Published in final edited form as:

Nat Immunol. 2018 September ; 19(9): 963–972. doi:10.1038/s41590-018-0176-1.

Epigenetic control of innate and adaptive immune memory

Colleen M. Lau¹, Nicholas M. Adams¹, Clair D. Geary¹, Orr-El Weizman¹, Moritz Rapp¹, Yuri Pritykin², Christina S. Leslie², and Joseph C. Sun^{1,3,*}

¹Immunology Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA.

²Computational and Systems Biology Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA.

³Department of Immunology and Microbial Pathogenesis, Weill Cornell Medical College, New York, NY, USA.

Abstract

Clonal expansion and immunological memory are hallmark features of the mammalian adaptive immune response and essential for prolonged host control of pathogens. Recent work demonstrates that natural killer (NK) cells of the innate immune system also exhibit these adaptive traits during infection. Here we demonstrate that differentiating and ‘memory’ NK cells possess distinct chromatin accessibility states and that their epigenetic profiles reveal a ‘poised’ regulatory program at the memory stage. Furthermore, we elucidate how individual STAT transcription factors differentially control epigenetic and transcriptional states early during infection. Finally, concurrent chromatin profiling of the canonical CD8⁺ T cell response against the same infection demonstrated parallel and distinct epigenetic signatures defining NK cells and CD8⁺ T cells. Overall, our study reveals the dynamic nature of epigenetic modifications during the generation of innate and adaptive lymphocyte memory.

Clonal expansion leading to immunological memory is a hallmark of the adaptive immune system and thus has been a feature that was traditionally attributed to antigen-specific T cells and B cells. However, recent studies have challenged this dogma by providing functional evidence that NK cells possess adaptive immune features during viral infection^{1,2}. In particular, mouse cytomegalovirus (MCMV) activates NK cells bearing the activating receptor Ly49H (which binds the MCMV-encoded glycoprotein m157)^{3,4} and results in clonal expansion and contraction of NK cells to generate a long-lived pool of memory cells that are capable of protective recall responses^{5–7}.

Reprints and permissions information is available at www.nature.com/reprints.

*Correspondence and requests for materials should be addressed to J.C.S. sunj@mskcc.org.

Author contributions

C.M.L. and J.C.S. designed the study. N.M.A., C.D.G., O.-E.W., and M.R. performed the experiments. C.M.L. performed the bioinformatic analyses. C.S.L. and Y. P. consulted on the bioinformatic analyses. C.M.L. and J.C.S. wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41590-018-0176-1>.

Although previous work has highlighted distinct transcriptional profiles of NK cells during MCMV infection⁸, we currently do not understand how transcription is controlled at the epigenetic level in NK cells as they transition between naive, effector, and memory states. Therefore, we have performed parallel chromatin accessibility analysis via the assay for transposase-accessible chromatin using high-throughput sequencing (ATAC-seq)⁹ and transcriptional profiling by RNA-seq on Ly49H⁺ NK cells during MCMV infection to elucidate how chromatin modifications dictate transcriptional fates. Furthermore, through parallel analysis of the chromatin landscape of MCMV-specific CD8⁺ T cells, our findings suggest that NK cells and T cells share common epigenetic programs during their transition from naive to memory cells.

Results

NK cell chromatin dynamics during infection.

Using ATAC-seq, we generated a kinetic profile of chromatin accessibility within the Ly49H⁺ NK cell population throughout the course of MCMV infection (Fig. 1a). NK cells were sorted as shown in Supplementary Fig. 1a, and samples displayed expected distributions of fragment lengths after processing (Supplementary Fig. 1b). Tabulation of pairwise changes showed that differentiating NK cells underwent considerable epigenetic changes of varying magnitude (Supplementary Fig. 1c), with putative enhancer regions (intronic and intergenic) showing the greatest numbers of high-fold change ($\log_2(\text{fold change}) > 1$) differentially accessible (DA) peaks (Fig. 1b) and vice versa when compared to all DA regions (Fig. 1c). In contrast, promoter regions, which generally showed higher baseline levels of accessibility (Supplementary Fig. 1d), underwent more subtle changes, as a majority of these DA peaks showed less than 0.5 $\log_2(\text{fold change})$ in accessibility across each sequential timepoint (Fig. 1b). Notably, analysis of DA peaks revealed the greatest global changes during the first week of virus infection (day 0 (d0) to d2, d2 to d4, and d4 to d7) and relatively little epigenetic modulation between d14 and d35 (Supplementary Fig. 1c). Hierarchical clustering of high-fold change regions revealed different waves of accessibility that exhibited various degrees of stability when comparing memory (d35) to naive cells (d0; Fig. 1d and Supplementary Fig. 1e). Clusters 1 and 6 had the highest proportion of stable changes that remained either closed or open, respectively, in the memory timepoint (Fig. 1d and Supplementary Fig. 1e). Regions near or within the gene loci of *Socs3*, *Cish*, *Pdcd1*, *Dnmt3a*, and *Il10* were among the top 10% most modulated regions within these clusters. Remaining clusters showed transient changes in chromatin accessibility (i.e., peaks that changed early during infection, but returned to baseline or near-baseline in memory cells). Most variable regions within these clusters included those found near *Tbx21*, *Klrg1*, *Ifng*, and *Zbtb32*.

Pathway analysis of regions within each high-fold change cluster showed enrichment in several biological processes familiar to NK cell and lymphocyte function. For example, among the most enriched pathways was the NK cell-mediated cytotoxicity pathway, found in clusters that showed spikes in chromatin accessibility of gene loci at d2, d4, or d7 (Fig. 1e and Supplementary Fig. 1f). In addition, a majority of these pathways involved genes related to cytokine signaling or antigen-receptor signaling, encompassing both the innate and adaptive features of NK cell immunity; these pathways include the interleukin 2 (IL-2)- and

IL-12-mediated signaling events pathway, Jak–STAT signaling pathway, and T cell antigen receptor signaling pathway (Fig. 1e and Supplementary Fig. 1f). Overall, we found that virus-specific NK cells underwent substantial changes in chromatin accessibility throughout the course of infection and reflected biological pathways related to the effector function of both innate and adaptive lymphocytes.

Early transcriptional and epigenetic coordination.

To better understand how changes in chromatin accessibility are associated with gene expression in differentiating NK cells, we generated a parallel kinetic profile of the transcriptional landscape during viral infection (Fig. 1a and Supplementary Fig. 1a). Comparison of principal component analyses (PCA) of chromatin accessibility and mRNA expression revealed that both profiles displayed similar patterns early after infection (between d0 and d7), but that these transitions diverged during later stages (d7 to d35; Fig. 2a). Notably, the majority of the chromatin remodeling appeared completed by d14, as d14 and d35 clustered tightly by ATAC-seq (Fig. 2a,b), suggesting that commitment to the memory cell fate may occur far earlier in differentiating lymphocytes than the transcriptome would dictate¹⁰. However, as in CD8⁺ T cell memory generation, NK cells transitioning from d14 to d35 continued to undergo changes in gene expression (Fig. 2a).

We next determined specific pathways (Supplementary Fig. 1f) in which chromatin accessibility and transcription were coordinated. Consistent with what we observed in the PCA analysis, all significant (FDR-adjusted $P < 0.05$) correlations were restricted to early transition timepoints and were more frequently observed in promoter regions than intergenic or intronic regions (Fig. 2c). Among the most strongly correlated pathways were type I and type II interferon (IFN) signaling at the earliest timepoints after infection (Fig. 2d), which are critical for anti-viral NK cell responses and generation of memory¹¹. For example, we observed a coordinated upregulation of *Ifng* expression and locus accessibility with a concomitant downregulation of its receptor, *Ifngr1*. In addition, there was a synchronized upregulation of well-characterized IFN-inducible genes, such as *Irf9*, *Mx1*, and *Isg20*. During the subsequent d2 to d4 transition, additional pathways included transcription factor NFAT signaling, which has been implicated in mediating NK cell production of IFN- γ ^{12–14}, as well as transcription factor ATF2–AP-1 signaling and T cell-related signaling (Fig. 2d). These observations collectively reveal a dynamic relationship between transcription and chromatin accessibility that changes throughout infection, with the most correlative relationship occurring during early stages of infection.

