



ORIGINAL RESEARCH

Landscape and Dynamics of the Transcriptional Regulatory Network During Natural Killer Cell Differentiation



Kun Li^{1,#}, Yang Wu^{1,#}, Young Li¹, Qiaoni Yu¹, Zhigang Tian^{1,2},
Haiming Wei^{1,2,*}, Kun Qu^{1,2,3,*}

¹ Department of Oncology, The First Affiliated Hospital of USTC, Division of Molecular Medicine, Hefei National Laboratory for Physical Sciences at Microscale, School of Basic Medical Sciences, Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei 230021, China

² CAS Center for Excellence in Molecular Cell Sciences, The CAS Key Laboratory of Innate Immunity and Chronic Disease, University of Science and Technology of China, Hefei 230027, China

³ School of Data Science, University of Science and Technology of China, Hefei 230026, China

Received 1 August 2018; revised 10 December 2018; accepted 4 March 2019

Available online 30 December 2020

Handled by Yi Xing

KEYWORDS

NK cell;
ATAC-seq;
Programmed differentiation;
FOSL2;
EGR2;
Dynamic regulatory network

Abstract Natural killer (NK) cells are essential in controlling cancer and infection. However, little is known about the dynamics of the transcriptional regulatory machinery during **NK cell** differentiation. In this study, we applied the assay of transposase accessible chromatin with sequencing (**ATAC-seq**) technique in a home-developed *in vitro* NK cell differentiation system. Analysis of ATAC-seq data illustrated two distinct transcription factor (TF) clusters that dynamically regulate NK cell differentiation. Moreover, two TFs from the second cluster, FOS-like 2 (**FOSL2**) and early growth response 2 (**EGR2**), were identified as novel essential TFs that control NK cell maturation and function. Knocking down either of these two TFs significantly impacted NK cell differentiation. Finally, we constructed a genome-wide transcriptional regulatory network that provides a better understanding of the regulatory dynamics during NK cell differentiation.

Introduction

Natural killer (NK) cells are innate lymphocytes that survey the environment and protect the host from infected or cancerous cells. As their name implies, NK cells represent an early line of defense during pathogen invasion by directly killing infected cells and secreting inflammatory cytokines [1]. Additionally, NK cell-based immunotherapy has become an

* Corresponding authors.

E-mail: ustcwhm@ustc.edu.cn (Wei H), qkun@ustc.edu.cn (Qu K).

Equal contribution.

Peer review under responsibility of Beijing Institute of Genomics, Chinese Academy of Sciences and Genetics Society of China.

<https://doi.org/10.1016/j.gpb.2020.12.003>

1672-0229 © 2020 The Authors. Published by Elsevier B.V. and Science Press on behalf of Beijing Institute of Genomics, Chinese Academy of Sciences and Genetics Society of China.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

emerging force in cancer treatment and will play an essential role in the future treatment [2–6]. For instance, adoptive NK cell immunotherapy has become increasingly popular due to its capability to induce graft-versus-leukemia effects without causing graft-versus-host disease in patients [7]. Therefore, major efforts are currently underway to fully utilize the anti-tumor properties of NK cells in the clinic. In addition, a variety of methods have been implemented to expand NK cells, establish a microenvironment that favors NK cell activity, and redirect the activity of NK cells onto tumor cells [8]. On the other hand, many researchers use umbilical cord blood (UCB) CD34⁺ cells to produce abundant NK cells [9–18] for clinical application without feeding cells by adding various cytokines [10]. In our previous work, we have developed a method to obtain sufficient functional NK cells by simply adding a mixture of cytokines and provide a mechanism by which NK cells can be used to treat leukemia [19]. The use of NK cells for immunotherapy relies on the presence of a great number of NK cells with optimal cytotoxic activity [4]; therefore, a comprehensive understanding of the regulatory circuits during NK cell differentiation is particularly important for boosting the efficacy of clinical treatments. However, the mechanisms underlying NK cell differentiation are not well understood. Studies have shown that transcriptional factors (TFs) play an essential role in driving NK cell maturation, and many TFs have been well studied in this process [1]. Additionally, it is known that different TFs play various roles at distinct stages of differentiation [1]. For example, PU.1 is a TF that is known to drive hematopoietic stem cell (HSC) differentiation into the earliest myeloid and lymphoid progenitors [20], whereas T-bet is an essential TF in the control of NK cell maturation and IFN- γ secretion [21]. However, how TFs work in concert to enforce the NK cell phenotype is not clear.

Assay of transposase accessible chromatin with sequencing (ATAC-seq), a newly developed epigenomic profiling technique [22], has been widely used to profile the epigenetic landscapes of cells at specific stages of interest and thereby delineates the underlying regulatory mechanisms of gene expression. For example, previous reports using ATAC-seq identified an exhaustion-specific enhancer that regulates PD-1 expression, thereby elucidating the regulatory mechanism of gene expression in exhausted CD8⁺ T cells [23,24]. Alternately, ATAC-seq analysis of pure cancer cell populations of human small cell lung cancer (SCLC) identified a novel TF, nuclear factor I B (NFIB), which is necessary and sufficient to drive the metastatic ability of SCLC cells [25]. ATAC-seq has also enabled researchers to track the epigenomic state changes in patient-derived immune cells [26], and to survey how the personal regulomes of the cutaneous T cell lymphoma patients determine their responses towards histone deacetylase inhibitor (HDACi) anti-cancer drugs [27]. More recently, ATAC-seq was applied to better understand the control of NK cells in innate immune memory during infection [28], illustrating the importance of the topic as well as the power of the technique used in this study.

Here, we have developed a systematic method to characterize the chromatin accessibility and regulatory network dynamics during NK cell differentiation from UCB CD34⁺ HSCs based on ATAC-seq. Motif and enrichment analyses from the Hypergeometric Optimization of Motif Enrichment (HOMER) [29] and Genomica [27] show that many TFs play important roles during NK cell differentiation. By integrating

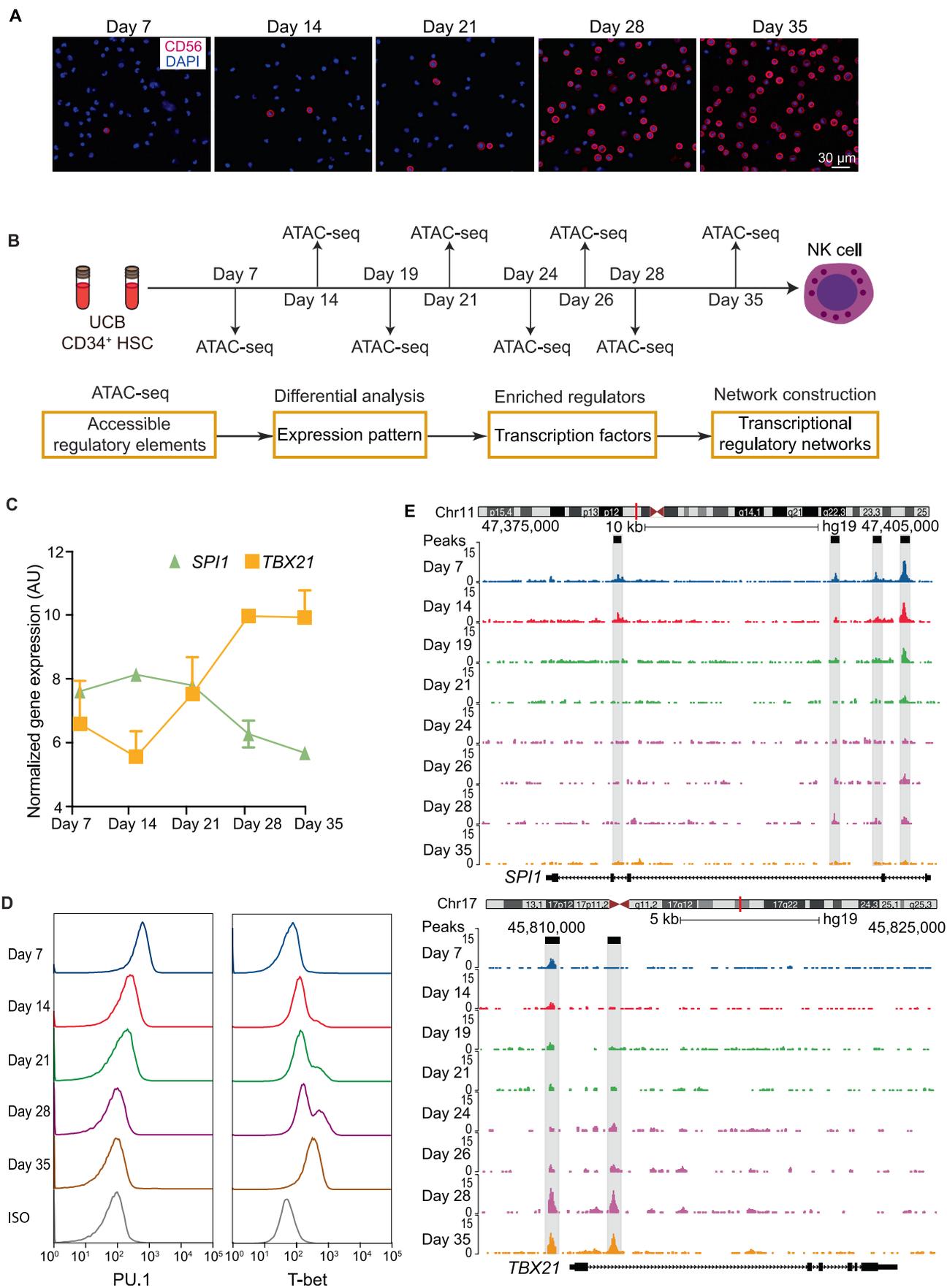
gene expression profiles from our previous study [19], two novel TFs, FOS-like 2 (FOSL2) and early growth response 2 (EGR2), were identified to be essential to drive NK cell maturation.

