CELL BIOLOGY

Temporal expression of MOF acetyltransferase primes transcription factor networks for erythroid fate

Cecilia Pessoa Rodrigues[1,2,3], Josip Stefan Herman[2,3,4], Benjamin Herquel[1], Claudia Isabelle Keller Valsecchi[1], Thomas Stehle[1], Dominic Grün[4,5], Asifa Akhtar[1,5]*

Self-renewal and differentiation of hematopoietic stem cells (HSCs) are orchestrated by the combinatorial action of transcription factors and epigenetic regulators. Here, we have explored the mechanism by which histone H4 lysine 16 acetyltransferase MOF regulates erythropoiesis. Single-cell RNA sequencing and chromatin immunoprecipitation sequencing uncovered that MOF influences erythroid trajectory by dynamic recruitment to chromatin and its haploinsufficiency causes accumulation of a transient HSC population. A regulatory network consisting of MOF, RUNX1, and GFI1B is critical for erythroid fate commitment. GFI1B acts as a *Mof* activator which is necessary and sufficient for cell type-specific induction of *Mof* expression. Plasticity of *Mof*-depleted HSCs can be rescued by expression of a downstream effector, *Gata1*, or by rebalancing acetylation via a histone deacetylase inhibitor. Accurate timing and dosage of *Mof* expression act as a rheostat for the feedforward transcription factor network that safeguards progression along the erythroid fate.

INTRODUCTION

A healthy adult individual produces on average 2×10^{11} red blood cells (RBCs) per day (*1*). For this, hematopoietic progenitor cells (HSPCs) have to be constantly engaged in erythrocyte differentiation, where failure may lead to severe defects such as anemia or leukemia (*2*, *3*). Adult erythropoiesis begins with an early phase in which hematopoietic stem cells (HSCs) or multipotent progenitor cells (MPPs) commit toward the erythroid branch (*4*–*6*). This is followed by a later phase of terminal differentiation, where the chromatin of erythroid cells undergoes a gradual condensation, followed by nuclear, mitochondrial, and endoplasmic reticulum shedding (*1*). In the hierarchical model of erythroid development, the bipotent megakaryocyte-erythroid progenitor cells (MEPs) give rise to either megakaryocytes and platelets or erythrocytes, depending on external cues such as oxygen tension, cytokines, glucocorticoids, and intrinsic processes such as cell cycle and chromatin environment (*1*). For instance, during terminal erythroid differentiation, histone post-translational modifications, such as H3K9ac, H3K79me2, and H4K16ac, have been shown to promote β-globin locus control region elongation and activation of β-globin gene loci (*7*–*9*). While the role of chromatin marks at such specific loci has been studied, whether and how erythropoiesis is orchestrated by histone modifiers at the genome-wide level is largely unexplored.

Because of the small number of progenitor cells, the earlier steps of erythroid fate commitment are less well defined. Experimental approaches using sorting based on surface markers are limited to a predefined set of cells and thus will not provide the full continuum of cell populations along the erythroid trajectory. Genome-wide single-cell techniques can, in part, overcome these issues and have raised questions about traditional pictures of cellular differentiation. For example, HSCs and MPPs appear more heterogeneous than previously anticipated (*10*, *11*), while erythroid cells may arise not only from MPP2 but also from MEPs (*6*). In addition, single-cell RNA sequencing (scRNA-seq) analysis revealed global fluctuation of mRNA levels during erythropoiesis (*12*), which was proposed to be due to the activity of epigenetic regulators in early progenitors and lineage-committed cells.

Similarly, long-term HSCs (LT-HSCs) display increased number of assay of transposase accessible chromatin sequencing (ATAC-seq) peaks when compared to short-term HSCs (*13*) and high levels of histone H4K16ac (*14*), which is known for directly affecting chromatin accessibility in vitro (*15*). Nevertheless, several questions remained unanswered: What factors control these global chromatin accessibility changes? And to what degree are they instructive and physiologically relevant to the erythroid lineage? The answer to those questions might reveal how chromatin accessibility, proper RNA levels, and its correct temporal expression affect erythropoiesis.

Here, we show that as part of the KAT8-associated non-specific lethal (KANSL) complex, the lysine acetyltransferase (KAT) MOF is responsible for orchestrating chromatin accessibility changes along erythropoiesis. *Mof* haploinsufficiency ($Mof^{+/-}$) and *Kansl2* and *Kansl3* deletion cause impaired RBC development that is associated with decreased levels of H4K16ac and altered chromatin accessibility dynamics in HSCs and erythroid progenitors (MPP2 and MEP). $Mof^{+/-}$ HSCs fail to sustain erythroid formation in vitro and in vivo. Chromatin immunoprecipitation sequencing (ChIP-seq) and scRNA-seq reveal that MOF-guided trajectories are markedly perturbed in the absence of correctly timed regulation of *Mof*. As a consequence of defective differentiation, $Mof^{+/-}$ animals accumulate a previously uncharacterized HSCs population, which expresses high levels of mRNA characteristic of both dormant (dHSCs) and active HSCs (aHSCs). We show that expression of *Mof* is tightly regulated along erythropoiesis by the transcription factor (TF) GFI1B, which is responsible for ensuring appropriate *Mof* levels in a cell type-specific fashion. Last, we demonstrate that aberrant trajectories can be rescued by ectopic expression of *Mof*, reestablishment of the erythroid program, or by recovering acetylation levels through use of a histone deacetylase inhibitor (HDACi). Therefore, we propose that *Mof* regulates the

[1]Department of Chromatin Regulation, Max Planck Institute of Immunobiology and Epigenetics, 79108 Freiburg, Germany. [2]University of Freiburg, Faculty of Biology, Schaenzlestrasse 1, 79104 Freiburg, Germany. [3]International Max Planck Research School for Molecular and Cellular Biology (IMPRS-MCB), Freiburg, Germany. [4]Laboratory of Quantitative Single-Cell Biology, Max Planck Institute of Immunobiology and Epigenetics, 79108 Freiburg, Germany. [5]CIBSS—Centre for Integrative Biological Signaling Studies, University of Freiburg, Freiburg, Germany.
*Corresponding author. Email: akhtar@ie-freiburg.mpg.de

erythropoiesis by orchestrating the interplay between chromatin accessibility and correct temporal gene expression.

RESULTS

The Non-specific lethal complex in erythropoiesis

Histone acetylation is controlled by KATs and histone deacetylases (HDACs), where a hyperacetylated state typically results in chromatin decompaction (*16*, *17*). To investigate which KATs and HDACs might be orchestrating the dynamic gene expression observed during erythropoiesis, we analyzed published transcriptome-wide data (*18*). While most enzymes (e.g., *Tip60*, *Ep300*, or *Sirt1*) were stably expressed throughout the erythroid trajectory, we observed that *Mof* expression was dynamic (Fig. 1A). We also evaluated the erythroid, myeloid, and lymphoid fate-bias probabilities from c-Kit$^+$ cells [data from (*12*), in relation to *Mof* expression (fig. S1, A and B)]. Similar to known erythroid markers (e.g., *Gata1*, *Kfl1*, *Gypa*, *Hba-a1*, and *Alas2*), cells expressing *Mof* displayed a higher propensity to develop along the erythroid lineage compared to the myeloid or lymphoid lineages (Fig. 1B).

We next wanted to determine whether loss of *Mof* affects erythropoiesis. Constitutive knockout (KO) mouse models of *Mof* are embryonic lethal (*19*). However, *Mof* heterozygous (*Mof*$^{+/-}$) mice are viable and do not show obvious phenotypes or altered life span. Hematopoietic progenitors of heterozygous animals display a 60% reduction in *Mof* mRNA (fig. S1C) and bulk reduction of MOF protein (fig. S1D) and H4K16ac levels (fig. S1E). We found that *Mof*$^{+/-}$ mice show low RBCs and hemoglobin (HB) and hematocrit (HCT) levels (Fig. 1C), classifying them as anemic. To pinpoint the origin of defective blood formation, we performed extensive characterization of the cellular repertoire of the bone marrow (BM) in *Mof*$^{+/-}$ mice (fig. S2 and note S1) and independently validated the phenotype using a conditional *Vav1*-iCre *Mof*-KO model (fig. S3, A to F, and note S1) (*20*). *Mof*$^{+/-}$ animals displayed a significant decrease of erythroid progenitors and MEPs (fig. S2B), accompanied by increased numbers of myeloid progenitors (fig. S2B). These mice also accumulate RBCs with morphological defects and had an increased percentage of circulating reticulocytes (fig. S2, C to F). However, these changes were not accompanied by any significant differences in total BM cellularity nor in the number of lineage$^-$cKit$^+$ cells. Furthermore, we also did not observe any major changes in splenic mass and/or cellularity and differences in frequency of apoptosis nor of cell cycle abnormalities in the erythroblast. Yet, there was a marked reduction of the erythroid compartment, and we detected an increased number of neutrophils in the periphery (fig. S2, G to M). Overall, these analyses suggest that reduction in *Mof* levels leads to impaired erythroid development accompanied by myeloid skewing.

