Title

Dynamics of Chromatin Accessibility During Hematopoietic Stem Cell Differentiation Into Progressively Lineage-Committed Progeny

Journal

Stem Cells, 41(5)

ISSN

1066-5099

Authors

Martin, Eric W
Rodriguez y Baena, Alessandra
Reggiardo, Roman E
et al.

Publication Date

2023-05-15

DOI

10.1093/stmcls/sxad022

Peer reviewed

Stem Cells, 2023, **41**, 520–539
https://doi.org/10.1093/stmcls/sxad022
Advance access publication 22 March 2023
Original Research

Dynamics of Chromatin Accessibility During Hematopoietic Stem Cell Differentiation Into Progressively Lineage-Committed Progeny

Eric W. Martin[1], Alessandra Rodriguez y Baena[1], Roman E. Reggiardo[2], Atesh K. Worthington[1], Connor S. Mattingly[2], Donna M. Poscablo[1], Jana Krietsch[2], Michael T McManus[3], Susan Carpenter[1], Daniel H. Kim[2], E. Camilla Forsberg[2,*]

[1]Institute for the Biology of Stem Cells, Department of Molecular, Cell and Developmental Biology, University of California Santa Cruz, Santa Cruz, CA, USA
[2]Institute for the Biology of Stem Cells, Department of Biomolecular Engineering, University of California Santa Cruz, Santa Cruz, CA, USA
[3]Department of Microbiology and Immunology, Diabetes Center, W.M. Keck Center for Noncoding RNAs, University of California San Francisco, San Francisco, CA, USA

*Corresponding author: Camilla Forsberg, 1156 High Street, Santa Cruz, CA 95064, USA. Tel: +1 831 459 2111; Email: cforsber@ucsc.edu

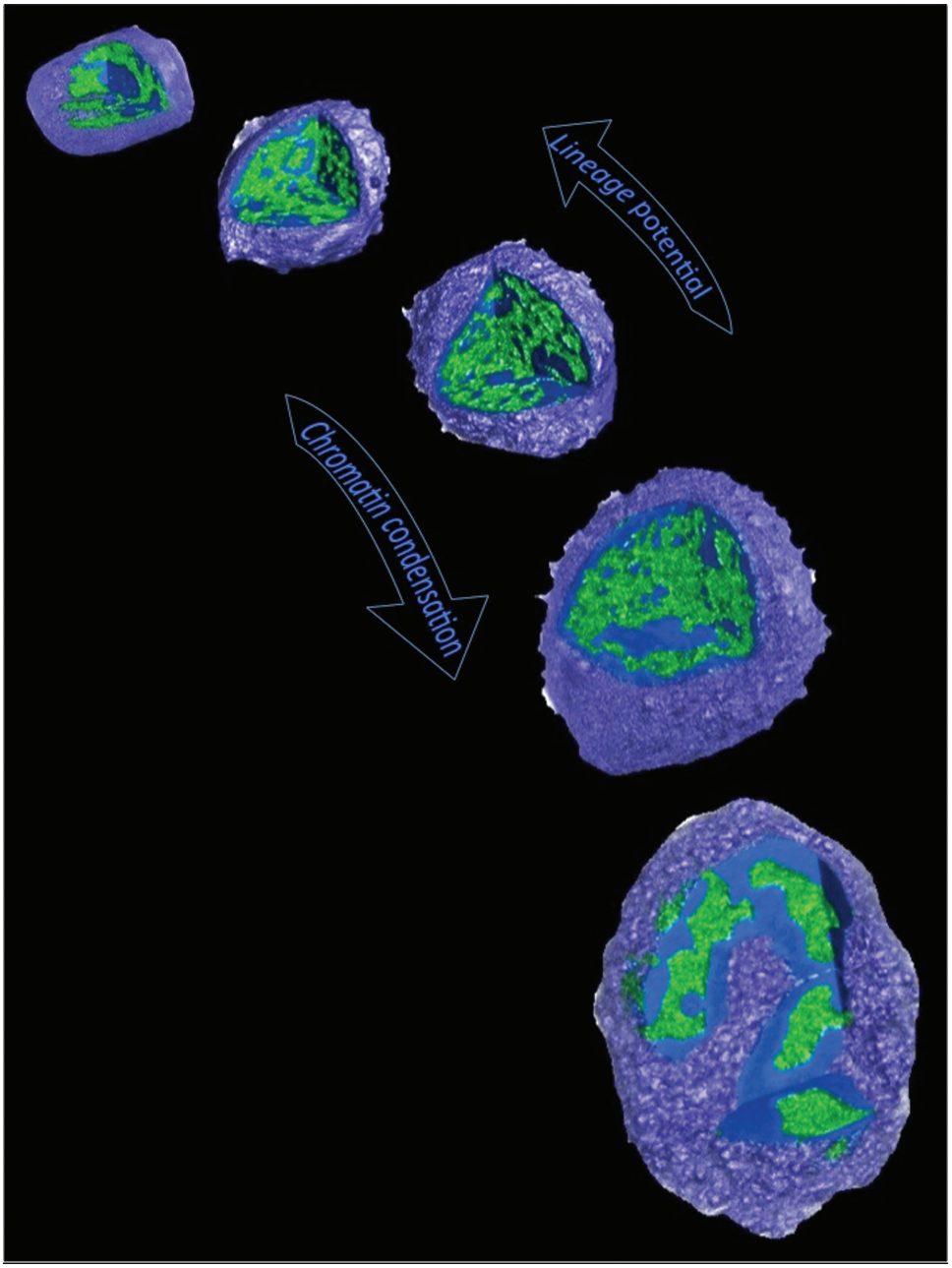
Abstract

Epigenetic mechanisms regulate the multilineage differentiation capacity of hematopoietic stem cells (HSCs) into a variety of blood and immune cells. Mapping the chromatin dynamics of functionally defined cell populations will shed mechanistic insight into 2 major, unanswered questions in stem cell biology: how does epigenetic identity contribute to a cell type's lineage potential, and how do cascades of chromatin remodeling dictate ensuing fate decisions? Our recent work revealed evidence of multilineage gene priming in HSCs, where open *cis*-regulatory elements (CREs) exclusively shared between HSCs and unipotent lineage cells were enriched for DNA binding motifs of known lineage-specific transcription factors. Oligopotent progenitor populations operating between the HSCs and unipotent cells play essential roles in effecting hematopoietic homeostasis. To test the hypothesis that selective HSC-primed lineage-specific CREs remain accessible throughout differentiation, we used ATAC-seq to map the temporal dynamics of chromatin remodeling during progenitor differentiation. We observed epigenetic-driven clustering of oligopotent and unipotent progenitors into distinct erythromyeloid and lymphoid branches, with multipotent HSCs and MPPs associating with the erythromyeloid lineage. We mapped the dynamics of lineage-primed CREs throughout hematopoiesis and identified both unique and shared CREs as potential lineage reinforcement mechanisms at fate branch points. Additionally, quantification of genome-wide peak count and size revealed overall greater chromatin accessibility in HSCs, allowing us to identify HSC-unique peaks as putative regulators of self-renewal and multilineage potential. Finally, CRISPRi-mediated targeting of ATACseq-identified putative CREs in HSCs allowed us to demonstrate the functional role of selective CREs in lineage-specific gene expression. These findings provide insight into the regulation of stem cell multipotency and lineage commitment throughout hematopoiesis and serve as a resource to test functional drivers of hematopoietic lineage fate.

Key words: hematopoietic stem and progenitor cells; epigenetics; chromatin accessibility; cell fate decisions; hematopoiesis.

Received: 6 February 2022; Accepted: 27 February 2023.

Graphical Abstract



Highlights

- HSCs displayed higher chromatin accessibility than any progeny population.
- Epigenetic branch points were evident between CMPs and CLPs and between MkPs and EPs.
- Lineage priming was selectively maintained throughout differentiation.
- HSC-unique accessible chromatin regions were highly enriched for regulatory elements of erythrocyte differentiation.
- CRISPRi-mediated targeting of primary HSCs identified functionally significant CREs.

Introduction

Hematopoiesis is the process by which multipotent hematopoietic stem cells (HSCs) undergo orchestrated epigenetic and transcriptional changes to produce increasingly lineage-restricted progenitors. According to classical models of hematopoiesis, progressively restricting cell fate decisions allow the differentiation of HSCs into multipotent progenitors (MPPs), which further differentiate into common lymphoid progenitors (CLPs) and common myeloid progenitors (CMPs).[1-3] Lymphopoiesis further

results in unipotent progenitors, ProB and ProT, of B and T cells, respectively. In myelopoiesis, granulocyte-macrophage progenitors (GMPs) generate primarily mature granulocytes and macrophages (GMs), while megakaryocytic-erythroid progenitors (MEPs), megakaryocyte progenitors (MkPs), and erythroid progenitors (EPs) produce primarily platelets and red cells.[2,4-6] Though it is clear that hematopoiesis is incredibly dynamic with variable flux within and between cell populations, this well-characterized mammalian hematopoietic system serves as a superb model for the analysis of factors responsible for the development of functionally distinct progenitors and mature cell populations from stem cells. Lineage-specific cell fate decisions are regulated through epigenetic remodeling of *cis*-regulatory elements (CREs), including promoters and enhancer regions. While proximal promoter sequences can suffice to assemble the Pol II transcriptional machinery, non-promoter CREs are often necessary to confer cell type-specific transcriptional regulation. These enhancer regions can be located far upstream or downstream of the target promoter and serve as sequence-specific binding sites for lineage-determining transcription factors (TFs) that regulate the expression of genes specifying cell identity.[7,8] While TFs are important contributors to cellular lineage specification and progressive lineage restriction, accessibility of enhancers to TFs is fundamental for spatiotemporal gene regulation during stem cell differentiation.[9-13]

In hematopoietic progenitors, there is evidence that multilineage priming of CREs precedes commitment to the different cell lineages.[14] "Priming," here defined as chromatin accessibility of a putative CRE despite lack of expression of its presumed target gene, likely contributes to stem and progenitor lineage potential. As differentiation of HSCs proceeds, genes involved in the target lineage are progressively upregulated in progenitor populations while genes involved in non-target lineages are repressed,[15-18] suggesting an essential role of epigenetic regulation in cell fate decisions. In our previous work, we showed evidence of multilineage priming in HSCs, where HSCs had increased global chromatin accessibility compared to their progeny[19] and where open CREs exclusively shared between HSCs and unipotent lineage cells were enriched for DNA binding motifs for known lineage-specific TFs.[20] These data led us to hypothesize that HSC-primed lineage-specific CREs remain accessible throughout differentiation into that specific lineage. Since CREs are often devoid of nucleosomes to allow TF binding,[21,22] we performed the Assay for Transposase Accessible Chromatin by high-throughput sequencing (ATAC-seq)[23,24] of seven, functionally well-characterized hematopoietic progenitor cell types[6,25] to understand CRE priming across hematopoiesis. Importantly, cell fate decisions, as well as lineage-selective expansion and apoptosis, appear to occur primarily in progenitor cell populations.[1,6,26,27] In this study, in-depth ATAC-seq investigation and comparative analysis of HSCs and 12 progeny populations of the 5 main hematopoietic cell lineages revealed potential multipotency, lineage-driving, and/or lineage-reinforcing regulatory elements and their corresponding transcription factors that orchestrate differentiation through epigenetic remodeling. As a proof-of-concept, we used CRISPRi-mediated silencing of CREs in primary HSCs isolated from a new transgenic CRISPRi mouse model to functionally link distal and proximal putative CREs to target genes.

Results

HSCs Had Greater Global Chromatin Accessibility Compared to Hematopoietic Progenitor Cell Types

To determine the dynamics of genome accessibility of multipotent and increasingly lineage-restricted hematopoietic progenitors, we purified 7 primary hematopoietic progenitor cell types (Fig. 1A) by fluorescent-activated cell sorting (FACS) and performed ATAC-seq. After careful quality control of individual and replicate samples (see below), we tested the hypothesis that multipotency is correlated with overall chromatin openness.[19,20,28] We reasoned that multipotent cell populations would have the highest level of accessibility relative to oligopotent cells and that unipotent progenitors would have the least. Thus, we ranked the relative overall accessibility of the hematopoietic progenitors relative to HSCs from our previous report.[20] We first combined the peak lists from each replicate (n = 2) using the irreproducible discovery rate (IDR)[29] for each cell type to quantify the number of peaks. HSCs had the highest number of peaks, followed by MPPs (Fig. 1B; Table 1). We also quantified global accessibility by calculating the cumulative normalized average signal over the master peak list for each cell type by generating histograms using HOMER.[30] HSCs had by far the largest peak signal of any progenitor cell type, while all the progenitors had a similar average signal (Fig. 1C). Although these 2 measurements are not completely independent, HSCs displayed both the highest number of peaks (Fig. 1B) and the cumulative greatest peak signal (Fig. 1C). Overall, these results are consistent with epigenetic stem cell priming and our previous reports[19,20] where HSCs have the greatest chromatin accessibility compared to their progeny and differentiated cells.

Chromatin Accessibility of Cell Type-Specific Genes Correlated With Known Expression Patterns in Hematopoietic Cells

We began the search for lineage-specific regulatory elements by using the gene expression commons (GEXC) expression database[31] to generate lists of genes that were expressed specifically in each progenitor cell type (examples shown in Fig. 1D). In parallel, we filtered the ATAC-seq peak lists of each progenitor cell type (HSC, MPP, CMP, GMP, MEP, CLP, ProB, and ProT) against each other to generate unique peak-lists for each cell type. We then intersected the unique peak lists with the uniquely expressed genes for each progenitor. For populations that had more than 10 unique promoter peaks (HSCs, MEPs, ProBs, and ProTs) we used HOMER[30] to calculate the normalized average signal centered at the promoter for peaks that overlapped with expressed genes (Fig. 1E). We observed cell type-specific read-count accumulation for each progenitor cell with minimal signal from other cell types, indicating that our strategy indeed resolved lineage-specific accessibility.