Distinct modes of STAT-induced chromatin changes.

STAT transcription factors play important roles during antiviral responses. Following MCMV infection, IL-12 signaling in NK cells induces phosphorylation of STAT4, which synergizes with IL-18 signaling to enhance production of IFN- γ and proliferation^{15,16}, while type-I IFN signaling activates STAT1, which complexes with STAT2 and IRF9 to promote granzyme production and protection from apoptosis^{11,17}. Notably, previous work has demonstrated a crucial role for both of these proinflammatory signaling pathways in the expansion and optimal formation of memory NK cells^{11,16} (Supplementary Fig. 2a). Given that these proinflammatory cytokine signaling pathways were among the pathways enriched

for coordinated epigenetic and transcriptional changes, we hypothesized that IL-12 and type-I IFN signaling through STAT4 and STAT1, respectively, may be affecting the epigenetic landscape in NK cells early after infection. Thus, we stimulated NK cells either with IL-12 plus IL-18 or with IFN- α and performed chromatin immunoprecipitation deep sequencing (ChIP-seq) analysis to assess STAT4 and STAT1 occupancy, respectively. As we recently described, IL-12 + IL-18-induced STAT4 occupancy was primarily located at putative enhancer sites (intronic and intergenic regions; Fig. 3a)¹⁸. In contrast, STAT1 primarily bound to promoter regions in response to IFN- α (Fig. 3a). To investigate whether these STAT-bound regions underwent epigenetic changes during MCMV infection, we overlapped peak regions identified through our ChIP-seq data with our time course of chromatin accessibility profiles. We found that the fate of STAT4- and STAT1-bound regions early during infection were very distinct. A large number of STAT4-bound regions showed highly dynamic transient increases in chromatin accessibility at d2, whereas STAT1-bound regions showed modest decreases in chromatin accessibility at d2 that continued through d4 (Fig. 3b). These modest changes at STAT1-bound promoter regions were consistent with a general trend of promoter regions showing fewer dynamic changes over time (Fig. 1b).

To directly test how STAT4 and STAT1 may be affecting both transcription and chromatin accessibility in NK cells following MCMV infection, we performed RNA-seq and ATAC-seq on wild-type versus *Stat4*^{-/-} or *Stat1*^{-/-} NK cells (Supplementary Fig. 2b,c). Because phosphorylation of STAT4 and STAT1 in NK cells occurs early after infection^{11,16} and changes in chromatin accessibility occurred as early as d2 (Fig. 3b), we focused on this early timepoint. Global analysis of chromatin profiles depicted different trends in response to STAT-deficiency (Fig. 3c). *Stat4*^{-/-} NK cells primarily showed decreases in chromatin accessibility at nonpromoter regions, with fewer and more balanced changes (i.e., similar numbers of gains and losses in accessibility) in promoter regions (Fig. 3c and Supplementary Fig. 3a). On the other hand, promoter regions in *Stat1*^{-/-} NK cells showed more losses than gains in accessibility, in contrast to nonpromoter regions, which showed a greater accumulation of gains (Fig. 3c and Supplementary Fig. 3a).

Type-I IFN signaling has been previously described as antagonizing IL-12-mediated IFN- γ production in a STAT1-dependent manner¹⁹. Consistent with these observations, we found that STAT1 deficiency increased *Ifng* expression (Supplementary Fig. 3b), similarly to IFNAR1 deficiency¹¹. Unexpectedly, comparison of STAT4- and STAT1-deficient transcriptional profiles revealed that a subset of STAT1-dependent genes were upregulated in the absence of STAT4 (Fig. 3d), including IFN-inducible *Mx1*, *Ifit2*, *Oas2*, and *Isg20* (Supplementary Fig. 3b). Regulation of type-I IFN signaling by STAT4 has been attributed to decreases in STAT1 expression²⁰. In agreement with those data, in the absence of STAT4 we observed an upregulation of *Stat1* and *Stat2* expression, and a lesser upregulation of *Stat6*, but did not observe upregulation in other STAT genes. In contrast, absence of STAT1 did not affect *Stat4* expression, but affected expression of other STAT genes to varying degrees (Supplementary Fig. 3c). These data support a cross-regulatory role of STAT1 and STAT4 through type-I IFN and IL-12 mediated signaling, respectively.

For both STATs, instances of STAT binding or changes in chromatin accessibility near gene loci were significantly enriched ($P < 2.09 \times 10^{-68}$ and $P < 3.97 \times 10^{-93}$ for STAT4; $P < 3.54$

$\times 10^{-16}$ and $P < 9.76 \times 10^{-31}$ for STAT1) among differentially expressed genes (Supplementary Fig. 3d). In contrast, only STAT4-bound peaks ($P < 2.76 \times 10^{-265}$), but not STAT1-bound peaks, were significantly enriched in DA regions (Supplementary Fig. 3d), suggesting that STAT4 may be playing a direct role in promoting changes in chromatin accessibility. To hone in on putative regions in which STAT4 may be actively influencing transcription, we combined all three datasets to search for commonly regulated gene loci (Fig. 3e). This analysis yielded 162 differentially expressed genes with a total of 207 peak regions near their loci that were simultaneously occupied by STAT4 and DA upon STAT4 deletion (Fig. 3e,f). For example, both *Fyn* and *Ifng* gene loci harbored putative enhancer regions that were occupied by STAT4 and decreased in chromatin accessibility in STAT4-deficient NK cells during MCMV infection (Fig. 3g). In contrast, similar STAT1 analysis revealed that only 28 genes were commonly regulated with regions that were both occupied by STAT1 as well as DA during infection (Fig. 3e), including IFN-inducible *Irgm1*, ISGF3 partner *Irf9*, and several MHC molecules (Supplementary Fig. 3e,f). Overall, our data suggest that STAT4 and STAT1 promote remodeling of the chromatin landscape in very different manners and that proinflammatory cytokine-induced STAT occupancy may play a direct role in affecting both local and distal changes in chromatin accessibility to promote transcription.

Epigenetic poisoning in memory NK cells.

Because of the pronounced distinction between chromatin accessibility profiles of memory and naive cells, we sought to better understand how memory cells are epigenetically different from their naive counterparts by first exploring the biological processes that may be affected. Pairwise analysis of memory and naive NK cells showed that a high proportion (~30%) of all accessible regions showed a significant $\log_2$ (fold change) of at least 0.5 (Supplementary Fig. 4a). To guide our analysis, we used our transcriptomic data to filter these regions down to those that were assigned to any genes that showed significant changes in expression at any sequential transition timepoint during infection, with the goal of enriching for the most relevant functional processes and active regions. Among more accessible regions, the most enriched pathway was NK cell-mediated cytotoxicity, as exemplified by the *Prf1* locus, where both distal and local regions showed increased chromatin accessibility in memory cells, correlating with increased transcription (Fig. 4a,b). Among the less accessible regions, the MAP kinase signaling pathway showed the highest ranking, as exemplified by downstream transcription factor *Mef2c*, in which all accessible regions within the gene body were downregulated in memory cells correlating with decreased transcript abundance (Fig. 4a,b).