MOF resides in two distinct chromatin-modifying complexes, the non-specific lethal (NSL; KANSL in mammals) complex and the male-specific lethal (MSL) complex. As part of the mammalian MSL complex, MOF participates in the fine-tuning of developmental genes (*21*, *22*). However, human individuals with loss-of-function mutations in *MSL3* do not show defects in blood cells (*21*). We therefore suspected that MOF exerts its function during erythropoiesis as part of the KANSL complex (*23*). We generated conditional *Kansl2* and *Kansl3* KO mice and characterized their blood composition upon deletion with *Vav1-iCre* (for details, see note S1 and fig. S3, G to O). Both mutants showed defects in erythroid development. *Vav1*-iCre/*Kansl2* KO mice more closely resemble *Mof*$^{+/-}$ mice in that they also have a normal life span but exhibit a pronounced decrease in erythroid progenitors and RBCs in vivo (fig. S3, H and I) and in vitro (fig. S3, J and K). Furthermore, sublethally irradiated wild-type mice receiving adoptive transfer of 200 *Vav1*-iCre/*Kansl2* KO HSCs (LSK$^+$CD34$^-$Flt3$^-$CD48$^+$CD150$^+$) also showed compromised erythropoiesis (fig. S3L). In contrast, *Vav1*-iCre/*Kansl3* KO triggered a stronger phenotype, characterized by embryonic lethality and loss of fetal liver erythroid progenitors, particularly in terminal erythroid differentiation populations (fig. S3, M to O). In contrast to *Mof*, the *Kansl2* and *Kansl3* conditional KOs did not show an increase in leukocytes (fig. S3I), which implies that the myeloid bias found in *Mof*$^{+/-}$ animals might arise from KANSL-independent functions.

Given the two expression peaks of *Mof* along the erythroid branch (Fig. 1A), we were interested in further exploring whether the anemic phenotype arises from functions in HSCs, progenitor cells, and/or committed cells. We first tested whether *Mof* loss affects the intrinsic capacity of HSCs to commit to the erythroid lineage. We sorted single HSCs from wild-type or *Mof*$^{+/-}$ mice into liquid culture wells containing cytokines that support myeloid, erythroid, and lymphoid differentiation [single-cell colony-forming unit (sc-CFU)]. We assayed the clonal output of 219 single colonies (72 wild-type and 147 *Mof*$^{+/-}$) by fluorescence-activated cell sorting (FACS) for markers of myeloid (c-Kit$^-$CD11b$^+$), erythroid (c-Kit$^-$Ter119$^+$), and lymphoid differentiation [c-Kit$^-$IgD$^+$; B cell population, since these conditions lack cell–cell interactions that are required for T cell priming (*24*)]. After 10 days in culture, we observed no difference in colony potency, as both wild-type and *Mof*$^{+/-}$ HSCs could give rise to multipotent (all three populations), bipotent (two populations), and oligopotent (only one population) colonies, with similar colony size and total cell number per colony (Fig. 1, D and E). However, we found the resulting cells to be skewed, where *Mof*$^{+/-}$ HSCs generated significantly more myeloid cells and less erythroid cells, while B cell numbers remained unchanged (Fig. 1F). To further validate the proliferation and differentiation capacities of these HSCs, we performed bulk CFU assay, wherein 100 FACS-sorted HSCs from wild-type or *Mof*$^{+/-}$ mice were plated in a methylcellulose-based medium supplemented with cytokines/growth factors that could sustain the growth of two types of erythroid progenitor cells [burst-forming unit erythroid (BFU-E) and CFU-E], granulocyte and/or macrophage (GM) progenitor cells (CFU-GM), and granulocyte, erythrocyte, monocyte, and megakaryocyte colony (CFU-GEMM). In agreement with the results obtained in the sc-CFU assay (Fig. 1, D and E), we observed decreased numbers of BFU-E and an expansion in CFU-GM from *Mof*$^{+/-}$ HSCs in the bulk CFU experiment (Fig. 1G).

Since lifelong depletion of *Mof* in heterozygous animals could cause secondary effects or lead to changes in expression of cell surface markers that may confound our analyses, we validated our conclusions using an inducible mouse model (*Cag-CreERT2*$^{T/+}$;*Mof*$^{fl/fl}$), in which *Mof* is only deleted upon 4-hydroxytamoxifen (4-OHT) treatment in culture (referred to as *Mof*-iKO). Inducing the *Mof* KO by 4-OHT treatment in HSCs resulted in increased numbers of myeloid colonies and less erythroid colonies (Fig. 1H), recapitulating the phenotype observed in *Mof*$^{+/-}$ HSCs. We observed that *Mof*$^{+/-}$ HSCs showed a marked reduction in the CFU-GEMM, which was much less pronounced in *Mof*-iKO mice, suggesting that this may be a secondary effect. Next, we performed serial plating CFU assays from *Mof*$^{+/-}$ and *Mof*-iKO HSCs. This assay revealed that whereas

