



Published in final edited form as:

Science. 2014 August 22; 345(6199): 943–949. doi:10.1126/science.1256271.

## Chromatin state dynamics during blood formation

David Lara-Astiaso<sup>1,\*</sup>, Assaf Weiner<sup>2,3,\*</sup>, Erika Lorenzo-Vivas<sup>1</sup>, Irina Zaretsky<sup>1</sup>, Diego Adhemar Jaitin<sup>1</sup>, Eyal David<sup>1</sup>, Hadas Keren-Shaul<sup>1</sup>, Alexander Mildner<sup>1</sup>, Deborah Winter<sup>1</sup>, Steffen Jung<sup>1</sup>, Nir Friedman<sup>2,3,†</sup>, and Ido Amit<sup>1,†</sup>

<sup>1</sup>Department of Immunology, Weizmann Institute of Science, Rehovot, Israel

<sup>2</sup>Institute of Life Sciences, The Hebrew University, Jerusalem, Israel

<sup>3</sup>School of Computer Science and Engineering, The Hebrew University, Jerusalem, Israel

### Abstract

Chromatin modifications are crucial for development, yet little is known about their dynamics during differentiation. Hematopoiesis provides a well-defined model to study chromatin state dynamics, however technical limitations impede profiling of homogeneous differentiation intermediates. We developed a high sensitivity indexing-first chromatin immunoprecipitation approach (iChIP) to profile the dynamics of four chromatin modifications across 16 stages of hematopoietic differentiation. We identify 48,415 enhancer regions and characterize their dynamics. We find that lineage commitment involves *de novo* establishment of 17,035 lineage-specific enhancers. These enhancer repertoire expansions foreshadow transcriptional programs in differentiated cells. Combining our enhancer catalog with gene expression profiles, we elucidate the transcription factor network controlling chromatin dynamics and lineage specification in hematopoiesis. Together, our results provide a comprehensive model of chromatin dynamics during development.

---

Chromatin plays a major regulatory role in cell type specific functions and response (1, 2). The current dogma of cellular differentiation suggests that there is a progressive closing of the regulatory potential of the genome. According to this model, differentiation is a gradual transition from an open chromatin state in multipotent stem cells to a compacted chromatin state in differentiated cells. However, genome-wide histone modification profiling of embryonic stem cells and terminally differentiated cells, are not fully compatible with this model (3–7). Resolving this discrepancy requires observation of chromatin dynamics during intermediate stages of differentiation.

An important model system for differentiation is hematopoiesis where a single hematopoietic stem cell gives rise to a large number of cell types (comprising the blood) through a series of characterized intermediate progenitor cells (8, 9). Chromatin regulation has a central role in hematopoiesis: mutations or loss of chromatin factors lead to hematopoiesis defects and diseases (10). Moreover, genome-wide chromatin profiling studies have revealed large-scale differences in the histone modifications and transcription

---

<sup>†</sup>To whom correspondence should be addressed: nir@cs.huji.ac.il (NF), ido.amit@weizmann.ac.il (IA).

\*These authors contributed equally to this work

factor binding maps between mature immune cell types (3–6, 11). Thus, dramatic chromatin reorganizations during hematopoiesis are critical for early hematopoietic decisions (12).

Comprehensive study of the chromatin events during hematopoiesis has been hampered by the low sensitivity and reproducibility for small cell numbers with current chromatin immunoprecipitation (ChIP) protocols. These protocols require several enzymatic steps with limited performance when the input DNA is below the nanogram range. While an average diploid mammalian cell has roughly 4–8 pg of DNA, losses following ChIP reduces the available DNA for analysis by 2–3 orders of magnitude, setting the lower limit for genome-wide chromatin analysis at 50,000 cells (13). Amplification of ChIP material partially alleviate this problem at the cost of introducing amplification biases (14, 15).

To profile chromatin dynamics of scarce *in vivo* cell populations, we developed an indexing-first chromatin IP approach (iChIP). In this protocol, barcoding is performed directly on the total cellular chromatin (Fig. 1A), thereby avoiding the low input enzymatic reactions occurring in conventional ChIP. Importantly, this enables multiple chromatin-barcoded samples to be pooled for ChIP in the same well, further reducing initial input requirements and increasing cross-sample reproducibility. To minimize centrifugation steps, cells are fixed before sorting, and sonication. Then, the sheared chromatin is immobilized on anti-H3 antibody-coated magnetic beads and indexed. The indexed chromatin is released from the H3-beads and pooled with chromatin from other samples. Finally, ChIP is performed with the desired antibody, and a single chromatin-barcoded pool can be split to multiple ChIPs for profiling various chromatin modifications.

The iChIP protocol is highly reproducible for low cell numbers (a few hundred cells) while increasing the sensitivity and throughput (Fig. 1). To benchmark iChIP, we barcoded decreasing amounts (10,000 to 500 cells) of chromatin isolated from bone-marrow-derived dendritic cells (BMDCs), in triplicate, and performed ChIP on the barcoded chromatin with an antibody for mono and tri-methylated histone H3 lysine 4 (H3K4me1 and H3K4me3) (Fig. 1B–D, fig. S1). We confirmed reproducibility of H3K4me3 peaks between replicates of 10,000 cells ( $r = 0.95$ ) down to 500 cells ( $r = 0.85$ ). iChIP profiles correlated with conventional ChIP-seq on 10 million BMDCs ( $r = 0.92$ ) (16). iChIP is applicable to profile other histone modifications and transcription factors for low cell numbers (Fig. 1C–D, fig. S1). The reproducibility and sensitivity for small cell numbers of iChIP open the way for *in vivo* characterization of chromatin dynamics during hematopoiesis.