De novo transcription factor binding motif analysis from these same filtered DA regions revealed an enrichment for an interferon-stimulated response element (ISRE)-like sequence in accessible memory peaks (Fig. 4c), consistent with enriched pathways involving IFN signaling and Jak-STAT (Fig. 4a). Additionally, we identified motifs most closely resembling putative binding sites for the TCF-LEF and NF- κ B transcription factor family members among DA regions that became less accessible in the memory state. Further motif analysis confirmed that, on average, our de novo ISRE-associated regions showed coordinated increases in accessibility in memory among all filtered DA regions, whereas our

de novo TCF-LEF- and NF- κ B-associated regions showed coordinated decreases in accessibility (Fig. 4d and Supplementary Fig. 4b). For example, an ISRE-like motif within a promoter region of *Klrg1* became more accessible in the memory timepoint, while TCF-LEF-like and NF- κ B-like motifs were found in less accessible regions of *Jdp2* and *Bcl2l1*, respectively (Fig. 4e). In addition, we scanned DA regions for matching known motifs, and found similar patterns and significant overlap (ISRE, $P < 4.26 \times 10^{-122}$; TCF, $P < 1.8 \times 10^{-65}$; NF- κ B, $P < 8.58 \times 10^{-109}$) between known and de novo motifs (Supplementary Fig. 4c,d). Notably, *Tcf7* and *Lef1* actually showed the highest expression among common TCF-LEF proteins in naive cells, and both *Tcf7* and *Nfkb1* showed decreases in expression during the memory timepoint, consistent with the decrease in accessibility to their genomic regions (Supplementary Fig. 4e,f). This down-regulation of NF- κ B accessibility may explain the dispensable role of NF- κ B-mediated IL-18 signaling in NK cell recall responses²¹. In summary, our data suggest that memory NK cells are epigenetically poised for specific pathway signaling and transcription factor binding that may modify effector function.

A common epigenetic signature of immune memory.

Several studies have provided insight on the dynamic nature of epigenetic regulation during CD8⁺ T cell differentiation in response to infection and tumor progression^{22–27}. Given the many parallel characteristics between NK cells and CD8⁺ T cells¹, we investigated whether these two cytolytic lymphocytes possess any similar epigenetic attributes during memory formation. To do so, we also performed RNA-seq and ATAC-seq on naive (d0), effector (d7), and memory MCMV-specific CD8⁺ T cells (d35) during MCMV infection (Supplementary Fig. 5a,b). To establish a common signature, we first limited our analysis to peaks that were commonly found in both NK cells and CD8⁺ T cells at any timepoint tested, which amounted to 53% of the entire merged atlas (Supplementary Fig. 6a). As expected, initial PCA of these common regions demonstrated cell-type specificity, as NK cells clustered closest to each other relative to CD8⁺ T cells (Supplementary Fig. 6b). However, we observed that transitions from naive to effector to memory followed similar trajectories in both cell types, and thus we re-analyzed our accessibility profiles using cell-normalized values that reflected relative changes in cell state within each cell type, rather than absolute values. Notably, PCA and hierarchical clustering of these normalized values revealed that naive NK cells clustered closer to memory CD8⁺ T cells than to either naive or effector CD8⁺ T cells (Fig. 5a,b), providing epigenetic evidence for prior head-to-head comparisons of these two cytotoxic lymphocyte states based on observed functionality and reactivity¹. Visualization of DA regions defined by CD8⁺ T cells showed that naive NK cells generally exhibited higher degrees of accessibility compared to naive T cells in regions that eventually opened up in effector and memory stages of CD8⁺ T cells (Fig. 5c).

Although memory CD8⁺ T cells were the closest T cell counterpart to naive NK cells, effector NK cells still clustered closest to effector CD8⁺ T cells, and memory NK cells clustered closest to memory CD8⁺ T cells (Fig. 5b), suggesting that there could be some common epigenetic trajectory during the generation of lymphocyte memory. Therefore, we performed a pairwise analysis between memory and naive lymphocytes for each cell type to find common patterns of chromatin accessibility. We first categorized these changes based on the nature of their differential accessibility: down in memory or up in memory

(Supplementary Fig. 6c). Among all DA regions, 21% of these peaks showed a chromatin accessibility pattern common to both NK and CD8⁺ T cells when comparing memory to naive cells (Fig. 6a and Supplementary Fig. 6c,d). To elucidate chromatin regions that may be influencing transcription, we performed a similar analysis on transcriptional profiles and found that only 9% of all differentially expressed genes also had similar patterns between NK and CD8⁺ T cells (Fig. 6a and Supplementary Fig. 6c,e). Overlap of accessible gene loci and differentially expressed genes yielded a total of 220 gene regions mapping to 94 genes that may play common roles in memory formation between NK cells and CD8⁺ T cells. Among highly expressed genes were transcription factors known to play roles in the generation of CD8⁺ T cell memory (*Bach2*, *Tcf7*, and *Zeb2*; Fig. 6b)^{28–31}. Furthermore, we identified previously described effector molecules (*Gzmb* and *Prfl*) and cell surface markers used to identify subsets of memory or memory precursor T cells (*Sell* and *Klrg1*). Notably, our analysis implicated a handful of molecules that have not been previously described to dictate memory formation, including *Tox* and *Themis2*, the former being expressed more highly in naive cells and the latter being induced only upon infection (Fig. 6b). Whether a novel role exists for these and additional factors (Fig. 6b) in the generation of both NK and T cell memory remains to be investigated.

The number of genes associated with common DA regions was disproportionately larger than the number of differentially expressed genes, even when only high-fold change DA regions were considered (Supplementary Fig. 6f; data not shown). This suggests that there may have been epigenetic changes that did not immediately result in transcriptional changes, and therefore may represent a poised chromatin profile. Thus, we performed de novo motif analysis to search for patterns of transcription factor binding potential among our subset of common DA regions, regardless of transcriptional regulation. We observed an enrichment of potential binding regions for Jun factors, members of the AP-1 transcription factor family known to play roles in T cell activation³², in regions that became more accessible in memory NK and CD8⁺ T cells (Fig. 6c,d). For example, an intergenic peak upstream of *Il10* showed an AP-1-like motif within a region that became highly accessible in the memory timepoint (Fig. 6e). Notably, in NK cells, *Junb* and *Jund* were the most highly expressed family members, whereas in CD8⁺ T cells *Junb* is highly expressed along with *Batf* (Supplementary Fig. 6g). Among regions that demonstrated losses in shared DA regions, we again found an enrichment of a TCF motif (Fig. 6c,d), consistent with results from our global analysis on NK cells alone, as exemplified by the *Tgfbr2* locus (Fig. 6e). Overall, we provide evidence that NK cells and CD8⁺ T cells possess shared underlying epigenetic signatures that may be commonly regulated by the AP-1 and TCF-LEF transcription factor families during generation of immune memory.

Discussion

As innate lymphocytes that share many functional features with their evolutionary cousins, the CD8⁺ T cells, NK cells have similarly acquired the ability to undergo clonal expansion and generate memory, even without diversification of their antigen receptor repertoire by gene rearrangement. We revealed that, similarly to CD8⁺ T cells, NK cells underwent dynamic changes in chromatin architecture during pathogen exposure, changes that were associated with genes that dictate both innate and adaptive effector functions. Notably, we

found that a majority of these epigenetic changes occurred by d14 after infection and that minimal changes continued to occur as differentiating NK cells transitioned between d14 and d35. While many of these changing regions returned to naive levels, a sizable portion of these regions (at least ~30%) remained at a substantially altered state of accessibility. Although it remains to be determined how specific poised features translate to NK cell longevity and function, we believe our analyses provide a resource and a guide for more precise functional studies to validate these observations.

Previous studies have highlighted the dynamic nature of distal regulatory regions in shaping cell-type specificity across multiple cell types^{33–35}. In NK cells, we observed a similar type of regulation across differentiation states as well, as our most dynamically regulated regions showed high representations of nonpromoter (i.e., putative enhancer) regions. It should be noted that our assay was limited in resolving potential changes in subsections within the promoter region, which tended to be wider in base pairs (data not shown) and show higher baseline levels of chromatin accessibility. ChIP-seq studies involving specific histone modifications will better expand our understanding of regulation at gene promoters within NK cells responding to viral infection.

Among altered chromatin regions, we also saw an overrepresentation of several transcription factor binding motifs, including those for high-mobility group transcription TCF transcription factors. Both TCF1 and LEF1 are often implicated in orchestrating lineage commitment of T cells during development and infection^{28,36–38}, as well as in various roles during innate lymphocyte development, including NK cells^{39,40}. Furthermore, TCF1 is a known regulator of the chromatin landscape^{38,41}. Thus, its role in NK cell memory formation warrants further investigation.