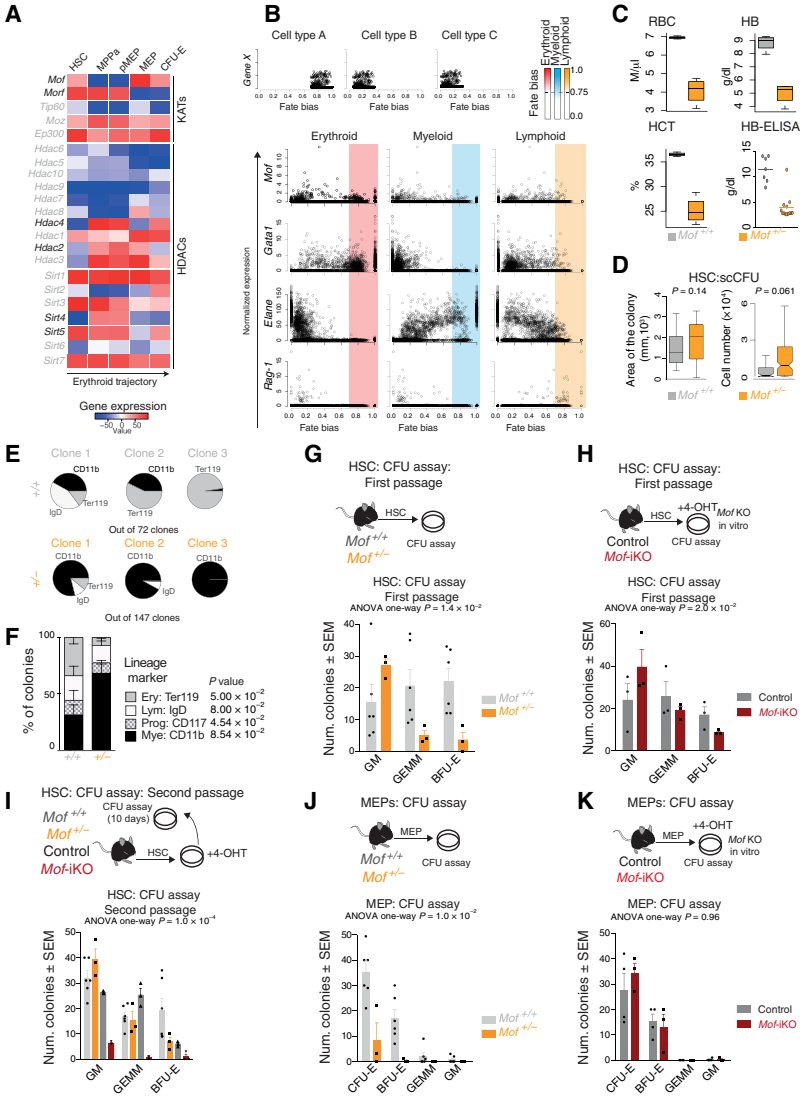

Fig. 1. *Mof* exhibits dynamic expression along erythropoiesis, and its reduction has a pronounced impact on erythroid lineage commitment. (**A**) Heatmap representing gene expression patterns of KATs and HDACs along the erythroid trajectory [data from (*18*)]. (**B**) Top: Schematic representation of fate-bias analysis, showing the cell fate-bias score (*x* axis) versus transcript expression (*y* axis). The fate-bias score indicates the probability of a given cell to belong to target cell types A, B, or C. A score of 1.0 indicates that the given cell has 100% probability of belonging to a target cell type. The example cell expressing *Gene X* has a 75 to 100% chance of belonging to cell type A but only a 0 to 20% chance of belonging to cell types B or C. In the analysis represented in the bottom panel, we selected erythroid, myeloid, and lymphoid cells as our target cell types. Bottom: Dot plot from 4763 cKit$^+$ cells, showing the cell fate-bias score determined by *FateID* plotted against *RaceID*-normalized transcript expression in erythroid [*Gata1*$^+$*Klf1*$^+$*Gypa*$^+$*Hba-a1*$^+$; cluster 3; erythroid versus *Mof* expression $r = +2 \times 10^{-2}$; Pearson correlation, $P = 4 \times 10^{-2}$), myeloid (*Mpo*$^+$*Elane*$^+$; cluster 1; myeloid versus *Mof* expression $r = -3.6 \times 10^{-2}$; Pearson correlation, $P = 6.3 \times 10^{-3}$), and lymphoid (*Ebf1*$^+$*Rag1*$^+$; cluster 7; lymphoid versus *Mof* expression $r = +6 \times 10^{-3}$; Pearson correlation, $P = 0.32$) lineages; original data from (*12*)]. For t-distributed stochastic neighbor embedding (t-SNE), maps, and cluster annotation see fig. S1 (A and B). (**C**) Box plot showing the total number of RBC, hemoglobin (HB) quantity, and hematocrit (HCT) percentage ($n = 4$ per genotype) and scatter plot showing the HB concentration measured using an enzyme-linked immunosorbent assay (ELISA)–based assay (wild type, $n = 7$; *Mof*$^{+/-}$, $n = 13$; $P < 0.001$). (**D**) Left: Box plot showing the total area of the colonies obtained from the single-cell colony-forming unit (sc-CFU) assay on fluorescence-activated cell sorting (FACS)–sorted HSCs (LSK$^+$Flt3$^-$CD34$^-$CD48$^-$CD150$^+$). Right: Box plot showing the total cell number per colony obtained in the sc-CFU assay. (**E**) Pie charts representing the lineage potency from the sc-CFU assay. (**F**) Stacked bar plot showing the fraction of lineage output from each clone. Final populations were defined on the basis of cell surface markers: myeloid (Mye; cKit$^-$CD11b$^+$, number of cells: 208,961 wild type and 905,455 *Mof*$^{+/-}$), erythroid (Ery; cKit$^-$Ter119$^+$, number of cells: 205,963 wild-type and 27,564 *Mof*$^{+/-}$), lymphoid [Lym; cKit$^-$IgD$^+$, number of cells: 65,093 wild-type and 63,815 *Mof*$^{+/-}$], and progenitor cells (Prog; cKit$^+$, number of cells: 98,988 wild-type and 124,690 *Mof*$^{+/-}$). Significant enrichment by Fisher's exact test was set as $P > 0.05$. sc-CFU data from 72 wild-type and 147 *Mof*$^{+/-}$ colonies from three independent animals. (**G**) Bar plots showing the CFU capacity output of 100 FACS-sorted HSCs (wild-type, $n = 6$; *Mof*$^{+/-}$, $n = 3$). (**H**) Same as (G), but using HSCs sorted from Cag-Cre:ERT2$^{Tg/+}$(control) or *Mof*$^{fl/fl}$Cag-Cre:ERT2$^{Tg/+}$transgene (*Mof*-iKO) animals. After sorting, cells were cultured with 4-hydroxytamoxifen (4-OHT) to induce *Mof* depletion in vitro. Bar plots show CFU capacity output of HSCs (control, $n = 3$; *Mof*-iKO, $n = 3$). (**I**) Serial CFU assay scheme and serial colony formation capacity from (G) or (H). (**J**) CFU assay scheme and CFU capacity output of FACS-sorted MEPs (Lin$^-$cKit^{high}Sca-1$^-$IL-7R$^-$CD34$^-$FcRgII/III$^-$) from wild-type or *Mof*$^{+/-}$ (wild-type, $n = 6$; *Mof*$^{+/-}$, $n = 3$) animals. (**K**) same as (J), but from control or *Mof*-iKO–sorted MEPs, wherein *Mof* depletion was induced in vitro after sorting and plating. Error bars represent means ± SEM, and biological replicates are represented as the overlaid dots. Experimental significance was determined by one-way analysis of variance (ANOVA), $P < 0.05$. Related to figs. S1 to S4.

Mof$^{+/-}$ HSCs maintained their bias toward the myeloid fate after replating, *Mof*-iKO HSCs formed significantly fewer colonies overall than controls (Fig. 1I), suggesting that the complete *Mof* KO impaired HSC self-renewal capacity.

Given that *Mof*$^{+/-}$ mice showed not only defective early erythroid differentiation but also signs of impaired terminal erythroid differentiation (fig. S2, C to I, and note S1), we investigated the consequences of *Mof* reduction in HSCs in comparison with later stages. To this end, we first conducted the CFU assay on MEPs (Lin$^-$cKit^{high}Sca-1$^-$CD34$^-$FcγgII/III$^-$) sorted from either wild-type or *Mof*$^{+/-}$ animals and evaluated their colony-forming capacity in culture.

MEPs from *Mof*$^{+/-}$ animals, where MOF had been chronically altered during differentiation, displayed a marked reduction in CFU-E and BFU-E colonies from *Mof*$^{+/-}$ MEPs (Fig. 1J). To understand the consequences of *Mof* ablation in already committed erythroid/megakaryocyte progenitors, we sorted MEPs and treated the cells with 4-OHT in vitro after sorting and performed CFU assay. A different result was obtained when we used the *Mof*-iKO model, in which *Mof* deletion in already primed progenitors (MEPs) did not affect CFU-E nor BFU-E formation (Fig. 1K). These data support the idea that *Mof* function is important in early stages of erythropoiesis and the observed terminal differentiation.

To analyze HSC lineage commitment in vivo, we conducted adoptive transfers of *Mof*$^{+/-}$ HSCs. Eight weeks after transplantation, we observed a significant increase in spleen size (fig. S4A), and although we did not observe overall differences in cell death, we noticed a remarkable leukocyte imbalance characterized by decreased numbers of lymphoid but increased numbers of myeloid cells (fig. S4, B and C). In contrast to wild-type recipient spleens, *Mof*$^{+/-}$ recipient spleens failed to recover their tissue architecture, wherein we observed defective follicle structures and an overall decrease in the red pulp (fig. S4D). Furthermore, in line with their genotype *Mof*$^{+/-}$, transplanted cells (CD45.2$^+$) showed decreased H4K16ac levels (fig. S4E). *Mof*$^{+/-}$ recipient mice displayed decreased frequency of splenic erythroid progenitors and a significant decrease in circulating mature RBCs (fig. S4, F and G). Collectively, these data suggested that the increase in spleen mass could be a reflection of myeloid expansion.

Next, we examined the BM where *Mof*$^{+/-}$ recipient animals showed a severe reduction in CD45.2$^+$ cells (fig. S4H). Further gating in HSPCs revealed an increased numbers of resident HSCs and decreased MPPs (fig. S4, I and J). To explore the HSC self-renewal and differentiation capacities, we next performed BM secondary transplantations. *Mof*$^{+/-}$ secondary recipients showed premature death due to severe anemia (fig. S4, K and L). This result also suggested that *Mof*$^{+/-}$ HSCs might acquire impaired self-renewal upon repetitive challenges, as noted for *Mof*-iKO in the CFU serial plating experiments (Fig. 1I).