Results

Landscape of DNA accessibility during NK cell differentiation

To elucidate the regulatory networks during NK cell differentiation, we developed a culture procedure to obtain differentiated NK cells from UCB CD34⁺ cells using a cocktail of cytokines [19]. The differentiation process took 35 days. Interestingly, after culturing of cord blood stem cells for 3 weeks, the proportion of NK cells rapidly increased from 5% on day 21 to approximately 60% on day 28, and peaked at close to 100% on day 35 (Figure 1A, Figure S1A) [19]. To elucidate the dynamic changes of transcriptional regulation during NK cell differentiation, we interrogated the landscape of chromatin accessibility using ATAC-seq at 8 different time points, with 2 replicates each along the process (Figure 1B). Multiple bioinformatics analyses (see Materials and methods) were then performed to obtain the differentially accessed chromatin sites, enriched TFs, and genome-wide regulatory elements. At least 50,000 cells were obtained from each sample, resulting in more than 78 million reads on average with a total of 1260 million reads for all time points (Table S1). From this dataset, 143,570 accessible DNA sites were identified ($P < 10^{-7}$, FDR $< 10^{-7}$). Transcription start site (TSS) enrichment score and ATAC-seq signal analysis indicated that the dataset was of high quality with a strong signal-to-background noise ratio and expected fragment length distribution (Figure S1B–E). Moreover, the dynamics of the chromatin accessibility around the known surface marker genes were consistent with changes of their gene expression levels during NK cell differentiation (Figure S1D) [30,31]. Pearson correlation analysis on the biological replicates at each time point also showed that our ATAC-seq profiles were highly reproducible (Figure S2A and B). Within all the accessible sites, only a very small portion (6.48%) were conserved across all stages (Figure S3A), while the majority of peaks changed over time, suggesting significant chromatin dynamics during NK cell differentiation. Peaks that were detected at all stages were enriched in Gene Ontology (GO) terms such as RNA transcription and other functions to maintain basic physiological activities (Figure S3B). In contrast, more and more immune response-related GO terms were enriched as the time went by, suggesting the activation of critical genes that govern NK cell function.

Several TFs are known to regulate NK cell differentiation, such as PU.1/Spi-1 proto-oncogene (SPI1) and T-bet/T-box transcription factor 21 (TBX21). The former is a NK cell repressor and the latter an activator. Gene expression profiles from our previous study (GEO: GSE87787) [19] indicated down-regulation of the *SPI1* gene and up-regulation of the *TBX21* gene during NK cell differentiation (Figure 1C). The levels of proteins encoded by these two genes were evaluated using flow cytometry, which suggested the same trends (Figure 1D). ATAC-seq analysis revealed four peaks at the intronic enhancers of *SPI1* and two peaks at the promoter and intronic enhancer of *TBX21*, respectively (Figure 1E). Accessible chromatin sites around the *SPI1* gene clearly became



unavailable and those around the *TBX21* gene gradually became available during the process, supporting the notion that the epigenetic dynamics of key regulators are consistent with their corresponding gene and protein expression levels. The consistency between the epigenetic and gene expression profiles of TFs known to drive NK cell differentiation, such as GATA binding protein 3 (*GATA3*) [1] and eomesodermin (*EOMES*) [1], is shown in Figure S1E, indicating the feasibility of predicting transcriptional regulatory networks from ATAC-seq profiles.

Epigenomic signatures of different stages during NK cell differentiation

To determine the differences in regulatory DNA activity among different stages during NK cell differentiation, we performed pairwise comparisons of the ATAC-seq signals between the corresponding samples, together with intrinsic analysis [32], a method that highlights the elements with varied accessibility across individuals but not between replicate samples from the same individual (File S1). We discovered 6401 peaks showing differentially accessed DNA sites across the genome, which were categorized into three distinct clusters via unsupervised hierarchical clustering (Figure 2A, Figure S4A and B). Principal component analysis (PCA) of all samples also illustrated three distinct clusters, which is consistent with the time process of NK cell differentiation representing the early, interim, and late stages of the entire process (Figure S4C). GO analysis of these peaks was performed in the Genomic Regions Enrichment of Annotations Tool (GREAT) [33]. Cluster I comprises 1584 elements that are more accessible at the early stage (days 7–21) of differentiation. Peaks in cluster I are significantly enriched in the GO terms of metabolic processes that are necessary for cell viability, proliferation, and differentiation (Figure 2B, upper panel; $10^{-8} < P < 10^{-4}$). Cluster II comprises 4461 elements that are highly accessible at the interim and late stages (days 24–35) of NK cell differentiation. GREAT analysis revealed that peaks in cluster II are strongly enriched ($P < 10^{-50}$) in immune-relevant GO terms, such as immune system process, immune response, and immune system development, among others (Figure 2B, lower panel), indicating that the epigenetic states of functional immune cell-specific genes were activated throughout the process. Cluster III consists of only 356 peaks, with no enrichment for GO terms, and were therefore not included for downstream analyses.

We next examined whether the DNA accessibility signatures at different stages are correlated with the expression pro-

files of corresponding genes. By comparing the ATAC-seq profiles with the genome-wide microarray data during NK cell differentiation, we found that genes that gained chromatin accessibility (cluster II) showed significant increases in expression (Figure 2C, right panel; $P < 0.001$), while genes that lost chromatin accessibility (cluster I) had decreased expression (Figure 2C, left panel; $P < 0.001$), indicating a high correlation between epigenetic and RNA profiles.

Chromatin structures and epigenetic modifications are known to regulate gene expression [34]. However, the chronological order of the dynamics of the chromatin accessibility, gene expression, and cell phenotype has not yet been well studied. Here, we integrated the ATAC-seq signals, microarray profile, and percentage of NK cell counts to examine the temporal dynamics of these three features. Interestingly, we noticed that the accessibility of mature NK cell-specific peaks (cluster II in Figure 2A) started to increase on day 14, the expression of NK cell-specific genes was turned on approximately two days later, while the percentage of NK cell counts started to increase after day 21 (Figure 2D). These results suggest a clear chronological order of the changes in chromatin structure, gene expression, and cell population during NK cell differentiation.

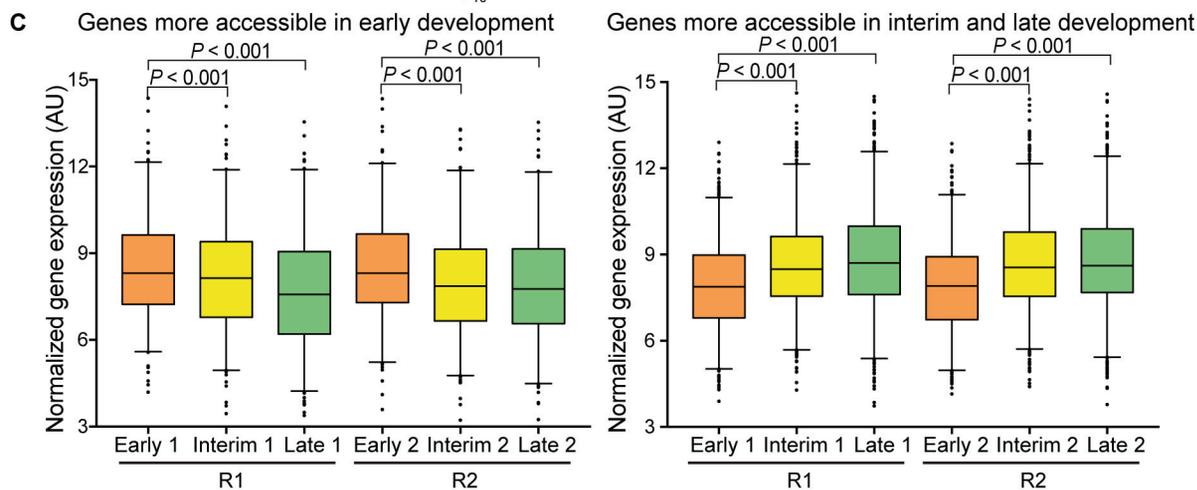
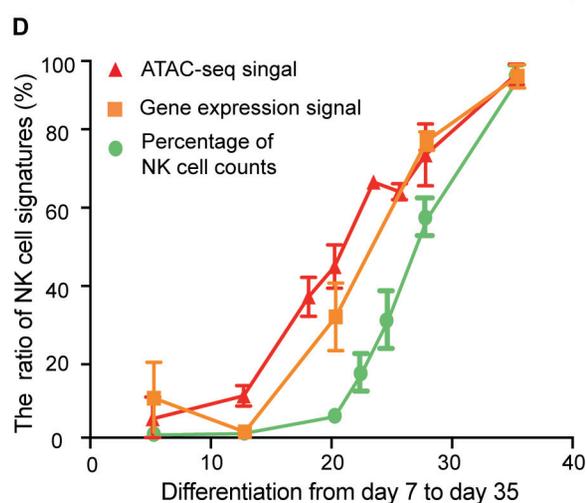
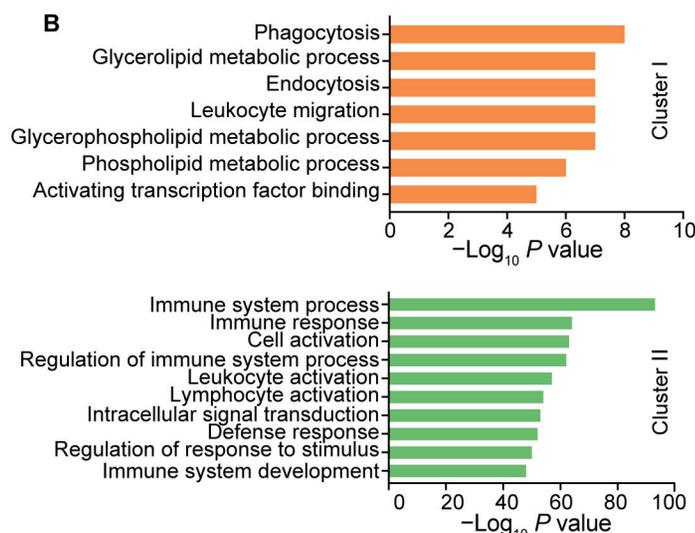
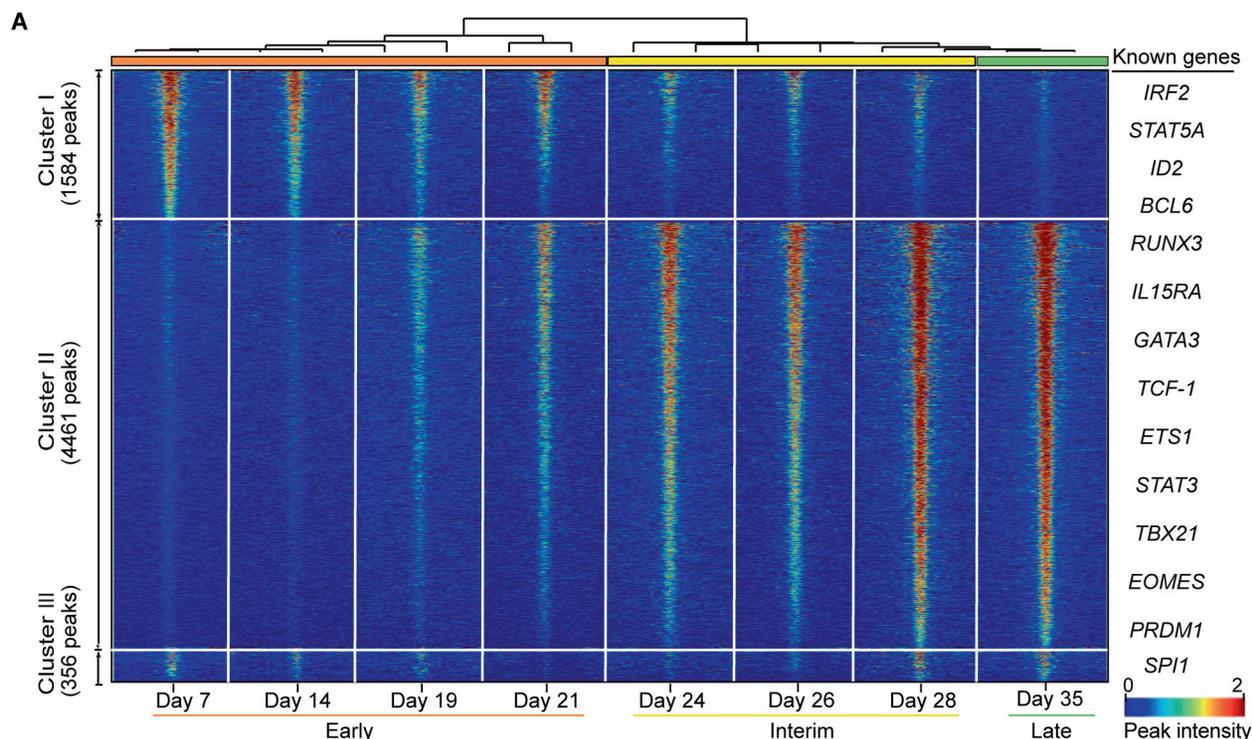
TF occupancy network during NK cell differentiation