We profiled four histone modifications (H3K4me1, H3K4me2, H3K4me3 and H3K27ac) (17) in three *in vivo* replicates at 16 developmental stages of hematopoietic commitment using 5,000 cells per mark. The cell populations chosen comprise all multipotent progenitor stages and the common lineage oligopotent progenitors from each of the major blood lineages (myeloid, lymphoid and erythroid) as well as several terminally differentiated cell types from each of the lineages (<https://www.immgen.org>) (16) (Fig. 2A). Replicates of histone mark ChIP and RNAseq displayed high reproducibility (average  $r = 0.95$ ; (16); fig. S2, table S1–2). We used histone marks to define promoter (high H3K4me3) and enhancer (high H3K4me1/2 and low H3K4me3) regions resulting in analysis of 48,415 enhancers and 17,923 promoters (16). Enhancer activity was defined from H3K27ac levels and RNA

expression values of the nearest gene (16–18). Genome-wide analysis of RNA expression and promoter activity (H3K4me3 signal) identified four major patterns for progenitors, lymphoid, myeloid and erythroid cells (Fig. 2B, fig. S2–3). In contrast, H3K4me1 signal revealed that lineage-progenitors were more similar to the differentiated cells within their lineage than to progenitors from other cell lineages (Fig. 2B, fig. S2–3). This distinction suggests that enhancer establishment is initiated in early lineage commitment and can reveal the differentiation potential of progeny prior to the execution of the RNA expression program.

Changes in chromatin marks during hematopoiesis (particularly H3K4me1 and H3K4me2) suggest lineage specific activity of regulatory elements. Of the 48,415 hematopoietic enhancers, 90% (43,428) changed state during hematopoiesis (16). We found that 60% (26,393) of these dynamic enhancers show the stereotypical behavior in which they are initially marked in hematopoietic stem cells (HSCs), but are maintained only in the relevant lineage (Fig. 2C (Gata2), fig. S4A (Meis1 and CD34)). Surprisingly, we discovered that a significant proportion of dynamic enhancers (40%, 17,035) are established *de novo* during the differentiation process (16). For example, enhancers present in myeloid gene loci IL-1b, CD14, S100a8 and F7 (coagulation factor VII) are either established *de novo* or become more prominent in the myeloid lineage (Fig. 2C, fig. S4A). Similarly, Ebf1 and Cr2 enhancers are established in B cells; Bcl11b and CD3g enhancers in T cells; Granzyme A and Ncr1 enhancers in NK cells; and Gata1, Gypa (Glycophorin A) and Cpox (Coproporphyrinogen Oxidase) enhancers in the erythroid lineage (Fig. 2, fig. S3B, S4A).

The establishment of *de novo* lineage specific enhancers occurs mainly at the root of the commitment point in the first progenitor of the lineage, whereas closing of enhancers occurs more gradually (Fig. 2). For example, S100a8 and F7 – F10 loci display H3K4me1 signal specifically in the myeloid lineage with *de novo* establishment of the enhancers at the root of the myeloid commitment point in common myeloid progenitor (CMP) (Fig. 2C, fig. S3B, S4A). Similarly, the Cpox and Gypa loci display H3K4me1 signal specifically in the erythrocyte lineage with *de novo* establishment of the enhancers at the root of the erythrocyte commitment point in the megakaryocyte erythroid progenitor (MEP) (Fig. 2C). Importantly, we observe a step-wise acquisition of the different chromatin modifications in the lineage-specific enhancers during hematopoiesis (Fig. 2D, fig. S4B, S5). For example, in both S100a8 (myeloid lineage) and Gypa (erythroid lineage) loci, H3K4me1/2 appear first in the root lineage progenitor (CMP and MEP, respectively), while H3K27ac (locus activation) is acquired together with active RNA transcription only once the cells (granulocytes and erythrocyte B) are terminally differentiated (Fig. 2D, fig. S4B). Globally, 32% of the activated (H3K27ac) enhancers in terminally differentiated cells are initially poised (H3Kme1 only) in the lineage progenitors (16).

Clustering of all 48,415 H3K4me1 peaks by their dynamic profiles during hematopoiesis revealed nine major clusters, consistent with the underlying biology of the system (Fig. 3A, fig. S6A, B) (16). H3K4me2 signal shows similar patterns in all nine clusters (fig. S6A). Cluster I comprises enhancers shared throughout hematopoiesis. Clusters II–IV group lineage-specific enhancers already marked in HSCs and shared with hematopoietic progenitors. Finally, clusters VI (Fig. 3B middle), VII VIII, IX (Fig. 3B bottom) group *de*

*novo* enhancers that are specific to a particular lineage and were not marked in HSCs; with 6382 myeloid, 5834 lymphoid and 4819 erythroid enhancers. We also found a group of 6612 enhancers (Cluster V) shared exclusively among progenitors (Fig. 3B top). Interestingly, the erythrocyte and progenitor enhancer clusters exhibit relatively high sequence conservation (16) in the mammalian clade, with the myeloid and lymphoid enhancers displaying lower conservation. These data suggests differential rates of evolutionary dynamics in these *cis* regulatory regions (16) (Fig. 3A).

