysis between NKG2C$^+$ and NKG2C$^-$ populations in HCMV$^+$ and HCMV$^-$ donors, cells were downsampled to equal numbers per population and donor group (6,000 cells each), and only genes that were detected with a minimum count of 100 in both objects were included. Further, NKG2C$^+$ NK cells from HCMV$^+$ donors were extracted using hashtag counts, followed by UMAP embedding. Signatures from the integrated analysis were assessed to confirm the cluster identities.

Integration of scRNA-seq and scATAC-seq data

Fully processed and donor-integrated datasets of both modalities were integrated by anchor-based integration as implemented in Seurat[35,69] using the integrated RNA assay as reference and the gene scores calculated from scATAC-seq data as query. Integration anchors were used to impute RNA expression of clusters defined by chromatin accessibility and inspected for consistency of key marker genes. Imputed RNA expression was assessed for correlation with chromatin accessibility to identify links between individual chromatin regions and gene expression.

Statistical analysis and reproducibility

Reported *P* values were corrected for multiple testing. No statistical methods were used to predetermine sample sizes, but our sample sizes are similar to those reported in previous publications[20,22]. Data distribution was assumed to be normal, but this was not formally tested. No randomization was performed. Data collection and analysis were not performed blind to the conditions of the experiments. No data points were excluded.

Reporting summary

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

Data availability

scRNA-seq and scATAC-seq BAM files and processed data as fragment files, gene count table and antibody count tables have been deposited in the Gene Expression Omnibus under accession code GSE197037. Single-nucleotide variants and indels were removed from BAM files as described in the Methods. To enable reproducibility of the clonotype analysis based on mitochondrial mutations, the mgatk results were deposited for each experiment. scATAC-seq data were mapped to the GRCh38 (GCF_000001405.39) reference genome, and scRNA-seq data were mapped to the hg19 (GCF_000001405.25) reference genome, as supplied by 10x Genomics. Source data are provided with this paper.

Code availability

All code has been deposited in GitHub (https://github.com/timorueckert/Clonal_NK).

References

56. Borrego, F., Ulbrecht, M., Weiss, E. H., Coligan, J. E. & Brooks, A. G. Recognition of human histocompatibility leukocyte antigen (HLA)-E complexed with HLA class I signal sequence-derived peptides by CD94/NKG2 confers protection from natural killer cell-mediated lysis. *J. Exp. Med.* **187**, 813–818 (1998).
57. Cossarizza, A. et al. Guidelines for the use of flow cytometry and cell sorting in immunological studies (third edition). *Eur. J. Immunol.* **51**, 2708–3145 (2021).
58. Milush, J. M. et al. Functionally distinct subsets of human NK cells and monocyte/DC-like cells identified by coexpression of CD56, CD7, and CD4. *Blood* **114**, 4823–4831 (2009).
59. Ziegenhain, C. & Sandberg, R. BAMboozle removes genetic variation from human sequence data for open data sharing. *Nat. Commun.* **12**, 6216 (2021).
60. Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34**, 525–527 (2016).
61. Melsted, P. et al. Modular, efficient and constant-memory single-cell RNA-seq preprocessing. *Nat. Biotechnol.* **39**, 813–818 (2021).
62. Stuart, T., Srivastava, A., Madad, S., Lareau, C. A. & Satija, R. Single-cell chromatin state analysis with Signac. *Nat. Methods* **18**, 1333–1341 (2021).
63. Korsunsky, I. et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).
64. Zhang, Y. et al. Model-based analysis of ChIP–seq (MACS). *Genome Biol.* **9**, R137 (2008).
65. Ntranos, V., Yi, L., Melsted, P. & Pachter, L. A discriminative learning approach to differential expression analysis for single-cell RNA-seq. *Nat. Methods* **16**, 163–166 (2019).
66. Fornes, O. et al. JASPAR 2020: update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* **48**, D87–D92 (2020).
67. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
68. Roelli, P., bbimber, Flynn, B., santiagorevale & Gui, G. Hoohm/CITE-seq-Count: 1.4.2. Zenodo https://doi.org/10.5281/zenodo.2590196 (2019).
69. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587 (2021).
70. Lun, A. T. L., McCarthy, D. J. & Marioni, J. C. A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor. *F1000Res.* **5**, 2122 (2016).

Acknowledgements

We thank members of the Romagnani group, G. Gasteiger, D.R. Withers, T. Schüler and T.N. Haschler for critically reading the manuscript. We are grateful for support by the DRFZ flow cytometry core facility for assistance with cell sorting. We thank O. Knight and P. Durek for help with setting up bioinformatic pipelines and the DRFZ lab managers, M. Klemm and G. Guerra, for technical support. This work was funded by the ERC Advanced Grant 'MEM-CLONK' (101055157) to C.R., the Deutsche Forschungsgemeinschaft (DFG) grants SPP 1937 (RO3565/4-2) and SFB TRR241 B02 to C.R., Leibniz-Science Campus Chronic Inflammation, Leibniz-Kooperative Exzellenz K259/2019 to C.R. and Berlin Health Innovations (BHI) Validation Fund to C.R. and T.R. Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

Author contributions

Conceptualization, T.R. and C.R.; methodology, T.R. with input from C.R., L.S.L. and C.A.L.; formal analysis, T.R. and C.A.L.; investigation, T.R.; resources, C.R. and M.-F.M.; writing the original draft, T.R. and C.R. with input from all authors; visualization, T.R. with input from

C.R., L.S.L. and C.A.L.; supervision, C.R.; project administration, C.R.; funding acquisition, C.R.

Funding
Open access funding provided by Deutsches Rheuma-Forschungszentrum Berlin (DRFZ)

Competing interests
T.R. and C.R. are listed as inventors on a patent series covering NKG2C-activating peptides. The other authors declare no competing interests.

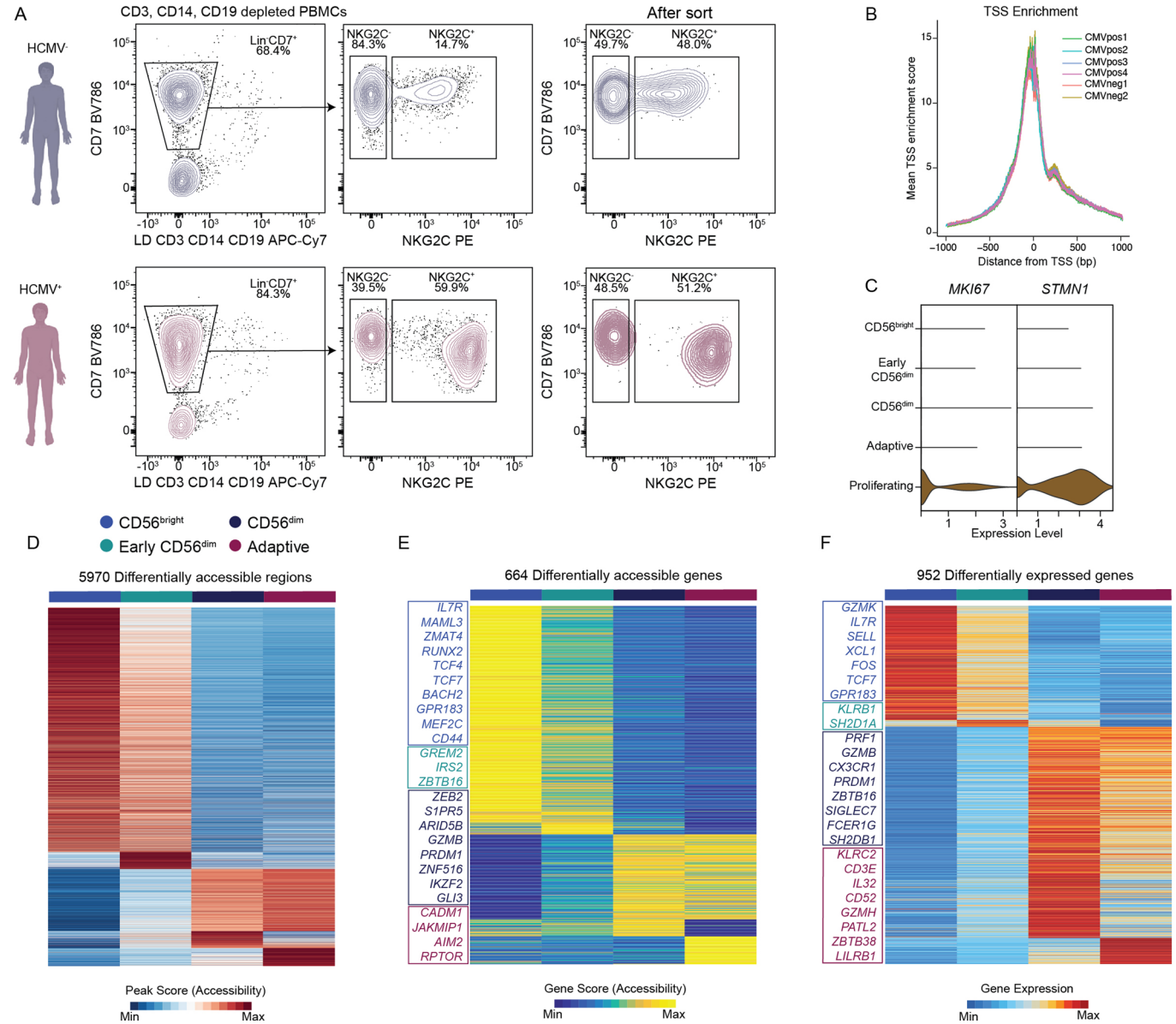
Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41590-022-01327-7.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41590-022-01327-7.