Together, these data suggest that loss of one allele of *Mof* perturbs HSC lineage commitment in a cell-intrinsic manner, with a preference toward the granulocytic/myeloid lineage at the expense of the erythroid lineage. Furthermore, we found that deletion of *Mof* after lineage specification, i.e., in MEPs, does not lead to erythroid colony-formation defects.

Accumulation of a transient progenitor population due to a perturbed erythroid program in *Mof*$^{+/-}$ animals

To understand the impaired erythroid lineage commitment in vivo, we performed scRNA-seq experiments on FACS-sorted progenitor cells from wild-type and *Mof*$^{+/-}$ mice [792 LT-HSCs, 576 LSK^{high}CD34$^-$Flt3$^-$, 864 LSK^{high}, 576 LSK^{low}, 768 cKit^{high}, 48 megakaryocyte progenitors, 24 pre-granulocyte/macrophage progenitors (pre-GMP), 24 granulocyte/macrophage progenitors (GMP), 48 pre–Meg-E, 48 pre–CFU-E, 48 MEP, and 24 Pro-E] (fig. S5, A and B, and see Materials and Methods). Viability dye was used to ensure that only viable cells were sorted. We applied RaceID (*25*) to generate clusters (fig. S5, C and D) based on gene expression and used various approaches, including the expression of known marker genes and transcriptome entropy, to validate the cluster profiles (Fig. 2, A to C, and fig. S5, E and F). We observed a pronounced deregulation of the erythroid program in *Mof*$^{+/-}$ erythroid cells, e.g., *Gata2*, *Hbb-bt*, *Hba-a1*, *Trib2*, and *Aqp1* in MEPs (cluster 6) and *Mpo*, *Car-1*, *Car-2*, and *Hba-a1* in erythroblasts (cluster 3) (Fig. 2D).

We were next interested in understanding how the cellular fate is perturbed in *Mof*$^{+/-}$ cells and performed fate-bias analyses using *FateID* (Fig. 2E and fig. S5, G and H) (*25*). These analyses confirmed that cells expressing *Mof* have a higher probability of becoming erythroid in comparison to cells expressing other KATs (fig. S5G). We next defined two different trajectory end points (GMPs, cluster 8; erythroblasts, cluster 3) and extracted the probability of all cells in other progenitor populations or of dHSCs to differentiate to these two end points. This revealed that *Mof*$^{+/-}$ cells have a significantly lower probability of following the erythroid lineage in comparison with the GMP lineage (Fig. 2E and fig. S5H). Consistent with the prediction of lower erythroid differentiation, we observed decreased erythroblast populations in the BM of *Mof*$^{+/-}$ mice (Fig. 2F), despite observing no significant changes in the BM cellularity in these animals (fig. S2B). This reduction in erythroblasts was not associated with increased apoptosis nor decreased proliferation (fig. S5I). However, it was consistent with a defective erythroid program as we observed an altered *Gata1-Pu.1* balance in sorted MEPs (fig. S5J), therefore validating our scRNA-seq findings.

We classified our t-distributed stochastic neighbor embedding (t-SNE) map based on genotype and detected a significant enrichment of wild-type over *Mof*$^{+/-}$ cells in dHSCs (cluster 2) and MPPs (cluster 4), whereas an otherwise rare HSC population (cluster 10) was enriched in *Mof*$^{+/-}$ animals (Fig. 2G and note S2). We confirmed this by FACS, where we observed increased HSCs, but significantly decreased numbers of MPP2 (fig. S6A) in *Mof*$^{+/-}$ BM, with no changes in cell death (HSC to MPP4), reactive oxygen species (ROS) production (HSCs), or cell cycle (LSK$^+$) (fig. S6, B to D). Taking into consideration that *Mof*$^{+/-}$ HSCs were less prone to forming erythroid colonies (Fig. 1, E and H), we suspect that accumulation of cells in cluster 10 could be a consequence of impaired differentiation.

Intrigued by this population, we extensively characterized its features (fig. S6, E to L, and note S2). We detected canonical dHSC expression patterns such as decreased regulation of the *Myc* and *Tal1* network, an active *Erg* and *Hoxa9* network, and strong correlations with the respective molecular overlap [*MoLO* genes (*11*)] and high retinoic acid signatures. Cluster 10 also simultaneously displayed features of aHSCs such as expression of *Meis1*, *Cdk4*, *Cdk6*, and the active *Stat1* pathway. Given the presentation of both active and dHSC features, we hypothesize that *Mof*$^{+/-}$ HSCs represent an intermediate HSC state. Together, our analyses suggest that correct execution of the erythroid program requires the action of MOF in progenitor cells.

MOF-mediated chromatin accessibility dynamics during erythropoiesis

We next sought to investigate the molecular mechanism by which MOF directs the erythroid program. Since MOF deposits H4K16ac