Lymphoid Commitment Displayed More Extensive Chromatin Remodeling Compared to Myelopoiesis

Next, we sought to pinpoint epigenetic changes at a main branchpoint in hematopoiesis, where the multipotent stem and progenitor cells differentiate into either erythromyeloid- or lymphoid-committed CMPs or CLPs, respectively.[5] We compared the peaks gained and lost between multipotent HSCs and MPPs (combined as "KLS" peaks) and CMPs or CLPs (Fig. 2A). First, we determined the number of peaks

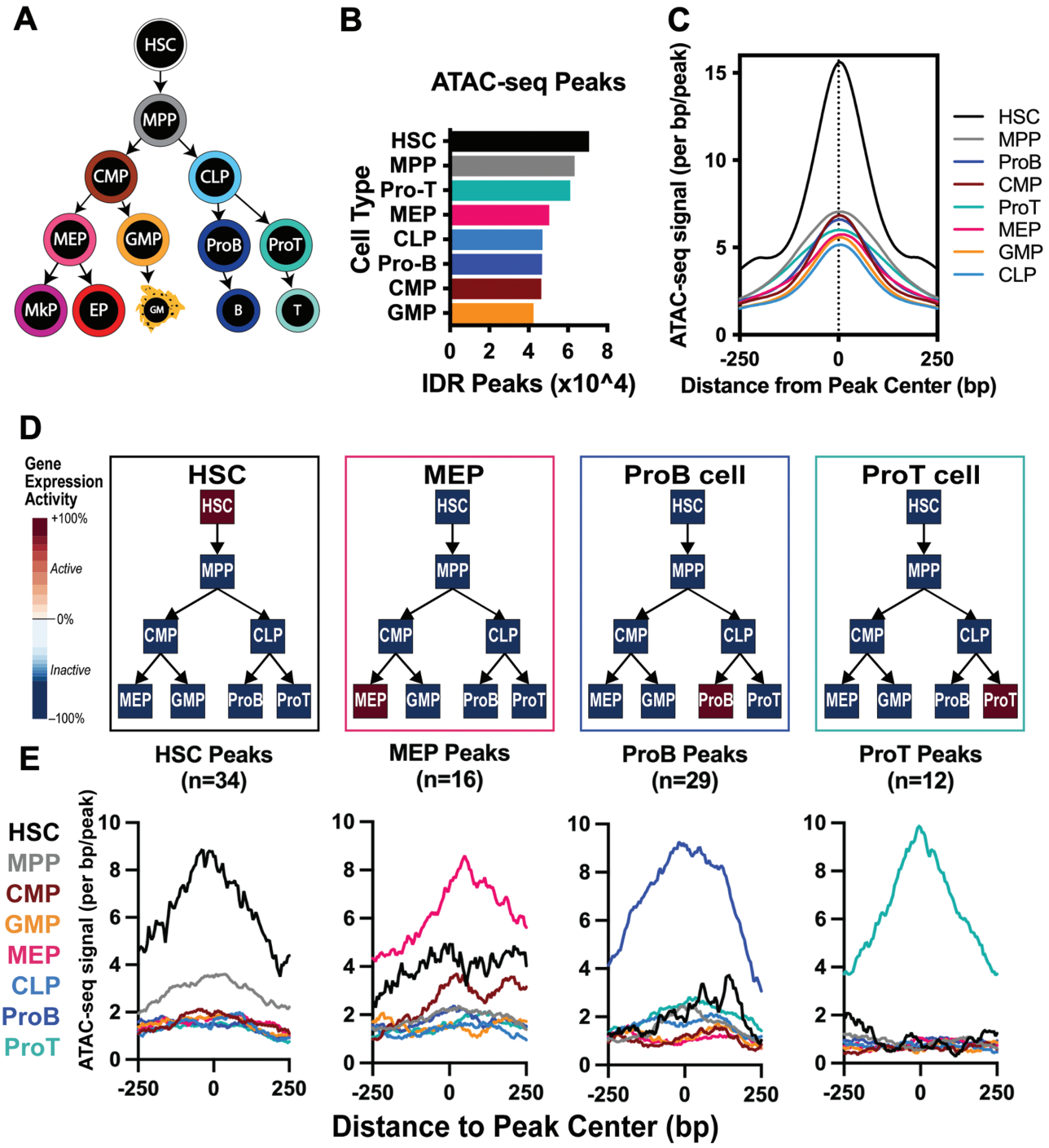


Figure 1. ATAC-seq analysis of hematopoietic progenitor cell populations revealed progressive and lineage-specific chromatin condensation. (**A**) Schematic diagram of the hematopoietic cells analyzed in this study. Thirteen cell populations, representing snapshots of a highly dynamic differentiation process, were investigated: multipotent Hematopoietic stem cells (HSCs) and multipotent progenitors (MPPs); lineage-restricted/oligopotent common myeloid progenitors (CMPs), common lymphoid progenitors (CLPs), Granulocyte macrophage progenitors (GMPs), megakaryocyte erythrocyte progenitors (MEPs); unilineage megakaryocyte progenitors (MkPs), Erythroid progenitors (EPs), B cell progenitors (ProBs), T cell progenitors (ProTs), and mature granulocyte/macrophages (GMs), B cells, and T cells. ATAC-seq profiles for HSCs and unilineage MkPs, EPs, GMs, B and T cells were reported previously[20]; data were integrated in selective analyses of the new data for intermediate progenitors for a comprehensive perspective of hematopoiesis. (**B**) HSCs had the highest number of peaks of all hematopoietic progenitor cell types. The total number of irreproducible discovery rate (IDR) peaks per cell type are displayed. HSCs had the highest number of peaks, followed by MPPs and then lineage-committed progenitors. (**C**) HSCs had the highest average signal across all peaks. Average cumulative signal across the peak-list for each population was determined by the -hist function of HOMER annotatePeaks.pl. Multipotent HSCs and MPPs had the highest average peak signal, whereas lineage-restricted progenitors had overall lower signal. (**D**) Lineage-specific gene expression patterns used to find examples of genes selectively expressed within each indicated cell type. The level of expression (+100% = high; −100% = low/not expressed) was obtained from the gene expression commons (GEXC) database. (**E**) Promoter accessibility correlated with cell type-specific gene expression in the corresponding progenitor cell types. Plots depict HOMER histograms of the average cumulative signal across the cell type-specific promoters for HSCs (34 peaks), MEPs (16 peaks), ProBs (29 peaks), and ProTs (12 peaks). MPPs, CMPs, GMPs, and CLPs were not displayed as each of these populations had fewer than 10 promoter peaks of uniquely expressed genes.

Table 1. Peak counts and peak distribution relative to protein-coding gene promoters in each cell type.

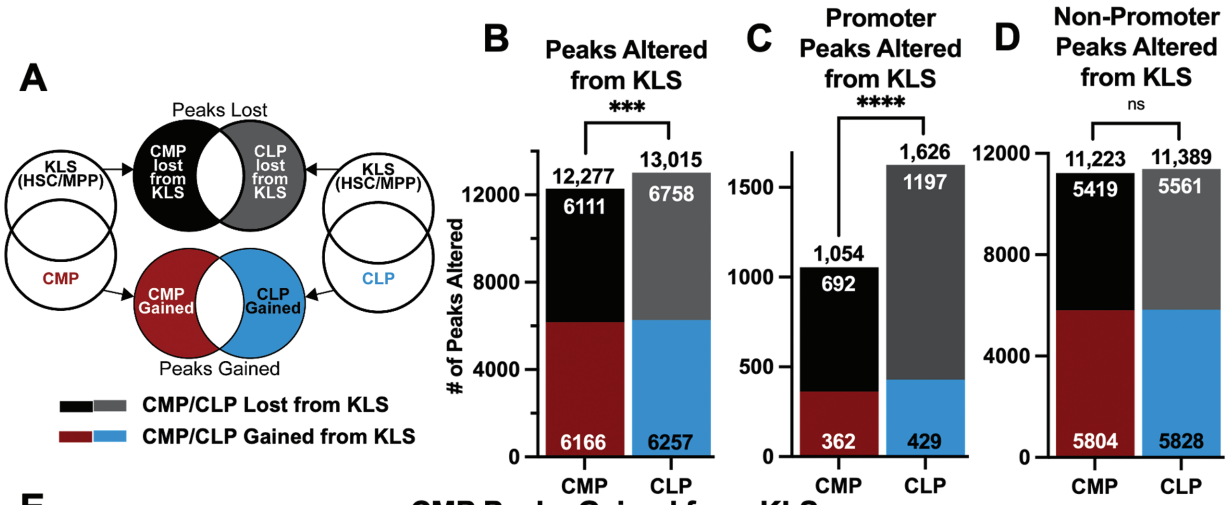
Cell type	ATAC peaks	Promoter peaks (±500bp of TSS)	Sum of all non-promoter peaks	Non-promoter peaks		
				coding (exons+TTS+TSS)	Introns	Intergenic
Master Peak-list	92 842	12 702	80 140	5 543	38 090	36 507
HSC	70 731	27 973	42 758	4 166	18 931	19 661
MPP	63 349	23 415	39 934	3 685	17 872	18 377
CLP	47 054	21 467	25 587	2 126	11 414	12 047
CMP	46 431	21 648	24 783	1 951	10 701	12 131
GMP	42 447	20 823	21 624	1 939	9 317	10 368
GM	30 529	15 559	14 970	1 440	6 697	6 833
MEP	50 483	26 064	24 419	2 281	10 803	11 335
EP	38 007	23 243	14 764	2 014	7 040	5 710
MkP	47 363	23 998	23 365	2 013	10 036	11 316
ProB	46 790	24 837	21 953	2 003	9 355	10 595
B	70 358	24 596	45 762	4 461	21 210	20 091
ProT	61 141	27 073	34 068	2 796	14 950	16 322
T	51 832	25 103	26 729	2 016	11 929	12 784

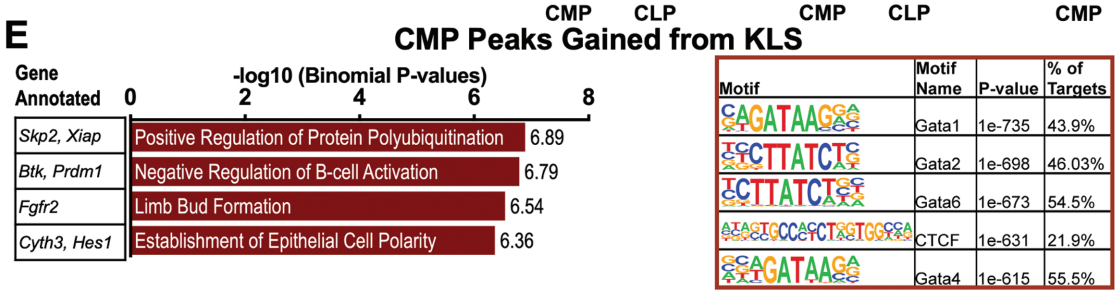
Breakdown of the peak numbers for either the master peak-list, or 1 of the 13 cell types described in this study. Peaks are categorized by: ATAC peaks (all peaks called/filtered for each cell type/master peak-list), promoter peaks (peaks that fall within 500 base pairs of the annotated transcription start site), non-promoter peaks (all peaks that fall outside of the promoter peak window), coding peaks (peaks called as exonic, transcription start site, or transcription termination site), intronic, and intergenic.

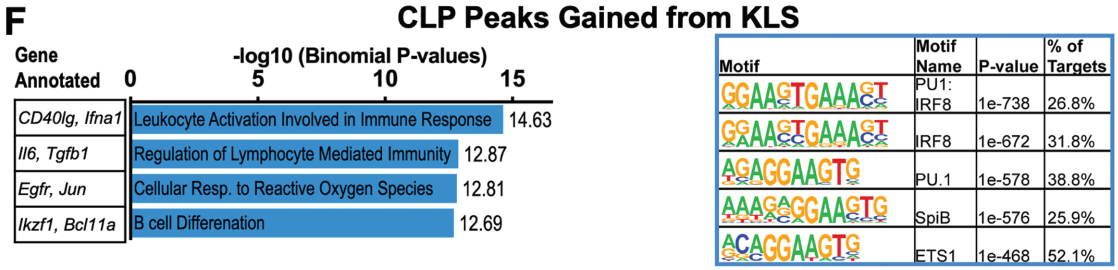
either CMPs or CLPs gained or lost from KLS. The CMP and CLP peaks were filtered against each other to focus only on peaks that were uniquely altered in either cell type. At a global level, CLPs had a higher number of peaks altered from KLS cells compared to CMPs with a significant difference in the distribution of peaks gained/lost between the 2 cell types (Fig. 2B). When categorizing peaks into promoter vs. non-promoter, we observed more promoter peaks altered in CLPs than CMPs (Fig. 2C), whereas similar numbers of non-promoter peaks were altered in both progenitors (Fig. 2D). We annotated the peaks that were gained and lost using Genomic Regions Enrichment of Annotations Tool (GREAT)[32,33] and reported the top 4 biological process gene ontology (GO) terms enriched, along with example genes in each GO term. We also performed motif enrichment by HOMER[30] (Fig. 2E–2H). The peaks gained in CMPs included "Negative Regulation of B-cell Activation," and the annotated genes of all 4 GO terms have known roles in myeloid differentiation, such as *Prdm1*[34] and *Btk*.[35] Gata1/2 motifs were among the highest enriched sequences (Fig. 2E). The CLP peaks gained were enriched for GO terms that pertained mainly to immune response and immunity, with genes *Ikzf1, Il6,* and *Jun* present within the top 4 GO terms, and were notably enriched with IRF8 and SpiB motifs (Fig. 2F). Peaks lost from KLS to CMPs were related to immune system activation and proliferation (Fig. 2G), with known immune development genes such as *CD180, Ikzf1, and Gata3*. In addition, there were enriched motifs from ETS/ERG transcription factors as well as SpiB, a known factor in immune development (Fig. 2G). In CLPs, peaks lost from KLS were enriched for GO terms related to the immune response, including genes expressed by innate immune cells that have known roles in complement activation (eg, *CD55*[36]), antigen presentation (eg, *Nod2, Lamp1*[37]), and activation of innate immune system pathways (eg, *Fcgr3, Ifng*[38,39]) (Fig. 2H). CLPs also lost peaks lined to *CD44*, *Gata2*, and *Gata3,* genes which all have known roles in HSC maintenance, engraftment, and self-renewal.[40–43] The CLP-lost peaks were enriched in motifs for erythromyeloid-specific Gata factors as well as CTCF motifs (Fig. 2H). These analyses suggest that at the first branchpoint, both myeloid and lymphoid differentiation require a combination of silencing of both HSC maintenance and alternative lineage genes and *de novo* activation of lineage drivers for the induced fate. Quantitatively, lymphoid differentiation appears to require more chromatin remodeling than myeloid differentiation, particularly in promoter regions.