Correspondence and requests for materials should be addressed to Timo Rückert or Chiara Romagnani.

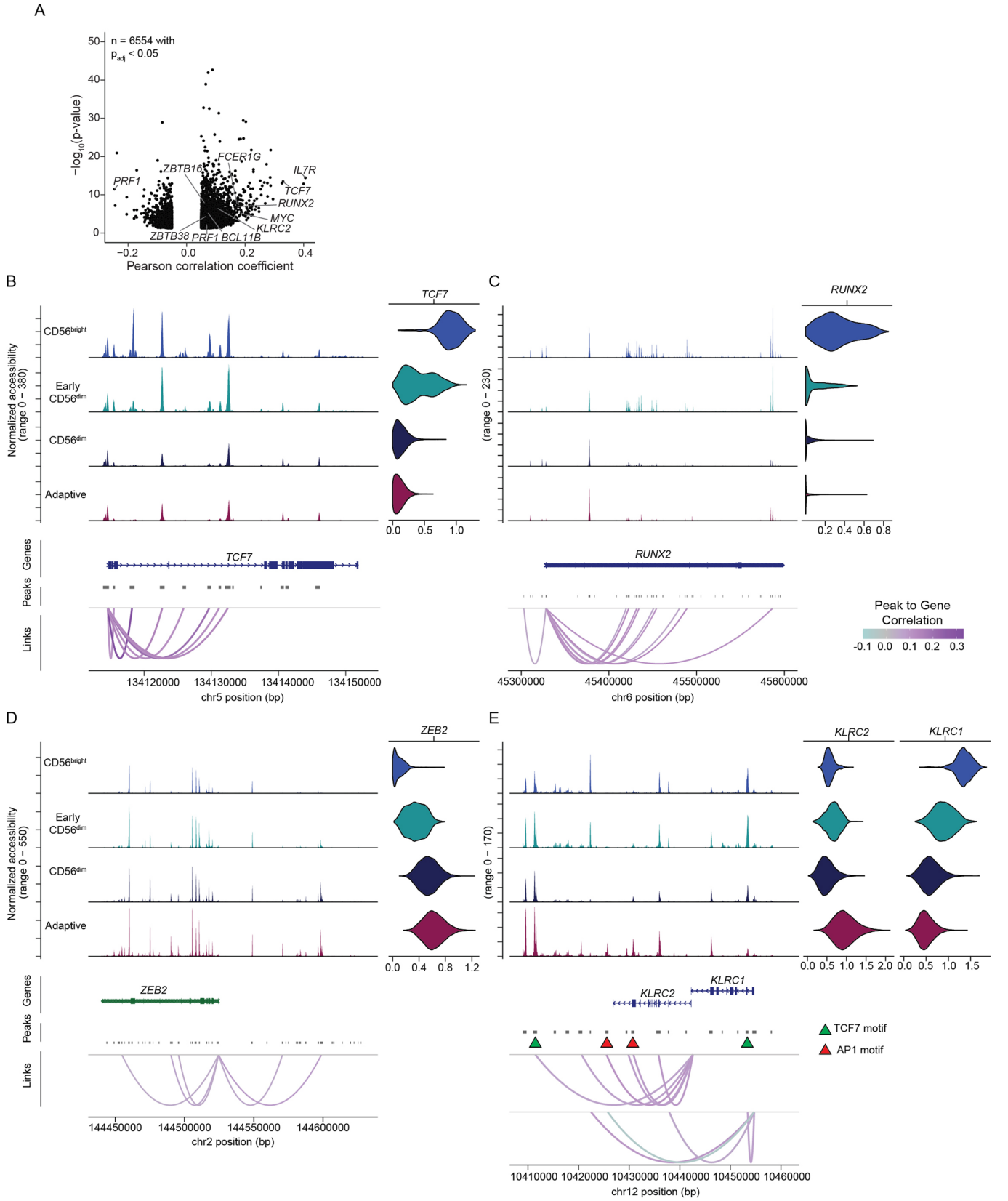
Peer review information *Nature Immunology* thanks the anonymous reviewers for their contribution to the peer review of this work. N. Bernard was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

https://doi.org/10.1038/s41590-022-01327-7



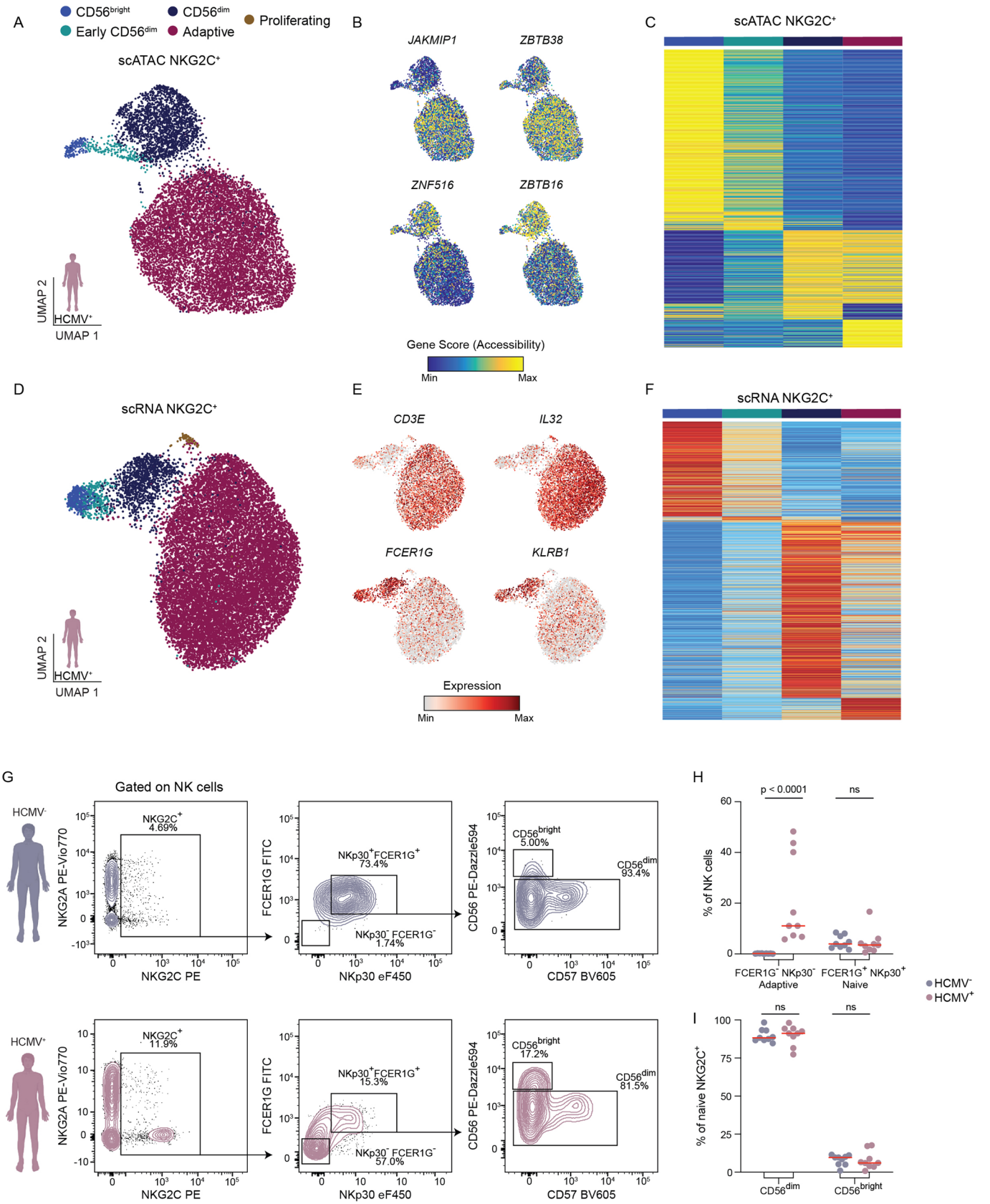
Extended Data Fig. 1 | Mapping NK cell subsets onto transcriptional and epigenetic landscapes. (a) Sorting strategy for one representative HCMV$^-$ and HCMV$^+$ donor, respectively. **(b)** Transcription start site (TSS) enrichment plot overlaid for all donors. **(c)** Expression of proliferation genes *MKI67* and *STMN1*. **(d-f)** Row-scaled differentially accessible regions (D), genes (E) and differentially expressed genes (F) per cluster. Exemplary cluster-associated genes are listed in the respective colors. **(h)** Gene accessibility, imputed expression (violins), and their Pearson correlation (indicated as links) for *RUNX2* and *ZEB2*. Created with BioRender.com.