Given the importance of STAT transduction in human disease^{42,43}, our multifaceted genomic and transcriptomic profiling provides better insight on the complexity of STAT signaling. Through our transcriptomic data, we found that STAT4 and STAT1 can converge to play antagonistic roles on each other's downstream network in activated NK cells during MCMV infection. During early infection, STAT4 can actively suppress STAT1 expression, whereas STAT1 can control IFN1- γ production induced by STAT4. This observation has been described in CD8⁺ T cells and NK cells during LCMV infection²⁰, and a similar phenomenon likely occurs in differentiating NK cells during MCMV infection.

We also observed that STAT4 and STAT1 greatly differed in their interaction with the epigenetic landscape. IL-12 + IL-18-induced STAT4 binding in NK cells coincided with changes in accessibility, which suggests that direct STAT4 binding at putative enhancers promoted local changes in chromatin accessibility that drove transcription during early stages of infection. Conversely, STAT1 ChIP-seq revealed that STAT1 interacted with the genome primarily within promoter regions. Notably, STAT1 binding did not significantly coincide with regions that showed significant changes in chromatin accessibility in the absence of STAT1. This observation could be reflective of the above caveats of promoter accessibility described above, or it could suggest that these changes were indirect and/or driven by mechanisms independent of type-I IFN responses. Consistent with the latter observation, STAT1 has been described as having differential binding in the context of type-I

versus type-II IFN-induced signaling⁴⁴, and perhaps type-II IFN signaling may be driving these changes in accessibility. Furthermore, our current understanding of how early STAT activation may impact later alterations in the chromatin landscape is still incomplete. Although we found that STATs had critical roles early in infection and promoted early epigenetic and transcriptional changes, they may have diverse stage-specific roles that cumulatively optimize proper memory formation and maintenance.

Finally, we have discovered that NK cells and CD8⁺ T cells possess shared epigenetic and transcriptional programs underlying host immunological memory. At baseline, we discovered common accessible chromatin regions between naive NK cells and antigen-experienced memory CD8⁺ T cells, providing epigenetic evidence for the ability of NK cells to undergo rapid activation and mediate effector function during infection¹. Furthermore, analysis of differentially accessible chromatin and transcription between naive versus memory NK cells and CD8⁺ T cells yielded a list of genes that may play a shared role in the generation of innate and adaptive lymphocyte memory. Although this shared signature was represented by a relatively small subset of their global profile, it was enriched for regions associated with genes known to mediate CD8⁺ T cell differentiation and memory formation, including *Bach2*, *Tcf7*, and *Zeb2*^{28–31}. Thus, our analysis revealed several gene candidates that may regulate the generation of both NK and T cell memory, as well as corresponding putative enhancer regions that may fine-tune gene expression. Future studies will be aimed at dissecting the functional relevance of these genes and their associated regulatory regions in the context of cytomegalovirus infection. Overall, our study expands our current understanding of the molecular mechanisms underlying immunological memory, and may unveil novel models for designing targeted vaccines against infectious disease.

Methods

Methods, including statements of data availability and any associated accession codes and references, are available at <https://doi.org/10.1038/s41590-018-0176-1>.

Methods

Animals.

All mice used in this study were bred at Memorial Sloan Kettering Cancer Center in accordance with the guidelines of the Institutional Animal Care and Use Committee (IACUC). The following strains were used, all on the C57BL/6 genetic background: WT CD45.2, WT CD45.1, *Stat4*^{-/-45}, *Stat1*^{-/-46}, and *Klra8*^{-/-} (*Ly49H*^{-/-})⁴⁷.

Mixed bone marrow chimeras and adoptive transfers.

Wild-type (WT) host CD45.1⁺CD45.2⁺ mice were lethally irradiated with 900 Gy and reconstituted with a 1:1 mixture of bone marrow cells from WT (CD45.1⁺) and *Stat4*^{-/-} or *Stat1*^{-/-} donor (CD45.2⁺) mice, co-injected with anti-NK1.1 (clone PK136) to deplete any residual donor or host mature NK cells. Adoptive transfer studies were performed as previously described⁷ using CD45.1⁺ *Klra8*^{-/-} recipients.

Viral infection.

Naive mice and mixed-bone-marrow chimeric mice were infected with MCMV (Smith strain) by intraperitoneal (i.p) injection of 7.5×10^3 plaqueforming units (PFU) in 0.5 mL. Experimental mice in adoptive transfer studies into *Klra8*^{-/-} recipients were infected with MCMV by i.p. injection of 7.5×10^2 PFU in 0.5 mL on the day following cell transfer.

Isolation of lymphocytes.

Spleens were dissociated using glass slides and filtered through a 100- μ m strainer. To isolate bone marrow lymphocytes, cleaned femur and tibia bones were ground with mortar and pestle and the resulting solution was filtered through a 100- μ m strainer. Red blood cells in spleen and bone marrow were lysed using ACK lysis buffer.

Flow cytometry and cell sorting.

Cell surface staining of single-cell suspensions from bone marrow and spleen was performed using fluorophore-conjugated antibodies. Intracellular staining was performed by fixing and permeabilizing with the eBioscience Foxp3/Transcription Factor Staining Set. The following antibodies were used for flow cytometry: CD3e (17A2, Tonbo, #25-0032), TCR β (H57-597, BioLegend, #109220), CD19 (ID3, BioLegend, #115530), F4/80 (BM8.1, BioLegend, #123117), NK1.1 (PK136, Tonbo, #65-5941), CD49b (DX5, BioLegend, #108918), Ly49H (3D10, eBioscience/Thermo Fisher Scientific, #11-5886-81), CD45.1 (A20, BioLegend, #110729), CD45.2 (104, BioLegend, #109821), CD69 (H1.2F3, BioLegend, #104524), IFN- γ (XMG1.2, Tonbo, #20-7311), CD8 α (53-6.7, BioLegend, #100730), CD62L (MEL-14, BioLegend, #104424), CD44 (IM7, BioLegend, #103006). MHC class I tetramers were generated by conjugating D^b/HGIRNASFI (m45) monomers (NIH Tetramer Facility) to streptavidin-phycoerythrin or streptavidin-APC (BioLegend, #405204 and #405207, respectively). Flow cytometry and cell sorting were performed on LSR II and Aria II cytometers (BD Biosciences), respectively. Data were analyzed with FlowJo software (Tree Star).

Sample preparation for ChIP-seq.

STAT4 chromatin immunoprecipitation (ChIP) was performed as previously described¹⁸. For STAT1 ChIP, $5\text{--}10 \times 10^6$ NK cells (TCR β ⁻CD19⁻CD3e⁻F4/80⁻NK1.1⁺CD49b⁺) were sorted from spleens of pooled C57BL/6 mice and incubated overnight (16–18 h) with 100U/mL IFN- α with or without 10 ng/mL IL-2. DNA and proteins were cross-linked for 8 min using 1% formaldehyde. ChIP was performed as previously described^{48,49} using 10 μ g of rabbit polyclonal anti-STAT1 antibody (Santa Cruz, sc-592, clone M-22). Fragments between 100 and 600 bp were size-selected, and Illumina HiSeq libraries were prepared using the Kapa DNA library preparation chemistry (Kapa Biosystems) and 12–15 cycles of PCR. Adaptors were diluted 1/10 or 1/50 depending on the amount of starting material available. Barcoded libraries were run on HiSeq 2500 1 T in a 50-bp/50-bp paired-end run, using a TruSeq SBS Kit v3 (Illumina). On average, 57 million paired reads were generated per sample.

Sample preparation for RNA-seq.