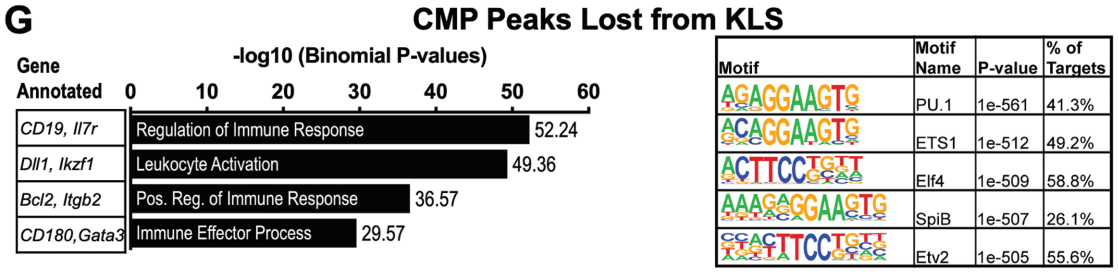
Differential Chromatin Dynamics at the Megakaryocyte-Erythroid Fate Branch

Using a similar strategy, we investigated epigenetic changes at another branchpoint in the hematopoietic hierarchy where MEPs differentiate in either MkPs or EPs (Fig. 3A). We identified peaks gained or lost between bipotent MEPs and unipotent MkPs or EPs, and then filtered them against each other to specifically focus on uniquely altered peaks upon fate determination (Fig. 3B). Interestingly, of the ~18 000 peaks altered, the distribution of peaks uniquely gained or lost from MEPs was opposite for MkPs and EPs: ~77% of peaks altered in MkPs were gained peaks, while a similar proportion (~61%) of peaks altered in EPs were lost (Fig. 3C). Of these, the large majority were non-promoter peaks (Fig. 3D, 3E). The gained and lost peaks were annotated using GREAT, for which we report the top 4 enriched GO Terms for "Mouse Phenotype" (Fig. 3F, 3G; Supplementary Fig. S1) and motif enrichment (Fig. 3H, 3I; Supplementary Fig. S1). The mutation/deletion of genes associated with the MkP-specific gained peaks led to phenotypes characterized under abnormal or decreased inflammation, as with multiple genes within these categories having known roles in megakaryopoiesis and platelet function/physiology (Fig. 3F). The mutation/deletion of genes associated with the EP-specific gained peaks led to phenotypes characterized under erythroid lineage, function,









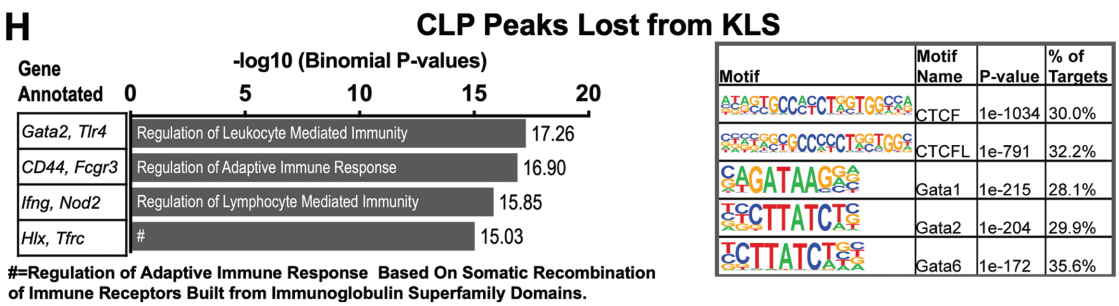


Figure 2. Comparisons of peak dynamics as multipotent HSCs and MPPs differentiate into CMPs or CLPs revealed quantitatively differential gain and loss of accessibility. (**A**) Schematic of the comparisons made between multipotent HSCs and MPPs (ckit+Lin-Sca1+; KLS) to lymphoid- or erythromyeloid-committed CLPs or CMPs. First, the peaks from HSCs and MPPs were combined using *bedtools merge* and then compared to CLPs or CMPs. The altered peak lists from the CMP and CLP comparisons were then intersected against each other to generate CMP- or CLP-specific peaks that were either gained or lost from KLS. (**B-D**) CLPs had more peak alterations than CMPs. The number of peaks gained and lost in each cell type are displayed. Compared to CMPs, CLPs had more total number of peaks gained/lost (B), promoter peaks altered (C), and similar numbers of non-promoter peaks altered (D). The distribution of peaks between CMPs and CLPs was significant by Chi-square for the total number of peaks (B) (***$P < .001$) and promoter peaks (C) (****$P < .0001$); and not significant for non-promoter peaks (D) ($P = 0.42$). **(E-H)** Cis-regulatory element analysis,

and morphology, with example genes including *Gata1* and *Slc4a1* (Fig. 3G). MkP (Fig. 3H) or EP (Fig. 3I) gained peaks were enriched for motifs of transcription factors known to be involved in the respective lineage. As examples of putative megakaryopoiesis-promoting CREs, we identified 2 MkP-specific peaks gained from MEPs in the gene *Alox5ap* (Fig. 3J). The promoter peak along with the enhancer just downstream of the promoter were called only in MkPs, while the second putative enhancer was called in both MEPs and MkPs, indicating possible priming in MEPs from this site. No peaks were called for this gene in EPs. The enhancer-gene map from ENCODE 3 showed interaction between the promoter of *Alox5ap* and 2 putative enhancer regions (as shown by the interaction lines below the signal track). Concordant with its chromatin accessibility, *Alox5ap* is highly expressed in MkPs but not MEPs or EPs (Fig. 3K).

Mapping of Chromatin Accessibility Throughout Hematopoiesis Identified Distinct Erythromyeloid and Lymphoid Clusters

To test our hypothesis that CREs primed in HSCs maintained accessibility throughout hematopoiesis, we needed to determine the dynamics of genome accessibility and further characterize lineage selective CREs throughout the whole continuum of hematopoiesis. To do so, we combined the ATAC-seq data from the 7 progenitors cell types with our previously reported HSCs and 5 unilineage cell types[20] (Fig. 1A). A master peak list of 92 842 peaks was produced by combining and filtering the peaks from 2 biological replicates for each of the 13 cell types using chromVAR[20,44] (Table 1). Principal component analysis (PCA) of the peak profiles of our 13 populations revealed a high concordance of replicates, as well as a distinct bifurcation of erythromyeloid and lymphoid populations, with the multipotent HSCs and MPPs landing within the erythromyeloid fraction (Fig. 4A). CMPs, MEPs, EPs, and MkPs all clustered together high on PC2, while CLPs, ProBs, and B cells clustered together, with ProTs and T cells grouped on the same PC1 scale but with higher PC2. HSCs and MPPs, together with GMPs and GMs, fell between the main myeloid and lymphoid groups. As a complement to PCA analysis, we performed uniform manifold approximation and projection (UMAP) using components derived from PCA of normalized ATAC-seq peak counts (Fig. 4B). We observed a similar bifurcation between erythromyeloid and lymphoid cell types with the multipotent HSCs and MPPs falling within the erythromyeloid quadrant. Additionally, we performed hierarchical clustering using the chromVAR output which similarly grouped the 13 populations into 2 distinct clusters, 1 erythromyeloid and 1 lymphoid (Fig. 4C). All biological replicates clustered directly next to each other, except for the 2 HSC samples which were separated by MPPs and GMPs. We ruled out batch-effects as closely associated samples were processed independently; the separation of the HSC replicates may instead reflect the presence of primed CREs of all lineages ([20]; also see below). In all 3 clustering analyses, the multipotent HSCs and MPPs associated near each other and within the erythromyeloid cluster, indicating a similar accessibility profile of these cell types. Overall, clustering analysis confirmed a high degree of reproducibility. Regardless of the method used, we observed distinct clustering based on similar accessibility profiles of lymphoid cell types, erythromyeloid cell types, multipotent HSCs, and MPPs, and of unipotent/mature cells with their presumed immediate upstream progenitor. The bifurcation of lymphoid and erythromyeloid lineages observed in the PCA (Fig. 4A), UMAP (Fig. 4B), and hierarchical clustering (Fig. 4C) is consistent with the erythromyeloid/lymphoid separation found in models of classical hematopoiesis, and with analyses independent of the purity of multi- and oligo-potent progenitor cells.[20] Of note, the observed similarity between HSCs/MPPs and erythromyeloid cells provides a potential epigenetic basis for the previously reported erythroid functional bias, where HSCs and MPPs predominantly produce red blood cells over all other cell types.[6]

Visualization and Comparison of ATAC-seq Data Generated in This Study Correlated With Known Expression Patterns at Two Well-Characterized Loci

To determine whether our mapping could detect known CREs, we visualized our ATAC-seq data across 2 well-characterized loci: the mouse β-globin cluster (Fig. 5A) and the mouse *Rag* locus (Fig. 5B, 5C). At the β-globin cluster (chr7: 103 792 027-103 879 340; mm10), we observed expected EP-selective accessibility of the HS3 site in the locus control region (LCR) and *β-minor (βmin)* promoter[45,46] (Fig. 5A). At the 3′ end of this gene we observed a B-lineage-selective peak. Further investigation revealed that the peak contains Pax5 binding motifs. It is possible that this site serves as a binding site for Pax5-mediated repression of globin gene expression in lymphoid cells[47]. Alternatively, the site could be a distal CRE for a gene regulating lymphoid development.[48] We observed erythroid-lineage specific accessibility (HSCs, MEPs, and EPs) of the β-major *(βmaj)* promoter as well as DNase I hypersensitive sites (HS1,2,4,6) of the LCR that are known to regulate erythroid-specific expression of the genes in this locus. This observation could indicate a "permissive" chromatin state in these erythroid-competent progenitor cells (HSCs, CMPs, MEPs, and EPs). Unexpectedly, we observed robust HS2 accessibility in GMPs, MkPs, and ProT cells, which are not currently known to have any erythroid cell potential. As expected, we did not observe any accessibility at the fetal-specific *epsilon Y globin (Ey)*, *β-h1 (βh1)*, *β-h2 (βh2)* genes,

GO term enrichment, and motif enrichment of the peaks that were altered between KLS and CLPs or CMPs, along with example target genes from each GO term. Briefly, each list of altered peaks was submitted to GREAT using the basal extension function with a parameter of 2kb upstream, 1kb downstream, and up to 1Mb extension. Example genes were extracted from the region-target association table for each GO term. The top 5 enriched known motifs from HOMER and corresponding transcription factors were also reported. (E) GREAT analysis of CMP-gained peaks contained the GO term "Negative Regulation of B cell Activation," and were enriched for motifs of Gata transcription factors. **(F)** Peaks gained by CLPs were primarily enriched in immune cell activation GO terms, with "Leukocyte Activation Involved in Immune Response" as the top hit. Peaks were enriched for motifs of ETS factor ETS1, as well as known lymphoid drivers IRF8 and SpiB. (G) CMP peaks that were lost from KLS cells all relate to immune cell processes, and were enriched with motifs for ETS factors and SpiB, similar to the peaks gained by CLPs. (H) CLP peaks lost from KLS contained GO terms that were immune related, such as "Regulation of Leukocyte Mediated Immunity" with *Gata2* and *Tlr4* as example genes. The peaks were enriched for Gata and CTCF/CTCFL transcription factor motifs. # the full title of this GO term is "Regulation of Adaptive Immune Response Based On Somatic Recombination of Immune Receptors Built from Immunoglobulin Superfamily Domains."

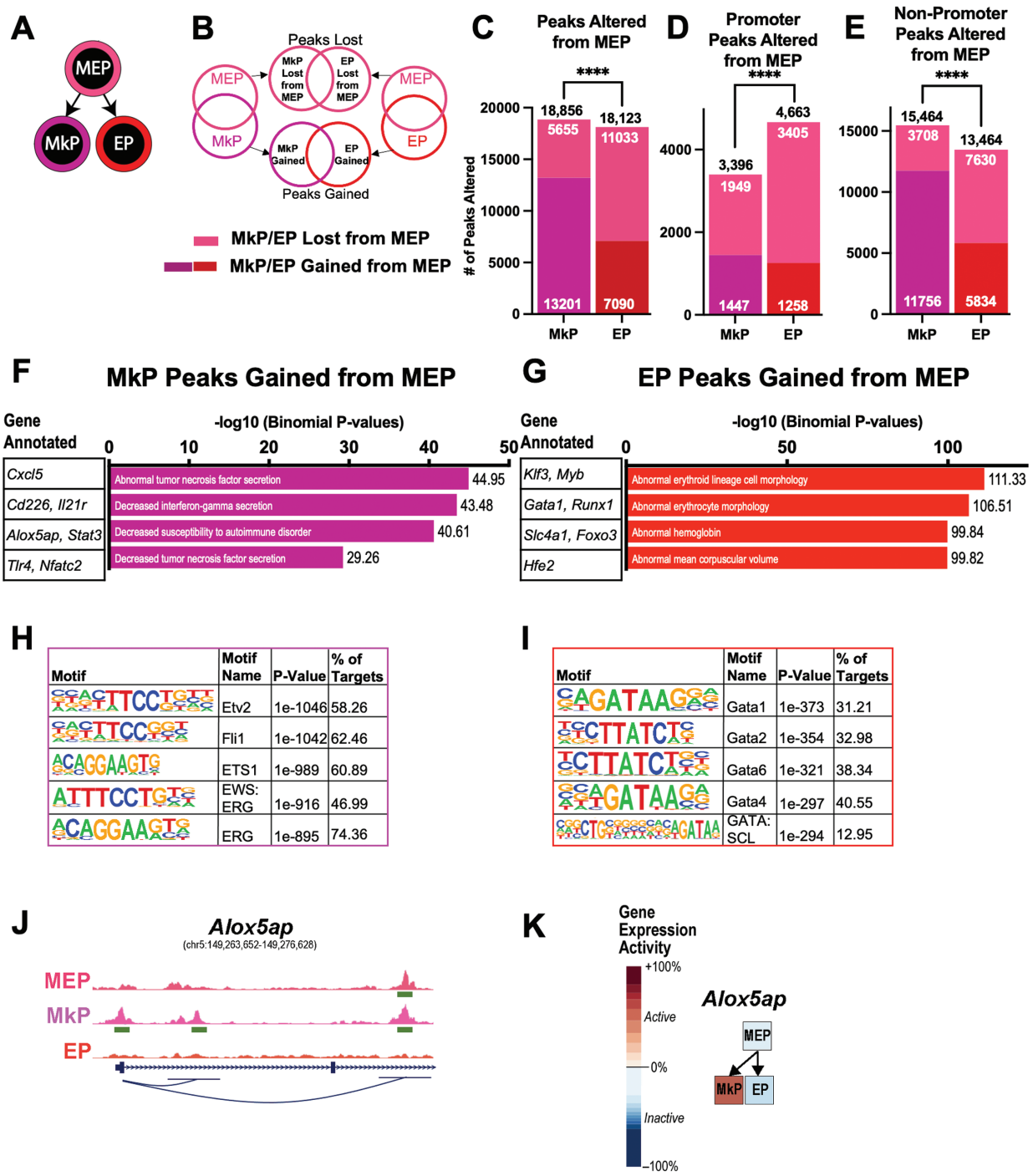


Figure 3. Comparison of peak dynamics as MEPs differentiate into MkPs or EPs revealed more gain of chromatin accessibility in MkPs and more loss in EPs. (**A**) Schematic of the differentiation branch analyzed for this figure, where MEPs differentiate into either MkPs or EPs. (**B**) Schematic of the comparisons made between MEPs and MkPs or EPs, similar to Fig. 2A. The peak profile of MEPs was compared to MkPs and EPs to assess which peaks were uniquely altered (gained or lost from MEPs) by MkPs or EPs during differentiation. (**C-E**) MkPs and EPs had a similar number of peaks altered. The number of peaks gained and lost in each cell type are also displayed. (D) Compared to EP, MkPs had a lower number of promoter peaks altered with a greater percentage of promoter peaks gained and (E) a greater number of non-promoter peaks gained. The distribution of peaks between MkPs and EPs was significantly different by Chi-square for (C-E) (**** $P < .0001$). **(F-G)** The lists of peaks gained from MEPs for each cell type were submitted to GREAT for functional annotation. The top 4 over-represented categories in Mouse Phenotype are reported, containing information about genotype-phenotype associations. Examples genes with known roles in MkPs (and/or platelets/megakaryopoiesis) or EPs (and/or red blood cells/erythropoiesis) were extracted from the term's genomic region-gene association tables. (F) The MkP gained peaks were enriched for genes whose alterations generate phenotypes related to inflammation. (G) The EP gained peaks were enriched for genes whose alterations generate phenotypes related erythroid cell lineage, function, and morphology. **(H-I)** Motif enrichment analysis by HOMER was performed on the lists of peaks gained from MEPs for each cell type and the top 5 transcription factor motifs were reported. (H) Peaks gained in MkPs were enriched for transcription factors known to be key players in the megakaryocytic lineage, such as Fli-1 and Erg. (I) Peaks gained in EPs were enriched for transcription factors required for erythropoiesis, including various Gata family members. **(J-K)** Example gene extracted from the lists of gained peaks in MkPs: Alox5ap. (J) ATAC-seq signal tracks for MEPs, MkPs, and EPs at the Alox5ap locus (12 000 bps shown). Peaks highlighted by green boxes represent called peaks by IDR at the promoter and putative enhancers for Alox5ap. (K) GEXC expression data reported high expression of Alox5ap in MkPs but not in MEPs or EPs.