https://doi.org/10.1038/s41590-022-01327-7



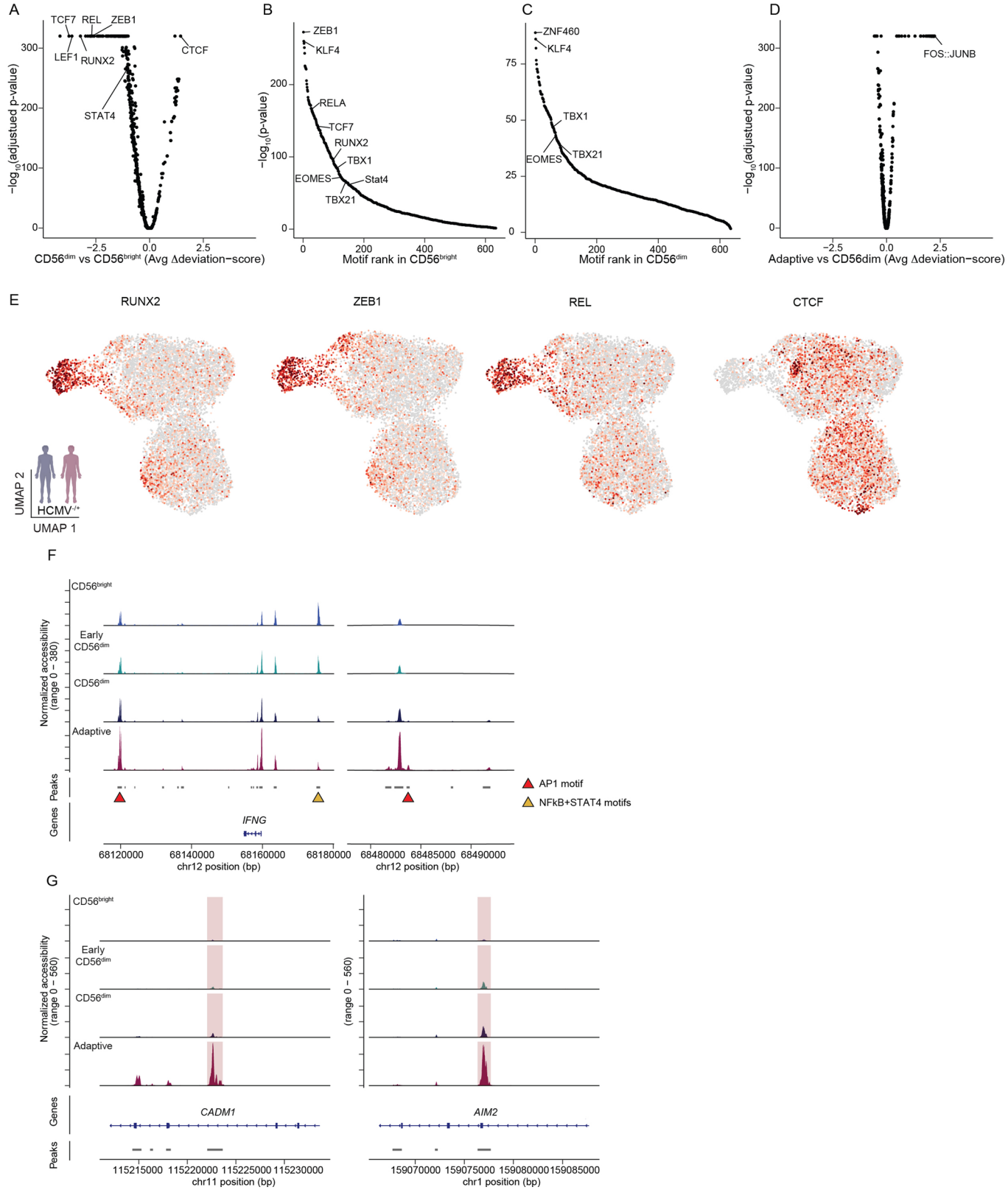
Extended Data Fig. 2 | Integration of chromatin accessibility and gene expression. (a) Pearson correlation between open chromatin regions and gene expression, p-value from one-sided z-test of background-normalized correlation-coefficient as implemented in Signac[62]. **(b-e)** Chromatin accessibility, imputed expression (violins) and their Pearson correlation (indicated as links) for indicated genes *TCF7* (B), *RUNX2* (C), *ZEB2* (D), *KLRC2* and *KLRC1* (E).

https://doi.org/10.1038/s41590-022-01327-7



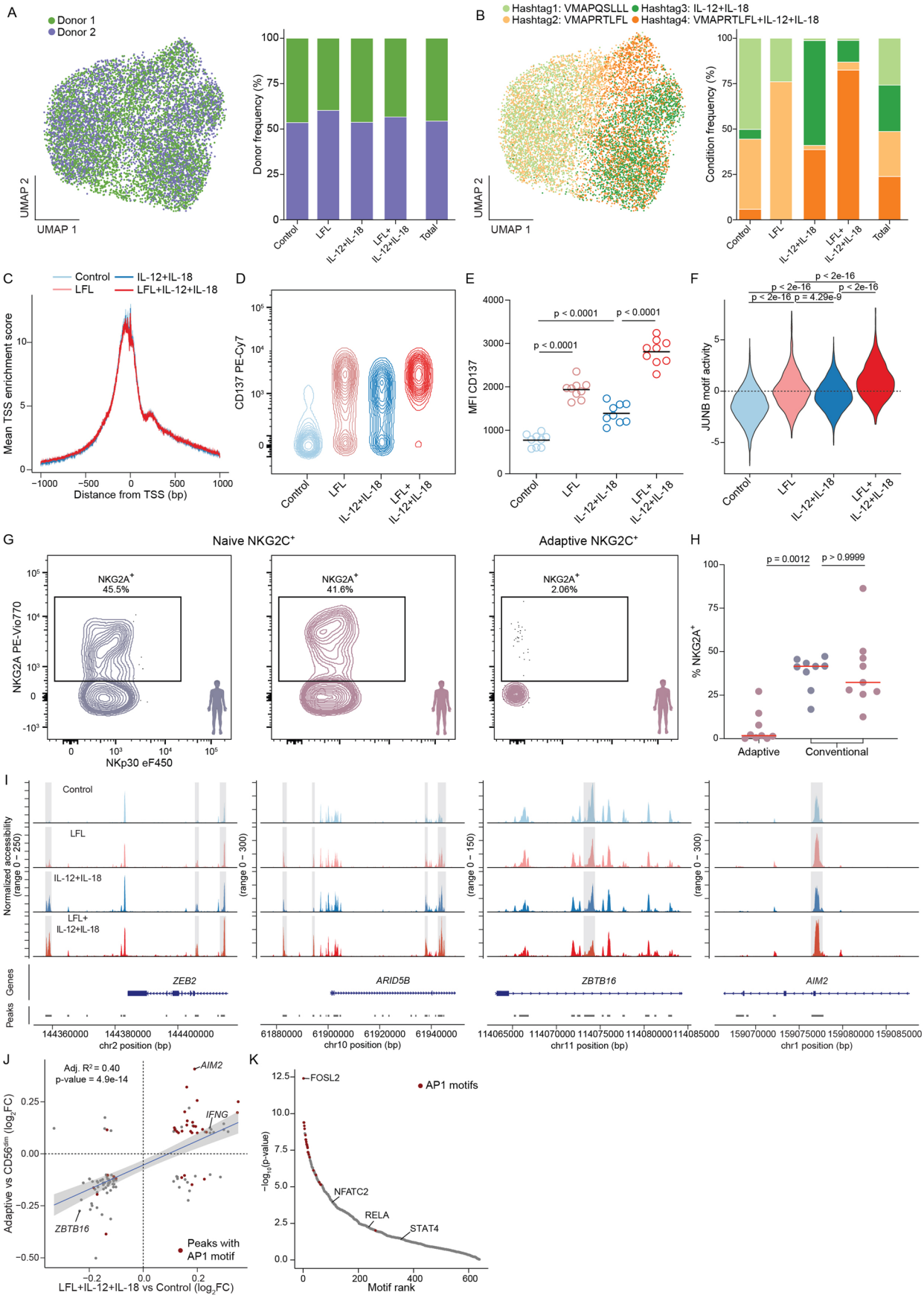
Extended Data Fig. 3 | Distinct signatures dissect NKG2C⁺ NK cells into naïve and adaptive. (a + d) UMAP embedding of scATAC-seq (A) and scRNA-seq (D) of NKG2C⁺ NK cells from HCMV⁺ individuals. **(b + e)** Accessibility (B) and expression (E) of key genes regulated in adaptive NK cells. **(c + f)** Row-scaled accessibility (C) and expression (F) of NK cell subset-specific signatures from fully integrated datasets (Fig. 1) in NKG2C⁺ NK cell clusters from HCMV⁺ donors. **(g–i)** Representative expression (G) and frequencies (H) of FCER1G⁺ NKp30⁺ ('Naive') and FCER1G⁻ NKp30⁻ ('Adaptive') NKG2C⁺ NK cells and distribution of naïve NKG2C⁺ NK cells into CD56^{bright} and CD56^{dim} subsets (I) in HCMV⁻ (n = 9) and HCMV⁺ (n = 9) donors. Kruskal-Wallis test with Dunn's post-hoc test. Created with BioRender.com.