TFs bind to cognate DNA sequences with patterns termed motifs. Therefore, we could predict TF occupancy on chromatin using DNA accessibility data from ATAC-seq for constructing regulatory networks [26]. To identify potential drivers of NK cell differentiation, we screened TFs that were enriched at accessible elements in clusters I and II using HOMER in the search modes with default (Figure 3A) or setting background (Figure S5A). Our finding suggests that NK cell differentiation and maturation require a variety of TFs. TFs enriched at cluster I peaks are potential regulators at the early stage of NK cell differentiation, while those enriched at cluster II peaks could be critical regulators at the interrim and late stages. We found several TF families that were significantly enriched ($P < 10^{-10}$). Many of them are known to be important for NK cell differentiation, such as the Runt-related transcription factor (*RUNX*) families, ETS proto-oncogene (*ETS*) families [20,35], and CCAAT enhancer-binding protein (*CEBP*) families [36], supporting the reliability of this method to detect critical regulators. For instance, PU.1 is widely expressed and controls multiple stages of bone marrow and lymphocyte differentiation in a variety of hematopoietic-derived lineages [37–41]. An decrease in the number of NK-



Figure 1 DNA accessibility during NK cell differentiation

A. Confocal microscopy images of membrane CD56 (red) in the cultured cells at day 7, day 14, day 21, day 28, and day 35. Scale bar, 30 μm . Nuclei are stained with DAPI. **B.** Schematic representation of the overall experimental design of this study. Chromosome opening at different time points were assessed using ATAC-seq data. The bioinformatics pipeline for data analysis is shown in the bottom. **C.** The gene expression profiles of *SPI1* and *TBX21* at different stages of NK cell differentiation. **D.** Flow cytometric measurement of PU.1/*SPI1* and T-bet/*TBX21* expression in cultured cells during a 35-day time course. The X-axis indicates the fluorescence intensity, and the Y-axis indicates the cell number density. **E.** Normalized ATAC-seq profiles of the *SPI1* (top) and *TBX21* (bottom) gene loci at different stages during NK cell differentiation. ATAC-seq signals were obtained from the UCSC Genome Browser. NK, natural killer; UCB, umbilical cord blood; HSC, hematopoietic stem cell; ATAC-seq, assay of transposase accessible chromatin with sequencing; *SPI1*, Spi-1 proto-oncogene; *TBX21*, T-box transcription factor 21; AU, arbitrary unit; ISO, isotype control.



cell progenitor (NKP) and immature NK (iNK) cells has been detected in chimeric mice [20], indicating that PU.1 may play an important role during the early stage of NK cell differentiation. Several known TFs, such as RUNX [35], E2A immunoglobulin enhancer binding factors E12/E47 (E2A) [42], T-bet [21], signal transducer and activator of transcription 5 (STAT5) [43], and EOMES [21], are also enriched at the interim and late stages. The most enriched TFs at cluster II peaks belong to the RUNX family: RUNX1, RUNX, and RUNX2 (Figure S5A), which are key regulators of lymphocyte lineage-specific gene expression [44].

Motif search of differential ATAC-seq peaks could provide information about the transcriptional regulatory network. However, one caveat of this method is the inability to distinguish similar binding motifs of TF family members. Thus, we sought to apply the Genomica's module map algorithm and "TF footprint" analysis to better predict TF occupancy on accessible sites by integrating the ATAC-seq and gene expression microarray profiles (Figure 3).

In total, 242 vertebrate TF motifs were obtained from the JASPAR database [45]. We then used Genomica [27] to produce a time point-specific TF occupancy network (Figure 3B). This analysis revealed distinct patterns of TF accessing to DNA at different time points. Many known TFs were enriched, such as STAT5, which is an IL-15 downstream signaling molecule and is indispensable throughout the lifetime of NK cells. Deficiency of STAT5a/b has been reported to result in complete elimination of NK cells, demonstrating the important and non-redundant effects of STAT5 [43]. The JAK-STAT pathway has also been shown to be an important signaling pathway used by various cytokines and growth factors [46]. The interferon regulatory factor (IRF) family regulates a variety of processes, including hematopoietic development, tumorigenesis, host immunization, and pathogen defense [47,48]. IRF2 is required to maintain the normal differentiation of NK cells in a cell-intrinsic manner [49,50], and IRF2-deficient NK cells show reduced levels of maturation markers and IFN- γ production during stimulation [49,50]. T-bet and EOMES are members of the T-box family and are known to control different aspects of NK cell differentiation and maturation [21,51,52].

In addition, several novel TFs were also enriched, such as FOSL2 and EGR2, suggesting that they are potential regulators of NK cell differentiation. These TFs were assessed along with the other enriched TF families from the motif analysis to identify which family members were expressed or differentially

expressed during NK cell differentiation to further filter out true regulators. At each time point, we plotted both the expression value (colored from red to green to represent high to low expression in the gene microarray profile) and the motif enrichment score (shown by the circle size representing the $-\log P$ value of the enrichment) in the same figure (Figure 3C). We observed that known regulators were highly expressed and enriched at different stages, same as FOSL2 and EGR2. Through the integral analyses (Figure 3A–C), we predict FOSL2 and EGR2 as potential regulators.