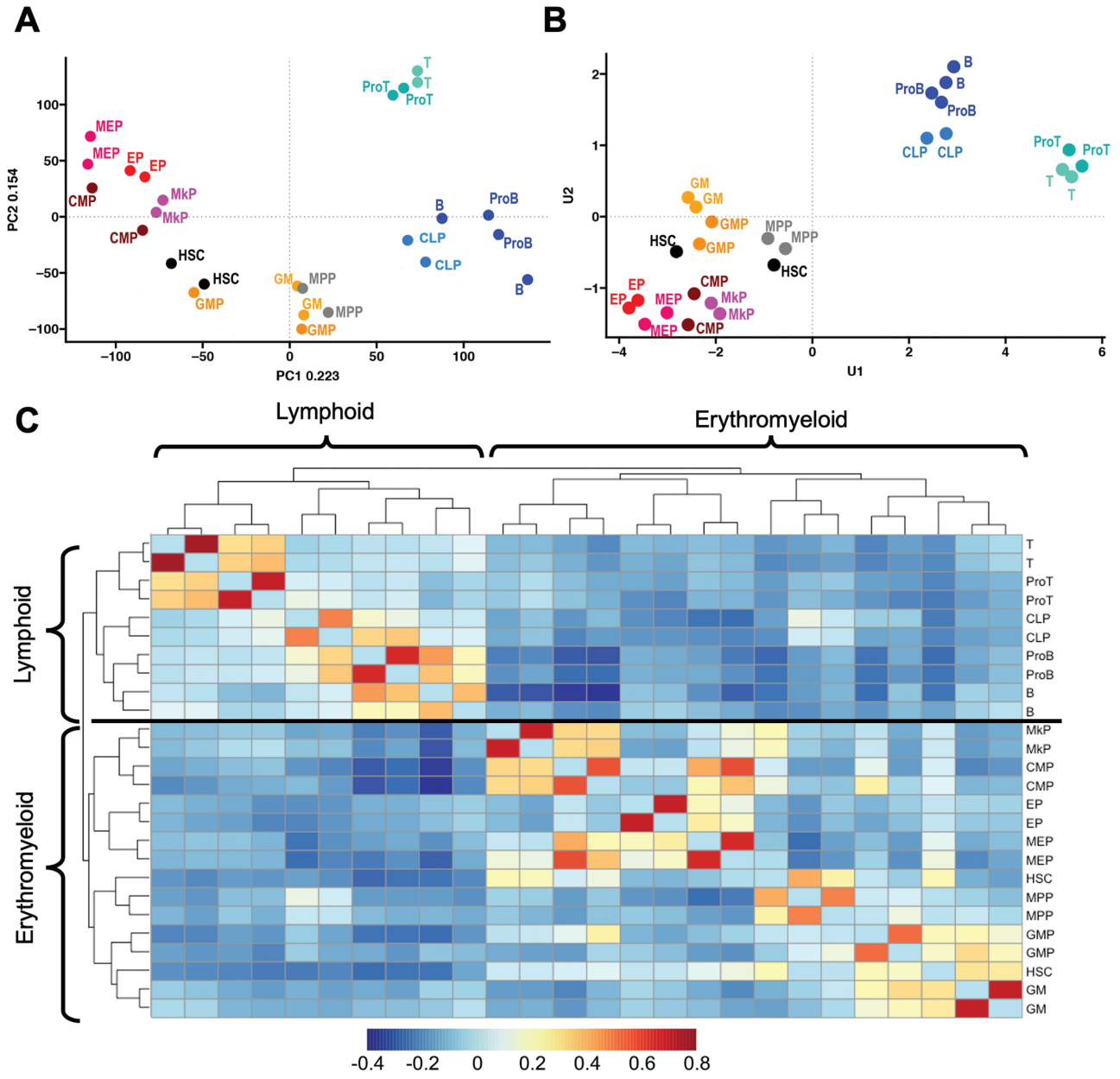


Figure 4. ATAC-seq maps of hematopoietic cell populations revealed distinct erythromyeloid and lymphoid clusters. (**A**) Principal Component Analysis (PCA) of chromVAR-normalized ATAC-seq peak counts revealed high concordance of replicates, and distinct erythromyeloid and lymphoid quadrants. Percent of total variance explained by each component are displayed on respective axes. (**B**) Uniform manifold approximation and projection (UMAP) using components derived from PCA generated distinct erythromyeloid and lymphoid clusters with the multipotent HSCs and MPPs associated with the erythromyeloid quadrant, similar to the PCA. (**C**) Hierarchical clustering of all 13 cell types revealed high concordance of replicates and distinct clusters consistent with classical models of hematopoiesis (Fig. 1A). Two primary associations were revealed: one erythromyeloid cluster and one lymphoid cluster. Multipotent HSCs and MPPs were designated to the erythromyeloid cluster. Additionally, there were four distinct sub-clusters: MkPs with CMPs; MEPs with EPs; ProBs with B cells and CLPs; and ProTs with T cells.

or HS5, and no accessibility was observed at any of these sites in GMs, CLPs, ProBs, and ProTs. Taken together, we observed expected accessibility in the β-globin locus in progenitors that give rise to cells that express β-globin genes, and little to no accessibility in progenitors that do not give rise to cells that express β-globin genes.

Similar specificity was observed for the *Rag* gene locus (chr2: 101 542 312-101 656 796; mm10) which consists of 4 CREs (Ep, D3, Erag, and ASE) and the gene bodies for *Rag1* and *Rag2*. Both *Rag1* and *Rag2* have lymphoid-specific gene expression patterns (Fig. 5B), and we observed lymphoid-specific accessibility of both *Rag1* and *Rag2* promoters (Fig. 5C). The progenitor- and myeloid-selective peak in the *Rag2* locus corresponds to the promoter for the *Iftap* gene that is expressed selectively by those cell types from the opposite strand of the *Rag* genes[49] (Fig. 5C). The CREs Ep and Erag, which have been characterized to be enhancers in B-cell lines,[50,51] exhibited CLP and B-cell specific (ProBs and B cells) accessibility. D3 has been characterized to act as a lymphoid-specific enhancer[51,52] and was accessible in all lymphoid cell types, while the previously characterized anti-silencing element (ASE), important for T-cell differentiation,[53,54] was only accessible in ProTs (Fig. 5C). In conclusion, our data demonstrated cell type-specific accessibility for

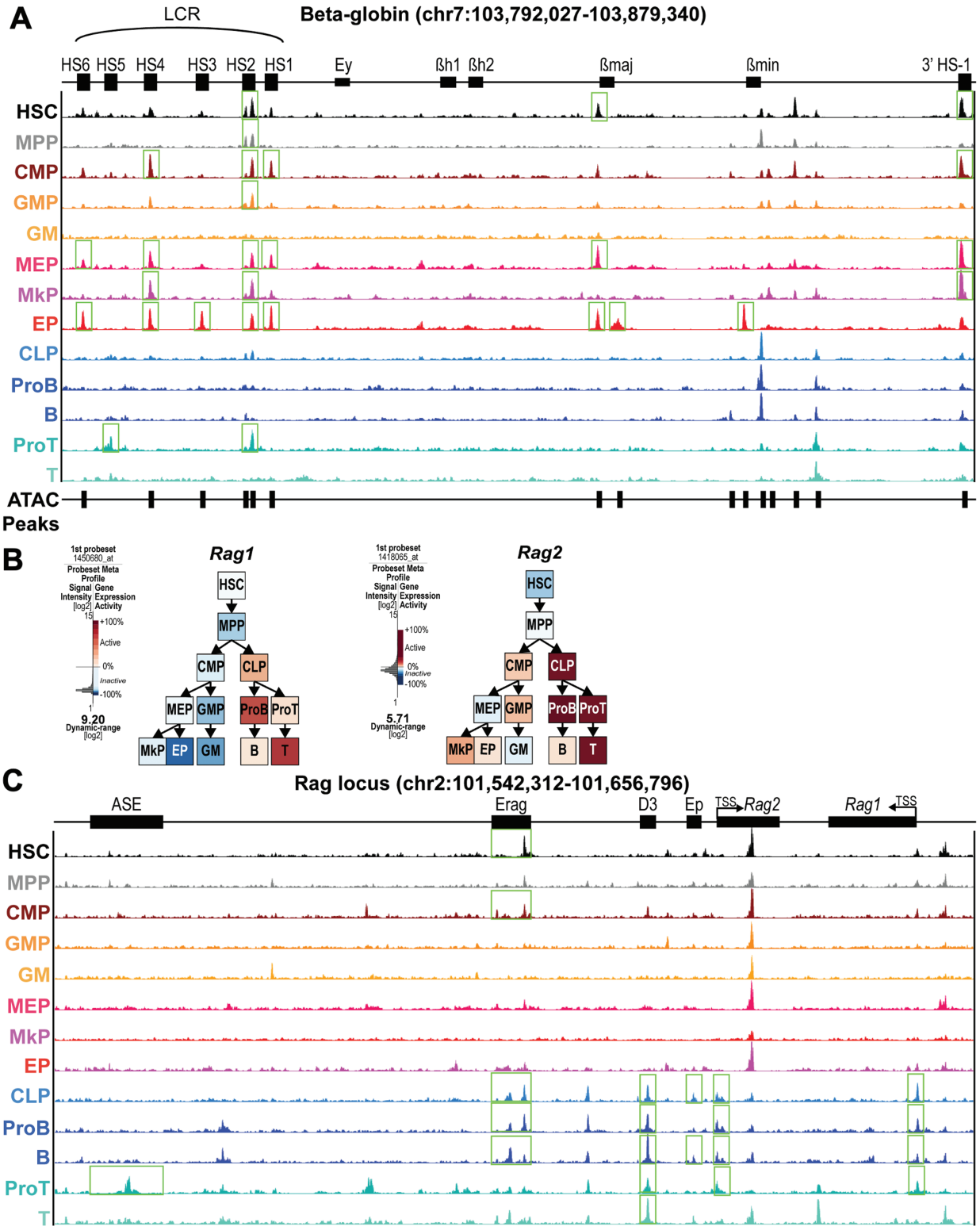


Figure 5. Accessibility correlated with known regulatory elements of well-characterized cell type-specific genes. (**A**) Chromatin accessibility of the β-globin locus revealed expression-selective patterns at known *cis*-regulatory elements (CREs). ATAC-seq signal tracks at the β-globin cluster (chr7: 103 792 027-103 879 340; mm10) of the thirteen cell types are shown. Peaks highlighted by boxes represent called peaks by Irreproducible Discovery Rate (IDR) at known CREs for each cell type. (B) Lymphoid-selective expression of *Rag1* and *Rag2*. GEXC expression data reported expression of *Recombination activating gene 1 (Rag1)* and *Recombination activating gene 2 (Rag2)* in CLPs, ProBs, ProTs, B, and T cells. *Rag2* expression in non-lymphoid cell types (CMPs, GMPs, MkPs, and EPs) is due to the *Iftap* promoter on the opposite strand of the *Rag* genes in the second intron of *Rag2*.[49] (C) Lymphoid-selective accessibility of the *Rag* locus. ATAC-seq signal tracks of the thirteen cell types in this study at the lymphoid-selective *Rag* gene locus (chr2: 101 542 312-101 656 796; mm10). The *Rag* gene locus consists of four previously characterized CREs (Ep, D3, Erag, ASE) and the gene bodies for *Rag1* and *Rag2*. The promoter for both *Rag1* and *Rag2* had accessibility only in lymphoid cell types (CLPs, ProBs, B cells, ProTs, and T cells). The lymphoid specific D3 CRE had expected lymphoid-only accessibility, and the B-cell specific CREs Ep and Erag had accessibility only in CLPs, ProBs, and B cells. The T-cell development specific anti-silencing element (ASE) only exhibited accessibility in ProT cells.

multiple progenitors and recapitulated the dynamics of regulatory element priming throughout differentiation at 2 well-characterized erythroid and lymphoid loci, suggesting that our dataset is sufficiently robust and accurate to also reveal novel CREs that can be functionally tested using the CRISPRi strategies described below.

A Subset of Lineage-Specific CREs Was Primed in HSCs as Well as in Select Progenitors

Previously, we reported evidence of multilineage priming in HSCs of CREs specific for each unipotent lineage.[20] We hypothesized that lineage-primed CREs are maintained throughout differentiation. To test this, we first compared the average cumulative accessibility of the lineage-specific peaks (Supplementary Tables S1–S5) primed in HSCs to all 13 cell types (Fig. 6A). As expected, we observed strong signals from HSCs and the corresponding unipotent progenitor cell type for each lineage-specific primed peak list. MPPs had a discernable peak in 4 out of the 5 primed peak lists, with a less distinct signal in EP-primed peaks. Notably, each unilineage region displayed accessibility signal in the presumed immediate upstream progenitor (MEPs in EP-primed peaks; GMPs in GM-primed peaks; ProB in B-cell peaks; and ProTs in T-primed peaks), except for MkPs, which lacked MEP signal and instead had notable accessibility in MPPs and CMPs. These observations revealed that lineage priming of a sizeable proportion of CREs persists throughout differentiation for every lineage.

To assess the distribution of the primed peaks in each progenitor population, we performed a *bedtools* intersect of the lineage-specific peaks primed in HSCs and determined the number of overlapping peaks with each progenitor. Interestingly, all progenitors from every lineage contained peaks from all 5 primed peak lists (Fig. 6B, Supplemental Tables S1-S5). The distribution of primed peaks of all 5 lineages was about equally distributed at ~20% each in HSCs, with similar distribution in MPPs, and CLPs (Chi-square > 0.01). Clear lineage bias was evident in other populations: erythromyeloid progenitors (CMPs, MEPs) were significantly enriched for EP- and MkP-primed peaks, GMPs were enriched for GM-primed elements, while the unipotent lymphoid progenitors (ProBs and ProTs) were significantly enriched with peaks from their immediate downstream progeny (B and T cells).