https://doi.org/10.1038/s41590-022-01327-7



Extended Data Fig. 4 | HCMV infection leaves a persistent chromatin footprint enriched for AP1 motifs. (a) Differentially active motifs between CD56^{dim} and CD56^{bright} NK cells by Wilcoxon rank sum test. **(b-c)** Motif enrichment in chromatin regions specifically accessible in CD56^{bright} (B) and CD56^{dim} NK cells as determined by hypergeometric test. **(d)** Differentially active motifs between adaptive and CD56^{dim} NK cells by Wilcoxon rank sum test. **(e)** Activity of the respective TF motifs projected onto UMAP embedding. **(f)** Accessibility of *IFNG*-associated open-chromatin regions containing AP1 or STAT4 and NFκB motifs. **(g)** Per cluster accessibility of AP1-motif containing chromatin regions in the proximity of *CADM1* and *AIM2*. Created with BioRender.com.

https://doi.org/10.1038/s41590-022-01327-7

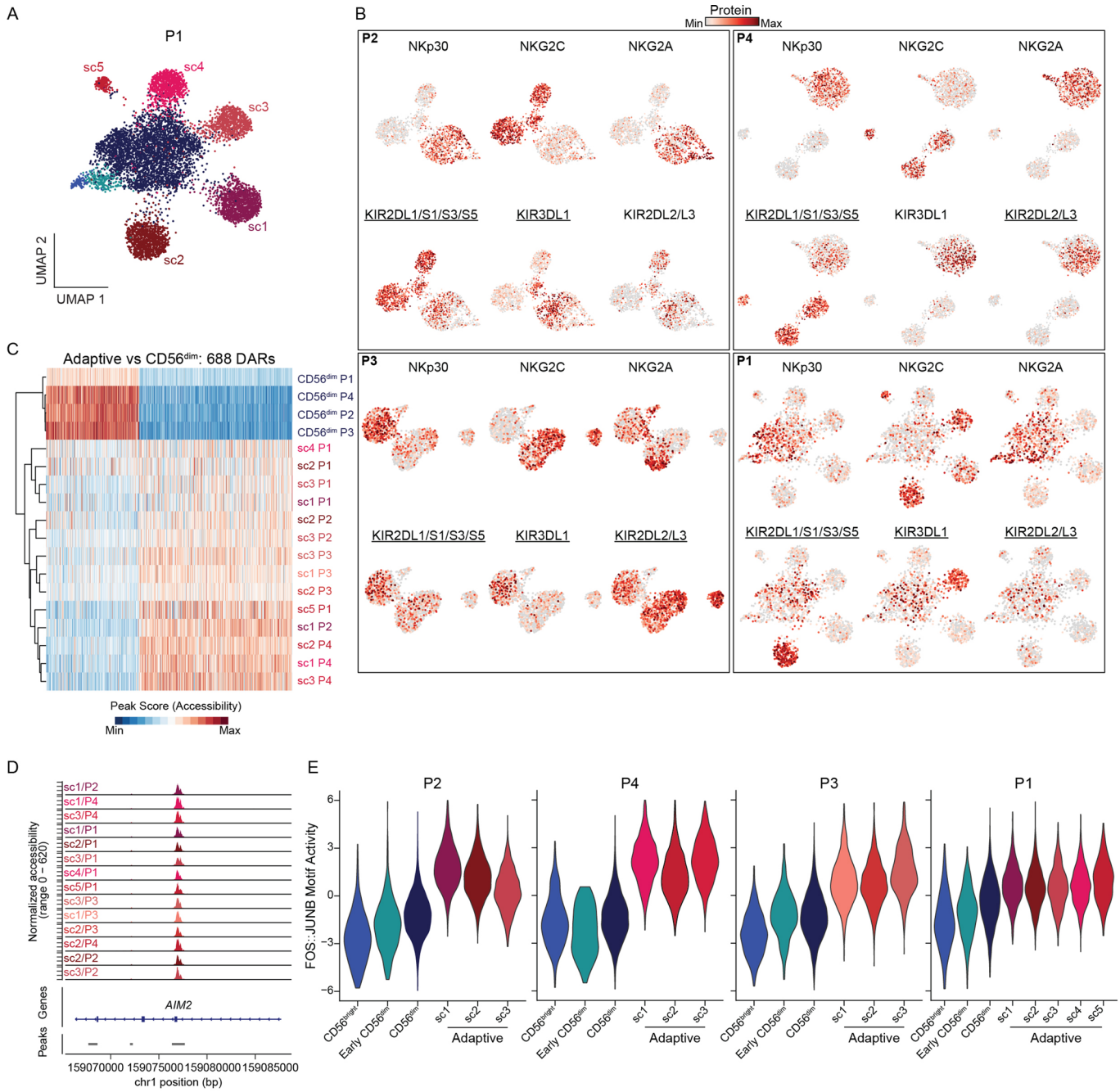


Extended Data Fig. 5 | See next page for caption.