We isolated $2\text{--}5 \times 10^4$ sorted splenic lymphocytes for RNA-sequencing. For wild-type NK cell time courses, Ly49H⁺ NK cells (TCR β ⁻CD19⁻CD3e⁻F4/80⁻NK1.1⁺Ly49H⁺) were sorted from wild-type mice at d0, d2, and d4 postinfection (PI) or from recipients of adoptive transfers at d7, d14, and d35 PI. CD8⁺ T cells were sorted from wild-type mice at d0 (naive: TCR β ⁺CD8⁺CD62L^{hi}CD44^{lo}), d7, or d35 PI (effector and memory: TCR β ⁺CD8⁺CD62L^{lo}CD44^{hi}Tetramer⁺). WT and STAT-deficient Ly49H⁺ NK cells (TCR β ⁻CD19⁻CD3e⁻F4/80⁻-NK1.1⁺Ly49H⁺) from bone marrow chimeras were sorted at d0 or d2 PI. Each sample represents 1–3 mice. RNA was isolated from sorted cell populations using Trizol (Invitrogen), and total RNA was amplified using the SMART-seq V4 Ultra Low Input RNA kit (Clontech). Subsequently, 10 ng of amplified cDNA was used to prepare Illumina HiSeq libraries with the Kapa DNA library preparation chemistry (Kapa Biosystems) using eight cycles of PCR. Samples were barcoded and run on HiSeq 2500 1 T, in a 50-bp/50-bp paired-end run, using the TruSeq SBS Kit v3 (Illumina). On average, 47 million paired reads were generated per sample, and the percent of mRNA bases was over 72% on average.

Sample preparation for ATAC-seq.

We sort-purified $0.7\text{--}5.0 \times 10^4$ cells using the same strategy as described for RNA-seq above. Profiling of chromatin was performed as described previously⁹. Briefly, fresh cells were washed in cold PBS and lysed. Transposition occurred at 42 °C for 45min. DNA was purified using the MinElute PCR purification kit (Qiagen) and amplified for five cycles. Additional PCR cycles were evaluated by real-time PCR. The final product was cleaned by Ampure Beads at a 1.5 $\times$ ratio. Libraries were sequenced on a HiSeq 2500 1 T in a 50-bp/50-bp paired-end run, using the TruSeq SBS Kit v3 (Illumina). On average, 51 million paired reads were generated per sample.

Genomic alignment of sequencing reads and differential analysis.

For all sequenced data, paired-end reads were trimmed for adaptors and removal of low-quality reads using Trimmomatic (v.0.36)⁵⁰. Trimmed reads were mapped to the *Mus musculus* genome (mm10 assembly) using Bowtie2 (v2.2.9)⁵¹. Read counts for features (peak regions for ATAC-seq and exons for RNA-seq) were generated using the summarizeOverlaps function from the GenomicAlignments package (v1.10.1)⁵². Differential analyses were executed with DESeq2 (v1.14.1)⁵³ using the UCSC Known Gene models as reference annotations. Features were considered differential if they showed a *P*value (*P*_{adj}) < 0.05, adjusted for multiple hypothesis correction.

ATAC-seq peak-calling and atlas generation.

For peak calling, all positive-strand reads were shifted 4 bp downstream and all negative-strand reads were shifted 5 bp upstream to center the reads on the transposon binding event⁹. Shifted, concordantly aligned paired mates were used for peak calling by MACS2 (v2.1.1.20160309)⁵⁴ at a *P*value of 0.01. Irreproducible discovery rate (IDR)⁵⁵ calculations using scripts provided by the ENCODE project (<https://www.encodeproject.org/software/idr/>; v2.0.2 and v2.0.3) were performed on all pairs of replicates using an oracle peak list

called from merged replicates for each condition, keeping only reproducible peaks showing an IDR value of 0.05 or less. Overlapping peaks from different conditions were merged using the following conditions: if a peak overlapped with more than one other peak, the peak with the highest overlap (calculated as the percentage of overlap width out of the total width of the narrower peak) was merged; if the overlap of two peaks was > 75%, nonoverlapping portions were removed; if the overlap was 75% or less, then both peaks were merged to create a single peak encompassing both peak ranges. A union of both merged overlapping peaks and unique reproducible peaks for each condition was used to generate the final atlas and was used for further analysis. For NK cell time course atlas, reproducible peaks from d0, d2, d4, d7, d14, and d35 were combined and generated a total of 46,579 peak regions. For *Stat4*^{-/-} and *Stat1*^{-/-} atlas, reproducible peaks from d0 and d2 in WT and KO conditions were combined and generated a total of 45,409 peak regions. For NK cell and CD8⁺ T cell atlas, naive (d0), effector (d7), and memory (d35) reproducible peaks were first merged together for each cell type to generate an NK cell-specific and CD8⁺ T cell-specific atlases. These cell-type-specific peak lists were then combined and generated a total of 64,097 peak regions. Merged overlapping peaks were designated as common atlas peaks, as shown in Supplementary Fig. 6a.

Peak annotation.

Peak assignment was done using ChipPeakAnno⁵⁶. Promoter regions were defined as peaks that overlapped a region that was + 2 kb to -0.5 kb from the transcriptional start site (TSS). Intragenic (intronic and exonic) peaks were defined as any peak that overlapped with annotated intronic and exonic peaks, respectively, based on the annotation database. Intergenic peaks were defined as any nonpromoter or nonintragenic peaks and were assigned to the gene of the nearest TSS based on the distance from the start of the peak. Priority was given to transcripts that were canonical, based on the UCSC Known Canonical database.

ATAC-seq data analysis.

For wild-type NK cell time course, two types of pairwise comparisons were performed: one for all timepoints against d0 (d2 vs. d0, d4 vs. d0, d7 vs. d0, d14 vs. d0, d35 vs. d0) and one for transition timepoints (d2 vs. d0, d4 vs. d2, d7 vs. d4, d14 vs. d7, d35 vs. d14). A peak was deemed differentially accessible (DA) if it had $P_{\text{adj}} < 0.05$ in both the d0 comparison and at least one of the transition timepoints. Using these criteria, 35,979 peaks out of a total of 46,579 called peaks were designated as DA and categorized by magnitudes of $\log_2$ fold change. A DA peak was designated as a high-fold change peak if it showed an absolute $\log_2(\text{fold change}) > 1$ at least once at any transition timepoint comparison. For *Stat4*^{-/-} and *Stat1*^{-/-} analyses, peaks were DA if they showed $P_{\text{adj}} < 0.05$. For NK cell versus CD8⁺ T cell analysis, common atlas peak regions were used for counting and for determining differential accessibility. Pairwise analysis was performed on memory and naive timepoints for each cell type using DESeq2. Peaks were DA if they showed $P_{\text{adj}} < 0.05$ and absolute $\log_2(\text{fold change}) > 0.5$. A total of 34,173 peaks were considered, with some being filtered out by independent filtering performed by DESeq2.

Pathway analysis.

BED files of peak regions were used as input for pathway analysis by GREAT⁵⁷. Results were retrieved either by using rGREAT (v1.6.0) for analyses on high-fold change peaks, or using the online web interface at <http://bejerano.stanford.edu/great/public/html/index.php> (v3.0.0) for analysis on d35 vs. d0. All analyses used default settings with the whole genome as background, considering analyses from the MSigDB Canonical Pathway databases. Only pathway terms with a minimum of 25 genes were considered and used for multiple-hypothesis correction. Enriched pathways were filtered for those that showed $P_{\text{adj}} < 0.05$ for both binomial and hypergeometric P values as calculated by GREAT and for those that had more than 10 or 20 gene hits. For memory versus naive pathway analysis, GREAT was performed on DA regions that (i) showed an absolute $\log_2(\text{fold change}) > 0.5$; (ii) were assigned to genes that had a base mean normalized gene expression count of > 50 across all timepoints; and (iii) showed significant modulation of gene expression ($P_{\text{adj}} < 0.05$, absolute $\log_2(\text{fold change}) > 1$) in at least one transition timepoint (d2 vs. d0, d4 vs. d2, d7 vs. d4, d14 vs. d7, d35 vs. d14). The resulting peak set that was used for gene ontology had a total of 5,118 peak regions (2,212 up; 2,906 down). Pathway results were further filtered on those that showed a HyperFoldEnrichment of more than 2, as calculated by GREAT.