To accomplish a direct longitudinal analysis of priming through multiple differentiation stages, we intersected the 5 HSC-primed peak lists with every assumed intermediate progenitor between HSCs and the unipotent lineage (ie, for HSC/EP shared peaks, we intersected MPPs, CMPs, and MEPs with the shared peak list, as those populations are in the HSC-to-EP lineage as shown in Fig. 1A). From those intersections, we identified and quantified the number of peaks that maintained accessibility throughout differentiation for each lineage and reported each peak as a heatmap throughout the expected differentiation trajectory (Fig. 6C; Supplemental Table S6). Surprisingly, even though ubiquitously primed CREs were detected for every lineage, this was far from the norm, as no lineage had more than 25% of the HSC-primed peaks maintain openness throughout differentiation. About 10% of the persistently primed peaks were promoters, with the B-cell lineage-specific *BAFF-R* and T-cell-specific *CD28* as examples that have known functional roles in those cell types.[55,56] Next, we examined 2 example CREs that were primed throughout differentiation for the GM (Fig. 6D) and T-cell (Fig. 6E) lineages. The GM-specific *Fcnb* gene is expressed only in GMPs and GMs, while the putative CRE associated with *Fcnb* was accessible in HSCs, MPPs, CMPs, GMPs, and GMs (Fig. 6D). The T-cell-specific *Wnt8b* is only expressed in T cells, while its putative CRE is accessible in HSCs, MPPs, CLPs, ProTs, and T cells (Fig. 6E). These findings support that lineage priming observed in HSCs is maintained throughout differentiation for certain CREs. Unexpectedly, most of the peaks primed in HSCs did not exhibit persistent priming in every intermediate progenitor. These observations, combined with the bias in signal and peak counts in progenitors could suggest preferred lineages at specific branch points. For example, EP-primed peaks had a high average signal and made up most of the overlapping peaks in MEPs which could suggest that MEPs are biased toward EPs over MkPs, or reinforce fate decisions initiated in upstream progenitors.

HSC-Unique Peaks Indicated an Erythropoiesis-Primed Chromatin State

Because HSCs are the only cell type in the hematopoietic tree that is capable of long-term reconstitution, we reasoned that HSC-unique peaks would be enriched in elements that promote self-renewal and/or engraftment. To test this, we identified and examined HSC-unique peaks (Fig. 7A). We found 3026 HSC-unique peaks, 92.7% of which were classified as non-promoter (Fig. 7B). To identify transcription factor motifs enriched within the HSC-unique peaks, we performed *de novo* motif finding and enrichment using the HOMER package and reported the top 10 results sorted by *P*-value (Fig. 7C). ELF3 (E74 Like ETS Transcription Factor 3) was the top-ranked motif, followed by CTCFL. There were 3 instances of CTCF-like motifs in the top 10 *de novo* motifs, while single instances of NF-E2, RUNX, HIC1, Gata6, Foxo1, and IRF4 rounded out the enriched motifs. Next, we annotated the CREs to nearby genes using GREAT. The top GO term was *definitive erythrocyte differentiation* (Fig. 7D), comprised of 14 peaks linked to 4 genes: *Ncor1, Tgfbr3, Zfpm1,* and *Smarca4*. All 4 genes have known roles in hematopoiesis, with knock-out studies presenting severe defects in erythropoiesis, or the entire hematopoietic compartment,[57–61] consistent with important roles in HSCs. We then visualized the ATAC accessibility of 3 example peaks enriched in the *definitive erythrocyte differentiation* GO term, along with their respective linked motif enrichment (Fig. 7E–7G). The CRE linked to *Ncor1* contained the NF-E2 and Foxo1 motif (Fig. 7E). The CRE linked to *Zfpm1* contained the motif for ELF3 (Fig. 7F), while the CRE linked to *Tgfbr3* contained CTCFL and Foxo1 motifs (Fig. 7G). Taken together, the unique HSC peaks are enriched for elements that prime erythroid cell fate in HSCs, such as NF-E2 binding sites and the 14 peaks that are linked to *Ncor1, Tgfbr3, Zfpm1,* and *Smarca4,* all of which have known roles in erythropoiesis.

CRISPRi-Mediated Targeting Functionally Linked CREs to Gene Expression

An immediate extension of the chromatin accessibility mapping accomplished in this study is to determine the functional role of putative CREs. Linking regulatory elements to the expression of specific genes is essential for understanding epigenetic regulation of fate decisions, but has proven persistently challenging, in particular in the context of native chromatin

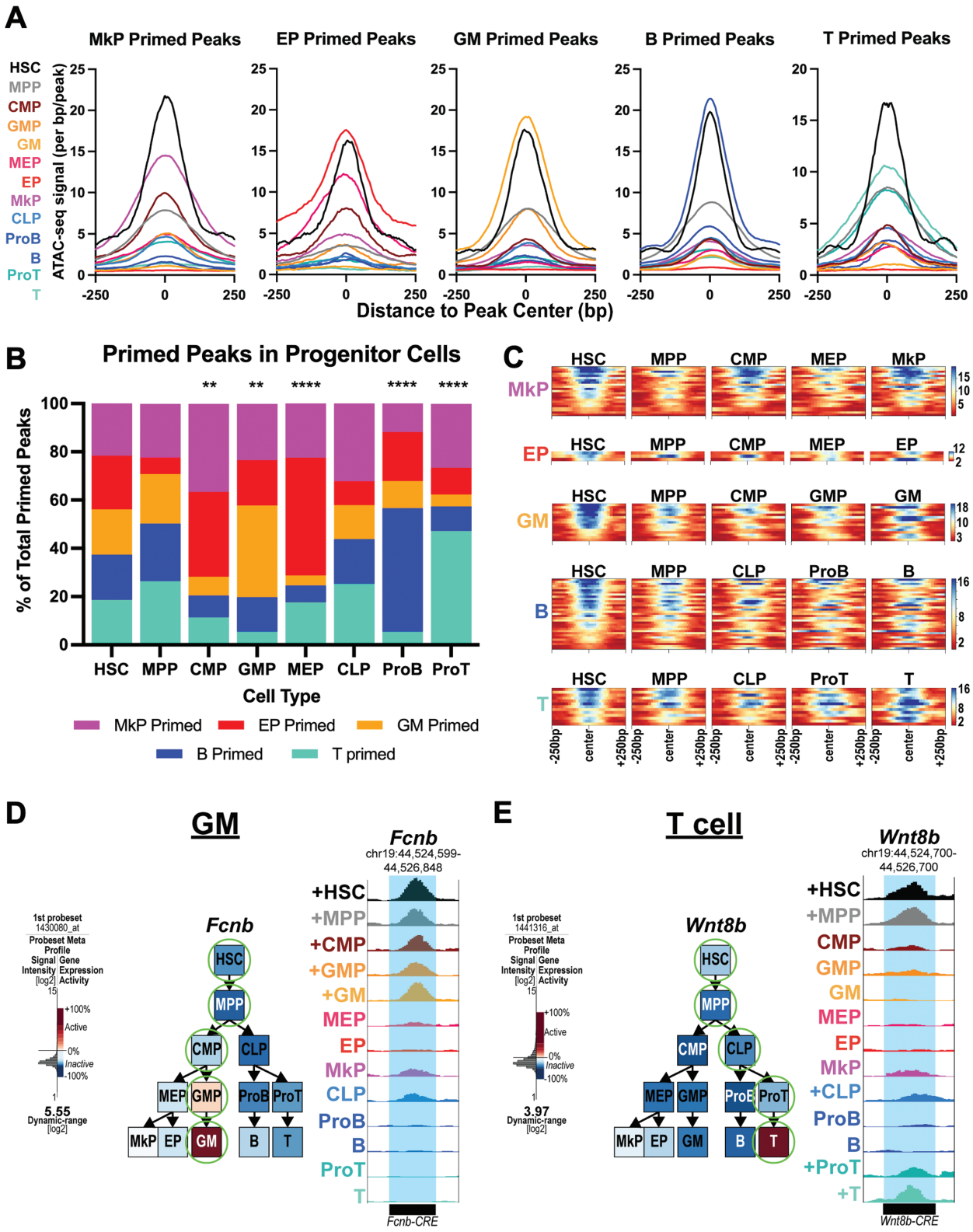


Figure 6. CREs of lineage-specific genes primed in HSCs also displayed accessibility in progenitors. (**A**) Lineage-specific peaks primed in HSCs also displayed selective enrichment in intermediate progenitors. HOMER histograms of the average cumulative accessibility in each of the 13 cell types in each lineage-primed peak-list. MkP lineage peaks that were primed in HSCs were also enriched in MPPs and CMPs, but less so in GMPs, CLPs, ProB, and ProTs; EP peaks were selectively enriched in MEPs and CMPs; GM peaks were enriched primarily in MPPs and GMPs; B cell peaks were enriched in ProBs and MPPs, and T cell peaks were enriched in ProTs and MPPs. (**B**) Peak distribution analysis revealed lineage skewing within progenitors. The distribution of lineage-primed peaks was displayed for each progenitor cell type. All progenitors contained lineage-primed peaks representing unique peaks of each of the five lineages, but at different proportions. HSCs had an almost equal distribution of peaks from all five lineages that did not deviate from an expected equal distribution (Chi-square, $P = .97$). MPPs and CLPs had similar peak distributions and were not significantly different when compared pairwise to HSCs (Chi-square, $P \geq .01$). In contrast, pairwise comparison of the distribution of peaks between HSCs and progenitors revealed

structure of primary, multipotent cells. To begin to tackle this, we employed a new genetic mouse model that ubiquitously expresses the dCas9-KRAB repressor protein from the safe harbor H11 locus (CRISPRi mouse) (Fig. 7H). We isolated hematopoietic stem and progenitor cells (HSPCs) from these CRISPRi mice and first, as proof-of-concept, transduced them with lentivirus expressing a single guide RNA (sgRNA) targeting the promoter of the cell surface protein CD81.[62] HSPCs were cultured under non-differentiation conditions for 2 days, then analyzed for CD81 cell-surface expression by flow cytometry. We observed a significant reduction in CD81 expression in HSPCs transduced with CD81 sgRNA compared to a scrambled control sgRNA (Fig. 7I), thereby demonstrating efficient and selective gene silencing in HSPCs.

Next, we tested whether CRISPRi targeting initiated in HSCs is maintained upon differentiation into mature cell and if silencing of putative distal CREs, like promoter-proximal CREs, can repress gene expression. We designed 2 sgRNAs for both the promoter and a putative CRE of the myeloid-associated cell surface proteins CD115 and CD11b, based on their ATAC peak profiles (Fig. 7J) and cloned the top 2 sgRNAs for each putative CRE into a dual-guide lentiviral vector.[63] HSCs from CRISPRi mice were isolated, transduced with the dual-guide lentivirus, and then cultured in a myeloid differentiation media for 5 days before surface expression of CD115 and CD11b was quantified via flow cytometry. As expected, the vast majority of cells infected with scrambled control sgRNAs expressed CD115 (Fig. 7K) or CD11b (Fig. 7L). We observed a significant reduction in the frequency of CD115+ cells when HSCs were transduced with sgRNAs targeting either the promoter or putative CRE targeting guides (Fig. 7K). Interestingly, we only observed a significant decrease in the frequency of CD11b+ cells with sgRNAs targeting the CD11b promoter, but not the putative CD11b CRE (Fig. 7L). These data establish the new CRISPRi mice as a powerful tool for identifying functional CREs based on ATAC-seq peaks. Specifically, we showed that CRISPRi-mediated targeting initiated in HSCs led to selective gene repression upon differentiation into myeloid cells and that the accessible sequence in the CD115 locus is essential for its expression, whereas targeting of the putative CD11b CRE does not significantly alter CD11b expression under these conditions.

Discussion

Global Chromatin Accessibility Throughout Hematopoiesis Is Highly Dynamic

Here, we mapped accessible loci in 7 hematopoietic progenitor cell types with distinct functional capacities. Integration of these new data with HSCs and mature progeny revealed epigenetic-based cell clustering into erythromyeloid and lymphoid branches (Fig. 4) and robust identification of known regulatory elements (Fig. 5). Consistent with previous evidence by us and others that stem cells have relatively decondensed chromatin structure,[19,20,64,65] we found that both the ATAC peak number and a cumulative signal was greatest for HSCs (Fig. 1). This study advances previous reports by pinpointing the location both of all putative CREs genome-wide in each population, as well as HSC-specific putative CREs (Fig. 7) and those associated with the major erythromyeloid/lymphoid branch points (Figs. 2 and 3). Importantly, only a subset of HSC peaks remained accessible throughout differentiation, substantially focusing the search for sequences serving as lineage priming elements (Fig. 6). Likewise, our CRISPRi experiments demonstrated selective functionality of putative distal CREs in regulating gene expression (Fig. 7). Although the models of hematopoiesis and phenotypes of progenitors are constantly reshaped and refined,[66,67] our maps and strategies derived from pre-defined cell types at approximate branchpoints will serve as valuable resources to identify, characterize, and functionally interrogate *cis*-regulatory elements and their roles in gene regulation, stem cell self-renewal, and fate decisions.

Defining Differential Chromatin Accessibility at Major Lineage Branchpoints

Functional studies have suggested that differential epigenetic priming may be evident at major branchpoints.[6,27] Indeed, we identified thousands of differential CREs at both the CMP/CLP and MkP/EP branchpoints. We found that HSC/MPP-descendant CLPs had a greater proportion of lost peaks compared to CMPs (Fig. 2). Interestingly, though most of the peaks altered were classified as non-promoter, the differential was mainly driven by promoter peak changes, possibly indicating that multipotency priming, but not implementation of non-lymphoid programs, may remain present in CLPs. Emergence of CMPs was associated with significantly less promoter remodeling, loss of accessibility of lymphoid lineage drivers, and gain of accessibility at negative regulators of lymphoid differentiation (Fig. 2C, 2E, 2G). Further down the erythromyeloid trajectory, MkPs gained significantly more peaks from MEPs compared to EPs, with gained peaks in MkPs and EPs linked to genes important for megakaryopoiesis and erythropoiesis, respectively (Fig. 3). In contrast to the CMP/CLP branchpoint, the divergence of MkPs versus EPs occurred primarily at non-promoter peaks. Interestingly, the specification of both CMPs (compared to CLPs) and EPs (compared to MkPs) entailed significantly reduced proportions of "peaks lost," indicating that relatively larger fractions of CREs were already accessible in their respective progenitor (MPPs and MEPs). This is consistent with strong erythroid priming in HSCs (also see below) and the predominant erythroid cell

significant differences in CMPs, GMPs, MEPs, ProBs, and ProTs by Chi-square. CMPs had a relative expansion primarily of erythromyeloid (MkP, EP) peaks; GMPs had primarily GM-unique peaks; MEPs were enriched for EP-unique peaks; whereas ProBs had more B cell peaks, and ProTs had mainly T-cell peaks. $^{**}P < .01$, $^{****}P < .0001$. (**C**) Heatmaps of primed peaks that maintain accessibility throughout the expected differentiation trajectory for each lineage. Each line is one peak, with accessibility indicated in blue centered around the peak +/−250 bp. Less than 30% of the primed peaks for each lineage followed the expected trajectory by maintaining accessibility throughout differentiation. 17% of MkP peaks, 11% of EP peaks, 13% of GM peaks, 12% of B cell peaks, and 26% of T-cell peaks maintained priming throughout differentiation. (**D**) A *cis* regulatory element (CRE) predicted by GREAT to be associated with *Fcnb* maintained accessibility ("priming") throughout differentiation into GMs. GEXC reported expression of *Fcnb* selectively in GMPs and GMs. Green circles indicate which cell type contained a called peak. Genome track snapshot of the *cis* regulatory element of *Fcnb* reported accessibility in HSCs, MPPs, CMPs, GMPs, and GMs. A "+" sign designated which cell type contained a called peak. (**E**) A CRE predicted by GREAT to be associated with *Wnt8b* maintained accessibility throughout differentiation into T cells. GEXC reported expression of *Wnt8b* selectively in T cells only. Green circles indicate which cell type contained a called peak. Genome track snapshot of the *cis* regulatory element of *Wnt8b* reported accessibility in HSCs, MPPs, CLPs, ProTs, and T cells. A "+" sign designated which cell type contained a called peak.