Newly formed enhancers could either be established at a specific branching point or established gradually during the development process. To further examine these dynamics, we generated a catalog of enhancers that are dynamic (gained or lost) during the process of differentiation from the long-term hematopoietic stem cells (LT-HSCs) to mature, terminally differentiated cells (Fig. 3C). We determined, for each enhancer, the stage of gain/loss along the differentiation path. In erythroid differentiation 65% of the *de novo* enhancer repertoire are gained in the MEP stage (Fig. 3D). Similarly, 40–50% of the *de novo* myeloid enhancers are gained in the first step of myeloid commitment, during the multipotent progenitor (MPP) to CMP transition, while the CMP to GMP transition involves fewer gains (15–30%). Together, the CMP and GMP stages are responsible for 63–80% of gained enhancers in terminally differentiated myeloid cells (Fig. 3D), consistent with the global enhancer similarity of CMP and GMP with the myeloid lineage (Fig. 2B). In contrast, enhancer loss is a more gradual process that initiates in the CMP or MEP stage (for myeloid and erythrocyte development, respectively) with a large proportion (40–50%) of enhancers lost in the last and definitive differentiation step to mature cells (fig. S6C).

Genome wide studies show that, while H3K4me1 marks both poised and active enhancers, H3K27ac marks only active enhancers (19, 20). Since poised enhancers represent potential gene expression programs, the ratio of these enhancers in a given cell type approximates the current regulatory potential of the cell (20). Analysis of the regulatory potential in hematopoiesis shows that progenitor cells are more plastic than differentiated cell types, with erythrocytes utilizing most (78%) of their enhancers, whereas CMPs use only 33% of their enhancer potential (Fig. 3E–F, fig. S6D) (16). Within the progenitor group, CMP and GMP are more plastic than the multipotent stem cells (33–37% versus 62–65% of enhancer use); likely due to the *de novo* expansion of myeloid enhancers (Fig. 2B, 3A and F). Notably there is a wide spectrum of plasticity across the terminally differentiated cells, with the myeloid lineage (macrophages, monocytes and granulocytes) showing higher degrees of plasticity than erythrocytes, B, and NK cells; this is consistent with the higher functional versatility of myeloid cells in comparison to other hematopoietic cell types (21).

*De novo* H3K4me1 establishment in hematopoiesis is concomitant to increase in chromatin accessibility (Fig. 4A, B). We measured chromatin accessibility, ‘open chromatin’ (22), during the developmental process in ten hematopoietic cell types using the assay for transposase-accessible chromatin followed by sequencing (ATAC-seq) (23). In erythroid (Gypa) and myeloid (F7 and F10) gene loci chromatin accessibility follows the temporal pattern observed for H3K4me1 (Fig. 4A). As expected, ATAC signal is enriched both in active promoters and H3K4me1-positive enhancers, as well as in insulators and other regulatory regions (23). To evaluate the proportion of *de novo* enhancers that also display *de*

*nov* establishment of open chromatin, we plotted the ATAC-seq signal in the regions from the H3K4me1 enhancer catalog (fig. S7A). We compared the ATAC-seq signal to H3K4me1 and H3K27ac intensities. We observed a similar pattern between ATAC and H3K4me1 signal ( $r=0.75$ ) and a weakened agreement with H3K27ac ( $r=0.62$ ) (Figs. 4B, fig. S7B) suggesting that the process of gain or loss of H3K4me1 mark on enhancers occurs concomitantly with formation of open chromatin sites.

Establishment of lineage specific enhancers is regulated by the activity of lineage specific transcription factors (24) (Fig. 4C–E). Using the ATAC peaks (16) and our enhancer catalog we searched for enriched transcription factor binding motifs in each cell type (Fig. 4C, fig. S8) (16) in order to identify lineage-determining factors. We found that, in line with their identified functions, PU.1, Gata1, and Foxo1 can be classified as potential regulators of myeloid, erythroid and lymphoid enhancers, respectively (3, 11, 25, 26).

To systematically identify potential regulators, we generated a logistic regression model to predict enhancer activity at each stage from the DNA binding motif scores and TF expression (Fig. 4D, E, fig. S8–11) (16). Our logistic model accurately predicts enhancer cell-type specific activity with 75% accuracy (with 70% sensitivity and 80% specificity; fig. S10B). Importantly, the model allowed us to elucidate the transcription factors controlling chromatin dynamics and lineage specification in hematopoiesis (25, 26). Our model identified the known myeloid lineage determining factors, PU.1, Cebpb and Cebpa as regulators of myeloid enhancers; additionally, our model suggests a hierarchy between the Cebp factors with Cebpa active in the progenitors (CMP/GMP) and Cebpb replacing Cebpa in the differentiated cell types (Fig. 4E, fig. S11C). Similarly, we identify Meis1, Hoxa9 and Erg as potential regulators of stem cell enhancers, Pax5 in B cells, Klf1 in erythroid cells and Ets1 in lymphoid cells (Fig. 4C–E, fig. S11). We identify many transcription factors that have been implicated in lineage development but have not previously been associated with chromatin regulation of lineage determination like Irf1 and Irf2 in B cells and Cebpe in granulocytes (27, 28). We also highlight new potential regulators of hematopoietic lineages: ATF3 in monocytes and Tcf7l2, Mef2a and Runx2 in NK cells (Fig. 4E). All together, our findings show that chromatin is highly dynamic during hematopoiesis, orchestrated by a defined set of transcription factors.