ChIP-seq data processing.

STAT4 ChIP-seq data was processed as described previously¹⁸. For STAT1 ChIP-seq data, peak-calling was performed using MACS2 with the arguments “-p 1e-2 -m 2 50 --to-large”. IDR was calculated on replicates as with ATAC-seq, and reproducible peaks passing a threshold of 0.05 or less was used as a reference peak list. This yielded a total of 3,471 peak regions. For comparison to ATAC-seq data, some overlapping ChIP-seq peaks were split to match peak regions of chromatin accessibility.

RNA-seq data analysis.

Genes were considered differentially expressed if they showed $P_{\text{adj}} < 0.05$. For NK cell versus CD8⁺ T cell analysis, genes were further filtered on those that showed an absolute $\log_2(\text{fold change}) > 1$. Expression is shown as normalized counts (untransformed and $\log_2$ -transformed) calculated by DESeq2, or $\log_2$ values of transcripts per million (TPM) plus a pseudocount of 1. For calculation of transcript abundance, cDNA FASTA files from *Mus musculus* GRCm38 transcriptome was downloaded from <https://www.ensembl.org/info/data/ftp/index.html> and used to index and calculate TPM using kallisto⁵⁸ with default arguments.

Hierarchical clustering and principal component analysis.

Blind, normalized $\log_2$ values calculated by DESeq2 were used for principal component analysis and to calculate Euclidean distances for hierarchical clustering using Ward's method. For heatmaps in Fig. 1d and Supplementary Fig. 1e, these normalized $\log_2$ values of all high-fold change peaks were used to hierarchically cluster peak regions into seven clusters, with the top 10% most variable regions (based on calculated variance across all samples) within each cluster shown in Fig. 1d to keep proportional representation. For NK cell versus CD8⁺ T cell analysis, values were normalized for cell type by centering peak

counts by each cell-type baseline, represented by the mean. In other words, for each peak, the mean of all NK cell sample counts was subtracted from each individual NK cell sample count, while the mean of all CD8⁺ T cell sample counts was subtracted from each individual CD8⁺ T cell sample count. The resulting values thus reflect relative deviations, independent of baseline accessibility.

Correlation analysis between accessibility and transcription.

Gene lists of MSigDB Canonical Pathways (C2) were downloaded as gmt files from <http://software.broadinstitute.org/gsea/downloads.jsp> and read into R (v.3.3.2 and v3.3.3) using the GSEABase package (v1.36.0). Gene IDs were converted from human to mouse using conversions downloaded from <http://www.genenames.org/cgi-bin/hcop>. All pathways discovered from GREAT analysis on high fold change clusters (Supplementary Fig. 1f) were tested. Peaks were matched to gene sets based on the genes they were assigned to. For each gene loci at each transition timepoint, we calculated the mean log₂(fold change) across all peaks of the same peak type, regardless of differential accessibility, to ultimately generate three summarized log₂(fold change) values per gene (intronic, intergenic, promoter). Each of these summarized values was paired with its corresponding RNA-seq log₂(fold change) and used to calculate Spearman correlations for each gene set. FDR-adjusted *P* values were calculated using the Benjamini and Hochberg method for all hypotheses. Correlated pathways were filtered on those showing a FDR-adjusted *P* < 0.05.

Motif analysis.

De novo analysis and known-motif search were performed using HOMER⁵⁹ using the findMotifsGenome function, with the parameters “-size given -len 6,8,10,12 -mset vertebrates -mask” for de novo analysis. For NK cell time courses, we used the entire ATAC-seq atlas background. For NK versus CD8⁺ T cell analysis, we used common peaks found in both NK and CD8⁺ T cells from the combined NK/CD8⁺ T cell atlas. Results show top motif hits with a motif score of at least 0.9. Output from de novo analysis and precompiled motif files provided by the HOMER software were used as input to search for motif instances across the genome. For metacoverage plots, normalized ATAC-seq reads were counted across 2-bp bins spanning ± 1 kb from the center of each motif instance. Median values of subset peaks are plotted for each bin.

Chromatin accessibility heatmaps and gene tracks.

For peak-centered heatmaps, average sample counts normalized using size factors calculated by DESeq2 were plotted in 40-bp bins across a 2-kb window. To improve visualization, binned counts greater than the 75th percentile + (3 × the interquartile range) were capped at that value, unless otherwise indicated. The log₂(fold changes) displayed in the heatmap in Fig. 6 are similarly capped. Gene tracks were generated by converting BAM files to bigWig files using bedtools2 (v.2.26.0)⁶⁰ and UCSC's bedGraphToBigWig (v.4) and were visualized using the Gviz R package (v.1.18.2). All tracks show a 1- to 10-kb scale on the *x* axis and normalized tag counts on the *y* axis.

Reporting Summary.

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

Data availability.

All data generated and supporting the findings of this study are available within the paper. The RNA-seq, ChIP-seq, and ATAC-seq data have been deposited at GEO with accession number [GSE106139](#).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank members of the Sun laboratory for comments, discussions, technical support, and experimental assistance. We thank S. Chhangawala, L. Fairchild, and C. Krishna for discussions and technical support. The Integrated Genomics Operation Core was supported by Cycle for Survival and the Marie-Josée and Henry R. Kravis Center for Molecular Oncology. C.M.L. was supported by a T32 award from the NIH (CA009149). N.M.A. was supported by a Medical Scientist Training Program grant from the National Institute of General Medical Sciences of the NIH under award number T32GM007739 to the Weill Cornell/Rockefeller/Sloan-Kettering Tri-Institutional MD-PhD Program and an F30 Predoctoral Fellowship from the NIH National Institute of Allergy and Infectious Diseases (F30 AI136239). M.R. was supported by a fellowship from the German Academic Exchange Service (DAAD; Germany). C.S.L. was supported by NIH grant U01 HG007893. J.C.S. was supported by the Ludwig Center for Cancer Immunotherapy, the Burroughs Wellcome Fund, the American Cancer Society, and grants from the NIH (AI100874, AI130043, and P30CA008748).

References

1. Sun JC & Lanier LL NK cell development, homeostasis and function: parallels with CD8⁺ T cells. *Nat. Rev. Immunol* 11, 645–657 (2011). [PubMed: 21869816]
2. Vivier E et al. Innate or adaptive immunity? The example of natural killer cells. *Science* 331, 44–49 (2011). [PubMed: 21212348]
3. Arase H, Mocarski ES, Campbell AE, Hill AB & Lanier LL Direct recognition of cytomegalovirus by activating and inhibitory NK cell receptors. *Science* 296, 1323–1326 (2002). [PubMed: 11950999]
4. Smith HR et al. Recognition of a virus-encoded ligand by a natural killer cell activation receptor. *Proc. Natl Acad. Sci. USA* 99, 8826–8831 (2002). [PubMed: 12060703]
5. Dokun AO et al. Specific and nonspecific NK cell activation during virus infection. *Nat. Immunol* 2, 951–956 (2001). [PubMed: 11550009]
6. Daniels KA et al. Murine cytomegalovirus is regulated by a discrete subset of natural killer cells reactive with monoclonal antibody to Ly49H. *J. Exp. Med* 194, 29–44 (2001). [PubMed: 11435470]
7. Sun JC, Beilke JN & Lanier LL Adaptive immune features of natural killer cells. *Nature* 457, 557–561 (2009). [PubMed: 19136945]
8. Bezman NA et al. Molecular definition of the identity and activation of natural killer cells. *Nat. Immunol* 13, 1000–1009 (2012). [PubMed: 22902830]
9. Buenrostro JD, Giresi PG, Zaba LC, Chang HY & Greenleaf WJ Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* 10, 1213–1218 (2013). [PubMed: 24097267]
10. Kaech SM, Hemby S, Kersh E & Ahmed R Molecular and functional profiling of memory CD8 T cell differentiation. *Cell* 111, 837–851 (2002). [PubMed: 12526810]
11. Madera S et al. Type I IFN promotes NK cell expansion during viral infection by protecting NK cells against fratricide. *J. Exp. Med* 213, 225–233 (2016). [PubMed: 26755706]