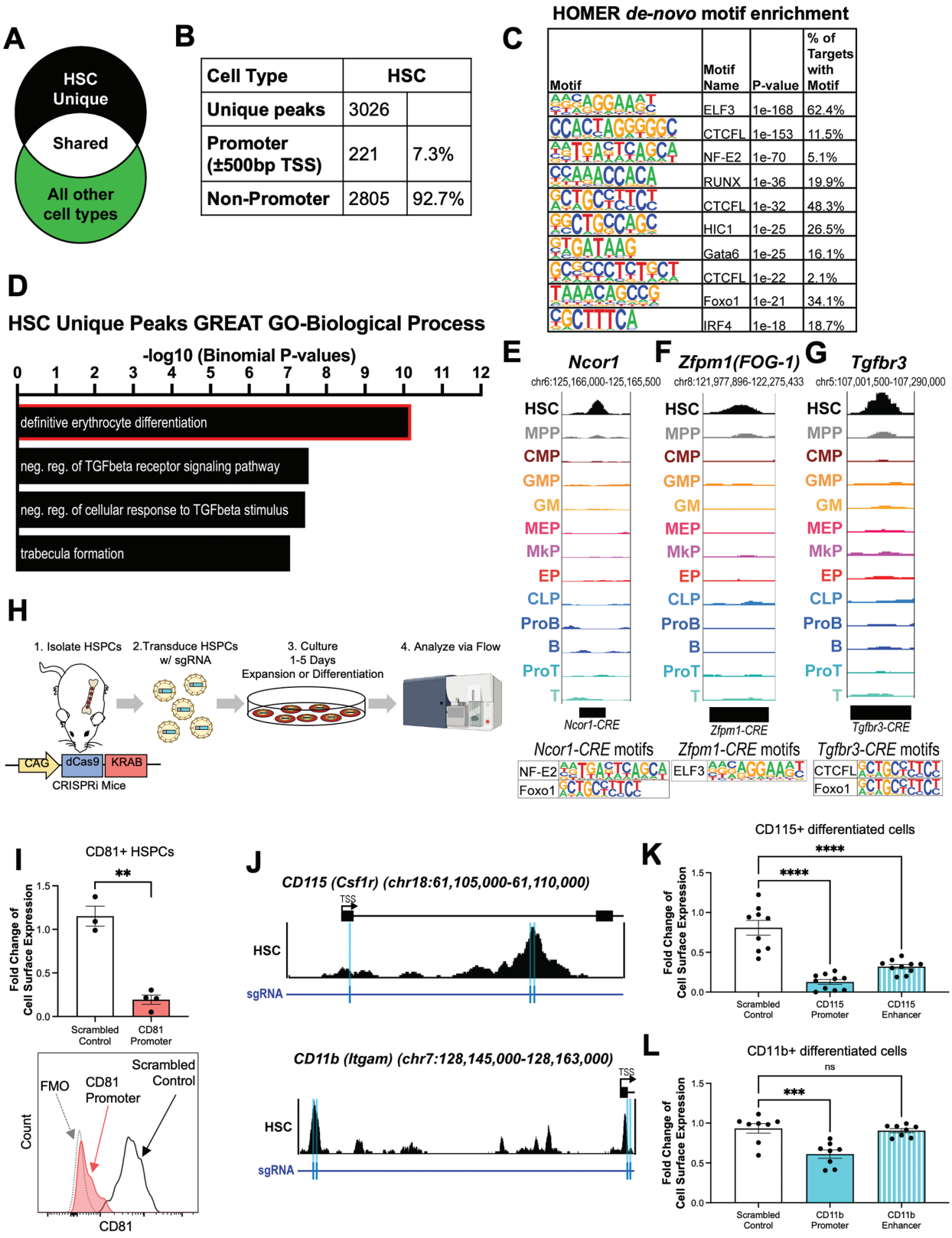


Figure 7. HSC-unique *cis*-regulatory elements are primarily enriched for transcription factors that drive erythropoiesis. (**A**) The HSC-unique peak-list was generated by filtering HSC peaks against the peak lists of the other 12 hematopoietic cell types. (**B**) HSC-unique peaks are primarily non-promoter peaks. Table of the composition of the HSC-unique peaks and percentage of non-promoter and promoter peaks. **(C)** *De novo* motif enrichment of HSC-unique peaks revealed binding sites for known hematopoietic regulators. ELF3, CTCFL, NF-E2, and Runx motifs were the top 5 enriched *de novo* motifs. (**D**) "Definitive erythroid differentiation" was the top enriched GO term from GREAT annotation and analysis of the unique HSC peaks. The resulting graphs are GO Biological Process terms and the −log10 *P*-value for the top four terms. (**E-G**) Three examples of putative CREs for target genes that were enriched in "definitive erythrocyte differentiation" and displayed unique HSC accessibility. (E) A putative CRE for *Ncor1* was unique to HSCs and contained motifs that closely match NF-E2 and Foxo1 binding sites. (F) A putative *Zfpm1* CRE contained the binding motif that closely matches

production observed in quantitative functional assays.[6] Collectively, the differential priming throughout differentiation uncovered here and previously[20,68] is now available for functional testing by CRISPRi or analogous strategies to determine the CREs and chromatin remodeling events required for balanced hematopoiesis.

Lineage Priming Was Selectively Maintained Throughout Differentiation

Our previous study identified CREs that were exclusively shared between unipotent lineage cells and HSCs.[20] Here, we found that a limited subset of those primed CREs maintained accessibility throughout differentiation in intermediate progenitors (Fig. 6). Lineage priming was also detected at the β-globin locus, where the strongest enhancer, HS2, was primed in HSCs and MPPs, with additional accessibility of HS1 and HS4 in CMPs, and then also HS3 and HS6 in MEPs and EPs (Fig. 5). The global enrichment of peaks within the intermediate progenitors reflected the distinct bifurcation found in hematopoiesis, with erythromyeloid-primed peaks enriched in erythromyeloid progenitors and lymphoid-primed peaks enriched in lymphoid progenitors. Surprisingly, CLPs did not significantly deviate from the distribution of peaks in HSCs and MPPs (Fig. 6B), potentially indicating "inherited" priming that is not implemented in vivo[69] but can be reignited in vitro.[70] We also tracked the accessibility of the primed peaks throughout differentiation and found that the majority of peaks do not maintain accessibility in every intermediate progenitor throughout differentiation (Fig. 6C, 6D). Collectively, these findings provide insight into the dynamics of CRE accessibility throughout differentiation and support a model where lineage priming in HSCs guides lineage competence during differentiation, while the gain and loss of accessibility at certain intermediate progenitors could regulate or reinforce differentiation in specific lineages.

HSC-Unique Peaks Were Highly Enriched for CREs That Drive Erythroid Differentiation

While HSCs are capable of producing all blood cell lineages, several studies have suggested lineage-specific priming within HSCs.[6,71–74] From these studies we hypothesized that CREs within HSCs would uncover drivers of erythro- and/or megakaryopoiesis. Our GREAT analysis of HSC-unique peaks revealed "definitive erythrocyte differentiation" as the top GO-Biological Process hit (Fig. 7D), and we observed HSC-specific accessibility in the CREs linked to genes that have known roles in erythropoiesis (Fig. 7E–7G). Furthermore, we observed de novo enrichment of transcription factor motifs in the HSC-unique peaks that are known to be key regulators of hematopoiesis, such as NF-E2 and Runx.[75–77] This suggests that the establishment of developmental competence for erythropoiesis in HSCs may occur primarily in CREs that are uniquely accessible in HSCs.

In summary, we present evidence that multilineage priming is present in HSCs and selectively maintained, or repressed, throughout differentiation. In addition, the observation that HSCs harbor the most ATAC-seq peaks of all hematopoietic cell types (Fig. 1) is consistent with previous findings that linked multipotency with global epigenetic regulation and the presence of poised loci that are distal to promoters in stem cells.[19,64,65] We also found that accessibility, especially of distal CREs, is highly dynamic, and that some, but not all, serve as functional CREs (Fig. 7I–7L). Our results provide insight into how lineage fate is reinforced at branchpoints through the collective action of specific transcription factors at these CREs. Future investigations using CRISPR-based technologies paired with in vivo methods ([78,79] Fig. 7I–7L) will allow us to determine which of these CREs are a consequence of differentiation and which elements drive differentiation into specific fates.

Experimental Procedures

Mice and Cells

All experiments were performed using 8- to 12-week-old C57BL/6 wild-type mice in accordance with UCSC IACUC guidelines. Hematopoietic stem, progenitor, and mature cells were isolated from BM of murine femurs, tibias, hips, and sternums as previously described.[1,25,80,81] Stem and progenitor cell fractions were enriched using cKit-coupled magnetic beads (Miltenyi). Cells were stained with unconjugated lineage rat antibodies (CD3, CD4, CD5, CD8, B220, Gr1, Mac1, and Ter119) followed by goat-α-rat PE-Cy5 (Invitrogen). Stem and progenitor cells were isolated using fluorescently labeled or biotinylated antibodies for the following antigens: cKit (2B8, Biolegend), Sca1 (D7, Biolegend), Slamf1 (CD150) (TC15-12F12.2, Biolegend), CD34 (RAM34, ebiosciences), FcγrII (93, Biolegend), and Il7rα (A7R34, Biolegend). Cells were sorted by fluorescence-activated cell sorting (FACS) with a target of 50 000 cells per cell type using a FACS Aria II (BD Bioscience). HSCs were defined as cKit$^+$ Lin$^-$ Sca1$^+$ Flk2$^-$ and Slamf1$^+$; MPPs as cKit$^+$Lin$^-$Sca1$^+$ Flk2$^+$ Slamf1$^-$ cells. CMPs were defined as cKit$^+$Lin$^-$Sca1$^-$ CD34^{mid} FcγrII^{mid}; GMPs as cKit$^+$Lin$^-$Sca1$^-$ CD34^{mid} FcγrII^{high}, MEPs as cKit$^+$Lin$^-$Sca1$^-$ CD34^{low} FcγrII^{low}. CLPs were isolated by lineage-depleting BM cells through staining of unconjugated lineage rat antibodies (CD3, CD4, CD5, CD8, B220, Gr1, Mac1, and Ter119) followed by sheep-α-rat Dynabeads (Life Technologies) and separation via EasySep magnet (STEMCELL Technologies). CLPs were isolated by Lin$^-$ Flk2$^+$ Il7rα$^+$ cKit^{mid} Sca1^{mid}. Lineage-restricted hematopoietic

ELF3. (G) A putative *Tgfbr3* CRE contained DNA motifs that closely matched CTCFL and Foxo1 binding sites. (**H**) Experimental setup using a CRISPRi model to functionally test putative CREs identified in this study. (**I**) CD81 expression was significantly reduced in HSPCs when CRISPRi HSPCs were transduced with lentivirus targeting the CD81 promoter. The fold change in the frequency of CD81+ cells of transduced cells compared to untransduced cells is represented in the bar graph and the representative histogram of CD81 expression in HSPCs transduced with CD81 promoter targeting sgRNA (red) compared to HSPCs transduced with a non-targeting scrambled sgRNA (blue) and CD81 FMO (grey dotted line). (**J**) ATAC-seq accessibility profiles of the CD115 (top) and CD11b (bottom) loci. The location of the single guide RNAs (sgRNA) designed to target the promoter or a putative CRE of each gene are denoted by blue bars below the respective locus. (**K**) CD115 expression was significantly reduced in differentiated cells when CRISPRi HSCs were transduced with lentivirus targeting either the CD115 promoter or enhancer. The fold change in the frequency of CD115+ cells of transduced cells compared to untransduced cells is represented in the bar graph. (**L**) CD11b expression was significantly reduced in differentiated cells when CRISPRi HSCs were transduced with lentivirus targeting only the CD11b promoter, but not the enhancer. The fold change in the frequency of CD11b+ cells of transduced cells compared to untransduced cells is represented in the bar graph.

progenitor and mature cells were isolated by the following markers: EPs, Lin(CD3, CD4, CD5, CD8, B220, Gr1, and Mac1)$^-$CD71$^+$Ter119$^{+/-}$; GMs, Lin(CD3, CD4, CD5, CD8, B220, and Ter119)$^-$Gr1$^+$Mac1$^+$ ("GM" cells were positive for both Gr1 and Mac1); T-progenitors (ProT), Lin(CD5, B220, Gr1, Mac1, and Ter119)$^-$CD3$^+$CD25$^+$; T cells, Lin(CD5, B220, Gr1, Mac1, and Ter119)$^-$CD25$^-$CD3$^+$CD4$^{+/-}$CD8$^{+/-}$; B-progenitors (ProB), Lin(CD3, CD4, CD8, Gr1, Mac1, and Ter119)$^-$CD43$^+$B220$^+$; B cells, Lin(CD3, CD4, CD8, Gr1, Mac1, and Ter119)$^-$CD43$^-$B220$^+$.

CRISPRi mice were generated by site-specific integration of CAG promoter-driven sequences for nuclease-deficient Cas9 protein (dCas9) fused to a zinc-finger protein 10 (ZNF10) Krüppel-Associated Box (KRAB) domain. The expression cassette was inserted into the H11 safe harbor locus of C57BL/6J mice using a site-specific integrase-mediated method.[82] The resulting CRISPRi mouse is similar to a previously generated CRISPRi mouse,[62] but lacks the mCherry and Puro resistance genes.