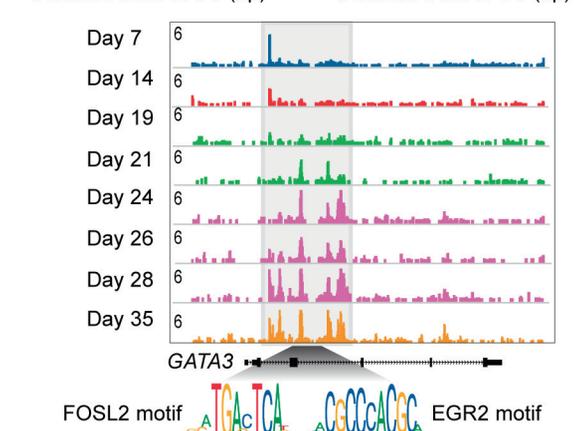
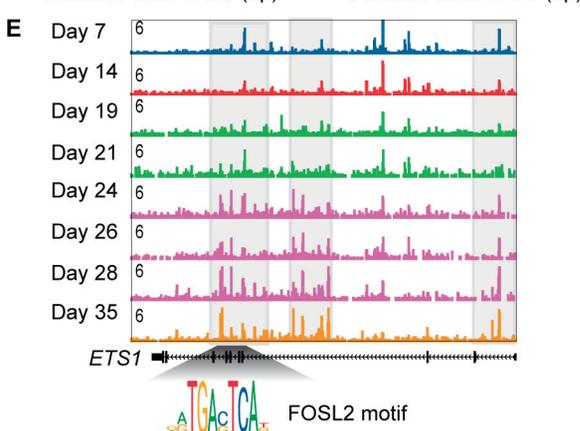
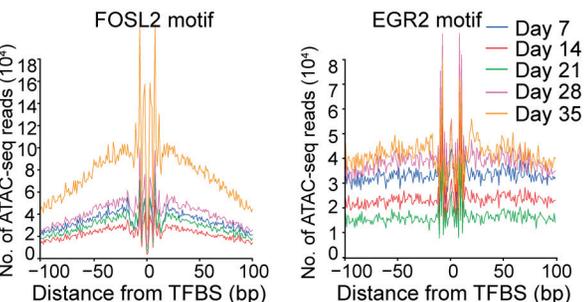
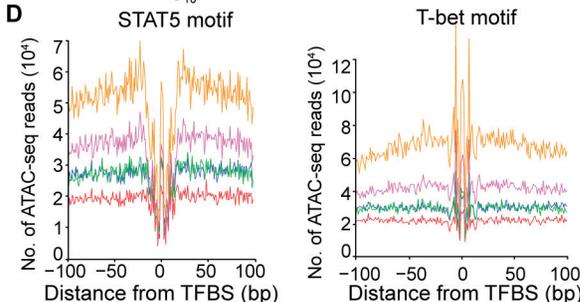
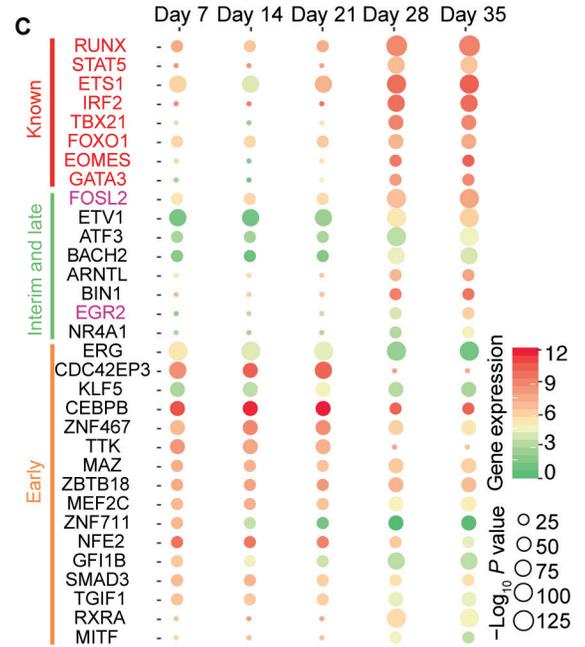
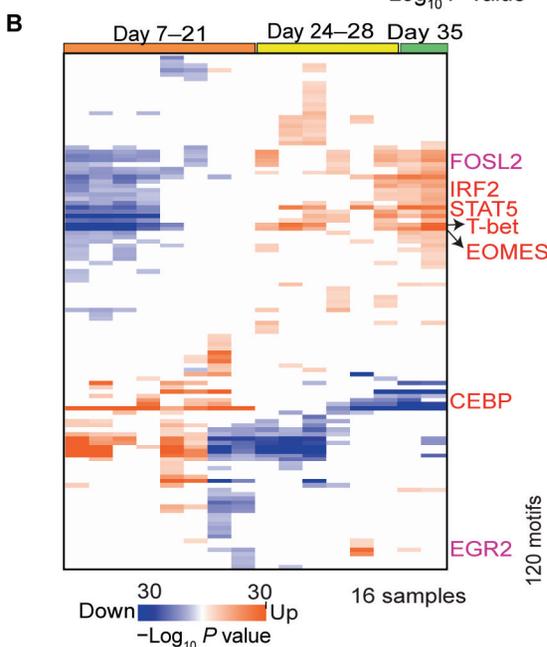
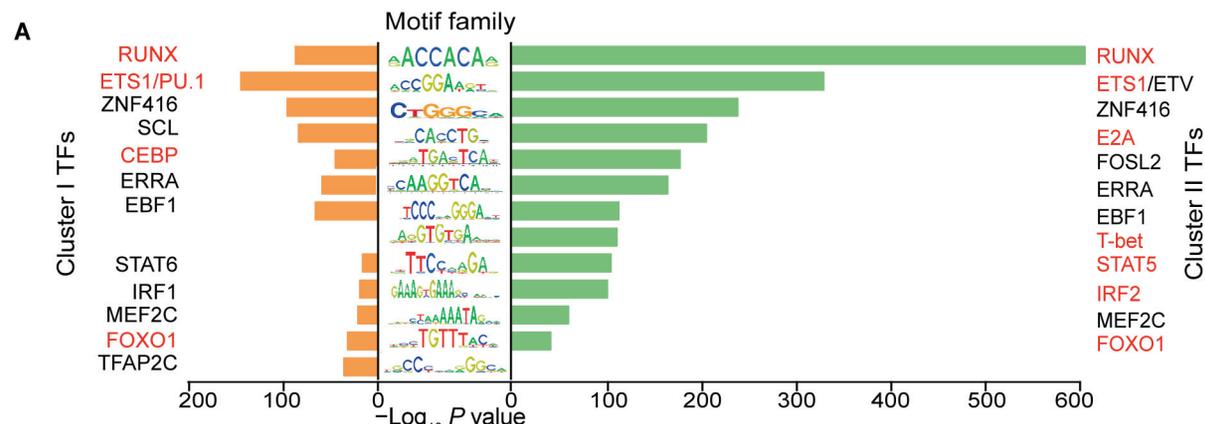
DNA sequences that are directly occupied by DNA-binding proteins are protected from transposition, and therefore the resulting sequence "footprint" reveals the presence of a DNA-binding protein at each site, analogous to DNase digestion footprints. The TF "footprint" analysis of our ATAC-seq profile provided further evidence of direct occupancy of a TF candidate on genomic DNA and thus refined the prediction of potential regulators. We illustrated the "footprints" of two known regulators, STAT5 and T-bet, and observed higher DNA accessibility and deeper "footprints" flanking their motifs in the interim and late stages than those in the early stage during NK cell differentiation (Figure 3D). Similarly, "footprints" of the TFs FOSL2 and EGR2 were also deeper and more accessible at the interim and late stages. These data suggest that not only the binding motifs of these two TFs are enriched at stage-specific peaks, but also these TFs most likely physically bind to these accessible chromatin sites, indicating that they are functional regulators of NK cell differentiation (Figure 3D). Overall, the results from footprint analysis agree with those from the HOMER and Genomica's motif enrichment analyses.

Genes that are regulated by any TF of interest can be predicted by combining the TF motif and "footprint" analyses. We thus predicted the genes regulated by TFs including RUNX, T-bet, FOSL2, and EGR2, and integrated the gene expression profiles at each time point (Figure S5B; Table S2). We found that genes regulated by each of these TFs were also highly expressed at the interim and late stages and were significantly enriched in GO terms of immune system construction and other related functions. TF ETS proto-oncogene 1 (ETS1) was reported to drive early stages of NK cell differentiation [53], and we found that there was a FOSL2-binding site in a dynamically accessible site on the gene body of *ETS1*, suggesting FOSL2 might regulate NK cell differentiation through *ETS1* (Figure 3E, left panel). TF GATA3 was found to regulate liver-specific NK cells, IFN- γ



Figure 2 Differential epigenetic regulation elements during NK cell differentiation

A. Heatmap of 6401 differentially accessed regulatory elements during NK cell differentiation. Each column is a sample, and each row is a peak. Samples and peaks are organized by two-dimensional unsupervised hierarchical clustering. Color scale shows the relative ATAC-seq peak intensity centered by each peak summit. Bottom: samples at all time points are categorized into three groups: early stage (days 7–21; orange), interim stage (days 24–28; yellow), and late stage (day 35; green). Samples from the same stage are labeled with the same color. Left: differential peaks are categorized into three clusters. **B.** The top 10 most significant GO terms enriched in cluster I (upper panel) and cluster II (lower panel) peaks. **C.** Box plots showing expression levels of genes in different clusters during NK cell differentiation. Left: genes in cluster I show higher expression in the early stage. Right: genes in cluster II show higher expression in the interim and late stages. R1 and R2 represent biological replicates 1 and 2, respectively. P values are estimated from one-way ANOVA. **D.** The changes in ATAC-seq signal (red), gene expression signal (orange), and percentage of NK cell counts (green) at different time points during NK cell differentiation.



production, and T-bet expression in mice [54]. Similarly, we found that both FOSL2 and EGR2 bind to the gene body of *GATA3* (Figure 3E, right panel), suggesting that the TFs FOSL2 and EGR2 may also regulate NK cell differentiation through regulation of *GATA3*.

FOSL2 and EGR2 affect NK cell differentiation

FOSL2 belongs to the Fos gene family, which encodes a leucine zipper protein and forms the AP-1 TF complex by dimerization with the JUN family. Thus, FOS proteins have been suggested as key regulators of transformation, differentiation, and cell proliferation. The GO annotations of *FOSL2* include sequence-specific DNA binding, TF activity, and RNA polymerase II specific DNA binding. A previous report has shown that FOSL2 is constitutively expressed in adult T-cell leukemia (ATL), up-regulates *CCR4* expression, and promotes ATL cell growth, together with JunD proto-oncogene, AP-1 transcription factor subunit (JUND) [55]. *EGR2* encodes a protein that contains three tandem C2H2-type zinc fingers. GO annotations of this gene include ligase activity, sequence-specific DNA binding, and TF activity. Previous reports have shown that EGR2 regulate T-cell and B-cell function in homeostasis and adaptive immune responses by controlling inflammation and promoting antigen receptor signaling [56,57].

Since their regulatory functions in NK cell differentiation have not been well characterized, we performed loss-of-function experiments to assess the effects of *FOSL2* and *EGR2* on NK cell differentiation. We first validated their expression levels with real-time PCR, and found that their

expression gradually increased in later stages of NK cell differentiation, consistent with our bioinformatics analysis (Figure S6A). Subsequently, we infected the cultured cells with TF-specific shRNA- and control shRNA-containing lentiviruses, which are represented by GFP expression (Figure 4A). We then sorted the GFP-positive cells and observed dramatically reduced expression levels of FOSL2 and EGR2 (Figure 4B), suggesting positive targeting of the TF-specific shRNAs. During NK cell differentiation, we observed a nearly 30% reduction in the proportion of differentiated NK cells in GFP-positive cells at day 28 and day 35 (Figure 4C and D). However, in the non-successfully transfected (GFP-negative) cells, the proportion of differentiated NK cells was not affected at the same time point (Figure 4E and F). These results indicate that knockdown of *FOSL2* and *EGR2* expression, but not viral infection, inhibits NK cell differentiation, suggesting that FOSL2 and EGR2 are necessary for NK cell differentiation. We then tested another important marker of NK cell maturation CD11b [1], and found that it was significantly reduced in GFP-positive NK cells, suggesting that FOSL2 and EGR2 might affect NK cell maturation (Figure 4G). Overall, we predicted these two TFs FOSL2 and EGR2 as key regulators based on ATAC-seq and gene microarray profile analyses, and then experimentally verified that they indeed affect NK cell differentiation.