https://doi.org/10.1038/s41590-022-01327-7

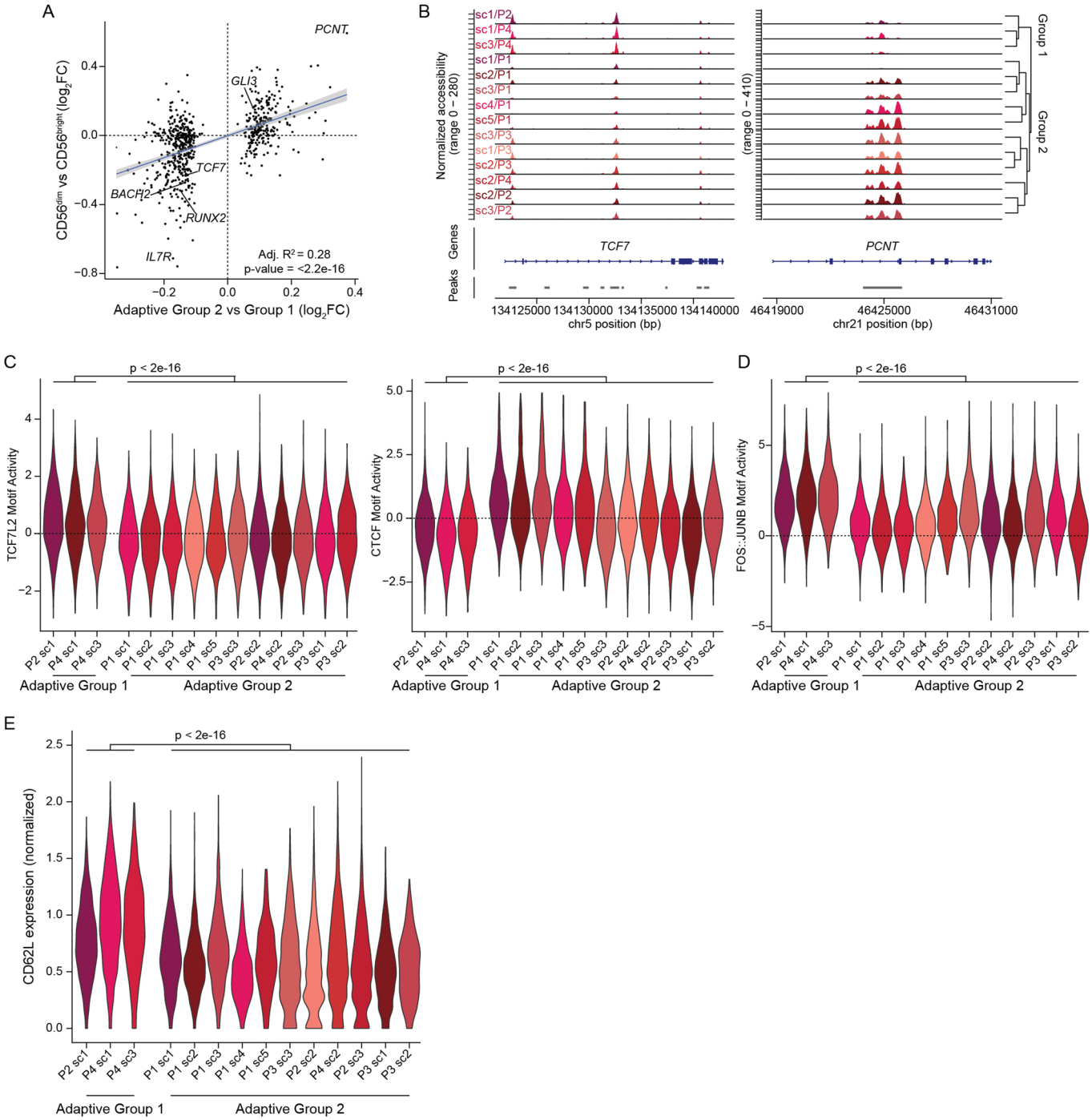
Extended Data Fig. 5 | Synergistic imprinting by HCMV peptides and pro-inflammatory cytokines. (a) Distribution of cells colored by donor origin and quantification per cluster. **(b)** Distribution of cells colored by culture condition and quantification per cluster, utilized for cluster annotation. **(c)** Transcription start site (TSS) enrichment plot overlaid for all clusters. **(d** + **e)** Representative (D) and mean (E) CD137 expression on NKG2C$^+$ NKG2A$^-$ NK cells from HCMV$^-$ donors (n = 9) cultured under the indicated conditions; One-way ANOVA with Holm-Šídák adjustment. **(f)** JUNB motif activity per cluster; Wilcoxon rank sum test with Benjamini-Hochberg adjustment. **(g-h)** Representative (G) and quantification (H) of NKG2A expression in naïve and adaptive NKG2C$^+$ NK cells from HCMV$^-$ (n = 9) and HCMV$^+$ (n = 9) donors; see Extended Data Fig. 2g for gating strategy; Kruskal-Wallis test with Dunn's post-hoc test. **(i)** Accessibility of regions near key adaptive NK cell-related genes regulated after stimulation. **(j)** Linear correlation between differentially accessible regions both induced by *in vitro* activation and observed *ex vivo* between adaptive and CD56^{dim} NK cells; error bands show 95 % confidence interval, p-value from two-sided F-test. **(k)** Motif enrichment in open chromatin regions displayed in (J) as determined by one-sided hypergeometric test with Benjamini-Hochberg adjustment. Created with BioRender.com.

https://doi.org/10.1038/s41590-022-01327-7



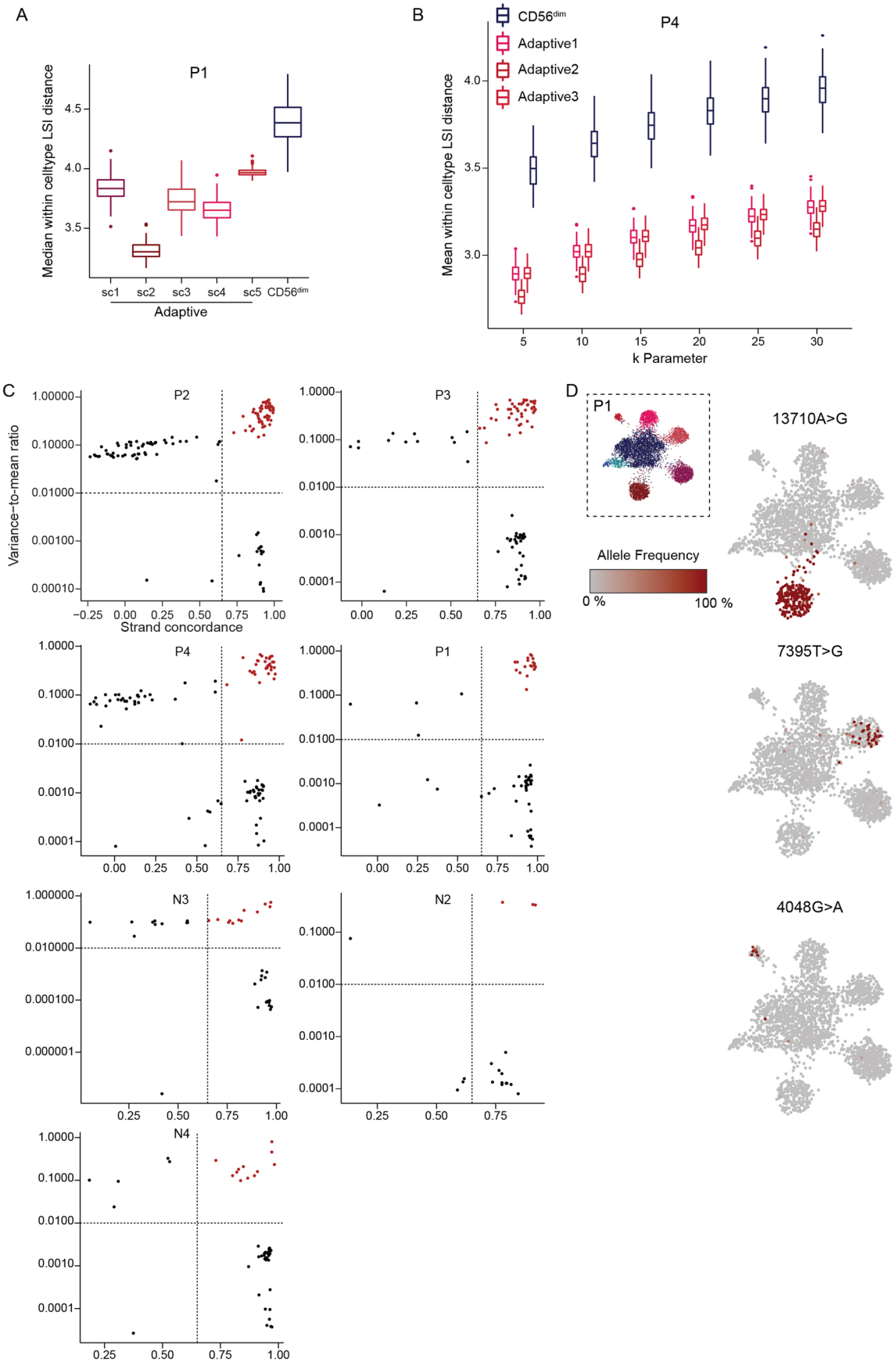
Extended Data Fig. 6 | Convergent and divergent epigenetic features of adaptive NK cells. (a) UMAP embedding of NK cells from HCMV^{+} individual P1 analyzed by scATAC-seq. **(b)** Surface expression of the indicated proteins projected on the UMAP embeddings for each donor. Educating KIRs are underlined, see Supplementary Table 2. **(c)** Column-scaled accessibility of differentially accessible regions (DARs) shared between at least three donors comparing the total adaptive and CD56^{dim} compartments. **(d)** Accessibility of adaptive NK-cell defining *AIM2* region within each donor and adaptive subcluster. **(e)** FOS::JUNB motif activity within each donor. **(f)** Accessibility of representative subcluster-specific chromatin regions for each donor.

https://doi.org/10.1038/s41590-022-01327-7



Extended Data Fig. 7 | Adaptive subcluster groups are defined by a maturation gradient. (a) Linear correlation between DARs comparing the two adaptive subcluster groups and CD56^{dim} to early CD56^{dim} NK cells; error bands show 95 % confidence interval, p-value from two-sided F-test. **(b)** Accessibility of adaptive group-defining regions within *TCF7* and *PCNT* for each donor and adaptive subcluster. **(c)** TCF7L2 and CTCF motif activity for each donor and adaptive subcluster; two-sided Wilcoxon test with Bonferroni adjustment. **(d)** FOS::JUNB motif activity for each donor and adaptive subcluster; two-sided Wilcoxon test with Bonferroni adjustment. **(e)** Normalized CD62L surface expression per cluster; two-sided Wilcoxon rank sum test with Bonferroni adjustment.