ATAC-seq

ATAC-seq was performed as previously described.[24] Briefly, cells were collected after sorting into microcentrifuge tubes, and centrifuged at 500 × *g* for 5 min at 4 °C to pellet the cells. The supernatant was aspirated, and the cells were washed with ice-cold 1xDPBS. Cells were centrifuged and the supernatant was discarded. Cells were immediately resuspended in ice-cold lysis buffer (10 mM Tris–HCl, pH 7.4, 10 mM NaCl, 3 mM MgCl2, and 0.1% IGEPAL CA-630) and centrifuged at 500 × *g* for 10 min. The supernatant was aspirated, and pellets were resuspended in transposase reaction mix (25 µL 2 × TD buffer, 2.5 µL transposase (Illumina), and 22.5 µL nuclease free water). The transposition reaction was carried out at 37 °C for 30 min at 600rpm in a shaking thermomixer (Eppendorf). Immediately after completion of the transposition reaction, the samples were purified using the MinElute Reaction Clean up kit (Qiagen) and eluted into 10 µL of EB. Samples were stored at −20 °C until PCR amplification step. PCR amplification was performed as previously described[24] using custom Nextera primers. After initial amplification (5 cycles), a portion of the samples was run on qPCR (ViiA7 Applied Biosystems) to determine the additional number of cycles needed for each library (typically 5-8 cycles). The libraries were purified using the MinElute Reaction Clean up kit (Qiagen), eluted into 20 µL EB and then size selected using AmpureXP (Beckman-Coulter) beads at a ratio of 1.8:1 beads/sample, and eluted into 40 µL of nuclease-free water. Library size distribution was determined by Bioanalyzer (Agilent) capillary electrophoresis and library concentration was determined by Qubit 3 (Life Technologies). Quality of libraries was checked by shallow sequencing (1 million raw reads) on a Miseq (Illumina) at 75 × 75 paired-end sequencing. Those libraries that appeared to have size distributions similar to previous reports were pooled together and deep sequenced on a HiSeq2500 (Illumina) at 100 × 100 reads at the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley.

Data Processing

Demultiplexed sequencing data were processed using the ENCODE ATAC-seq pipeline version 1.1.6 and 1.4.2 (https://github.com/ENCODE-DCC/atac-seq-pipeline) using the mm10 assembly and the default parameters. In version 1.4.2 changed: atac.multimapping = 0, atac.smooth_win = 150, atac.enable_idr = true, atac.idr_thresh = 0.1 to be consistent with the mapping/peak calling performed with previous versions.

Peak filtering and hierarchical clustering were performed using the chromVAR package (https://github.com/GreenleafLab/chromVAR). First, the optimal peak-list from the IDR output for each cell type was concatenated and sorted, then used as the peak input for chromVAR. The blacklist filtered bam files for each replicate (n = 2 for each cell type) were used as input along with the sorted peak file. The fragment counts in each peak for each replicate and GC bias was calculated, and then the peaks were filtered using filterPeaks function with the default parameters and nonoverlapping = TRUE. The master peak-list was extracted at this point, which contained 92 842 peaks, and used throughout the study. The deviations were calculated using every peak, and the tSNE and correlation functions were also performed using the deviations output and the default parameters.

Normalized chromVAR counts were log + 1 scaled, centered, and filtered to peaks that had above-median coefficient of variance. These filtered counts were used in principal component analysis (PCA) with the R package *prcomp*. Following this, the resulting components were used to calculate UMAP dimensions using the R package *uwot*. The component values were plotted using the R package *ggplot2*.

We noted that although showing strong similarity by PCA (Fig. 4A) and designating to the same main cluster by hierarchical clustering (Fig. 4C), the 2 HSC samples were the only replicate pair not localized immediately adjacent to each other. Analysis of additional HSC ATAC-seq samples did not resolve the hierarchical clustering in a meaningful way, as the results were similar with regard to intermixing within other cell types and failed to pinpoint one of the original HSC samples as an outlier.

Annotation of peaks, generation of histogram plot, merging of peaks, and motif enrichment were performed by HOMER (http://homer.ucsd.edu/homer/). Peaks were annotated using the annotatePeaks.pl function with the mm10 assembly and default parameters. Histogram was created by first shifting the bam files using DeepTools alignmentSieve.py with the flag –ATACshift. Next, tag directories were made using the Tn5 shifted bam files using HOMER makeTagDirectory. The histogram was made using the annotatePeaks.pl function with the default settings and the flags: -size -500, 500 and -hist 5. Peak lists were compared using the mergePeaks.pl function with default settings and the flags -d given, -venn, and for the unique peak lists -prefix. Motif enrichment was performed using the findMotifsGenome.pl package with default parameters using the flag -size given.

The GREAT tool (http://great.stanford.edu/public/html/) was used to annotate non-promoter peaks to target genes. The peak lists were reduced to BED4 files from the HOMER annotations output and used as input. The whole mm10 genome was used as the background regions, and the association rule settings were set as Basal plus extension, proximal window 2kb upstream, 1kb downstream, plus distal up to 1Mb and included curated regulatory domains. All genome track visualizations were made using the UCSC genome browser. Statistical analysis was performed using GraphPad Prism 9. Graphs were made in either Microsoft Excel or GraphPad Prism 9. Annotations to figures were performed using Adobe Illustrator CC and Adobe Photoshop CC.

CRISPRi Experiments

Single guide RNA (sgRNA) sequence targeting CD81 was from a previous study.[62] Dual targeting sgRNAs for promoter regions were non-overlapping guide sequences from previously published libraries.[83] The dual sgRNAs for candidate CREs were designed by generating a 200bp window centered on the peak summit and inputting those coordinates into the CRISPOR tool.[84,85] The top 2 non-overlapping guides were selected and cloned into pJR85 (Addgene plasmid #140095).[63] psPAX2 (Addgene plasmid #12260) and pMD2.G (Addgene plasmid # 12259) were combined with pJR85 and transfected into 293T cells by Lipofectamine 2000. Seventy-two hours after transfection, the supernatant was collected, 0.45 µ filtered, and concentrated by PEG precipitation. Concentrated lentivirus was resuspended in a minimal volume of Optimem.

HSPCs (Lin–, cKit+, Sca+ BM cells) or HSCs (Lin–, cKit+, Sca+, SLAM mid-hi, Flk2-) from CRISPRi mice were FACS isolated and plated at 6000 cells/well (CD81) or 200 cells/well (CD11b/CD115) in either HSC minimal media (IMDM; TPO 50 ng/mL; SCF 50 ng/mL; polybrene 5 ug/mL) or HSC maintenance media (IMDM; 20% FBS; TPO 50 ng/mL; SCF 50 ng/mL; IL-6 20ng/mL; IL-3 10 ng/mL, IL-11 20 ng/mL, Primocin, non-essential amino acids). After 24 h in culture, lentivirus containing sgRNAs was added to each well and spinoculated for 1 h, 400 × *g* at 32 °C. Twenty-four hours later, virus was washed out and the media was changed into HSC minimal media (CD81) or to a myeloid differentiation media (CD11b/CD115)[86] and cells were allowed to expand in culture for 2 days (CD81) or 5 days (CD11b/CD11b) before analysis by flow cytometry.

List of sgRNAs used:

CD81 promoter	ATGAGACGTAGGGTAGAGAA
CD115 promoter 1	GAGCGTGAGCCGATGCAGGT
CD115 promoter 2	GCCGATGCAGGTTGGAGAGT
CD11b promoter 1	GCTTCTGGTCACAGGTATGT
CD11b promoter 2	GGTAGGTGGGGAGAGATCAA
CD115 enhancer 1	GTGAGAGCCCAAGTGTCGAA
CD115 enhancer 2	CAATGTGTTTCCGCCCACAC
CD11b enhancer 1	AGTTGTCTATATCCGCTGTG
CD11b enhancer 2	GGTCTGAATCACTAAAGATA
Scrambled control (for CD81 expts.)	GTCCATACGCATAATCACCG
Scrambled control 1 (for CD115/CD11b expts.)	CTGTGCAATCCGCATGATAT
Scrambled control 2 (for CD115/CD11b expts.)	ATCTGGCACCTCACCCACGT

Acknowledgments

We thank Bari Nazario and the IBSC flow cytometry core (RRID:SCR_021149) for assistance and support; Sol Katzman for bioinformatic assistance; Joseph Replogle, Angela Pogson, and Jonathan Weissman for CRISPRi vectors and advice, and Forsberg lab members for comments on the manuscript. This work was supported by NIH/NHLBI (R01HL115158), NIDDK (R01DK100917) and NIA (R01AG062879) awards to E.C.F.; by NIH/NHLBI fellowship (F31HL144115) to E.W.M.; by CIRM SCILL grant TB1-01195 to E.W.M. via San Jose State University; by CIRM Training grant TG2-01157 to J.K.; by a UCSC Genomic Sciences Graduate Training Program from NIH/NHGRI (NIH T32 HG008345) and an F99 (5F99DK131504-02) to R.E.R., by a Tobacco-Related Disease Research Program (TRDRP) predoctoral award to A.R.y.B. (T31DT1690); by an NIH T32 Training Grant (T32GM8646), a TRDRP predoctoral award (T30DT0878) and NHLBI F31 predoctoral award (F31HL151199) to A.K.W., by an NIH IMSD Training Grant (R25GM058903), an American Heart Association Predoctoral Fellowship Award, and Howard Hughes Medical Institute Gilliam fellowships to D.M.P.; by the Baskin School of Engineering and the Ken and Glory Levy Fund for RNA Biology to D.H.K., and by CIRM Shared Stem Cell Facilities (CL1-00506) and CIRM Major Facilities (FA1-00617-1) awards to the University of California, Santa Cruz.

Conflict of Interest

S.C. is a consultant for NextRNA Therapeutics and received honoraria for NextRNA from the American Association of Immunologists. All the other authors declared no potential conflicts of interest.

Author Contributions

E.W.M.: conception and design, financial support, collection and/or assembly of data, data analysis and interpretation, manuscript writing, final approval of manuscript. A.R y B.: conception and design, financial support, collection and/or assembly of data, data analysis and interpretation, manuscript writing, final approval of manuscript. R.E.R: financial support, data analysis and interpretation, final approval of manuscript. A.K.W.: conception and design, financial support, collection and/or assembly of data, data analysis and interpretation, manuscript writing, final approval of manuscript. C.S.M: collection and/or assembly of data, final approval of manuscript. D.M.P.: financial support, collection and/or assembly of data, final approval of manuscript. J.K.: collection and/or assembly of data, final approval of manuscript. M.T.M.: provision of study material, final approval of manuscript. S.C.: provision of study material, final approval of manuscript. D.H.K.: financial support, final approval of manuscript. E.C.F.: conception and design, financial support, data analysis and interpretation, manuscript writing, final approval of manuscript.

Data Availability

The datasets generated in the current study are available in the Gene Expression Omnibus (GEO), accession number GSE184851. The exclusively primed peaks are provided in Supplementary Tables S1–S6. Previously published datasets are available at GSE162949. Signal tracks and called peaks for all 13 cell types are available as a custom session on the UCSC genome browser at: https://genome.ucsc.edu/s/ewmartin/atac_bw_mean_allpeaks

Supplementary Material

Supplementary material is available at *Stem Cells* online.

References

1. Boyer SW, Schroeder A. V, Smith-Berdan S, Forsberg EC. All hematopoietic cells develop from hematopoietic stem cells through Flk2/Flt3-positive progenitor cells. *Cell Stem Cell*. 2011;9(1):64-73. https://doi.org/10.1016/j.stem.2011.04.021.
2. Orkin SH, Zon LI. Hematopoiesis: an evolving paradigm for stem cell biology. *Cell*. 2008;132:631-644. https://doi.org/10.1016/j.cell.2008.01.025.
3. Pronk CJH, Rossi DJ, Månsson R, et al. Elucidation of the phenotypic, functional, and molecular topography of a myeloerythroid progenitor cell hierarchy. *Cell Stem Cell*. 2007;1:428-442. https://doi.org/10.1016/j.stem.2007.07.005.
4. Krause DS. Plasticity of marrow-derived stem cells. *Gene Ther*. 2002;9:754-758. https://doi.org/10.1038/sj.gt.3301760.
5. Seita J, Weissman IL. Hematopoietic stem cell: self-renewal versus differentiation. *WIREs Syst Biol Med*. 2010;2:640-653. https://doi.org/10.1002/wsbm.86.
6. Boyer SW, Rajendiran S, Beaudin AE, et al. Clonal and quantitative in vivo assessment of hematopoietic stem cell differentiation reveals strong erythroid potential of multipotent cells. *Stem Cell Rep*. 2019;12:801-815. https://doi.org/10.1016/j.stemcr.2019.02.007.
7. Surani MA, Hayashi K, Hajkova P. Genetic and epigenetic regulators of pluripotency. *Cell*. 2007;128:747-762. https://doi.org/10.1016/j.cell.2007.02.010.
8. Whyte WA, Orlando DA, Hnisz D, et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell*. 2013;153:307-319. https://doi.org/10.1016/j.cell.2013.03.035.
9. Creyghton MP, Cheng AW, Welstead GG, et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci*. 2010;107(50):201016071.
10. Heintzman ND, Hon GC, Hawkins RD, et al. Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature*. 2009;459:108-112. https://doi.org/10.1038/nature07829.
11. Koch CM, Andrews RM, Flicek P, et al. The landscape of histone modifications across 1% of the human genome in five human cell lines. *Genome Res*. 2007;17:691-707. https://doi.org/10.1101/gr.5704207.
12. Rada-Iglesias A, Bajpai R, Swigut T, et al. A unique chromatin signature uncovers early developmental enhancers in humans. *Nature*. 2011;470:279-283. https://doi.org/10.1038/nature09692.
13. Visel A, Blow MJ, Li Z, et al. ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature*. 2009;457:854-858. https://doi.org/10.1038/nature07730.
14. Hu M, Krause D, Greaves M, et al. Multilineage gene expression precedes commitment in the hemopoietic system. *Genes Dev*. 1997;11:774-785. https://doi.org/10.1101/gad.11.6.774.
15. Chambers SM, Boles NC, Lin KYK, et al. Hematopoietic fingerprints: an expression database of stem cells and their progeny. *Cell Stem Cell*. 2007;1:578-591. https://doi.org/10.1016/j.stem.2007.10.003.
16. Forsberg EC, Prohaska SS, Katzman S, et al. Differential expression of novel potential regulators in hematopoietic stem cells. *PLoS Genet*. 2005;1:e28. https://doi.org/10.1371/journal.pgen.0010028.
17. Phillips R, Ernst RE, Brian B, et al. The genetic program of hematopoietic stem cells. *Science. (80-)*. 2000;288:1635-1640.
18. Terskikh A. V, Miyamoto T, Chang C, Diatchenko L, Weissman IL. Gene expression analysis of purified hematopoietic stem cells and committed progenitors. *Blood*. 2003;102:94-101. https://doi.org/10.1182/blood-2002-08-2509.
19. Ugarte F, Sousae R, Cinquin B, et al. Progressive chromatin condensation and H3K9 methylation regulate the differentiation of embryonic and hematopoietic stem cells. *Stem Cell Rep*. 2015;5:728-740. https://doi.org/10.1016/j.stemcr.2015.09.009.
20. Martin EW, Krietsch J, Reggiardo RE, et al. Chromatin accessibility maps provide evidence of multilineage gene priming in hematopoietic stem cells. *Epigen Chromatin*. 2021;14:1-15.
21. Gross DS, Garrard WT. Nuclease hypersensitive sites in chromatin. *Annu Rev Biochem*. 1988;57:159-197. https://doi.org/10.1146/annurev.bi.57.070188.001111.
22. Heintzman ND, Stuart RK, Hon G, et al. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet*. 2007;39:311-318. https://doi.org/10.1038/ng1966.
23. Buenrostro J, Wu B, Chang H, Greenleaf W. ATAC-seq: a method for assaying chromatin accessibility genome-wide. *Curr Protoc Mol Biol*. 2015;109:21.29.1-21.29.9.
24. 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 Meth*. 2013;10:1213-1218.
25. Poscablo DM, Worthington AK, Smith-Berdan S, Forsberg EC. Megakaryocyte progenitor cell function is enhanced upon aging despite the functional decline of aged hematopoietic stem cells. *Stem Cell Rep*. 2021;16:1598-1613. https://doi.org/10.1016/j.stemcr.2021.04.016.
26. Mohrin M, Bourke E, Alexander D, et al Hematopoietic stemcell quiescence promotes error-prone DNA repair and mutagenesis. *Cell Stem Cell* 2010;7(2):174-185.
27. Rodriguez-Fraticelli AE, Wolock SL, Weinreb CS, et al. Clonal analysis of lineage fate in native haematopoiesis. *Nature*. 2018;553:212-216. https://doi.org/10.1038/nature25168.
28. Gaspar-Maia A, Alajem A, Meshorer E, Ramalho-Santos M. Open chromatin in pluripotency and reprogramming. *Nat Rev Mol Cell Biol*. 2011;12:36-47. https://doi.org/10.1038/nrm3036.
29. Li Q, Brown JB, Huang H, Bickel PJ. Measuring reproducibility of high-throughput experiments. *Ann. Appl. Stat*. 2011;5:1752-1779.
30. Heinz S, Benner C, Spann N, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell*. 2010;38:576-589. https://doi.org/10.1016/j.molcel.2010.05.004.
31. Seita J, Sahoo D, Rossi DJ, et al. Gene expression commons: an open platform for absolute gene expression profiling. *PLoS One*. 2012;7:e403211-e403211. https://doi.org/10.1371/journal.pone.0040321.
32. McLean CY, Bristor D, Hiller M, et al. GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol*. 2010;28:495-501. https://doi.org/10.1038/nbt.1630.
33. Tanigawa Y, Dyer ES, Bejerano G. WhichTF is functionally important in your open chromatin data? *PLoS Comput Biol*. 2022;18:e1010378. https://doi.org/10.1371/journal.pcbi.1010378.
34. Chang DH, Angelin-Duclos C, Calame, K. BLIMP-1: trigger for differentiation of myeloid lineage. *Nat Immunol*. 2000;1:169-176.
35. Schmidt U, van den Akker E, Parren-van Amelsvoort M, et al. Btk is required for an efficient response to erythropoietin and for SCF-controlled protection against TRAIL in erythroid progenitors. *J Exp Med*. 2004;199:785-795. https://doi.org/10.1084/jem.20031109.
36. Fang C, Miwa T, Song W-C. Decay-accelerating factor regulates T-cell immunity in the context of inflammation by influencing costimulatory molecule expression on antigen-presenting cells. *Blood*. 2011;118:1008-1014. https://doi.org/10.1182/blood-2011-04-348474.
37. Cooney R, Baker J, Brain O, et al. NOD2 stimulation induces autophagy in dendritic cells influencing bacterial handling and antigen presentation. *Nat Med*. 2010;16:90-97. https://doi.org/10.1038/nm.2069.
38. Hazenbos WL, Heijnen IA, Meyer D, et al. Murine IgG1 complexes trigger immune effector functions predominantly via Fc gamma RIII (CD16). *J Immunol*. 1998;161:3026-3032.
39. Ivashkiv LB. IFNγ: signalling, epigenetics and roles in immunity, metabolism, disease and cancer immunotherapy. *Nat Rev Immunol*. 2018;18:545-558. https://doi.org/10.1038/s41577-018-0029-z.
40. Cao H, Heazlewood SY, Williams B, et al. The role of CD44 in fetal and adult hematopoietic stem cell regulation. *Haematologica*. 2016;101:26-37. https://doi.org/10.3324/haematol.2015.135921.