In addition, we attempted to identify the signaling pathways through which FOSL2 and EGR2 were involved in driving NK cell differentiation by performing module map analysis of all the differential peaks. Eighteen modules were identified according to the patterns of their accessibility (Table S3), sev-



Figure 3 TFs enriched at different stages of NK cell differentiation

A. TF motifs enriched in cluster I (left) and cluster II (right) peaks, with enrichment *P* values estimated from HOMER. TFs known to regulate NK cell differentiation are colored in red. **B.** Enrichment of known TF motifs in differentially accessible elements in all samples. Each row is a TF motif and each column is a sample. Values in the matrix represent the significance levels expressed as $-\log_{10} P$ value of the enrichment estimated from Genomica. Red in the matrix indicates that the motif is enriched in the corresponding sample, whereas blue in the matrix indicates depletion. Red texts on the right indicate known key TFs that regulates NK cell development, while pink texts indicate new TFs whose functions will be experimentally tested later (see Figure 4). **C.** Prediction of TFs that may regulate NK cell differentiation. TFs known to regulate NK cell differentiation are shown at the top, TFs predicted to regulate the early stage of the differentiation process are shown in the bottom, and those predicted to regulate the interim and late stages are shown in the middle. The color of each circle represents the expression level of the gene encoding the corresponding TF, while the size of the circle represents the significance of the motifs estimated by *P* values. **D.** Visualization of the ATAC-seq footprints for STAT5, T-bet, FOSL2, and EGR2 motifs at five indicated time points during NK cell differentiation. ATAC-seq signals across all these motif binding sites in the genome were aligned on the motif and averaged. **E.** Normalized ATAC-seq profiles of the *ETS1* (left) and *GATA3* (right) gene loci at indicated time points during NK cell differentiation. ATAC-seq signals were obtained from the UCSC Genome Browser. The gray blocks indicate genomic loci that are gradually more accessible during NK cell differentiation. The locations of binding motifs for FOSL2 and EGR2 are depicted at the bottom. RUNX, Runt-related transcription factor; ETS1, ETS proto-oncogene 1; ETV, ETS variant transcription factor; ZNF416, zinc finger protein 416; SCL, TAL bHLH transcription factor 1; E2A, E2A immunoglobulin enhancer binding factors E12/E47; CEBP, CCAAT enhancer-binding protein; FOSL2, FOS-like 2; ERRA, estrogen related receptor alpha; EBF1, EBF transcription factor 1; STAT, signal transducer and activator of transcription; IRF, interferon regulatory factor; MEF2C, myocyte enhancer factor 2C; FOXO1, forkhead box O1; TFAP2C, transcription factor AP-2 gamma; EOMES, eomesodermin; EGR2, early growth response 2; GATA3, GATA binding protein 3; ATF3, activating transcription factor 3; BACH2, BTB domain and CNC homolog 2; ARNTL, aryl hydrocarbon receptor nuclear translocator like; BIN1, bridging integrator 1; NR4A1, nuclear receptor subfamily 4 group A member 1; ERG, ETS transcription factor ERG; CDC42EP3, CDC42 effector protein 3; KLF5, Kruppel like factor 5; CEBPB, CCAAT enhancer binding protein beta; ZNF467, zinc finger protein 467; TTK, TTK protein kinase; MAZ, MYC associated zinc finger protein; ZBTB18, zinc finger and BTB domain containing 18; ZNF711, zinc finger protein 711; NFE2, nuclear factor, erythroid 2; GFI1B, growth factor independent 1B transcriptional repressor; SMAD3, SMAD family member 3; TGIF1, TGFB induced factor homeobox 1; RXRA, retinoid X receptor alpha; MITEF, melanocyte inducing transcription factor; TFBS, TF binding site.

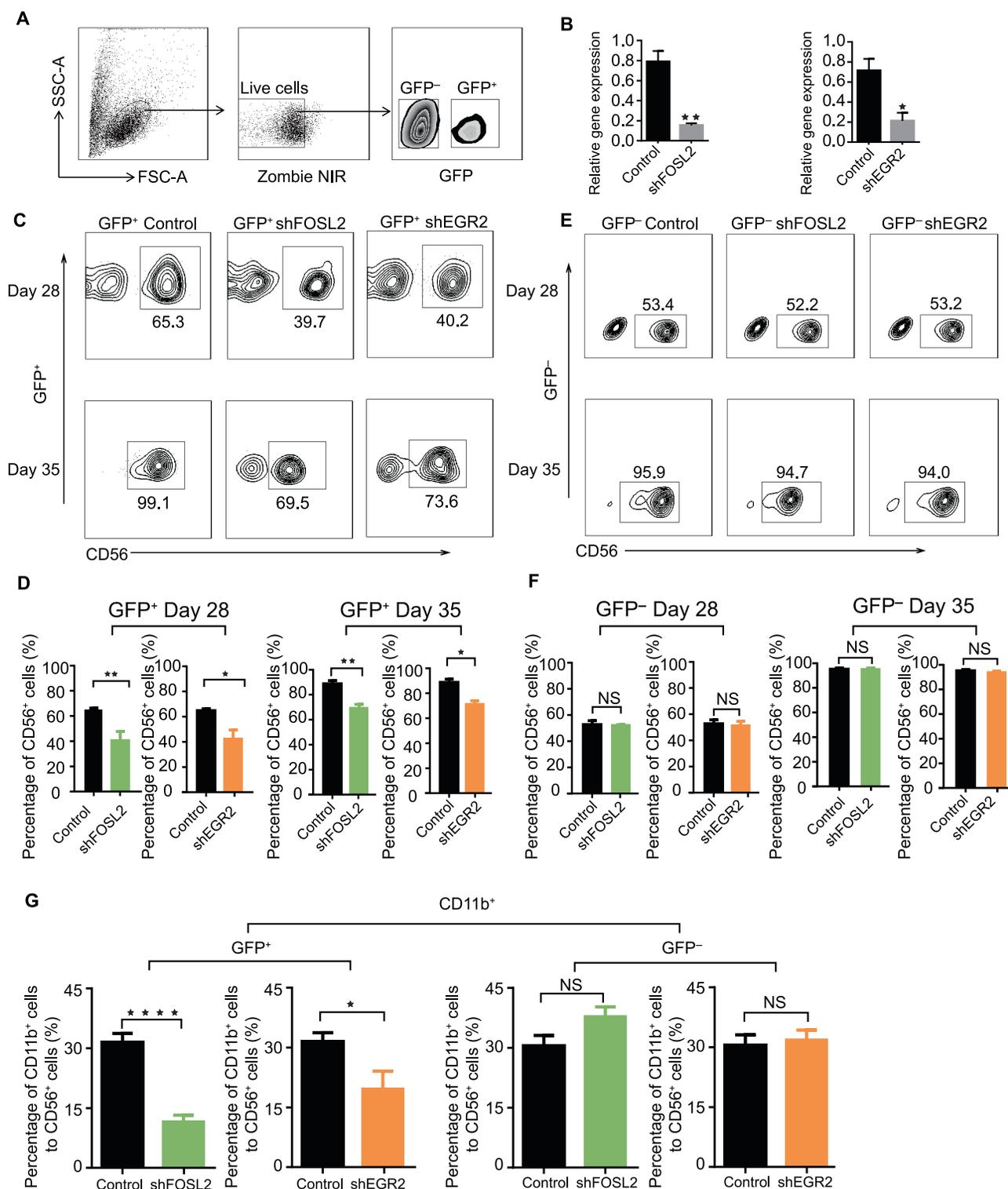


Figure 4 Function of EGR2 and FOSL2 during NK cell differentiation

A. The gating strategy for cultured cells transfected with lentiviruses expressing shFOSL2, shEGR2, or control shRNA via detection of GFP expression. **B.** Knockdown efficiency of *FOSL2* and *EGR2* with shFOSL2 and shEGR2. **C.** Flow cytometric analysis of CD56⁺ cells in the GFP⁺ cells at day 28 (top) and day 35 (bottom). GFP⁺ cells refer to cultured cells successfully transfected with lentiviruses expressing control shRNA (vector with a meaningless fragment), shFOSL2, or shEGR2. **D.** Quantification of CD56⁺ cells in the GFP⁺ cell population at day 28 (left) and day 35 (right). $n = 5$. **E.** Flow cytometric analysis of CD56⁺ in the GFP⁻ cells at day 28 (top) and day 35 (bottom). GFP⁻ cells refer to cultured cells not successfully transfected with lentiviruses expressing control shRNA, shFOSL2, or shEGR2. **F.** Quantification of CD56⁺ cells in the GFP⁻ cell population at day 28 (left) and day 35 (right). $n = 5$. **G.** Quantification of CD11b⁺ cells in the CD56⁺ cell population at day 35 ($n = 5$). Data are presented as the mean \pm SEM. *, $P < 0.05$; **, $P < 0.001$; ***, $P < 0.0005$; ****, $P < 0.0001$. P values were estimated from Student's t -test. SSC-A, side scatter area; FSC-A, forward scatter area; NIR, near infrared.

eral of which were significantly enriched with known TFs that regulate NK cell differentiation (Figure S6B), including *FOSL2* and *EGR2*, and several other genes from the JAK-STAT pathway, such as the *STAT* family, *IRF2*, *TBX21*, and *GATA3*. These results suggested that *FOSL2* and *EGR2* may regulate NK cell differentiation through the JAK-STAT pathway (Figure S6C).

Transcriptional regulatory network dynamics during NK cell differentiation

The dramatic chromatin accessibility differences during NK cell differentiation prompted us to check the time point-specific transcriptional regulatory network. Although some TFs have been discovered to regulate NK cell differentiation, little is known about the dynamics of the entire regulatory network during this process. Since the TF footprint pattern from the ATAC-seq reads could simultaneously and directly reveal the binding profiles of hundreds of TFs with known cognate motifs [22], together with gene expression profiles, we could construct a regulatory network at each time point and assess how it changes during NK cell differentiation. First, we used HOMER to identify enriched TFs that bind to the cluster I and cluster II peaks shown in Figure 3A ($P < 0.05$). We then examined the gene expression profiles of these TFs and found that 120 TFs were expressed at least one stage during the differentiated process. By applying differential analysis, we obtained 14–26 TFs that were distinctly expressed at each stage with fold change (FC) > 1.5 (Figure S7), and defined them as the nodes of the regulatory network. The connections (edges) between any two TFs were defined as follows: if the TF A's motif is on the promoter of TF B, then we say TF A regulates TF B and draw an arrow from TF A to TF B. Here, only TFs that were expressed at the specific time point were considered [58]. Using this method, we constructed the transcriptional regulatory network at each time point with both the enrichment (P value) and expression information for all the relevant TFs (Figure 5A–E). Interestingly, intensive interconnections appeared for day 7-specific TFs, and quickly vanished after two weeks (Figure 5A). In contrast, the day 35-specific network gradually grew through the induction of relevant TFs. Many known regulators, such as *EOMES*, *TBX21*, *ETS1*, *PR/SET domain 1 (PRDM1)*, and *GATA3*, as well as *FOSL2* and *EGR2*, were increasingly enriched in the network (Figure 5E). Similar phenomena were also observed on other networks (Figure 5B–D). The dynamics of the transcriptional regulatory networks seem to explain the increase in the proportion and the differentiation of NK cells.