https://doi.org/10.1038/s41590-022-01327-7

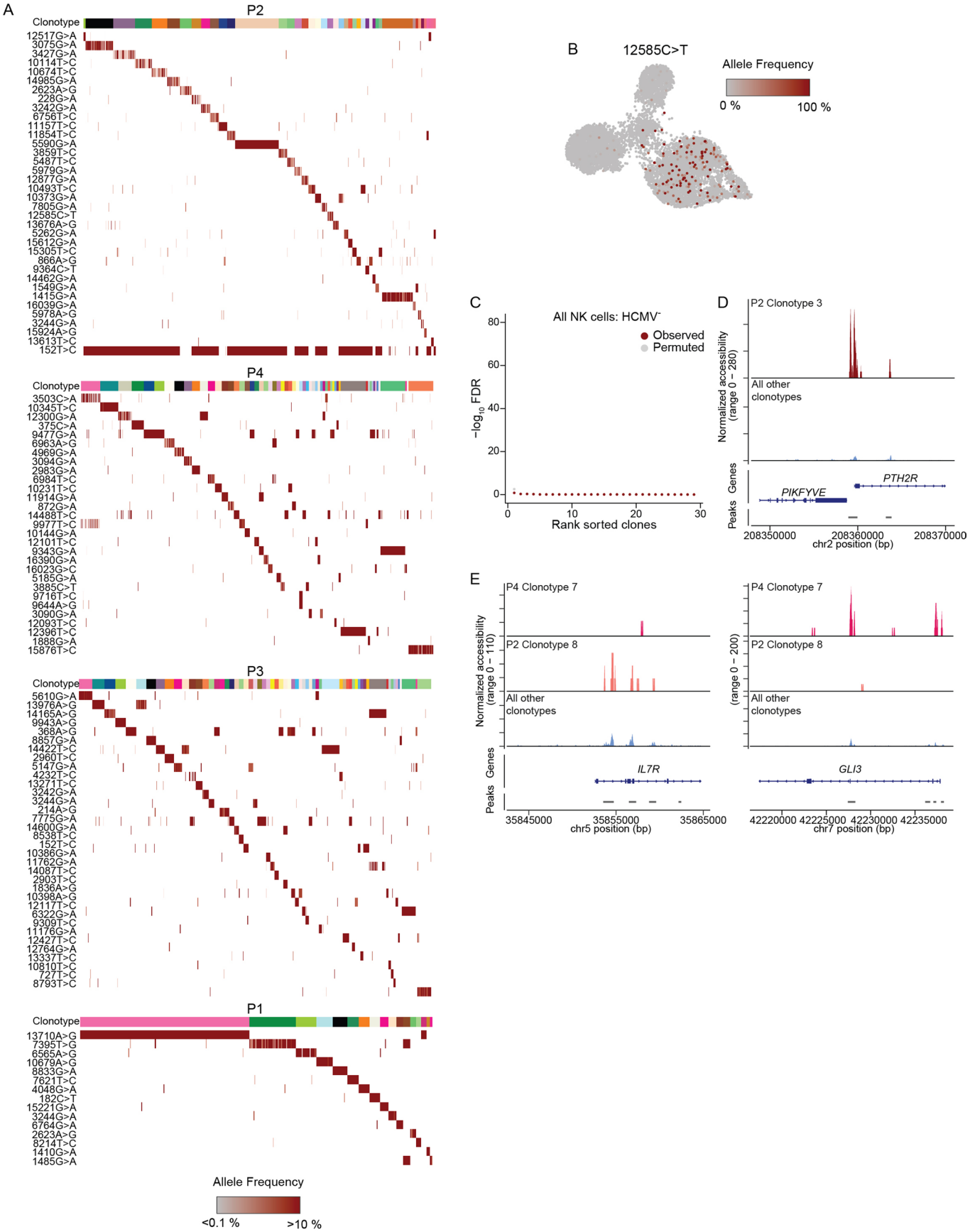


Extended Data Fig. 8 | See next page for caption.

https://doi.org/10.1038/s41590-022-01327-7

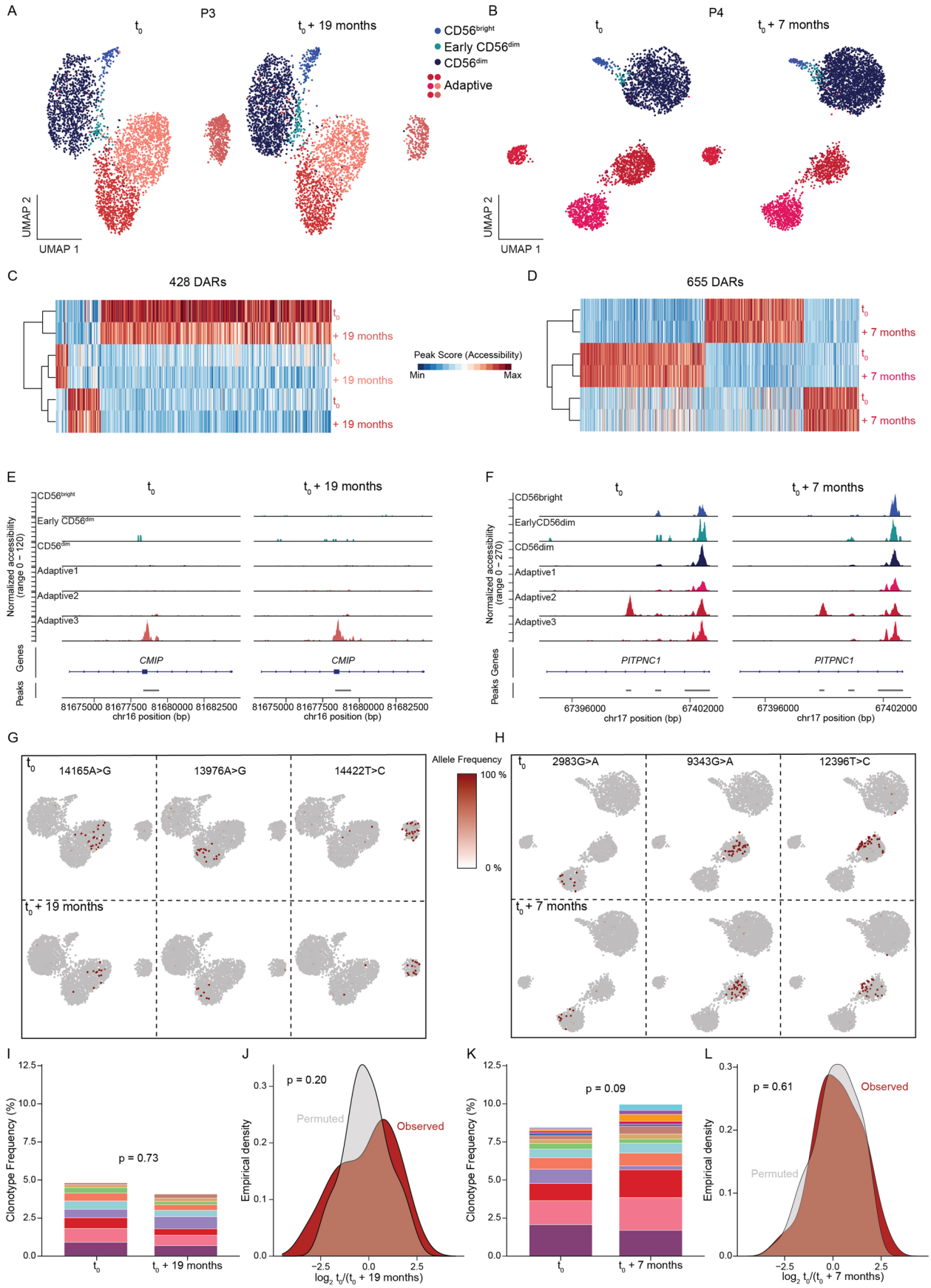
Extended Data Fig. 8 | Clonal expansion underlies divergent epigenetic signatures of adaptive NKG2C$^+$ NK cells. (a) Cluster heterogeneity of donor P1 NK cells as assessed by measuring the median distance of 200 randomly sampled cells to their k nearest neighbors and repeating this process 100 times. **(b)** The k parameter was varied from 5-30 for donor P4 to assess robustness of this heterogeneity metric. The upper and lower hinges of boxplots correspond to the first and third quartiles, respectively. The upper and lower whiskers extend to the largest/smallest value no further than 1.5 times the interquartile range from the hinges. Outliers beyond whiskers are displayed as individual points. **(c)** Strand-concordance and variance-to-mean ratio applied for extraction of informative mutations for each donor. **(d)** Allele frequency of representative somatic mtDNA mutations projected onto UMAP embedding for donor P1.

https://doi.org/10.1038/s41590-022-01327-7



Extended Data Fig. 9 | Adaptive clonotypes are associated to unique epigenetic identities. (A) Clonotypes defined by clustering on per-cell allele frequency for all high-confidence variants. **(b)** Allele frequency of representative somatic mtDNA mutation within the conventional compartment of donor P2. **(c)** Association of clonotypes to clusters defined by chromatin accessibility for all NK cells from HCMV⁻ donors (n = 3). **(d-e)** Representative open chromatin regions specifically associated to individual clonotypes.