41. Frelin C, Herrington R, Janmohamed S, et al. GATA-3 regulates the self-renewal of long-term hematopoietic stem cells. *Nat Immunol*. 2013;14:1037-1044. https://doi.org/10.1038/ni.2692.
42. Rodrigues NP, Tipping AJ, Wang Z, Enver T. GATA-2 mediated regulation of normal hematopoietic stem/progenitor cell function, myelodysplasia and myeloid leukemia. *Int J Biochem Cell Biol*. 2012;44:457-460. https://doi.org/10.1016/j.biocel.2011.12.004.
43. Yoshida T, Georgopoulos K. GATA-3 controls self-renewal in stressed HSCs. *Nat Immunol*. 2013;14:1032-1033. https://doi.org/10.1038/ni.2715.
44. Schep AN, Wu B, Buenrostro JD, Greenleaf WJ. ChromVAR: Inferring transcription-factor-associated accessibility from single-cell epigenomic data. *Nat Methods*. 2017;14:975-978. https://doi.org/10.1038/nmeth.4401.
45. Li Q, Peterson KR, Fang X, Stamatoyannopoulos G. Locus control regions. *Blood*. 2002;100:3077-3086. https://doi.org/10.1182/blood-2002-04-1104.
46. Palstra R, de Laat W, Grosveld, F.B.T.-A. Chapter 4 β-globin regulation and long-range interactions. In *Long-Range Control of Gene Expression*. Academic Press, 2008, pp. 107–142.
47. Linderson Y, Eberhard D, Malin S, et al. Corecruitment of the Grg4 repressor by PU.1 is critical for Pax5-mediated repression of B-cell-specific genes. *EMBO reports*. 2004;5:291-296. https://doi.org/10.1038/sj.embor.7400089.
48. Nutt SL, Eberhard D, Horcher M, Rolink AG, Busslinger M. Pax5 determines the identity of B cells from the beginning to the end of B-lymphopoiesis. *Int Rev Immunol*. 2001;20:65-82. https://doi.org/10.3109/08830180109056723.
49. Laszkiewicz A, Sniezewski L, Kasztura M, et al. Bidirectional activity of the NWC promoter is responsible for RAG-2 transcription in non-lymphoid cells. *PLoS One*. 2012;7:e44807. https://doi.org/10.1371/journal.pone.0044807.
50. Hsu L-Y, Lauring J, Liang H-E, et al. A conserved transcriptional enhancer regulates RAG gene expression in developing B cells. *Immunity*. 2003;19:105-117. https://doi.org/10.1016/s1074-7613(03)00181-x.
51. Wei X-C, Kishi H, Jin Z-X, et al. Characterization of chromatin structure and enhancer elements for murine recombination activating gene-2. *J Immunol*. 2002;169:873. https://doi.org/10.4049/jimmunol.169.2.873.
52. Kuo TC, Schlissel MS. Mechanisms controlling expression of the RAG locus during lymphocyte development. *Curr Opin Immunol*. 2009;21:173-178. https://doi.org/10.1016/j.coi.2009.03.008.
53. Yannoutsos N, Barreto V, Misulovin Z, et al. A cis element in the recombination activating gene locus regulates gene expression by counteracting a distant silencer. *Nat Immunol*. 2004;5:443-450. https://doi.org/10.1038/ni1053.
54. Yu W, Misulovin Z, Suh H, et al. Coordinate regulation of RAG1 and RAG2 by cell type-specific DNA elements 5′ of RAG2. *Science (80-.)*. 1999;285:1080-1084.
55. Dodson LF, Boomer JS, Deppong CM, et al. Targeted knock-in mice expressing mutations of CD28 reveal an essential pathway for costimulation. *Mol Cell Biol*. 2009;29:3710-3721. https://doi.org/10.1128/MCB.01869-08.
56. Shulga-Morskaya S, Dobles M, Walsh ME, et al. B cell-activating factor belonging to the TNF family acts through separate receptors to support B cell survival and T cell-independent antibody formation. *J Immunol*. 2004;173:2331-2341. https://doi.org/10.4049/jimmunol.173.4.2331.
57. Bultman, SJ, Gebuhr, TC, Magnuson, T. A Brg1 mutation that uncouples ATPase activity from chromatin remodeling reveals an essential role for SWI/SNF-related complexes in β-globin expression and erythroid development. *Genes Dev*. 2005;19:2849-2861.
58. Chi TH, Wan M, Lee PP, et al. Sequential roles of Brg, the ATPase subunit of BAF chromatin remodeling complexes, in thymocyte development. *Immunity*. 2003;19:169-182. https://doi.org/10.1016/s1074-7613(03)00199-7.
59. Jepsen K, Hermanson O, Onami TM, et al. Combinatorial roles of the nuclear receptor corepressor in transcription and development. *Cell*. 2000;102:753-763. https://doi.org/10.1016/s0092-8674(00)00064-7.
60. Stenvers KL, Tursky ML, Harder KW, et al. Heart and liver defects and reduced transforming growth factor β2 sensitivity in transforming growth factor β type III receptor-deficient embryos. *Mol Cell Biol*. 2003;23:4371-4385. https://doi.org/10.1128/mcb.23.12.4371-4385.2003.
61. Tsang AP, Fujiwara Y, Hom DB, Orkin SH. Failure of megakaryopoiesis and arrested erythropoiesis in mice lacking the GATA-1 transcriptional cofactor FOG. *Genes Dev*. 1998;12:1176-1188. https://doi.org/10.1101/gad.12.8.1176.
62. Oguri Y, Shinoda K, Kim H, et al. CD81 controls Beige fat progenitor cell growth and energy balance via FAK signaling. *Cell*. 2020;182:563-577.e20 https://doi.org/10.1016/j.cell.2020.06.021.
63. Replogle JM, Norman TM, Xu A, et al. Combinatorial single-cell CRISPR screens by direct guide RNA capture and targeted sequencing. *Nat Biotechnol*. 2020;38(8):954–961. https://doi.org/10.1038/s41587-020-0470-y
64. Lara-Astiaso D, Weiner A, Lorenzo-Vivas E, et al. Chromatin state dynamics during blood formation. *Science*. 2014;55:1-10.
65. Wang A, Yue F, Li Y, et al. Epigenetic priming of enhancers predicts developmental competence of hESC-derived endodermal lineage intermediates. *Cell Stem Cell*. 2015;16:386-399. https://doi.org/10.1016/j.stem.2015.02.013.
66. Challen GA, Pietras EM, Wallscheid NC, Signer RAJ. Simplified murine multipotent progenitor isolation scheme: establishing a consensus approach for multipotent progenitor identification. *Exp Hematol*. 2021;104:55-63. https://doi.org/10.1016/j.exphem.2021.09.007.
67. Laurenti E, Göttgens B. From haematopoietic stem cells to complex differentiation landscapes. *Nature*. 2018;553:418-426. https://doi.org/10.1038/nature25022.
68. Heuston EF, Keller CA, Lichtenberg J, et al. Establishment of regulatory elements during erythro-megakaryopoiesis identifies hematopoietic lineage-commitment points. *Epigen Chromatin*. 2018;11:1-18.
69. Schlenner SM, Madan V, Busch K, et al. Fate mapping reveals separate origins of T cells and myeloid lineages in the thymus. *Immunity*. 2010;32:426-436. https://doi.org/10.1016/j.immuni.2010.03.005.
70. Karsunky H, Merad M, Cozzio A, Weissman IL, Manz MG. Flt3 ligand regulates dendritic cell development from Flt3+ lymphoid and myeloid-committed progenitors to Flt3+ dendritic cells in vivo. *J Exp Med*. 2003;198:305-313. https://doi.org/10.1084/jem.20030323.
71. Carrelha J, Meng Y, Kettyle LM, et al. Hierarchically related lineage-restricted fates of multipotent haematopoietic stem cells. *Nature*. 2018;554:106-111. https://doi.org/10.1038/nature25455.
72. Ema H, Morita Y, Suda T. Heterogeneity and hierarchy of hematopoietic stem cells. *Exp Hematol*. 2014;42:74-82.e2. https://doi.org/10.1016/j.exphem.2013.11.004.
73. Yamamoto R, Morita Y, Ooehara J, et al. Clonal analysis unveils self-renewing lineage-restricted progenitors generated directly from hematopoietic stem cells. *Cell*. 2013;154:1112-1126. https://doi.org/10.1016/j.cell.2013.08.007.
74. Yamamoto R, Wilkinson AC, Ooehara J, et al. Large-scale clonal analysis resolves aging of the mouse hematopoietic stem cell compartment. *Cell Stem Cell*. 2018;22, 600-607.e4. https://doi.org/10.1016/j.stem.2018.03.013.
75. Gasiorek JJ, Nouhi Z, Blank V. Abnormal differentiation of erythroid precursors in p45 NF-E2–/– mice. *Exp Hematol*. 2012;40:393-400. https://doi.org/10.1016/j.exphem.2012.01.007.
76. Shivdasani RA, Orkin SH. Erythropoiesis and globin gene expression in mice lacking the transcription factor NF-E2. *Proc. Natl. Acad. Sci*. 1995;92:8690-8694.
77. Willcockson MA, Taylor SJ, Ghosh S, et al. Runx1 promotes murine erythroid progenitor proliferation and inhibits differentiation by preventing Pu.1 downregulation. *Proc. Natl. Acad. Sci*. 2019;116:17841-17847.
78. Rodriguez y Baena A, Manso BA, Forsberg EC. CFU-S assay: a historical single-cell assay that offers modern insight into

clonal hematopoiesis. *Exp Hematol*. 2021;104:1-8. https://doi.org/10.1016/j.exphem.2021.10.003.
79. Worthington AK, Forsberg EC. A CRISPR view of hematopoietic stem cells: moving innovative bioengineering into the clinic. *Am J Hematol*. 2022;97:1226-1235. https://doi.org/10.1002/ajh.26588.
80. Rodriguez y Baena A, Rajendiran S, Manso BA, et al. New transgenic mouse models enabling pan-hematopoietic or selective hematopoietic stem cell depletion in vivo. *Sci Rep*. 2022;12:3156. https://doi.org/10.1038/s41598-022-07041-6.
81. Worthington AK, Cool T, Poscablo DM, et al. IL7Rα, but not Flk2, is required for hematopoietic stem cell reconstitution of tissue-resident lymphoid cells. *Development*. 2022;149:dev200139. https://doi.org/10.1242/dev.200139.
82. Tasic B, Hippenmeyer S, Wang C, et al. Site-specific integrase-mediated transgenesis in mice via pronuclear injection. *Proc. Natl. Acad. Sci*. 2011;108:7902-7907.
83. Horlbeck MA, Gilbert LA, Villalta JE, et al. Compact and highly active next-generation libraries for CRISPR-mediated gene repression and activation. *Elife*. 2016;5:1-20.
84. Concordet J-P, Haeussler M. CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Res*. 2018;46:W242-W245. https://doi.org/10.1093/nar/gky354.
85. Haeussler M, Schönig K, Eckert H, et al. Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biol*. 2016;17:148. https://doi.org/10.1186/s13059-016-1012-2.
86. Forsberg EC, Serwold T, Kogan S, Weissman IL, Passegué E. New evidence supporting megakaryocyte-erythrocyte potential of Flk2/Flt3+ multipotent hematopoietic progenitors. *Cell*. 2006;126:415-426. https://doi.org/10.1016/j.cell.2006.06.037.