Discussion

NK cells are important innate immune cells that have been recognized as efficient effector cells to treat tumors. To better understand the differentiation machinery of NK cells and to identify new regulators in the process, we studied the landscape of active elements during NK cell differentiation using the sensitive ATAC-seq method. As a result, we find three distinct clusters of DNA accessible elements. Further analysis shows that the chromatin accessibility is correlated well with the expression levels of the corresponding genes. In short, this study provides an epigenomic landscape and dynamics of

NK cell differentiation and presents fundamental profiles for studying the relationship between chromatin accessibility, gene expression, and cell growth during this process (Figure 2).

TFs bind to their motifs and are often obligate nucleosome evictors, and therefore we can use ATAC-seq to predict critical regulators in NK cell differentiation [26]. By motif analysis in HOMER, we find that several known TFs are enriched at different stages during NK cell differentiation. Similar results have also been obtained using Genomica (Figure 3B). The discovery of known regulators strongly supports the reliability of our analysis. Furthermore, by integrating results from HOMER, Genomica, and motif footprint analysis, we have identified two novel TFs, *FOSL2* and *EGR2*, which are essential for NK cell maturation. Knockdown of either of these two TFs significantly inhibits NK cell transformation in the *in vitro* NK cell differentiation system. Module map analysis suggests that these two TFs may regulate NK cell differentiation through the JAK-STAT pathway, and therefore further studies of this pathway may facilitate the generation of NK cells and thus promote the NK cell-based immunotherapy. Overall, this study also provides a framework to identify new regulators from chromatin accessible data for NK cell differentiation.

TFs do not usually function alone, they always interact with other molecules to fulfill their unique roles. Hence, we have depicted the transcriptional regulatory networks at different stages during NK cell differentiation. In order to construct a stable transcriptional regulatory network, we have performed a rigorous screening of TFs to avoid stochastic fluctuations, and integrated both the enrichment (P value) and expression information for all the relevant TFs. Therefore, although alterations in either the enrichment or gene expression cutoff may result in different networks, the most critical TFs to the regulatory process still remain. However, since the screening of TFs mainly rely on the gene microarray, which is not as accurate as RNA-seq, the structure of these predicted networks may not be very robust.

From our previous study, we notice that with a minimal cytokine cocktail, we can generate sufficient number of functional NK cells that express the cytokines necessary for NK cells and have a high effect on tumor cells [19]. Although there may also be a certain proportion of other lymphocytes, *in vitro* expansion of NK cells from peripheral blood (PB) or UCB cells has been successfully performed and developed for several clinical strategies to treat cancers [4,5,59–61]. Therefore, a comprehensive understanding of the regulatory patterns at each differentiation stage of the *in vitro*-derived NK cells in this system will help to uncover the underlying mechanisms of NK cell differentiation. The transcriptional regulatory networks revealed in this study will lay the foundation for efficient *in vitro* production of effective NK cells, thus facilitating NK cell-based immunotherapy. Moreover, we have identified two novel TFs, *FOSL2* and *EGR2*, as essential regulators in controlling NK cell maturation and function. We have also predicted the potential signaling pathways in which these two TFs may be involved and have illustrated the dynamics of the transcriptional regulatory networks during NK cell differentiation. In spite of the advantages of our strategy, there are two main limitations of this study. 1) Although induced NK cells are very similar to those produced *in vivo*, these two types of NK cells are not identical. We have observed certain differences between the induced NK cells and the NK cells produced *in vivo* in terms of chromatin accessibility (data not shown).

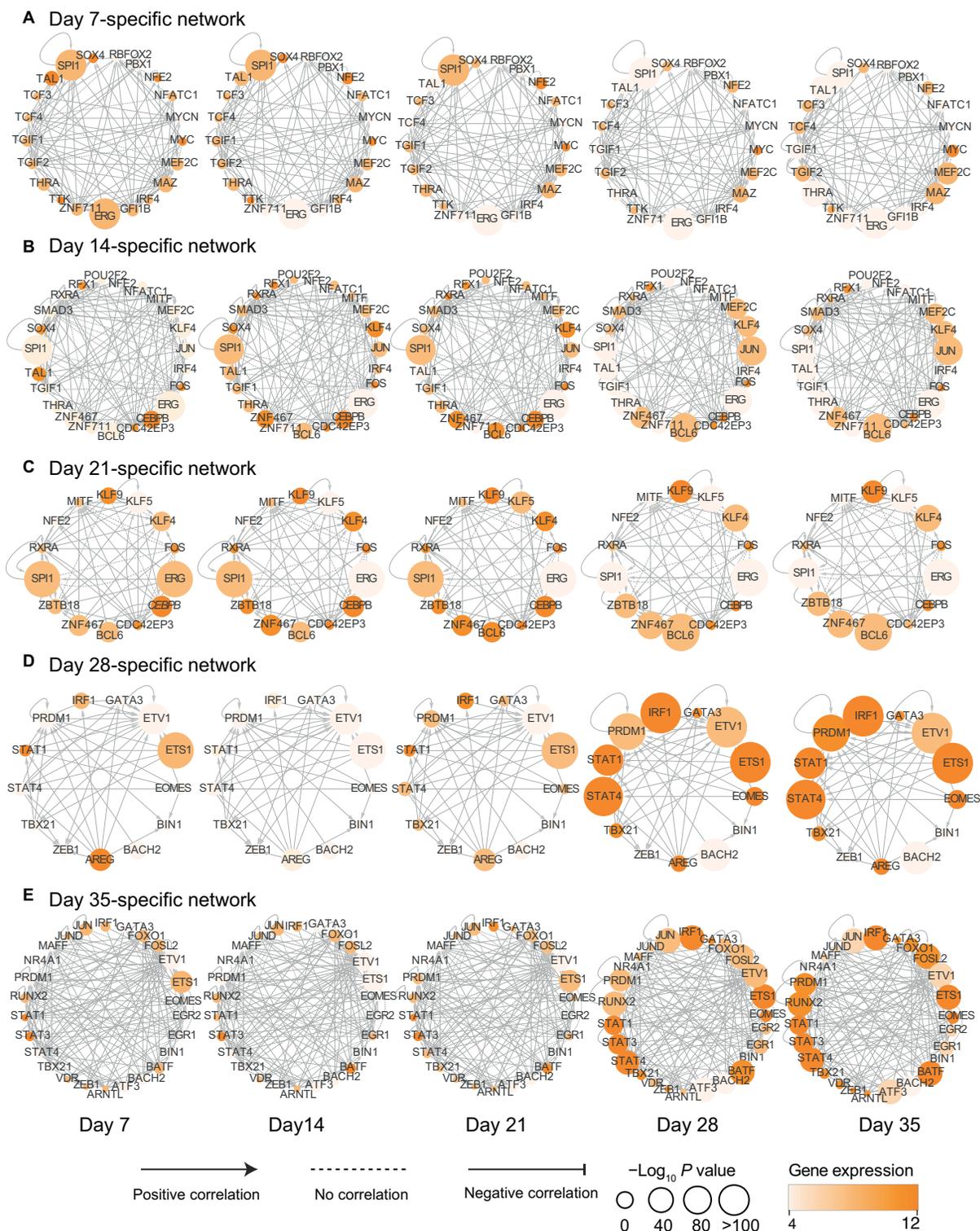


Figure 5 Transcriptional regulatory network during NK cell differentiation

A.–E. *Cis*-regulatory networks between TFs (nodes) enriched in ≥ 1 gene set and specifically expressed ($FC > 1.5$) at day 7 (A), day 14 (B), day 21 (C), day 28 (D), and day 35 (E). Nodes represent TFs with gene expression levels and TF enrichment scores at day 7, day 14, day 21, day 28, and day 35 (from left to right). The arrow at the edge from node A to node B indicates that TF A regulates TF B by binding to the promoter region of TF B. The color of each node indicates the expression level of the gene encoding that TF, and the size of the circle represents the significance of the motif enrichment according to the P value. The types of edges indicate the Pearson correlation between the gene expression profiles of the connected TFs: positively correlated ($PCC > 0.4$); negatively correlated ($PCC < -0.4$); no correlation ($-0.4 \leq PCC \leq 0.4$, dashed line). FC, fold change; PCC, Pearson correlation coefficient.

2) Before NK cells are fully developed, there is always a mixture of cell populations with stem cells, NKP cells, iNK cells, mature NK cells, and other cell types that we are unable to delineate at this moment, since any bulk cell-based analysis may neglect the huge heterogeneity between cells by default. Therefore, to fully uncover the regulatory mechanism, single cell technologies are required in the future to further delineate the cell-to-cell heterogeneity and regulatory dynamics at the single cell precision.

Materials and methods

Samples

UCB samples were collected at birth from women with normal, full-term deliveries at Anhui Provincial Hospital, Hefei, China, after receiving their written informed consent. The culture procedure for NK cell differentiation from UCB CD34⁺ cells has been previously described [19].

Immunofluorescence staining

The cultured cells were post-fixed in 4% paraformaldehyde, blocked with 10% goat serum (Catalog No. 31872; ThermoFisher Scientific, Grand Island, NY), before incubating with primary antibodies at 4 °C overnight. Secondary antibodies were then added and incubated at 37 °C for 1 h after washes. Cell nuclei were stained with DAPI for 5 min at room temperature. Confocal images were acquired using a Zeiss LS710 microscope.