https://doi.org/10.1038/s41590-022-01327-7



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

https://doi.org/10.1038/s41590-022-01327-7

Extended Data Fig. 10 | Adaptive NK cell clonotypes are stably maintained over time. (a) UMAP embedding of NK cells from HCMV$^+$ donor P3 and P4 analyzed by scATAC-seq at two different time points. **(c-f)** Stability of overall (C-D) and representative (E-F) subcluster-defining differentially accessible regions (DARs) over time; column-scaled. **(g-h)** Representative clonotype-defining mutations projected onto UMAP embeddings at the two time points. **(i + k)** Clonotype frequency of adaptive NK cell clonotypes within total adaptive NK cell compartment over time; Fisher's exact test with Monte Carlo simulation. **(j + l)** Observed and permuted distribution of clonotype $\log_2$ fold-changes between time points; two-sided Kolmogorov-Smirnov test.

nature portfolio

Corresponding author(s): Chiara Romagnani, Timo Rückert

Last updated by author(s): Aug 27, 2022

Reporting Summary

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

Statistics

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

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

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

Software and code

Policy information about availability of computer code

Data collection	For in vitro cultures NK cells were sorted using a FACS Aria II (BD Biosciences). Flow cytometry data was acquired using LSR Fortessa (BD Biosciences). Sequencing libraries were sequenced on a NextSeq500 or NovaSeq6000 sequencer (Illumina).
Data analysis	Flow cytometry data was analyzed using FlowJo v 10.7.1. Statistical analysis was performed GraphPad Prism v8.4.3. Analysis of sequencing data was performed as detailed in the methods section, using the following software packages: Python 3.7.9 cellranger-atac v1.2.0 cellranger v3.0.2 ASAP to kite v2 kallisto v0.46.0 bustools v0.39.2 bamboozle v0.5.0 CITE-seq count v1.4.1 mgatk v0.5.9 R v4.1.0 Seurat v4.0.6 Signac v1.3.0 Harmony v0.1.0 chromVAR 1.14.0 tidyheatmap v0.0.0.9000 pheatmap v1.0.12 eulerr v6.1.0 Homer v4.11

Analysis code was deposited at GitHub: https://github.com/timorueckert/Clonal_NK

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

Data

Policy information about availability of data

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

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

scRNA and scATAC-seq bam-files, as well as processed data as fragment files, gene count table and antibody count tables have been deposited at GEO (accession number GSE197037) and are publicly available as of the date of publication. Single-nucleotide variants and indels were removed from bam files as described in the detailed methods section. To enable reproducibility of the clonotype analysis based on mitochondrial mutations, the mgatk results were deposited for each experiment. scATAC-seq data were mapped to the GRCh38 (GCF_000001405.39), scRNA-seq data to hg19 (GCF_000001405.25) as supplied by 10x genomics.

Field-specific reporting

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

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

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

Life sciences study design

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

Sample size	No sample size calculation was performed but our sample sizes are similar to those reported in previous publications. For the scATAC- and scRNA-seq, we decided to include more HCMV+ donors as the focus of the study was on adaptive NK cells that were specifically present in HCMV+ donors whereas conventional NK cells were well represented in both HCMV+ and HCMV- donors.
Data exclusions	No data was excluded.
Replication	scATAC-seq ex vivo data were collected in 6 independent experiments. scRNA-seq ex vivo data were collected in four independent experiments. In vitro stimulation followed by scATAC-seq is from one experiment with two indepedently stimulated donors, read out by flow cytometry was performed in three independent experiments with three donors each, all of which successfully replicated the results.
Randomization	Donors were allocated into groups based on HCMV-serostatus. No additional group allocation or randomization was performed.
Blinding	Investigators were not blinded to group allocation. The same unbiased analysis strategy was applied for all comparative analysis of the HCMV+ and HCMV- groups, including equal enrichment of NKG2C+ and NKG2C- cells.

Reporting for specific materials, systems and methods

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

Materials & experimental systems

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

Methods

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

Antibodies

Antibodies used: TotalSeq™-A0084 anti-human CD56 (NCAM) Biolegend AB_2734445 (BioLegend Cat#392421) 1:200

Antibodies used

TotalSeq™-A0083 anti-human CD16 Biolegend AB_2734255 (BioLegend Cat#302061) 1:500
TotalSeq™-A0436 anti-Biotin Biolegend AB_2801086 (BioLegend Cat#409008) 1:100
TotalSeq™-A0147 anti-human CD62L Biolegend AB_2750365 (BioLegend Cat#304847) 1:100
TotalSeq™-A0168 anti-human CD57 Biolegend AB_2810588 (BioLegend Cat#393319) 1:100
TotalSeq™-A0367 anti-human CD2 Biolegend AB_2783172 (BioLegend Cat#309229) 1:1000
TotalSeq™-A0801 anti-human CD337 (NKp30) Biolegend AB_2800852 (BioLegend Cat#325221) 1:100
TotalSeq™-A0149 anti-human CD161 Biolegend AB_2749998 (BioLegend Cat#339945) 1:100
TotalSeq™-A0420 anti-human CD158 (KIR2DL1/S1/S3/S5) Biolegend AB_2800901 (BioLegend Cat#339515) 1:100
TotalSeq™-A0592 anti-human CD158b (KIR2DL2/L3, NKAT2) Biolegend AB_2800818 (BioLegend Cat#312615) 1:100
TotalSeq™-A0599 anti-human CD158e1 (KIR3DL1, NKB1) Biolegend AB_2800819 (BioLegend Cat#312723) 1:100
TotalSeq™-A0390 anti-human CD127 (IL-7Rα) Biolegend AB_2734366 (BioLegend Cat#351352) 1:400
TotalSeq™-A0902 anti-human CD328 (Siglec-7) Biolegend AB_2832662 (BioLegend Cat#339217) 1:1000
TotalSeq™-A0867 anti-human CD94 Biolegend AB_2814142 (BioLegend Cat#305521) 1:100
TotalSeq™-A0896 anti-human CD85j (ILT2) Biolegend AB_2814225 (BioLegend Cat#333723) 1:100
TotalSeq™-A0911 anti-phycoerythrin (PE) Biolegend AB_2820078 (BioLegend Cat#408109) 1:100
TotalSeq™-A0250 anti-mouse/human KLRG1 (MAFA) Biolegend AB_2800648 (BioLegend Cat#138431) 1:100
TotalSeq™-A0061 anti-human CD117 (c-kit) Biolegend AB_2734287 (BioLegend Cat#313241) 1:100
TotalSeq™-A1018 anti-human HLA-DR, DP, DQ Biolegend AB_2832712 (BioLegend Cat#361717) 1:400
TotalSeq™-A0355 anti-human CD137 (4-1BB) Biolegend AB_2783173 (BioLegend Cat#309835) 1:100
TotalSeq™-A0152 anti-human CD223 (LAG-3) Biolegend AB_2749999 (BioLegend Cat#369333) 1:100
TotalSeq™-A0366 anti-human CD184 (CXCR4) Biolegend AB_2800790 (BioLegend Cat#306531) 1:100
TotalSeq™-A0251 anti-human Hashtag 1 Biolegend AB_2750015 (BioLegend Cat#394601) 1:200-1:400
TotalSeq™-A0252 anti-human Hashtag 2 Biolegend AB_2750016 (BioLegend Cat#394603) 1:200-1:400
TotalSeq™-A0253 anti-human Hashtag 3 Biolegend AB_2750017 (BioLegend Cat#394605) 1:200-1:400
TotalSeq™-A0254 anti-human Hashtag 4 Biolegend AB_2750018 (BioLegend Cat#394607) 1:200-1:400
TotalSeq™-A0255 anti-human Hashtag 5 Biolegend AB_2750019 (BioLegend Cat#394609) 1:200-1:400
TotalSeq™-A0256 anti-human Hashtag 6 Biolegend AB_2750020 (BioLegend Cat#394611) 1:200-1:400
TotalSeq™-A0257 anti-human Hashtag 7 Biolegend AB_2750021 (BioLegend Cat#394613) 1:200-1:400
TotalSeq™-A0258 anti-human Hashtag 8 Biolegend AB_2750022 (BioLegend Cat#394615) 1:200-1:400
TotalSeq™-A0259 anti-human Hashtag 9 Biolegend AB_2750023 (BioLegend Cat#394617) 1:200-1:400
TotalSeq™-A0260 anti-human Hashtag 10 Biolegend AB_2750024 (BioLegend Cat#394619) 1:200-1:400
TotalSeq™-A0262 anti-human Hashtag 12 Biolegend AB_2750025 (BioLegend Cat#394623) 1:200-1:400
TotalSeq™-A0263 anti-human Hashtag 13 Biolegend AB_2750026 (BioLegend Cat#394625) 1:200-1:400
TotalSeq™-A0264 anti-human Hashtag 14 Biolegend AB_2750027 (BioLegend Cat#394627) 1:200-1:400
TotalSeq™-A0265 anti-human Hashtag 15 Biolegend AB_2750028 (BioLegend Cat#394629) 1:200-1:400
PE/Dazzle™ 594 anti-human CD56 Biolegend AB_2563564 (BioLegend Cat#318348) 1:200
PE/Cyanine7 anti-human CD137 (4-1BB) Biolegend AB_2207741 (BioLegend Cat#309818)
Brilliant Violet 605™ anti-human CD57 Recombinant Antibody AB_2728426 (BioLegend Cat. No. 393304) 1:25
BV786 Mouse Anti-Human CD7 BD Biosciences AB_2740589 (BD Biosciences Cat#740964) 1:25
CD159a (NKG2A) Antibody, anti-human, Biotin, REAfinity Miltenyi Biotec AB_2783969 (Miltenyi Biotec Cat#130-114-090) 1:50
CD159c (NKG2C) Antibody, anti-human, PE, REAfinity™ Miltenyi Biotec AB_2751866 (Miltenyi Biotec Cat#130-119-814) 1:100
CD159a (NKG2A) Antibody, anti-human, PE-Vio770, REAfinity Miltenyi Biotec AB_2655388 (Miltenyi Biotec Cat#130-105-647) 1:50
CD3 Monoclonal Antibody (SK7), APC-eFluor 780 Invitrogen AB_10717514 (ThermoFisher Cat#47-0036-42) 1:50
CD14 Monoclonal Antibody (61D3), APC-eFluor 780 Invitrogen AB_1834358 (Cat#47-0149-42) 1:50
CD19 Monoclonal Antibody (HIB19), APC-eFluor 780, Invitrogen AB_1582230 (Cat#47-0199-42) 1:50
Anti-FcεRI Antibody, γ subunit-FITC Merck Cat#FCABS400F 1:50
CD337 (NKp30) Monoclonal Antibody (AF29-4D12), eFluor 450 ThermoFisher AB_2574058 (Cat#48-3379-42) 1:25
BUV737 Mouse Anti-Human CD56 (NCAM-1) BD Biosciences AB_2871176 (BD Biosciences Cat#741842) 1:50
BUV805 Mouse Anti-Human CD3 BD Biosciences (BD Biosciences Cat#612896) 1:50
BUV496 Mouse Anti-Human CD16 BD Biosciences AB_2870224 (BD Biosciences Cat#612944) 1:50
BV421 Mouse Anti-Human CD337 (NKp30) BD Biosciences AB_2738171 (BD Biosciences Cat#563385) 1:25
BUV395 Streptavidin BD Biosciences AB_2869553 (BD Biosciences Cat#564176) 1:100