In conclusion, iChIP enables the execution of reproducible and sensitive ChIP on only a few hundred cells in a manner broadly applicable across organisms and tissues. We show that poised enhancers are established in lineage progenitors before their activation and precede RNA expression in subsequent lineage differentiation. These enhancers are established concomitantly with the formation of open chromatin sites. We then show that most of enhancer dynamics can be accounted for by the activity of known lineage-specific factors as well as new candidate regulators. These results suggest a new model for chromatin dynamics during differentiation (Fig. 4F) and show that development involves massive dynamic reorganization of the chromatin landscape. While some enhancers are preset in hematopoietic stem cells, as suggested by the conventional development model, a comparable number of enhancers appear to be established *de novo* during hematopoiesis. We believe that the establishment of newly poised enhancers in the early lineage commitment steps initiates regulatory programs that are subsequently applied in

differentiated cells, while closing of enhancers occurs during later differentiation stages. This suggests that cellular enhancer potential reaches its maximum not at the HSC stage, but during the oligopotent progenitor stages. Taken together, these observations reshape our understanding of the role of chromatin and pioneer factors during differentiation.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

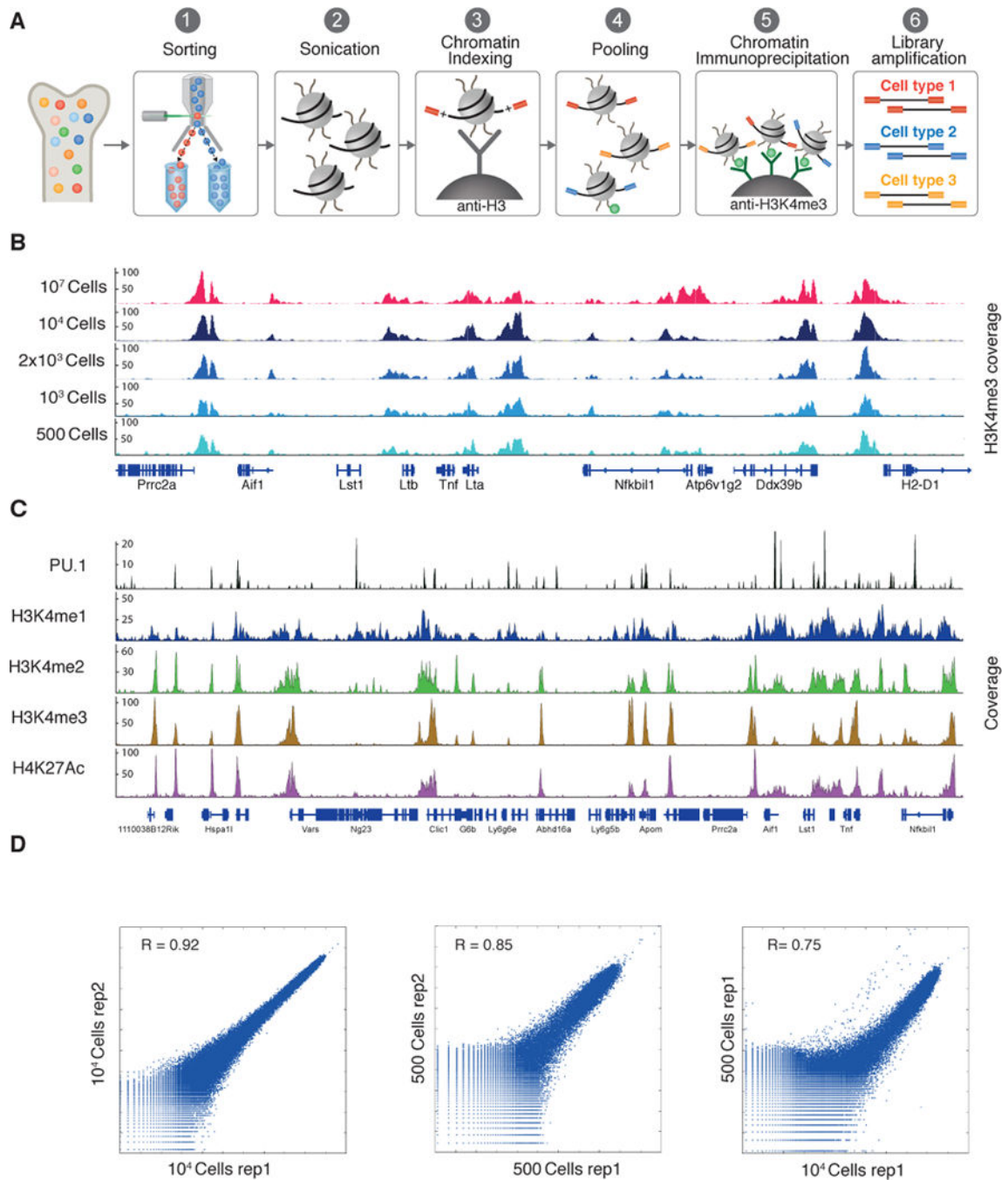
## Acknowledgments

We thank members of the Amit and Friedman labs for critical discussions. We thank G. Brodsky for artwork. Supported by grants from the European Research Council (#309788 and #340712), the Israeli Science Foundation (1782/11), The Human Frontiers Science Program, Career Development Award (CDA00028/2011-c) (IA), EMBO Young Investigator Award (IA), NHGRI Center for Excellence in Genome Science (1P50HG006193) and the Israeli Centers Of Research Excellence (I-CORE). RNA-seq, ChIP-seq and ATAC-seq data are deposited in GEO; accession number GSE59636. The data can be viewed from the following website: <http://compbio.cs.huji.ac.il/blood-chromatin/index.html>. A patent application for iChIP has been filed by Yeda, Weizmann Institute of Science, Israel.