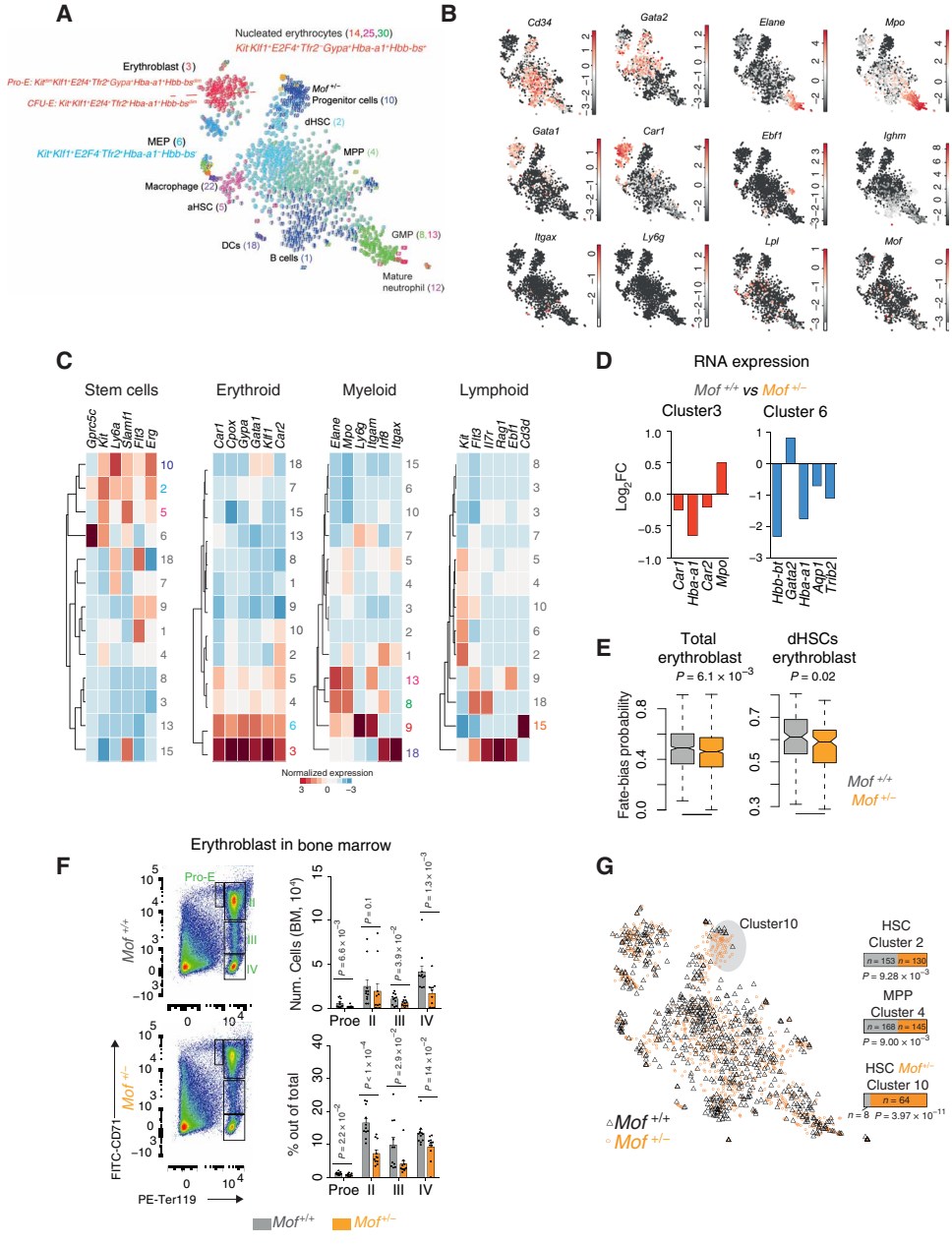

Fig. 2. *Mof*$^{+/-}$ erythroid progenitor cells show altered erythroid identity and low probability to belong to the erythroid trajectory. (**A**) t-SNE representation of transcriptome similarities between each cell. After normalization and filtering, 34,111 genes and 1842 cells were analyzed. Sorting strategy, scRNA-seq quality controls, and cluster characterization are shown in fig. S5 (A to D). t-SNE map shows the population annotation based on differentially expressed genes, transcriptome entropy, and key TFs generated by the *RaceID3*/*StemID2* algorithm. (**B**) Expression of *Mof* and representative marker genes is highlighted on the t-SNE map from (A). Multipotent stem cells (*Cd34* and *Gata2*), myeloid (*Elane* and *Mpo*), erythroid (*Gata1* and *Car2*), B cells (*Ebf1* and *Ighm*), dendritic cells (DCs; *Itgax*), neutrophils (*Ly6g*), and macrophage (*Lpl*). The scale bars show the normalized expression for each gene [for further markers, see figs. S5 (E and F) and S6 (E to K)]. (**C**) Heatmap showing the log-normalized expression of six key population markers each for "stem cells," "erythroid," "myeloid," and "lymphoid" across all clusters with $n > 5$ cells. Expression was scaled by genes. (**D**) Bar plot showing that the top four differentially expressed genes in clusters 3 and 6 are related to erythropoiesis. scRNA-seq data were generated from $n = 7$ animals in three independent experiments (see Materials and Methods for details). FC, fold change. (**E**) Box plots showing the erythroblast fate-bias probability of wild-type (gray) and *Mof*$^{+/-}$ (orange) erythroblast cells (cluster 3) and dormant HSCs (dHSCs; cluster 2). Statistical significance was determined by *t* test, $P < 0.05$. (**F**) Left: Representative FACS dot plot of resident BM erythroblasts. Erythroid populations are shown in green. Right: Bar plot showing total number or total frequency of erythroid progenitor cells in the BM. Data from $n = 11$ animals per genotype. After normal distribution evaluation, the *P* values were calculated by unpaired *t* test or two-tailed Mann-Whitney test. (**G**) Left: t-SNE projection depicting cells genotype, wild-type (black triangles) and *Mof*$^{+/-}$(orange circles) cells. The gray oval highlights cluster 10. Right: Bar plots showing clusters that are significantly enriched in one genotype. *P* value was determined by Fisher's exact test. Related to figs. S5 and S6.