Validation

All purchased antibodies were validated by their manufacturers and further in-house testing.
- Miltenyi Biotec, https://www.miltenyibiotec.com/DE-en/lp/antibody-validation-improved-reproducibility.html
Three pillars of antibody validation: 1. Antibody reproducibility and consistency (Pure antibody products & Lot-to-lot consistent performance); 2. Antibody specificity (Epitope competition assay,Knockout validation via targeted genome editing & RNAi knockdown); 3. Antibody sensitivity (Functional testing of every product prior to release, Performance comparison & Compatibility with fixation).

- BioLegend, https://www.biolegend.com/en-us/quality/quality-control
Flow Cytometry Reagents
Specificity testing of 1-3 target cell types with either single- or multi-color analysis (including positive and negative cell types). Once specificity is confirmed, each new lot must perform with similar intensity to the in-date reference lot. Brightness (MFI) is evaluated from both positive and negative populations.
Each lot product is validated by QC testing with a series of titration dilutions.
TotalSeq™ Antibodies
Bulk lots are tested by PCR and sequencing to confirm the oligonucleotide barcodes. They are also tested by flow cytometry to ensure the antibodies recognize the proper cell populations.
Bottled lots are tested by PCR and sequencing to confirm the oligonucleotide barcodes.

- BD Biosciences, https://www.biocompare.com/Antibody-Manufacturing/355107-Antibody-Manufacturing-Perspectives-BDBioscience/
We conduct quality control (QC) testing in primary model systems to ensure biological accuracy in an ISO 9001 certified facility. BD carefully selects and characterizes antibody content in product development and tests in relevant primary model systems to ensure

biological accuracy. BD conducts rigorous QC testing of each antibody lot tested side-by-side with a previously produced lot as reference. Our product development process includes testing on a combination of primary cells, cell lines and/or transfectant cell models with relevant controls using multiple immunoassays to ensure biological accuracy. We also perform multiplexing with additional antibodies to interrogate antibody staining in multiple cell populations. BD believes antibody validation is critical to ensure accurate scientific results. Both the consumer and the reagent provider share the responsibility for reproducible science.

- Invitrogen, https://www.thermofisher.com/de/de/home/life-science/antibodies/invitrogen-antibody-validation.html
To help ensure superior antibody results, we've expanded our specificity testing methodology using a 2-part approach for advanced verification. Part 1—Target specificity verification. This helps ensure the antibody will bind to the correct target. Our antibodies are being tested using at least 1 of the following methods to ensure proper functionality in researcher's experiments.
Part 1—Target specificity verification
This helps ensure the antibody will bind to the correct target. Our antibodies are being tested using at least one of the following methods to ensure proper functionality in researcher's experiments.

Knockout—expression testing using CRISPR-Cas9 cell models
Knockdown—expression testing using RNAi to knockdown gene of interest
Independent antibody verification (IAV)—measurement of target expression is performed using two differentially raised antibodies recognizing the same protein target
Cell treatment—detecting downstream events following cell treatment
Relative expression—using naturally occurring variable expression to confirm specificity
Neutralization—functional blocking of protein activity by antibody binding
Peptide array—using arrays to test reactivity against known protein modifications
SNAP-ChIP™—using SNAP-ChIP to test reactivity against known protein modifications
Immunoprecipitation-Mass Spectrometry (IP-MS)—testing using immunoprecipitation followed by mass spectrometry to identify antibody targets
Part 2—Functional application validation.
These tests help ensure the antibody works in a particular application(s) of interest, which may include
(but are not limited to):

Western blotting
Flow cytometry
ChIP
Immunofluorescence imaging
Immunohistochemistry

-Merck
Evaluated by flow cytometry using RBL cells.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	RMA-S/HLA-E were provided by J. Coligan, National Institutes of Health
Authentication	Expression of HLA-E was frequently validated by flow cytometry, no further authentication was performed.
Mycoplasma contamination	RMA-S/HLA-E were tested negative for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used in the study.

Human research participants

Policy information about studies involving human research participants

Population characteristics	5 HCMV+ and 4 HCMV- healthy blood donors were included in the sequencing study, age and gender are listed in Supplementary Table. 1.
Recruitment	Healthy participants were initially randomly recruited for screening of HCMV serostatus and presence of NKG2C+ NK cell expansions. HCMV+ donors were selected based on the presence of NKG2C+ NK cell expansions as judged by flow cytometry, which might have biased the donor selection towards donors with macroscopically visible adaptive NK cell expansions.
Ethics oversight	The Charité ethics committee approved the study (EA4/196/18 and EA4/059/17).

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

Flow Cytometry

Plots

Confirm that:

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

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

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

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

Methodology

Sample preparation	PBMCs were isolated from freshly drawn peripheral blood of healthy donors or from buffy coats obtained from DRK Blutspendedienst Nord-Ost, Dresden, Germany by density-gradient centrifugation (Ficoll Paque Plus, GE Healthcare) and either processed immediately or cryopreserved in FBS containing 10 % DMSO. Dead cells were excluded using Fixable Viability Dye eFluor780 (ThermoFisher), or Zombie Aqua Fixable Viability Kit (BioLegend). Cells were stained with fluorochrome conjugated antibodies for 15 min at room temperature, for sorting, cells were stained at 4 °C. Staining with nucleotide-labeled antibodies was performed for 30 min at 4 °C. For sequencing, NK cells were pre-enriched by magnetic depletion of cells expressing CD3, CD14 or CD19 (Miltenyi Biotec). For in vitro stimulation analyzed by flow cytometry, NK cells were pre-enriched with CD56 microbeads (Miltenyi Biotec).
Instrument	LSR Fortessa, FACSAria II (both BD Biosciences)
Software	FlowJo v 10.7.1
Cell population abundance	Purity of the sorted cell population was analyzed in a post-sort reanalysis and was consistently >90%. NKG2C+ and NKG2C- cells were equally enriched for HCMV+/- donor groups.
Gating strategy	For sequencing, NK cells were sorted as CD3- CD14- CD19- CD7+ NKG2C+/- cells. For in vitro stimulation and measurement by flow cytometry, NK cells were sorted as CD56+ CD3- from CD56-enriched fractions.

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