RNA isolation and real-time PCR

Cultured cells were lysed in TRIzol reagent (Catalog No. 15596026; ThermoFisher Scientific), and total RNA was extracted following the manufacturer's instructions. cDNA was synthesized with Moloney murine leukemia virus reverse transcriptase (Catalog No. 28025013; ThermoFisher Scientific) and oligo (dT) 20 primers. Then, SYBR Premix Ex Taq (Catalog No. RR420L; TaKaRa, Dalian, China) was used for real-time PCR using primers with sequences listed in Table S4. The data were analyzed using the 2^{- $\Delta\Delta C_t$} method.

Lentivirus production and transduction

To produce lentiviral particles, 293 T cells were transfected with the plasmids PLKO.1, pRRE, pREV, or pVSV-G using Lipofectamine 2000 (Catalog No. 11668-019; Invitrogen, Grand Island, NY) following the manufacturer's protocol. Then, we harvested the supernatants 48 h and 72 h post-transfection. To remove cell debris, supernatants were centrifuged at 3000 rpm for 10 min, and then, the lentivirus particles were concentrated by ultracentrifugation at 50,000 *g* for 2 h at 4 °C. Finally, the virus particles were gently resuspended in Hanks' balanced salt solution (HBSS) and stored at -80 °C. After UCB CD34⁺ cells were cultured with multiple cytokines for 14–18 days, we incubated the lentiviruses and the cultured cells with polybrene (Catalog No. H9268-10G; Sigma Aldrich, Shanghai, China) (5 μ g/ml) for 30 min and the mixture was subjected to centrifugation at 1000 rpm for 70 min at 10 °C.

Statistical analysis

Statistical significance was analyzed using unpaired two-tailed Student's *t*-test or one-way ANOVA. *P* values less than 0.05 were considered statistically significant.

ATAC-seq and analysis

ATAC-seq was performed as previously described [22], and 2 × 150 paired-end sequencing was performed on an Illumina HiSeq X-10 to yield, on average, 78 M reads/sample. Sample reads from biological replicates were then grouped together and divided into eight categories: day 7, day 14, day 19, day 21, day 24, day 26, day 28, and day 35. Intrinsic analysis and other ATAC-seq analysis was performed same as our previous work (see File S1) [26,62].

Differential analysis

All samples were grouped into eight categories (16 samples): day 7, day 14, day 19, day 21, day 24, day 26, day 28, and day 35. Data normalization and significance analysis were performed via pairwise comparison between the eight categories using DESeq2 [63] with *P* < 0.01, log₂ FC > 5, and FDR < 0.01, and intrinsic analysis [26] with a z-score > 1. We finally obtained 6401 differential peaks. Unsupervised clustering was performed using Cluster 3.0 and visualized in Treeview. GREAT [33] was used to predict enriched functions and GO terms.

Identification of stage-specific peaks

As shown in Figure S3, each stage (*e.g.*, day 7) consists of genes that were more highly expressed (FC > 1.5) at this stage compared with all the other stages, and we defined these genes as stage-specific genes. Similarly, we defined peaks that were more accessible (FC > 1.5) in one stage compare with all the other stages as stage-specific peaks. Peaks that were accessible at all stages were defined as conservative peaks.

TF motif enrichment analysis

HOMER [29] was used to perform the TF enrichment analysis with the following options: findmotifs.pl input.fasta output/ (Figure 3A) and findmotifs.pl input.fasta human uotputdir -fasta bg.fasta (Figure S5A). TF enrichment analysis was performed in Genomica. TF “footprint” analysis was performed in the same way as described in our previous work [27].

Gene module and STRING analysis

Gene module analyses were performed using WGCNA [64] with the options SoftPower = 20, minModuleSize = 30. Protein–protein interaction analyses were performed using STRING [65].

Construction of transcriptional regulatory networks

We used HOMER to find the TFs that bind to cluster I and cluster II peaks and obtained TFs that could regulate stage

differentially expressed elements ($P < 0.05$). We define the region 2 kb upstream of the transcription start site as a promoter. If TF A bound to the promoter of TF B, we defined TF A as a regulator of TF B and then constructed a transcriptional regulatory network. Networks of TFs were assembled from TFs that were expressed in at least one sample. An edge between TF A and TF B indicated that TF A binds to the promoter of TF B [58].

Ethical statement

This study was approved by the Ethics Committee of the University of Science and Technology of China.

Data availability

The raw ATAC-seq data have been deposited in the Genome Sequence Archive [66] at the National Genomics Data Center, Beijing Institute of Genomics, Chinese Academy of Sciences / China National Center for Bioinformation (GSA: CRA000846) and are publicly accessible at <http://bigd.big.ac.cn/gsa>.

CRedit author statement

Kun Li: Conceptualization, Methodology, Software, Writing - original draft, Visualization. **Yang Wu:** Conceptualization, Data curation, Validation, Writing - original draft. **Young Li:** Data curation, Validation. **Qiaoni Yu:** Methodology. **Zhigang Tian:** Supervision. **Haiming Wei:** Conceptualization, Writing - review & editing. **Kun Qu:** Conceptualization, Writing - review & editing, Visualization, Project administration, Funding acquisition. All authors read and approved the final manuscript.

Competing interests

All authors have no competing financial interests.

Acknowledgments

This work was supported by the National Key R&D Program of China (Grant No. 2017YFA0102900) to KQ, the National Natural Science Foundation of China (Grant Nos. 91640113, 91940306, 31771428, and 31970858 to KQ, as well as 81330071 to HW). It was also supported by the Fundamental Research Funds for the Central Universities, China (Grant Nos. YD2070002019, WK2070000158, and WK9110000141) to KQ. We thank the USTC Supercomputing Center and the School of Life Science Bioinformatics Center for providing supercomputing resources for this project. We thank the CAS interdisciplinary innovation team for helpful discussion.

Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.gpb.2020.12.003>.

ORCID

0000-0001-6550-4937 (Kun Li)
 0000-0002-4430-428X (Yang Wu)
 0000-0002-6595-2577 (Young Li)
 0000-0003-3272-7119 (Qiaoni Yu)
 0000-0003-2742-2361 (Zhigang Tian)
 0000-0002-1675-6502 (Haiming Wei)
 0000-0002-5555-8437 (Kun Qu)

References

- [1] Sun JC. Transcriptional control of NK cells. *Curr Top Microbiol Immunol* 2016;395:1–36.
- [2] Miller JS, Soignier Y, Panoskaltis-Mortari A, McNearney SA, Yun GH, Fautsch SK, et al. Successful adoptive transfer and *in vivo* expansion of human haploidentical NK cells in patients with cancer. *Blood* 2005;105:3051–7.
- [3] Miller JS, Verfaillie C, McGlave P. The generation of human natural killer cells from CD34⁺/DR⁻ primitive progenitors in long-term bone marrow culture. *Blood* 1992;80:2182–7.
- [4] Luevano M, Madrigal A, Saudemont A. Generation of natural killer cells from hematopoietic stem cells *in vitro* for immunotherapy. *Cell Mol Immunol* 2012;9:310–20.
- [5] Cheng M, Chen Y, Xiao W, Sun R, Tian Z. NK cell-based immunotherapy for malignant diseases. *Cell Mol Immunol* 2013;10:230–52.
- [6] Dolstra H, Roeven MWH, Spanholtz J, Hangalapura BN, Tordoir M, Maas F, et al. Successful transfer of umbilical cord blood CD34⁺ hematopoietic stem and progenitor-derived NK cells in older acute myeloid leukemia patients. *Clin Cancer Res* 2017;23:4107–18.
- [7] Ruggeri L, Capanni M, Urbani E, Perruccio K, Shlomchik WD, Tosti A, et al. Effectiveness of donor natural killer cell alloreactivity in mismatched hematopoietic transplants. *Science* 2002;295:2097–100.
- [8] Guilleray C, Huntington ND, Smyth MJ. Targeting natural killer cells in cancer immunotherapy. *Nat Immunol* 2016;17:1025–36.
- [9] Sivori S, Cantoni C, Parolini S, Marcenaro E, Conte R, Moretta L, et al. IL-21 induces both rapid maturation of human CD34⁺ cell precursors towards NK cells and acquisition of surface killer Ig-like receptors. *Eur J Immunol* 2003;33:3439–47.
- [10] Spanholtz J, Tordoir M, Eissens D, Preijers F, van der Meer A, Joosten I, et al. High log-scale expansion of functional human natural killer cells from umbilical cord blood CD34-positive cells for adoptive cancer immunotherapy. *PLoS One* 2010;5:e9221.
- [11] Kao IT, Yao CL, Kong ZL, Wu ML, Chuang TL, Hwang SM. Generation of natural killer cells from serum-free, expanded human umbilical cord blood CD34⁺ cells. *Stem Cells Dev* 2007;16:1043–52.
- [12] Haraguchi K, Suzuki T, Koyama N, Kumano K, Nakahara F, Matsumoto A, et al. Notch activation induces the generation of functional NK cells from human cord blood CD34-positive cells devoid of IL-15. *J Immunol* 2009;182:6168–78.
- [13] Yoon SR, Lee YS, Yang SH, Ahn KH, Lee JH, Lee JH, et al. Generation of donor natural killer cells from CD34⁺ progenitor cells and subsequent infusion after HLA-mismatched allogeneic hematopoietic cell transplantation: a feasibility study. *Bone Marrow Transplant* 2010;45:1038–46.
- [14] Shah N, Martin-Antonio B, Yang H, Ku S, Lee DA, Cooper LJ, et al. Antigen presenting cell-mediated expansion of human umbilical cord blood yields log-scale expansion of natural killer cells with anti-myeloma activity. *PLoS One* 2013;8:e76781.
- [15] Vitale C, Cottalasso F, Montaldo E, Moretta L, Mingari MC. Methylprednisolone induces preferential and rapid differentiation