12. Tassi I et al. NK cell-activating receptors require PKC-theta for sustained signaling, transcriptional activation, and IFN-gamma secretion. *Blood* 112, 4109–4116 (2008). [PubMed: 18784374]
13. Brand A et al. LDHA-associated lactic acid production blunts tumor immunosurveillance by T and NK Cells. *Cell Metab.* 24, 657–671 (2016). [PubMed: 27641098]
14. Freund-Brown J et al. Cutting edge: murine NK cells degranulate and retain cytotoxic function without store-operated calcium entry. *J. Immunol* 199, 1973–1978 (2017).
15. Andrews DM, Scalzo AA, Yokoyama WM, Smyth MJ & Degli-Esposti MA Functional interactions between dendritic cells and NK cells during viral infection. *Nat. Immunol* 4, 175–181 (2003). [PubMed: 12496964]
16. Sun JC et al. Proinflammatory cytokine signaling required for the generation of natural killer cell memory. *J. Exp. Med* 209, 947–954 (2012). [PubMed: 22493516]
17. García-Sastre A & Biron CA Type 1 interferons and the virus-host relationship: a lesson in détente. *Science* 312, 879–882 (2006). [PubMed: 16690858]
18. Rapp M et al. Core-binding factor β and Runx transcription factors promote adaptive natural killer cell responses. *Sci. Immunol* 2, eaan3796 (2017). [PubMed: 29222089]
19. Nguyen KB et al. Interferon alpha/beta-mediated inhibition and promotion of interferon gamma: STAT1 resolves a paradox. *Nat. Immunol* 1, 70–76 (2000). [PubMed: 10881178]
20. Gil MP et al. Regulating type 1 IFN effects in CD8 T cells during viral infections: changing STAT4 and STAT1 expression for function. *Blood* 120, 3718–3728 (2012). [PubMed: 22968462]
21. Madera S & Sun JC Cutting edge: stage-specific requirement of IL-18 for antiviral NK cell expansion. *J. Immunol* 194, 1408–1412 (2015). [PubMed: 25589075]
22. Pauken KE et al. Epigenetic stability of exhausted T cells limits durability of reinvigoration by PD-1 blockade. *Science* 354, 1160–1165 (2016). [PubMed: 27789795]
23. Sen DR et al. The epigenetic landscape of T cell exhaustion. *Science* 354, 1165–1169 (2016). [PubMed: 27789799]
24. Scott-Browne JP et al. Dynamic changes in chromatin accessibility occur in CD8⁺ T cells responding to viral infection. *Immunity* 45, 1327–1340 (2016). [PubMed: 27939672]
25. Yu B et al. Epigenetic landscapes reveal transcription factors that regulate CD8⁺ T cell differentiation. *Nat. Immunol* 18, 573–582 (2017). [PubMed: 28288100]
26. Ghoneim HE et al. De novo epigenetic programs inhibit PD-1 blockade-mediated T cell rejuvenation. *Cell* 170, 142–157.e19 (2017). [PubMed: 28648661]
27. Philip M et al. Chromatin states define tumour-specific T cell dysfunction and reprogramming. *Nature* 545, 452–456 (2017). [PubMed: 28514453]
28. Zhou X et al. Differentiation and persistence of memory CD8(+) T cells depend on T cell factor 1. *Immunity* 33, 229–240 (2010). [PubMed: 20727791]
29. Dominguez CX et al. The transcription factors ZEB2 and T-bet cooperate to program cytotoxic T cell terminal differentiation in response to LCMV viral infection. *J. Exp. Med* 212, 2041–2056 (2015). [PubMed: 26503446]
30. Omilusik KD et al. Transcriptional repressor ZEB2 promotes terminal differentiation of CD8+ effector and memory T cell populations during infection. *J. Exp. Med* 212, 2027–2039 (2015). [PubMed: 26503445]
31. Roychoudhuri R et al. BACH2 regulates CD8(+) T cell differentiation by controlling access of AP-1 factors to enhancers. *Nat. Immunol* 17, 851–860 (2016). [PubMed: 27158840]
32. Smith-Garvin JE, Koretzky GA & Jordan MS T cell activation. *Annu. Rev. Immunol* 27, 591–619 (2009). [PubMed: 19132916]
33. González AJ, Setty M & Leslie CS Early enhancer establishment and regulatory locus complexity shape transcriptional programs in hematopoietic differentiation. *Nat. Genet* 47, 1249–1259 (2015). [PubMed: 26390058]
34. Shih HY et al. Developmental acquisition of regulomes underlies innate lymphoid cell functionality. *Cell* 165, 1120–1133 (2016). [PubMed: 27156451]
35. Corces MR et al. Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. *Nat. Genet* 48, 1193–1203 (2016). [PubMed: 27526324]

36. Weber BN et al. A critical role for TCF-1 in T-lineage specification and differentiation. *Nature* 476, 63–68 (2011). [PubMed: 21814277]
37. Germar K et al. T-cell factor 1 is a gatekeeper for T-cell specification in response to Notch signaling. *Proc. Natl Acad. Sci. USA* 108, 20060–20065 (2011).
38. Xing S et al. Tcf1 and Lef1 transcription factors establish CD8(+) T cell identity through intrinsic HDAC activity. *Nat. Immunol* 17, 695–703 (2016). [PubMed: 27111144]
39. Yang Q et al. TCF-1 upregulation identifies early innate lymphoid progenitors in the bone marrow. *Nat. Immunol* 16, 1044–1050 (2015). [PubMed: 26280998]
40. Jeevan-Raj B et al. The transcription factor Tcf1 contributes to normal NK cell development and function by limiting the expression of granzymes. *Cell Rep.* 20, 613–626 (2017). [PubMed: 28723565]
41. Johnson JL et al. Lineage-determining transcription factor TCF-1 initiates the epigenetic identity of T cells. *Immunity* 48, 243–257.e10 (2018). [PubMed: 29466756]
42. O'Shea JJ et al. The JAK-STAT pathway: impact on human disease and therapeutic intervention. *Annu. Rev. Med* 66, 311–328 (2015). [PubMed: 25587654]
43. Villarino AV, Kanno Y & O'Shea JJ Mechanisms and consequences of Jak-STAT signaling in the immune system. *Nat. Immunol* 18, 374–384 (2017). [PubMed: 28323260]
44. Hartman SE et al. Global changes in STAT target selection and transcription regulation upon interferon treatments. *Genes Dev.* 19, 2953–2968 (2005). [PubMed: 16319195]
45. Kaplan MH, Sun YL, Hoey T & Grusby MJ Impaired IL-12 responses and enhanced development of Th2 cells in Stat4-deficient mice. *Nature* 382, 174–177 (1996). [PubMed: 8700209]
46. Meraz MA et al. Targeted disruption of the *Stat1* gene in mice reveals unexpected physiologic specificity in the JAK-STAT signaling pathway. *Cell* 84, 431–442 (1996). [PubMed: 8608597]
47. Fodil-Cornu N et al. Ly49h-deficient C57BL/6 mice: a new mouse cytomegalovirus-susceptible model remains resistant to unrelated pathogens controlled by the NK gene complex. *J. Immunol* 181, 6394–6405 (2008). [PubMed: 18941230]
48. Zheng Y et al. Genome-wide analysis of Foxp3 target genes in developing and mature regulatory T cells. *Nature* 445, 936–940 (2007). [PubMed: 17237761]
49. Beaulieu AM, Zawislak CL, Nakayama T & Sun JC The transcription factor Zbtb32 controls the proliferative burst of virus-specific natural killer cells responding to infection. *Nat. Immunol* 15, 546–553 (2014). [PubMed: 24747678]
50. Bolger AM, Lohse M & Usadel B Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120 (2014). [PubMed: 24695404]
51. Langmead B & Salzberg SL Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359 (2012). [PubMed: 22388286]
52. Lawrence M et al. Software for computing and annotating genomic ranges. *PLOS Comput. Biol* 9, e1003118 (2013). [PubMed: 23950696]
53. Love MI, Huber W & Anders S Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550 (2014). [PubMed: 25516281]
54. Zhang Y et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* 9, R137 (2008). [PubMed: 18798982]
55. Li QH, Brown JB, Huang HY & Bickel PJ Measuring reproducibility of high-throughput experiments. *Ann. Appl. Stat* 5, 1752–1779 (2011).
56. Zhu LJ et al. ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. *BMC Bioinforma.* 11, 237 (2010).
57. McLean CY et al. GREAT improves functional interpretation of cis-regulatory regions. *Nat. Biotechnol* 28, 495–501 (2010). [PubMed: 20436461]
58. Bray NL, Pimentel H, Melsted P & Pachter L Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol* 34, 525–527 (2016). [PubMed: 27043002]
59. 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). [PubMed: 20513432]