## References

1. Ho L, Crabtree GR. Chromatin remodelling during development. *Nature*. 2010; 463:474–484. [PubMed: 20110991]
2. Zhou VW, Goren A, Bernstein BE. Charting histone modifications and the functional organization of mammalian genomes. *Nat Rev Genet*. 2011; 12:7–18. [PubMed: 21116306]
3. Ghisletti S, et al. Identification and characterization of enhancers controlling the inflammatory gene expression program in macrophages. *Immunity*. 2010; 32:317–328. [PubMed: 20206554]
4. Mercer EM, et al. Multilineage priming of enhancer repertoires precedes commitment to the B and myeloid cell lineages in hematopoietic progenitors. *Immunity*. 2011; 35:413–425. [PubMed: 21903424]
5. Wei G. Global mapping of H3K4me3 and H3K27me3 reveals specificity and plasticity in lineage fate determination of differentiating CD4<sup>+</sup> T cells. *Immunity*. 2009; 30:155–167. [PubMed: 19144320]
6. Revilla-I-Domingo R, et al. The B-cell identity factor Pax5 regulates distinct transcriptional programmes in early and late B lymphopoiesis. *EMBO J*. 2012; 31:3130–3146. [PubMed: 22669466]
7. Mikkelsen TS, et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature*. 2007; 448:553–560. [PubMed: 17603471]
8. Seita J, Weissman IL. Hematopoietic stem cell: self-renewal versus differentiation. *Wiley Interdiscip Rev Syst Biol Med*. 2010; 2:640–653. [PubMed: 20890962]
9. Doulatov S, Notta F, Laurenti E, Dick JE. Hematopoiesis: a human perspective. *Cell Stem Cell*. 2012; 10:120–136. [PubMed: 22305562]
10. Shih AH, Abdel-Wahab O, Patel JP, Levine RL. The role of mutations in epigenetic regulators in myeloid malignancies. *Nat Rev Cancer*. 2012; 12:599–612. [PubMed: 22898539]
11. Heinz S, 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. [PubMed: 20513432]
12. Ji H, et al. Comprehensive methylome map of lineage commitment from haematopoietic progenitors. *Nature*. 2010; 467:338–342. [PubMed: 20720541]
13. Furey TS. ChIP-seq and beyond: new and improved methodologies to detect and characterize protein-DNA interactions. *Nat Rev Genet*. 2012; 13:840–852. [PubMed: 23090257]
14. Shankaranarayanan P, et al. Single-tube linear DNA amplification (LinDA) for robust ChIP-seq. *Nat Methods*. 2011; 8:565–567. [PubMed: 21642965]

15. Garber M, et al. A high-throughput chromatin immunoprecipitation approach reveals principles of dynamic gene regulation in mammals. *Mol Cell*. 2012; 47:810–822. [PubMed: 22940246]
16. Materials and methods are available as supplementary materials on Science Online.
17. Rivera CM, Ren B. Mapping human epigenomes. *Cell*. 2013; 155:39–55. [PubMed: 24074860]
18. Kaikkonen MU, et al. Remodeling of the enhancer landscape during macrophage activation is coupled to enhancer transcription. *Mol Cell*. 2013; 51:310–325. [PubMed: 23932714]
19. Rada-Iglesias A, et al. A unique chromatin signature uncovers early developmental enhancers in humans. *Nature*. 2011; 470:279–283. [PubMed: 21160473]
20. Creighton MP, et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci USA*. 2010; 107:21931–21936. [PubMed: 21106759]
21. Wynn TA, Chawla A, Pollard JW. Macrophage biology in development, homeostasis and disease. *Nature*. 2013; 496:445–455. [PubMed: 23619691]
22. Thurman RE, et al. The accessible chromatin landscape of the human genome. *Nature*. 2012; 489:75–82. [PubMed: 22955617]
23. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods*. 2013; 10:1213–1218. [PubMed: 24097267]
24. Cirillo LA, et al. Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. *Mol Cell*. 2002; 9:279–289. [PubMed: 11864602]
25. Cantor AB, Orkin SH. Transcriptional regulation of erythropoiesis: an affair involving multiple partners. *Oncogene*. 2002; 21:3368–3376. [PubMed: 12032775]
26. Ouyang W, Li MO. Foxo: in command of T lymphocyte homeostasis and tolerance. *Trends Immunol*. 2011; 32:26–33. [PubMed: 21106439]
27. Matsuyama T, et al. Targeted disruption of IRF-1 or IRF-2 results in abnormal type I IFN gene induction and aberrant lymphocyte development. *Cell*. 1993; 75:83–97. [PubMed: 8402903]
28. Yamanaka R, et al. Impaired granulopoiesis, myelodysplasia, and early lethality in CCAAT/enhancer binding protein epsilon-deficient mice. *Proc Natl Acad Sci USA*. 1997; 94:13187–13192. [PubMed: 9371821]
29. Jaitin DA, et al. Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. *Science*. 2014; 343:776–779. [PubMed: 24531970]
30. McCullagh, P.; Nelder, J. *Generalized Linear Models*. Second. Chapman & Hall; 1989. ed 2



**Figure 1. Indexing-first Chromatin Immunoprecipitation for profiling histone modifications and TF binding**

(A) Schematic diagram of the iChIP approach to chromatin immunoprecipitation involving an initial chromatin barcoding step prior to ChIP with the desired antibody. (B) Normalized H3K4me3 profiles of peaks found in a 100 Kb region in the TNF locus (genes are indicated below) obtained with iChIP of decreasing amounts of bone-marrow derived dendritic cells (BMDC). **Top**, in red, H3K4me3 profile obtained using conventional ChIP with 20 million cells (15). Below, in shades of blue, H3K4me3 profiles obtained with iChIP. (C)