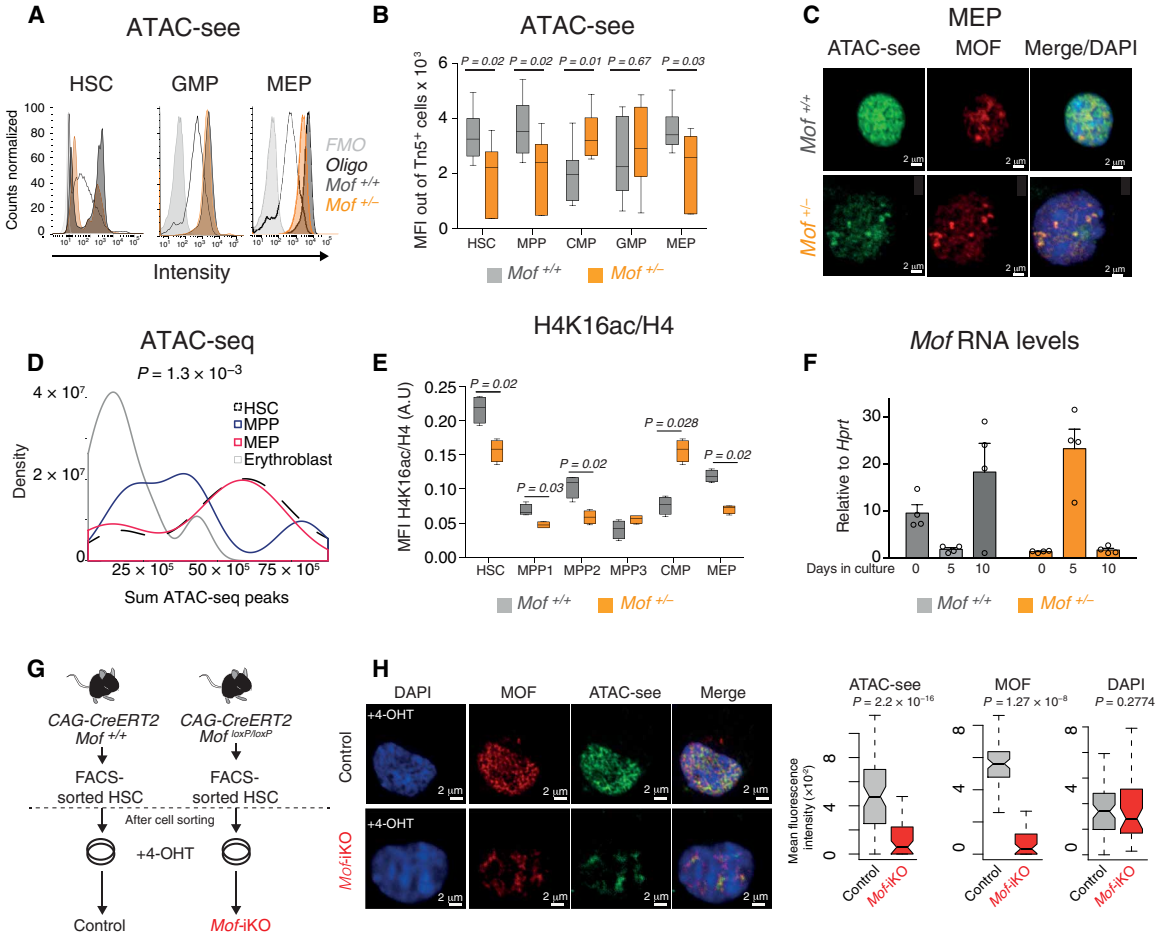

Fig. 3. *Mof* expression is associated with global chromatin accessibility and H4K16ac levels. (**A**) Representative flow cytometry histograms showing ATAC-see staining intensities. The dark-gray shading shows the ATAC-see in wild-type animals, and orange shows the ATAC-see for *Mof*$^{+/-}$animals. The black line marks ATAC-see oligo autofluorescence, and light gray shading shows the fluorescence minus one (FMO) control in which the ATAC-see oligo was omitted. (**B**) Summary box plot showing the quantification of the ATAC-see FACS median fluorescence intensities (MFIs) of HSCs (LSK^{+}CD34^{-}CD150^{+}), MPPs (LSK^{+}CD34^{+}), CMPs (Lin^{-}Kit^{high}Sca1^{-}CD34^{+}CD16/32^{dim}), GMPs (Lin^{-}Kit^{high}Sca1^{-}CD34^{+}CD16/32^{high}), and MEPs (Lin^{-}Kit^{high}Sca1^{-}CD34^{-}CD16/32^{-}) from wild-type (gray bars) and *Mof*$^{+/-}$ (orange bars) mice. Data represent $n = 7$ independent animals. *P* values were calculated by two-tailed Mann-Whitney test, with confidence level set as 95%. (**C**) Representative images of ATAC-see (green), MOF immunofluorescence (red), and 4′,6-diamidino-2-phenylindole (DAPI; blue) in sorted MEPs from $n = 3$ animals and two independent experiments. Scale bars, 2 µm (see also fig. S8, A and B). (**D**) Density plot showing the mean number of ATAC-seq peaks found in human HSCs (dash line), MPPs (blue line), MEPs (red line), and erythroblasts (gray line). Data from GSE74912. (**E**) Box plots showing FACS MFI signal of H4K16ac in HSC (LSK^{+}CD34^{-}Flt3^{-}CD48^{-}CD150^{+}), MPP1 (LSK^{+}Flt3^{-}CD34^{+}CD150^{+}CD48^{-}), MPP2 (LSK^{+}Flt3^{-}CD34^{+}CD150^{+}CD48^{+}), MPP3 (LSK^{+}CD34^{+}Flt3^{-}CD150^{-}CD48^{+}), CMP, and MEP populations [as in (B)] normalized by their total H4 levels in wild-type and *Mof*$^{+/-}$ mice. Number of animals $n = 4$ from two independent experiments. For representative histogram, see fig. S8C. Mean value is shown as a line. *P* value was determined by two-tailed Mann-Whitney test, with confidence level set as 95%. A.U, arbitrary units. (**F**) Ex vivo erythropoiesis assay. Bar plot showing reverse transcription quantitative polymerase chain reaction (RT-qPCR) analyses of *Mof* levels in HSCs isolated from wild-type (gray) or *Mof*$^{+/-}$(orange) animals and cultured in methylcellulose-based medium for the indicated number of days. *Mof* expression was determined by [CT-*Mof* – CT-*Hprt*, dCT] and then normalized by the smallest dCT (normalization was conducted considering each genotype and day of interest), ddCT. The fold change is 2^{ddCT}. Each overlaid data point represents the number (*n*) of independent animals. Error bars represent means ± SEM. Significance was determined after normal distribution analysis by two-tailed Mann-Whitney test. (**G**) Scheme of CFU experiment. Wild-type, Cag-Cre:ERT2$^{Tg/+}$(control), or *Mof*$^{fl/fl}$Cag-Cre:ERT2$^{Tg/+}$transgene (*Mof*-iKO) HSCs were sorted and cultured with 4-OHT in methylcellulose medium for 10 days. After 10 days in culture, ATAC-see signal was evaluated. (**H**) Left: Representative immunofluorescence image of MOF (red) and ATAC-see (green) from wild-type, Cag-Cre:-ERT2$^{Tg/+}$, or *Mof*$^{fl/fl}$ Cag-Cre:ERT2$^{Tg/+}$ (*Mof*-iKO) mice. Right: MFI of ATAC-see, MOF, and DAPI signal (number of animals, $n = 4$). After normality test, two-tailed Mann-Whitney test was applied for statistical significance. Related to fig. S7.

and this directly affects chromatin accessibility in vitro (*15*), we investigated chromatin compaction along erythropoiesis using assay of transposase-accessible chromatin with visualization (ATAC-see). Briefly, this method permits visualization of the accessible genome by transposase-assisted in situ imaging at single-cell resolution (*26*). We analyzed ATAC-see signal by FACS (Fig. 3, A and B) and microscopy (Fig. 3C and fig. S7, A and B). Collectively, these data revealed that *Mof*$^{+/-}$ animals display significantly decreased global chromatin accessibility in HSCs (LSK^{+}CD34^{-}CD150^{+}), MPPs (LSK^{+}CD34^{+}), and MEPs (Lin^{-}Sca-1^{-}cKit^{High}CD34^{-}FcgRII/III^{-}), but not GMPs (Lin^{-}Sca-1^{-}cKit^{High}CD34^{-}FcgRII/III^{high}), supporting a cell type-specific function.

Our ATAC-see data revealed global changes in chromatin accessibility along the process of erythroid differentiation in wild-type

cells. When we reanalyzed published human ATAC-seq data from (*27*), we also noticed that HSCs and MEPs show higher numbers of accessible regions than MPP and erythroblasts (Fig. 3D). In line with this, we found significantly increased levels of H4K16ac in HSCs, MEPs, and MPP2 in our profiles (Fig. 3E and fig. S7C). Collectively, these data suggest a direct connection between H4K16ac and global chromatin accessibility in vivo.

We therefore asked whether this dynamic pattern in chromatin accessibility is mirrored by expression changes in the *Mof* itself. For this purpose, we sorted HSCs and performed CFU assays and collected RNA at day 0 (after cell sorting), 5 days, and 10 days after plating. We observed that *Mof* RNA levels in wild-type cells followed the same dynamic pattern that we had already observed in published and our own in vivo genome-wide data (Figs. 3F and 1A and fig. S9A). In comparison with wild-type cells, colonies derived from $Mof^{+/-}$ HSCs were characterized by initially low *Mof* levels, followed by an aberrant induction of *Mof* RNA at 5 days, and then again by low levels after 10 days in methylcellulose-based medium culture (Fig. 3F). These fluctuations in *Mof* RNA were mirrored by the dynamics of both chromatin accessibility and H4K16ac in vivo, where we observed a peak in chromatin accessibility by ATAC-see and higher H4K16ac levels in $Mof^{+/-}$ common myeloid progenitors (CMPs) (Figs. 3, B and E). These data indicate that erythropoietic defects arise not only in the absence of *Mof* [e.g., in *Vav1*-iCre/*Mof* KO mice; see also (*28*)] but also when its temporal expression is perturbed (e.g., $Mof^{+/-}$ mice). To verify that these changes arise from cell-intrinsic functions of MOF, we repeated ATAC-see (Fig. 3, G and H) in *Mof*-iKO HSCs, which recapitulated the earlier conclusions in $Mof^{+/-}$ mice. Together, this suggests that global chromatin accessibility dynamics along erythropoiesis is determined by MOF-mediated H4K16ac deposition.

Distinct MOF genomewide binding profiles in HSCs and MEPs

To assess how MOF directs hematopoietic lineage commitment, we next generated ChIP-seq profiles from primary HSCs (15,000 cells; $LSK^+CD34^-Flt3^-$) and MEPs (30,000 cells) isolated from wild-type and $Mof^{+/-}$ animals (fig. S8A). After quality control assessment (FastQC), we called peaks and identified 20,096 MOF-bound regions in HSCs (see Supplementary Data) and 25,521 MOF-bound regions in MEPs (see Supplementary Data), of which 16,924 and 6817, respectively, occur in close vicinity to transcription start sites (TSSs) [Fig. 4A and fig. S8, B and C; see also (*29*, *30*)]. We observed not only dynamically bound (cluster 1, HSC > MEP; cluster 2, HSC < MEP) but also cell type-invariant peaks (cluster 3; Fig. 4B). Validating the specificity of our peaks, the MOF ChIP-seq signal was globally decreased in $Mof^{+/-}$ HSCs, but this effect was most pronounced in cell type-specific cluster 1 (fig. S8, D and E). As a complementary dataset, we generated MOF and H4K16ac ChIP-seq profiles in hematopoietic precursor cell-7 (HPC7) cells. These data are in great agreement with the HSC data and further corroborated the specific overlap between MOF and H4K16ac on cell type-specific target sites, where other KATs [e.g., p300 data from (*31*)] are mostly absent (Fig. 4C).

To validate whether MOF-bound regions were affected, we assessed their expression in our scRNA-seq dataset. Consistent with the activating role of MOF-mediated histone acetylation, we observed that RNA levels of MOF-bound genes, but not randomized control genes, were globally and significantly reduced in $Mof^{+/-}$ HSCs (scRNA-seq, clusters 2 and 5) and MEPs (scRNA-seq, cluster 6) but not MPPs (scRNA-seq, cluster 4) (Fig. 4D). Moreover, examples of HSC-specific MOF targets (cluster 1), e.g., *Runx1* and *Fos*, showed decreased RNA levels in sorted $Mof^{+/-}$ HSCs ($LSK^+CD34^-Flt3^-CD48^-CD150^+$) (Fig. 4E and fig. S8F), while we did not observe changes in RNA levels of regions that were still bound in $Mof^{+/-}$ cells (e.g., *Cdk8*) (Fig. 4E).

Next, we were interested in understanding which biological processes were associated with MOF-bound genes in the different ChIP clusters and performed gene ontology (GO) term analysis. Targets that were specifically enriched in HSCs but not in MEPs (cluster 1) appeared enriched for terms associated with HSC differentiation (fig. S8G), for example, cytoskeleton changes and TF binding (*32*). Accordingly, cluster 1 TF targets were implicated in biological processes such as cell fate commitment and cellular differentiation (fig. S8G). On the other hand, cluster 2 MOF peaks (MEP > HSCs; fig. S8H) were specifically enriched in terms such as "core promoter sequence-specific DNA binding." This was mirrored in the respective TF target enrichments revealing "erythrocyte differentiation" and "RUNX1 regulates genes." The connection between MOF and *Runx1* was further strengthened by the observation that MOF binds the *Runx1* promoter in HSCs (Fig. 4E). Together, these analyses suggest that a major fraction of cell type-specific MOF targets are TF genes in HSCs and MEPs.