60. Quinlan AR & Hall IM BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842 (2010). [PubMed: 20110278]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Fig. 1 | NK cells undergo dynamic chromatin changes during viral infection.

a, Schematic of timepoints assayed for time course of ATAC-seq ($n = 3$ or 4 samples per d) and RNA-seq profiling ($n = 2$ samples per d). **b**, Number of DA (false discovery rate) $(FDR) < 0.05$ regions that either gain (red) or lose (blue) chromatin accessibility at indicated transition timepoints. **c**, Absolute numbers and proportions of all DA regions versus high-fold change (FC; absolute $\log_2(FC) > 1$) regions. **d**, Shown are line graphs (left) and heatmap (right) of high-FC peaks. Line plots showing mean (red line) and s.d. (gray ribbon) of mean-centered normalized $\log_2$ values for each high-FC cluster. Heatmap is hierarchically clustered based on all high-FC $\log_2$ peak counts (see Supplementary Fig. 1e) and shows the top 10% most variable regions within each cluster, with stable and transient clusters as indicated. **e**, Heatmap of top 20 most enriched pathways of any high-FC cluster shown as $-\log_{10}$ raw binomial P values, as calculated by Genomic Regions Enrichment of Annotations Tool (GREAT). TCR, T cell antigen receptor.

Fig. 2 |. NK cells show epigenetic and transcriptional coordination during early stages of viral infection.

a, PCA of ATAC-seq and RNA-seq data on all atlas peaks and all genes. **b**, Hierarchical clustering on scaled Euclidean distances of normalized $\log_2$ values of ATAC-seq and RNA-seq data. **c**, Heatmap of significant Spearman correlations (FDR < 0.05) for indicated gene sets (y axis). All coefficients shown are greater than 0.4. **d**, Scatter plots of average ATAC-seq and RNA-seq $\log_2(\text{FC})$ for the top six gene sets ranked by Spearman coefficient, categorized by ATAC-seq peak type. ATAC-seq FC reflects a summarized value calculated from peak types and transition days (shown as numerator/denominator), as indicated within parentheses.

Fig. 3 | STAT4 and STAT1 mediate chromatin remodeling in distinct manners.

a, Bar plot showing proportions of STAT-occupied regions categorized by peak type. **b**, Number of STAT-occupied DA regions (FDR < 0.05) that either gain (red) or lose (blue) accessibility at indicated transition timepoints. **c**, MA plots of DA regions categorized by peak type comparing *Stat4*^{-/-} (blue) or *Stat1*^{-/-} (green) to wild-type (WT) 2 d postinfection (*n* = 2 or 3 per condition). **d**, Scatter plot of $\log_2(FC)$ of differentially expressed (DE) genes (FDR < 0.05) comparing *Stat4*^{-/-} (blue) or *Stat1*^{-/-} (green) to WT 2 d postinfection (*n* = 2 or 3 per condition). **e**, Venn diagram of overlapping gene loci that are occupied (top) by indicated STAT protein, and DE (left) and DA (right) upon STAT deficiency. Darker center circle indicates number of DE genes with overlapping DA and occupied peak regions. PI, postinfection. **f**, Heatmap of gene expression from overlapping genes highlighted in **e**. **g**, Genomic tracks of mapped STAT4 ChIP-seq and *Stat4*^{-/-} vs. WT ATAC-seq reads, zoomed

in on regions of interest. *Y* axis depicts normalized counts, while *x* axis displays genomic axis with scale bar. Bar plots quantify number of reads within highlighted peak regions (dashed lines), shown as mean $\pm$ s.d.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Fig. 4 |. Memory NK cells are epigenetically distinct from naive NK cells.

a, Enrichment of pathway terms on filtered DA regions ($FDR < 0.05$, absolute $\log_2(FC) > 0.5$, assigned to genes that show significant changes by RNA-seq at any transition timepoint) between memory and naive NK cells shown as bar plots of $-\log_{10}$ binomial raw P value. **b**, Select gene loci from reported pathways shown in **a**. Left: genomic tracks, with dashed lines indicating DA peak regions. Y axis depicts normalized counts, while x axis displays genomic axis with scale bar. Right: gene expression as normalized $\log_2$ counts. Line represents the mean. **c**, Enriched de novo motifs found by Hypergeometric Optimization of Motif Enrichment (HOMER) among filtered DA regions, as described in **a**. ISRE data from GSE23622; TCF-LEF data from SRA012054; NF- κ B data from MA0778.1 of JASPAR database (<http://jaspar.genereg.net/>); ISRE, interferon-stimulated response element; IRF, interferon-regulatory factor; HMG, high-mobility group transcription. **d**, Metacoverage of DA median normalized tag counts surrounding de novo motif regions listed in **c** found among all filtered DA regions. **e**, Select DA gene loci that contain motif instances, with actual sequence in black and de novo motif sequence below. Y axis depicts normalized counts, while x axis displays genomic axis with scale bar. BCR, B cell antigen receptor.

Fig. 5 |. Naive NK cells resemble memory CD8⁺ T cells.

a, PCA of ATAC-seq on all common cell-normalized log₂ values between Ly49H⁺ NK cells and m45-specific CD8⁺ T-cells ($n = 2$ or 3 per condition) during MCMV infection. **b**, Hierarchical clustering of Euclidean distances calculated from cell-normalized log₂ values. **c**, Heatmap of CD8⁺ T cells and NK cells depicting differentially accessible regions based on CD8⁺ T cell comparisons.

Fig. 6 |. Memory NK cells and CD8⁺ T cells possess common epigenetic signatures.

a, Heatmap of $\log_2(\text{FC})$ comparing chromatin accessibility (left) and transcription (right) of memory and naive timepoints for NK cells and CD8⁺ T cells. **b**, Fold change of transcription and accessibility from commonly regulated peaks comparing memory and naive cell states in both NK cells and CD8⁺ T cells. DA regions ($\text{FDR} < 0.05$, absolute $\log_2(\text{FC}) > 0.5$) are represented by number of dots, with colors showing magnitude and direction of FC. Extreme outliers of accessibility FC were capped at the 75th percentile $\pm 3 \times$ the interquartile range. Position of bottom circle represents transcriptional FC of differentially expressed genes ($\text{FDR} < 0.05$, absolute $\log_2(\text{FC}) > 1$). Shown are genes with average transcripts-per-million (TPM) counts greater than 5 across naive, effector, and memory timepoints. **c**, Table of top enriched motifs from all common DA regions shown in **a** that are either higher in memory or lower in memory compared to naive. AP-1 data from [GSE36099](#); TCF-LEF data from [MA0769.1](#) of JASPAR database; bZIP, basic leucine zipper domain. **d**, Metacoverage of normalized tag counts surrounding de novo motif regions listed in **c** among all common DA regions. **e**, Select DA gene loci that contain motif instances with genomic sequence in black

and de novo motif sequence below. *Y* axis depicts normalized counts, while *x* axis displays genomic axis with scale bar. Bar plots quantify number of reads overlapping the motif center (dashed line), shown as mean $\pm$ s.d.