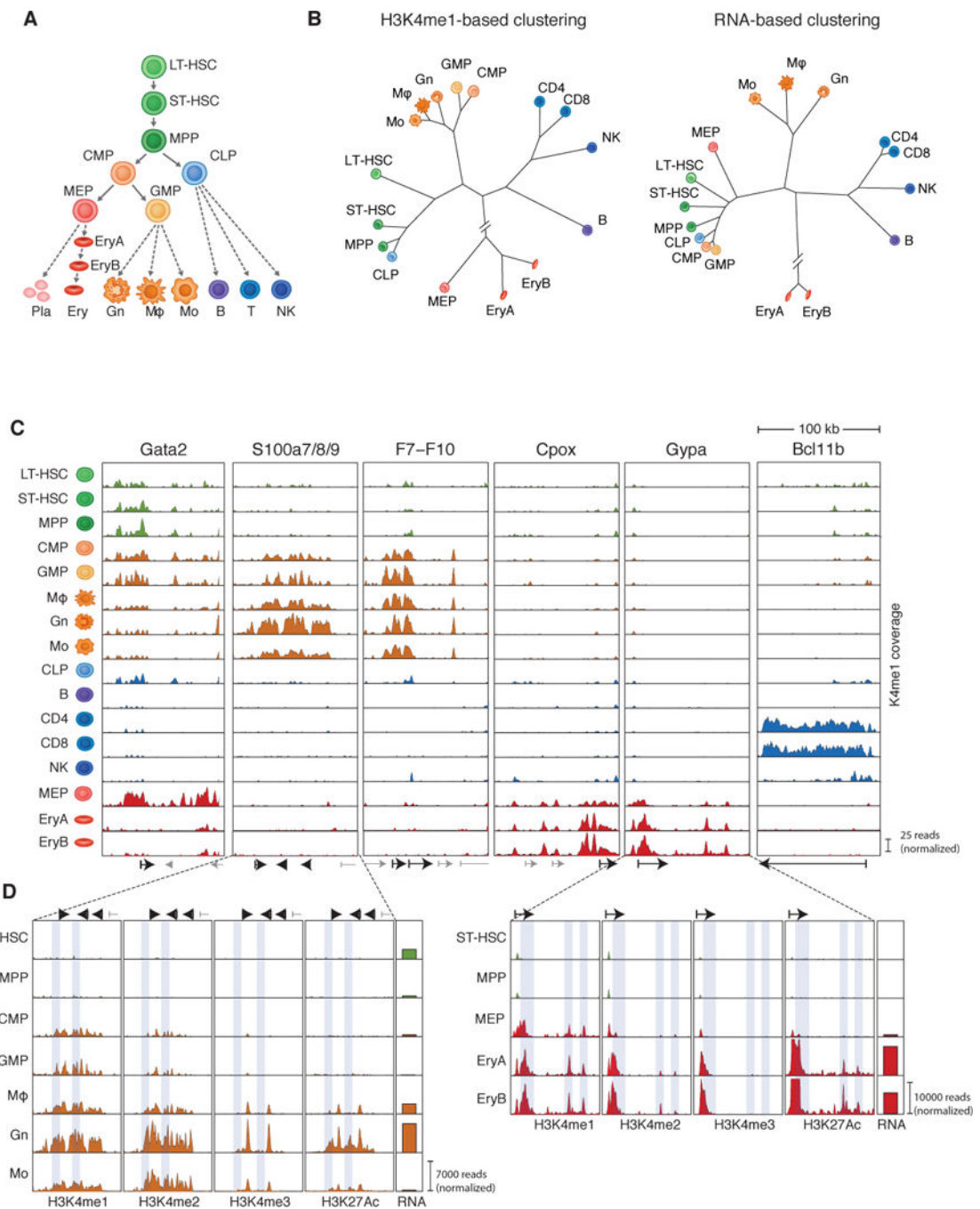
Normalized profiles of PU.1, H3K4me1, H3K4me2, H3K4me3 and H3K27ac found in a 250 Kb region in the TNF locus obtained with iChIP with  $10^4$  BMDC. **(D)** Scatter plots showing correlation between representative H3K4me3 iChIP replicates. From left to right: correlation between  $10^4$  cell replicates, 500 cell replicates, and a 500 compared to  $10^4$  cells.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