To characterize the TF networks that cooperate with MOF further, we performed TF affinity prediction [TRAP; (*33*)] analyses for the MOF peaks (Fig. 4F). Apart from an AT-rich motif in both clusters, the cell type-specific cluster 1 peaks were significantly enriched for *Elf5* and *Runx1* motifs, which are TFs involved in HSC differentiation (*13*, *34*). In cluster 2, we found motifs for the known erythroid regulators *Gata1* and *Arid3a* (*35*, *36*).

We next set out to solidify the link between MOF binding and changes in chromatin accessibility observed by ATAC-see (Fig. 3, A to D). For this, we analyzed bulk (*37*) and single-cell ATAC-seq (scATAC-seq) of hematopoietic cells (*38*). We found a significant enrichment of MOF peaks within open regions in both datasets. This was not observed in a control group of genes that displays on average similar expression levels, indicating that accessibility is not simply driven by gene activity per se but rather correlates with the presence of MOF (Fig. 4, G and H). This suggests that MOF-bound regions account for most accessible sites in HSCs and MEPs.

The results above prompted us to investigate whether regions made accessible by MOF in HSCs and MEPs were specifically enriched along the erythroid trajectory (fig. S8I and see also Materials and Methods). For this, we used a DNA motif-based approach starting with a pseudotime analysis of published scATAC-seq data. We extracted the *k*-mer information [flanking 7 bp of scATAC-seq peak, as published in STREAM (*39*)] from accessible sites in each cell along the erythroid trajectory (Fig. 4I). Next, we extracted the flanking *k*-mer from each MOF ChIP-seq TSS peak in HSCs and MEPs. We then asked whether the MOF-peak *k*-mers obtained in the two groups (HSCs or MEPs) are significantly enriched at the accessible sites and/or characteristic for a given population along the trajectory. HSC-MOF *k*-mers were specifically enriched at accessible motifs in the HSC-MPP trajectory (Fig. 4J). The MEP-MOF *k*-mers, instead, were increased in more committed cells in the MPP-MEP trajectory (Fig. 4K). Our analysis revealed significant enrichment of cell type-specific MOF *k*-mers in the HSC-MPP and MPP-MEP pseudotrajectory (Fig. 4, J to L).

Together, our analyses show that MOF binding sites in HSCs and MEPs are associated with open chromatin regions at those stages.

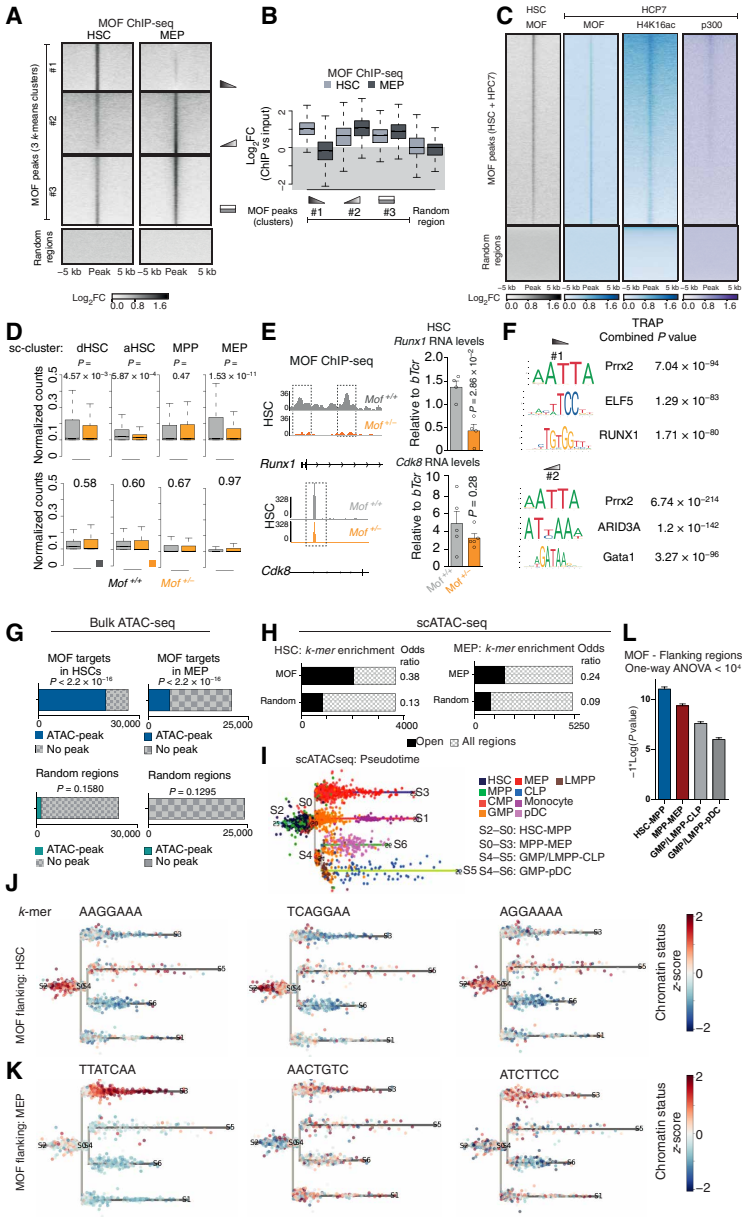

Fig. 4. Identification of cell type-specific MOF targets and its association with chromatin status. (**A**) Heatmap showing the $\log_2$(fold change) MOF ChIP-seq versus input enrichment on combined MACS2 peaks in HSCs and MEPs. Three unsupervised *k*-means clusters were generated, and the peak center was used as reference point ±5 kb (see Materials and Methods). The regions were ordered according to signal intensity and kept the same in HSCs and MEPs. Cluster 1 (#1) contains 13,150, cluster 2 (#2) 16,942, and cluster 3 (#3) 17,782 targets. (**B**) Box plot displaying the MOF ChIP enrichment intensities on all MACS2 peaks per *k*-means cluster and on random regions. Enrichment scores were calculated using deepTools2 multiBigwigSummary. For ChIP-seq peaks characterization, see fig. S8 (A to C), and for MACS2 regions, see Supplementary Data. (**C**) Heatmap showing the $\log_2$(fold change) ChIP versus input enrichment in HSCs and HPC7 cells. The peak center was used as the reference point while plotting the signal ±5 kb. ChIP-seq profiles were generated for MOF in HSC (gray) and HPC7 (blue), as well as those for H4K16ac in HPC7 (blue). p300 data from HPC7 was from (*31*) (purple). (**D**) Box plot showing the normalized transcript expression from MOF-bound genes or random genes across the scRNA-seq clusters. (**E**) Left: MOF ChIP-seq genome browser snapshots at *Runx1* and *Cdk8* loci (common bound peak) obtained from sorted HSCs of wild-type or *Mof*$^{+/-}$ mice. Right: *Runx1* and *Cdk8* RNA levels in sorted HSCs. dCT was determined by [CT-target gene − CT-*bTrc*] and then normalized by the smallest dCT (normalization was conducted considering each genotype), ddCT. The fold change is 2^{-ddCT}. After normality test, significance was evaluated by Mann-Whitney test (see fig. S8, D and F for MOF ChIP-seq analysis in *Mof*$^{+/-}$ HSCs). (**F**) TF affinity prediction (TRAP) motif analysis on MOF peaks from cluster 1 or cluster 2. The top-scoring motifs were chosen on the basis of combined *P* value. (**G**) Stacked plot showing the overlap between MOF target regions in HSCs, MEPs, or equally sized random regions enrichment in bulk ATAC-seq peaks (data from Encode). *P* value was calculated by Fisher's exact test using the BEDtools function *FisherBED*. (**H**) Overlap between extracted *k*-mers from single-cell ATAC-seq (scATAC-seq) HSCs or MEPs [data from (*38*)] and MOF-TSS flanking *k*-mers in HSC and MEP (see Materials and Methods for details). Enrichment was calculated by Fisher's exact test, and odds ratios are shown in the figure. (**I**) Subway map showing scATAC-seq pseudotime trajectory. Dots represent single cells, and colors represent cell population [data from (*38*)]. (**J**) Subway map showing three *k*-mers from MOF peaks in HSC or (**K**) MEP-associated *k*-mers in the pseudotime trajectory. Scale shows the scATAC-seq *z*-score. Pseudotime analyses were conducted by STREAM (*39*). (**L**) Bar plots showing the *P* value of MOF-peaks *k*-mers along HSC-MPP (S2-S0), MPP-MEP (S0-S3), GMP/LMPP-CLP (S4-S5), or GMP/LMPP-pDC (S4-S6) trajectories. One-way ANOVA determined experimental significance. Related to fig. S8.