- of CD34⁺ cord blood precursors toward NK cells. *Int Immunol* 2008;20:565–75.
- [16] Carayol G, Robin C, Bourhis JH, Bennaceur-Griscelli A, Chouaib S, Coulombel L, et al. NK cells differentiated from bone marrow, cord blood and peripheral blood stem cells exhibit similar phenotype and functions. *Eur J Immunol* 1998;28:1991–2002.
- [17] Yu Y, Hagihara M, Ando K, Gansuud B, Matsuzawa H, Tsuchiya T, et al. Enhancement of human cord blood CD34⁺ cell-derived NK cell cytotoxicity by dendritic cells. *J Immunol* 2001;166:1590–600.
- [18] Cany J, van der Waart AB, Spanholtz J, Tordoir M, Jansen JH, van der Voort R, et al. Combined IL-15 and IL-12 drives the generation of CD34⁺-derived natural killer cells with superior maturation and alloreactivity potential following adoptive transfer. *Oncoimmunology* 2015;4:e1017701.
- [19] Wu Y, Li Y, Fu B, Jin L, Zheng X, Zhang A, et al. Programmed differentiated natural killer cells kill leukemia cells by engaging slam family receptors. *Oncotarget* 2017;8:57024–38.
- [20] Colucci F, Samson SI, DeKoter RP, Lantz O, Singh H, Di Santo JP. Differential requirement for the transcription factor PU.1 in the generation of natural killer cells versus B and T cells. *Blood* 2001;97:2625–32.
- [21] Gordon SM, Chaix J, Rupp LJ, Wu J, Madera S, Sun JC, et al. The transcription factors T-bet and Eomes control key checkpoints of natural killer cell maturation. *Immunity* 2012;36:55–67.
- [22] Buenostro 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* 2013;10:1213–8.
- [23] Sen DR, Kaminski J, Barnitz RA, Kurachi M, Gerdemann U, Yates KB, et al. The epigenetic landscape of T cell exhaustion. *Science* 2016;354:1165–9.
- [24] Pauken KE, Sammons MA, Odorizzi PM, Manne S, Godec J, Khan O, et al. Epigenetic stability of exhausted T cells limits durability of reinvigoration by PD-1 blockade. *Science* 2016;354:1160–5.
- [25] Denny SK, Yang D, Chuang CH, Brady JJ, Lim JS, Gruner BM, et al. Nf1b promotes metastasis through a widespread increase in chromatin accessibility. *Cell* 2016;166:328–42.
- [26] Qu K, Zaba LC, Giresi PG, Li R, Longmire M, Kim YH, et al. Individuality and variation of personal regulomes in primary human T cells. *Cell Syst* 2015;1:51–61.
- [27] Qu K, Zaba LC, Satpathy AT, Giresi PG, Li R, Jin Y, et al. Chromatin accessibility landscape of cutaneous T cell lymphoma and dynamic response to HDAC inhibitors. *Cancer Cell* 2017;32:27–41.e4.
- [28] Lau CM, Adams NM, Geary CD, Weizman OE, Rapp M, Pritykin Y, et al. Epigenetic control of innate and adaptive immune memory. *Nat Immunol* 2018;19:963–72.
- [29] Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, et al. Simple combinations of lineage-determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities. *Mol Cell* 2010;38:576–89.
- [30] Yu J, Freud AG, Caligiuri MA. Location and cellular stages of natural killer cell development. *Trends Immunol* 2013;34:573–82.
- [31] Montaldo E, Del Zotto G, Della Chiesa M, Mingari MC, Moretta A, De Maria A, et al. Human NK cell receptors/markers: a tool to analyze NK cell development, subsets and function. *Cytometry A* 2013;83:702–13.
- [32] Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. *Nature* 2000;406:747–52.
- [33] McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, et al. GREAT improves functional interpretation of *cis*-regulatory regions. *Nat Biotechnol* 2010;28:495–501.
- [34] Grewal SI, Moazed D. Heterochromatin and epigenetic control of gene expression. *Science* 2003;301:798–802.
- [35] Ohno S, Sato T, Kohu K, Takeda K, Okumura K, Satake M, et al. Runx proteins are involved in regulation of CD122, LY49 family and IFN-gamma expression during NK cell differentiation. *Int Immunol* 2008;20:71–9.
- [36] Yamanaka R, Lekstrom-Himes J, Barlow C, Wynshaw-Boris A, Xanthopoulos KG. CCAAT/enhancer binding proteins are critical components of the transcriptional regulation of hematopoiesis (review). *Int J Mol Med* 1998;1:213–21.
- [37] Chang HC, Zhang S, Thieu VT, Slee RB, Bruns HA, Laribee RN, et al. PU.1 expression delineates heterogeneity in primary TH2 cells. *Immunity* 2005;22:693–703.
- [38] Dakic A, Metcalf D, Di Rago L, Mifsud S, Wu L, Nutt SL. PU.1 regulates the commitment of adult hematopoietic progenitors and restricts granulopoiesis. *J Exp Med* 2005;201:1487–502.
- [39] Nutt SL, Metcalf D, D'Amico A, Polli M, Wu L. Dynamic regulation of PU.1 expression in multipotent hematopoietic progenitors. *J Exp Med* 2005;201:221–31.
- [40] Carotta S, Dakic A, D'Amico A, Pang SH, Greig KT, Nutt SL, et al. The transcription factor PU.1 controls dendritic cell development and FLT3 cytokine receptor expression in a dose-dependent manner. *Immunity* 2010;32:628–41.
- [41] Iwasaki H, Somoza C, Shigematsu H, Duprez EA, Iwasaki-Arai J, Mizuno S, et al. Distinctive and indispensable roles of PU.1 in maintenance of hematopoietic stem cells and their differentiation. *Blood* 2005;106:1590–600.
- [42] Boos MD, Yokota Y, Eberl G, Kee BL. Mature natural killer cell and lymphoid tissue-inducing cell development requires ID2-mediated suppression of E protein activity. *J Exp Med* 2007;204:1119–30.
- [43] Eckelhart E, Warsch W, Zebidin E, Simma O, Stoiber D, Kolbe T, et al. A novel *Ncr1*-Cre mouse reveals the essential role of STAT5 for NK-cell survival and development. *Blood* 2011;117:1565–73.
- [44] Collins A, Littman DR, Taniuchi I. RUNX proteins in transcription factor networks that regulate T-cell lineage choice. *Nat Rev Immunol* 2009;9:106–15.
- [45] Mathelier A, Fornes O, Arenillas DJ, Chen CY, Denay G, Lee J, et al. JASPAR 2016: a major expansion and update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res* 2016;44:D110–5.
- [46] O'Shea JJ, Schwartz DM, Villarino AV, Gadina M, McInnes IB, Laurence A. The JAK-STAT pathway: impact on human disease and therapeutic intervention. *Annu Rev Med* 2015;66:311–28.
- [47] Tamura T, Yanai H, Savitsky D, Taniguchi T. The IRF family transcription factors in immunity and oncogenesis. *Annu Rev Immunol* 2008;26:535–84.
- [48] Taniguchi T, Ogasawara K, Takaoka A, Tanaka N. IRF family of transcription factors as regulators of host defense. *Annu Rev Immunol* 2001;19:623–55.
- [49] Lohoff M, Duncan GS, Ferrick D, Mittrucker HW, Bischof S, Prechtel S, et al. Deficiency in the transcription factor interferon regulatory factor (IRF)-2 leads to severely compromised development of natural killer and T helper type 1 cells. *J Exp Med* 2000;192:325–36.
- [50] Taki S, Nakajima S, Ichikawa E, Saito T, Hida S. IFN regulatory factor-2 deficiency revealed a novel checkpoint critical for the generation of peripheral NK cells. *J Immunol* 2005;174:6005–12.
- [51] Szabo SJ, Kim ST, Costa GL, Zhang X, Fathman CG, Glimcher LH. A novel transcription factor, T-bet, directs Th1 lineage commitment. *Cell* 2000;100:655–69.
- [52] Townsend MJ, Weinmann AS, Matsuda JL, Salomon R, Farnham PJ, Biron CA, et al. T-bet regulates the terminal maturation and homeostasis of NK and Valpha14i NKT cells. *Immunity* 2004;20:477–94.
- [53] Ramirez K, Chandler KJ, Spaulding C, Zandi S, Sigvardsson M, Graves BJ, et al. Gene deregulation and chronic activation in natural killer cells deficient in the transcription factor ETS1. *Immunity* 2012;36:921–32.

- [54] Samson SI, Richard O, Tavian M, Ranson T, Vosshenrich CA, Colucci F, et al. GATA-3 promotes maturation, IFN-gamma production, and liver-specific homing of NK cells. *Immunity* 2003;19:701–11.
- [55] Higuchi T, Nakayama T, Arao T, Nishio K, Yoshie O. *SOX* is a direct target gene of FRA-2 and induces expression of HDAC8 in adult T-cell leukemia/lymphoma. *Blood* 2013;121:3640–9.
- [56] Li S, Miao T, Sebastian M, Bhullar P, Ghaffari E, Liu M, et al. The transcription factors EGR2 and EGR3 are essential for the control of inflammation and antigen-induced proliferation of B and T cells. *Immunity* 2012;37:685–96.
- [57] Okamura T, Fujio K, Sumitomo S, Yamamoto K. Roles of LAG3 and EGR2 in regulatory T cells. *Ann Rheum Dis* 2012;71:i96–i100.
- [58] Novershtern N, Subramanian A, Lawton LN, Mak RH, Haining WN, McConkey ME, et al. Densely interconnected transcriptional circuits control cell states in human hematopoiesis. *Cell* 2011;144:296–309.
- [59] Suck G, Koh MB. Emerging natural killer cell immunotherapies: large-scale *ex vivo* production of highly potent anticancer effectors. *Hematol Oncol Stem Cell Ther* 2010;3:135–42.
- [60] Smyth MJ, Hayakawa Y, Takeda K, Yagita H. New aspects of natural-killer-cell surveillance and therapy of cancer. *Nat Rev Cancer* 2002;2:850–61.
- [61] Fujisaki H, Kakuda H, Shimasaki N, Imai C, Ma J, Lockey T, et al. Expansion of highly cytotoxic human natural killer cells for cancer cell therapy. *Cancer Res* 2009;69:4010–7.
- [62] Zuo Z, Jin Y, Zhang W, Lu Y, Li B, Qu K. ATAC-pipe: general analysis of genome-wide chromatin accessibility. *Brief Bioinform* 2019;20:1934–43.
- [63] Anders S, Huber W. Differential expression analysis for sequence count data. *Genome Biol* 2010;11:R106.
- [64] Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 2008;9:559.
- [65] Snel B, Lehmann G, Bork P, Huynen MA. STRING: a web-server to retrieve and display the repeatedly occurring neighbourhood of a gene. *Nucleic Acids Res* 2000;28:3442–4.
- [66] Wang Y, Song F, Zhu J, Zhang S, Yang Y, Chen T, et al. GSA: Genome Sequence Archive. *Genomics Proteomics Bioinformatics* 2017;15:14–8.