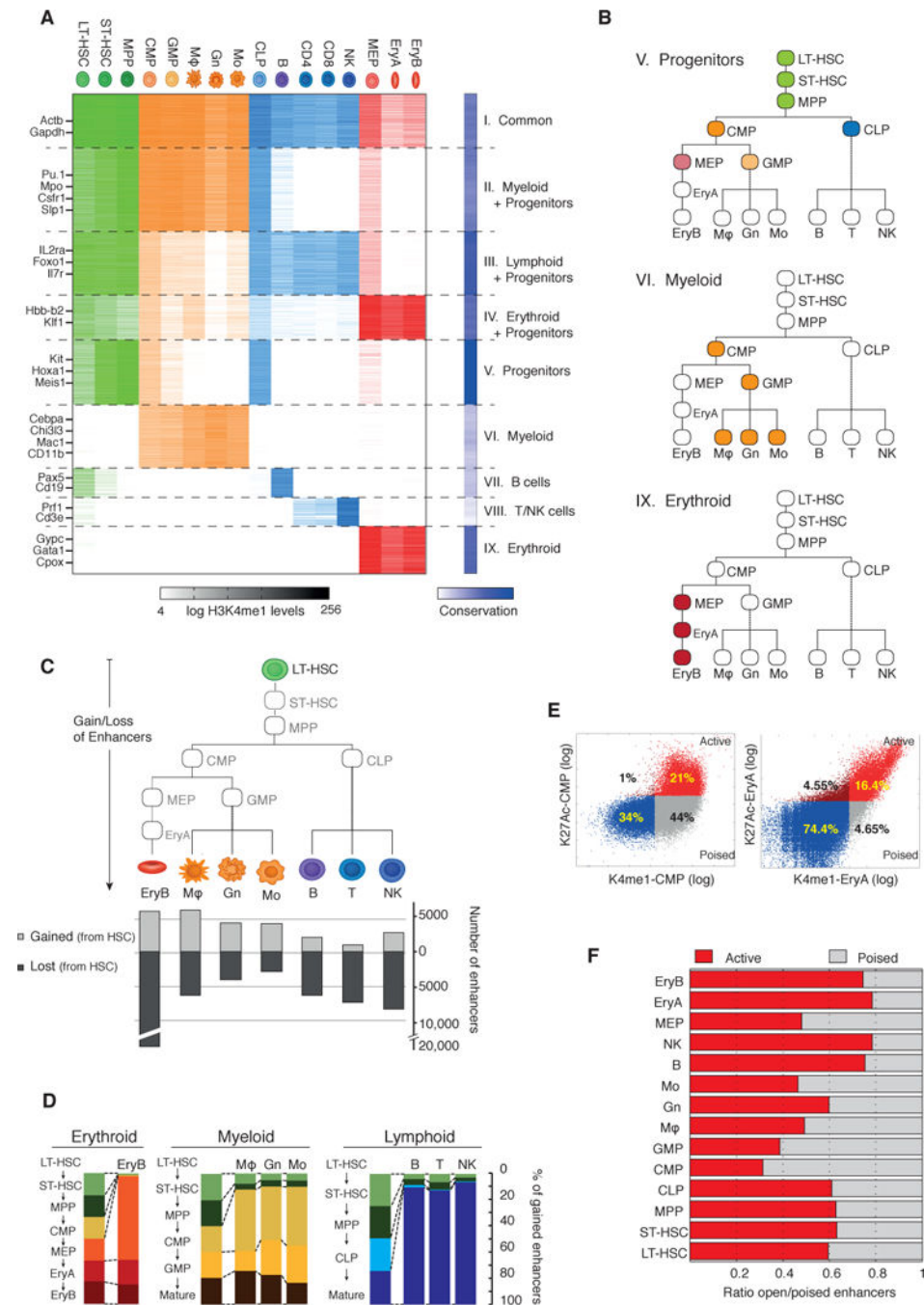


**Figure 2. Chromatin dynamics in hematopoiesis**

(A) Schematic diagram of the hematopoietic differentiation stages included in this study (platelets are displayed, but were not included). Dashed-arrow indicates intermediate progenitors not measured in this study. Color code: green for multipotent progenitors, orange for myeloid lineage (including the oligopotent progenitors CMP and GMP), blue for lymphoid lineage, red for erythroid lineage. (B) Clustering dendrogram of cell types based on H3K4me1 profiles (**left**) and RNA-seq levels (**right**) showing the differential association of lineage progenitors (CMP, GMP, MEP). (C) Representative examples of H3K4me1

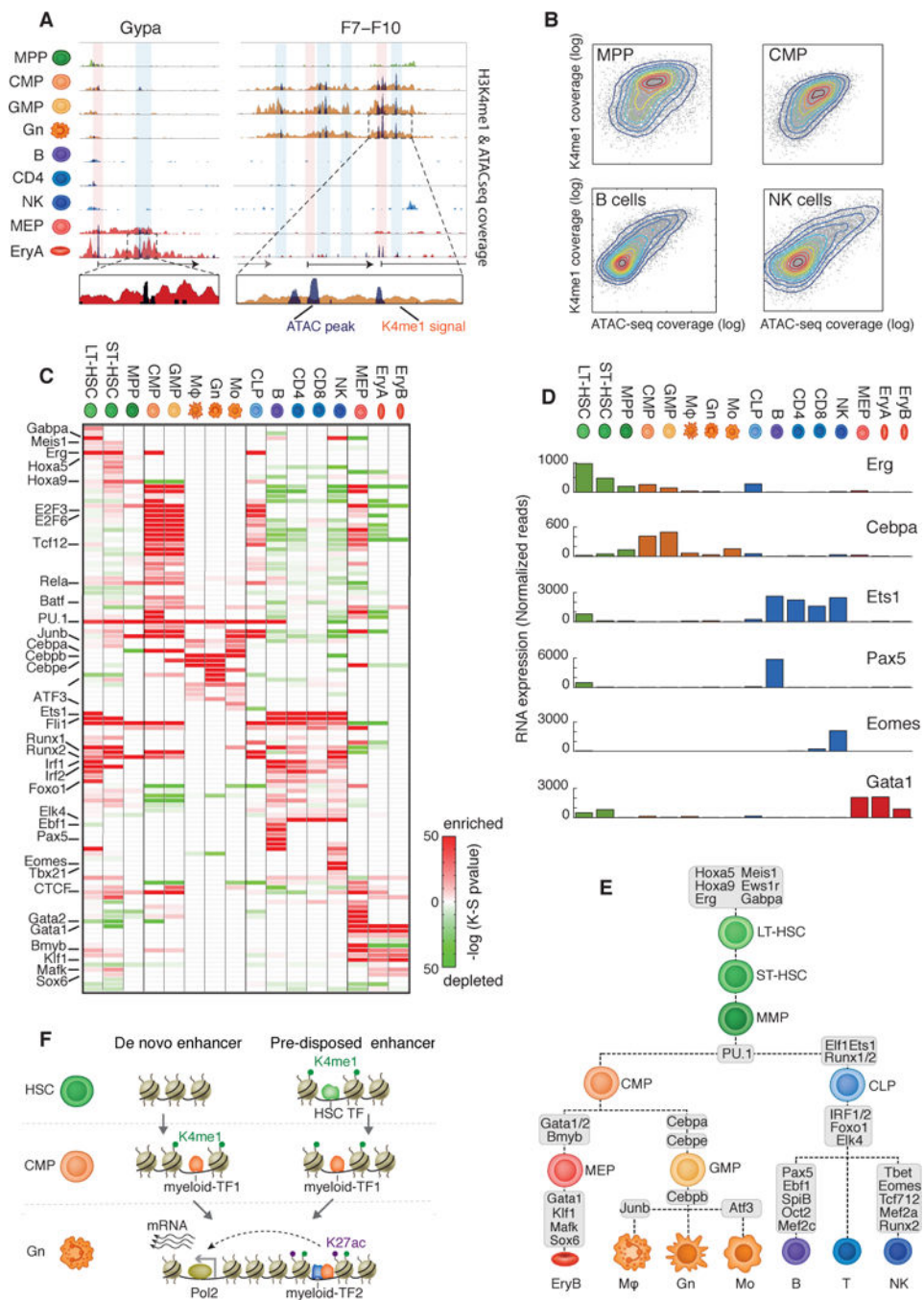
signal (cell types labeled at left) in several loci, (from left to right): Gata 2 for progenitors, F7–F10 and S100a8 for myeloid lineage, Gypa and Cpox for erythroid lineage, Bcl11b for T cells. Displayed are normalized reads coverage in a 100 Kb region around the gene body.

**(D)** Profiles of H3K4me1, H3K4me2, H3K4me3, H3K27ac modifications and RNA expression levels in two lineage specific gene loci: S100a8 in (myeloid) and Gypa (erythroid). Displayed are profiles for the lineage-specific cell types and multipotent progenitor cells over a 100 Kb region around the gene body. Putative lineage specific enhancers are shadowed in blue. Lineage specific genes are indicated below (black), as well as other genes in the loci (light gray).



**Figure 3. Hematopoiesis progresses through gain and loss of lineage specific enhancers**  
**(A)** Heatmap showing 48,415 hematopoiesis enhancers clustered with K-means ( $K=9$ ) of average reads count within H3K4me1 regions (16) annotated with the genomic sequence conservation of the enhancers, right (blue scale;  $I6$ ). Annotations for important lineage specific genes loci are shown on the left. **(B)** Schematic tree-view of three representative enhancer clusters; progenitors (V), myeloid (VI) and erythroid (IX). Color fill represents stages with H3K4me1 mark at the enhancer **(C)** Bar plot showing the number of enhancers gained (top, light gray) and lost (bottom, dark gray) during the development from HSC to

mature hematopoietic cells along each lineage. **(D)** Bar plot showing the percent of *de novo* enhancers (from C) established at each developmental stage from HSC to mature hematopoietic cells. **(E)** Scatter plots showing H3K27ac signal versus H3K4me1 signal in both CMPs and EryA cells. Active enhancers (H3K4me1 positive, H3K27ac positive) are colored in red, poised enhancers (H3K4me1 positive, H3K27ac negative) in light gray, inactive enhancers (H3K4me1 negative, H3K27Ac negative) in blue. **(F)** Proportions of active (red) and poised (light gray) enhancers in each of the cell types studied.



**Figure 4. Lineage-specific enhancers are associated with transcription factor cohorts**  
**(A)** ATAC-seq signal co-occurs temporally with H3K4me1 in lineage-specific enhancers. Shown are normalized profiles in 9 hematopoietic cells for ATAC-seq (Black) and H3K4me1 in erythroid-specific, Gypa (Red) and myeloid-specific, F7–10 loci (Orange). Displayed are peaks found in a 100 Kb region around the gene body. Putative enhancers (K4me1) shadowed in blue, transcriptional start site shadowed in red. Zoom in of ATAC-seq peak is shown below. **(B)** Scatter plot showing correlation between H3K4me1 signal and enhancer restricted ATAC-seq signal in four cell types: B cells, NK, MPP and CMPs (colors

represent density or points). **(C)** Heatmap showing the p-values of transcription factor motif (Kolmogorov-Smirnov test) for the indicated cell-type-specific enhancers (16). Red indicates significant enrichment ( $p < 1e-5$ ) of motif associated to the labeled TF, green indicates motif depletion ( $p < 1e-5$ ). White indicates either no significant enrichment or no RNA expression. **(D)** Bar plots showing gene expression profiles across the hematopoietic cell types for representative transcription factors (from C). **(E)** Schematic of the hematopoietic tree showing the representative transcription factors regulating the lineage enhancers at each cell type as predicated by the logistic model. **(F)** Schematic of establishment of lineage-specific enhancers in the myeloid lineage by a conventional mechanism (right) as well as via TF-mediated establishment of de novo enhancers (left).