This MOF-mediated chromatin accessibility, in turn, directs erythroid fate.

Cell type-specific control of *Mof* expression by GFI1B

Considering the dynamic expression and genome-wide occupancy of MOF along the erythroid trajectory, we predicted that the *Mof* gene is itself dynamically controlled by lineage-specific TFs. Therefore, we next focused our attention on the regulation of the *Mof* promoter. We performed individual motif occupancy predictions within the mouse and human *Mof* promoters and identified a conserved enrichment for two consecutive motifs for the TF GFI1B (Fig. 5A). GFI1B appeared enriched at the *Mof* promoter region in HPC7 cells in published GFI1B ChIP-seq profiles (Fig. 5B) (*40*). As *Gfi1b* is a direct downstream target of RUNX1 (*41*), which is itself a MOF target, we hypothesized that the TFs *Runx1* and *Gfi1b* participate in a regulatory loop with *Mof* that progressively triggers erythroid fate commitment. To validate targeting of these factors, we performed ChIP–quantitative polymerase chain reaction (qPCR) experiments in HPC7 cells, which reproducibly confirmed the binding of RUNX1 to the *Gfi1b* promoter and of GFI1B to the *Mof* promoter (Fig. 5C).

Mof mRNA levels show two peaks during erythropoiesis (Fig. 3F). Our scRNA-seq dataset uncovered that the first in vivo expression peak of *Mof* coincides with *Runx1* expression and the second one with increased *Gfi1b* levels (fig. S9A). Furthermore, *Mof*$^{+/-}$ animals also displayed misregulation of *Runx1* and *Gfi1b* (Figs. 4E and 5D and fig. S9A). To investigate whether *Gfi1b* is required for *Mof* expression, we performed small interfering RNA (siRNA) knockdowns (KDs) of *Gfi1b* in HPC7 cells, which resulted in markedly reduced MOF levels by immunofluorescence (Fig. 5E). To dissect this in further mechanistic detail, we next performed luciferase reporter assays to compare the wild-type *Mof* promoter with a mutant *Mof* promoter in which the 39-bp GFI1B binding region was shuffled (Fig. 5F). The assays were performed in HPC7 cells (reflecting a more naïve state) and K562 cells (reflecting a more erythroid-committed state). We observed a pronounced *Mof* promoter activity in HPC7 and K562 cells that was completely abolished upon either GFI1B binding site mutation or *Gfi1b* KD. The activity of this reporter construct was boosted upon ectopic expression of *Gfi1b* in K562 cells, whereas in HPC7 cells, this effect was rather subtle. Moreover, this regulation of the *Mof* promoter by GFI1B appeared to be highly cell type specific, as we did not observe a decrease in luciferase activity upon promoter mutation in peripheral blood mononucleated cells (Wehi3) or human embryonic kidney (HEK) 293T (nonhematopoietic origin) cells.

To explore the functional relevance of *Runx1* and *Gfi1b* for erythropoiesis, we single cell-sorted *Runx1* KD or *Gfi1b* KD HSCs ($LSK^+CD34^-Flt3^-CD48^-CD150^+$) into methylcellulose-based medium precoated wells (Fig. 5G). After 10 days in culture, we observed that *Gfi1b* KD HSCs showed substantially decreased colony formation. *Runx1* KDs in HSCs resulted in impaired BFU-E capacity but were able to sustain CFU-GM formation (Fig. 5, H and I). In agreement with our earlier observations, *Gfi1b* KD led to decreased expression of *Mof*. In turn, *Runx1* KD led to decrease in *Mof* and *Gfi1b* levels (Fig. 5J). Collectively, these data strongly support that GFI1B positively and directly regulates *Mof* through binding of its promoter. Hence, *Gfi1b* and *Runx1* together participate in a regulatory network, controlling the expression of a chromatin accessibility regulator (MOF) that is crucial for erythropoiesis. Collectively, we have found that the sequential action of lineage-specific TFs and MOF-mediated histone acetylation contributes to erythroid fate commitment.

Rescue of aberrant erythroid trajectories by acetylation rebalance and enforced erythroid fate

Given these intricate and precisely timed events, we wanted to understand how *Mof* regulates expression dynamics along erythropoiesis in further detail. We therefore performed pseudotemporal analysis of our scRNA-seq data. We found the highest probability for the erythroid fate in cluster 3 (erythroblast) (Fig. 6A, fig. S9B, and Materials and Methods). Using this cluster as the end point, we calculated the fate bias and extracted the cells belonging to the erythroid trajectory from the principal curve [see (*25*)]. In total, we ordered 264 wild-type and 215 *Mof*$^{+/-}$ cells and analyzed the expression of 14,279 genes along the erythroid trajectory. Genes with similar expression patterns were aggregated into nodes (Pearson's correlation coefficient, >0.85; list of genes/nodes is provided in Supplementary Data). This analysis allowed us to identify the timing of gene expression activities along the trajectory. Overall, we detected a notable difference in the gene expression patterns in wild-type compared with *Mof*$^{+/-}$ animals (Fig. 6B). For instance, a total of 3449 genes in nodes 1 to 16 are normally expressed during early stages, but in *Mof*$^{+/-}$ animals, they appear much later in pseudotime (Fig. 6, B and C). Genes that are normally expressed in more lineage-committed stages (nodes 45 to 49, 2966 genes), e.g., *Zfpm1*, did not display a uniform expression pattern at later stages of differentiation upon *Mof* haploinsufficiency. *Mof* itself appeared in node 42, which in *Mof*$^{+/-}$ cells showed a single, pronounced peak at lineage-committed stages. Node 42 also contained master regulators of erythropoiesis, such as *Klf1*, *Gata1*, *Gfi1b*, and *Epor*, which, similar to *Mof*, displayed a marked up-regulation at abnormally late stages in *Mof*$^{+/-}$ mice relative to the wild type. Furthermore, this cluster displayed enrichment for known erythroid TFs by GO term analysis (fig. S9C and see Supplementary Data). These notable differences found in the *Mof*$^{+/-}$ trajectory were also scored in trajectory-based differential gene expression analyses by *Monocle*, where we scored 1063 genes ($P < 0.05$; see Supplementary Data) as significantly deregulated in *Mof*$^{+/-}$ cells when compared to wild type. Together, these analyses suggest that the few HSCs that manage to become erythroid in *Mof*$^{+/-}$ mice do so in a completely aberrant manner.

These findings prompted us to test whether we could rescue erythroid lineage specification by restoring H4K16ac deposition. For this, we used *Mof*-iKO HSCs, which we transfected with an expression vector encoding either wild-type *Mof* (*Mof*) or a catalytically inactive mutant (*Mof*^{E350Q}) (Fig. 6D). By ectopically expressing *Mof*, we restored not only *Runx1* and *Gfi1b* levels but also the formation of erythroid colonies. By contrast, the *Mof* catalytic mutant was unable to rescue these features (Fig. 6, E and F, and fig. S9D). Next, we wondered whether these erythropoietic perturbations could also be ameliorated by expression of a master erythroid TF or by boosting acetylation levels. For this, we transfected *Mof*-iKO HSPCs with an expression vector for *Gata1* and sorted HSCs by FACS (fig. S9E). In parallel, *Mof*-iKO or *Mof*$^{+/-}$ HSCs were treated with the histone deacetylation inhibitor Ex-527 (*42*, *43*) to raise global H4K16ac levels. Colony formation assays revealed that erythroid trajectories were completely rescued in both of these scenarios (Fig. 6, G to M, and fig. S9, F to H). Notably, ectopically expressing *Mof* in wild-type cells also showed perturbed colony formation, arguing that *Mof* dosage plays a vital role during HSC differentiation.

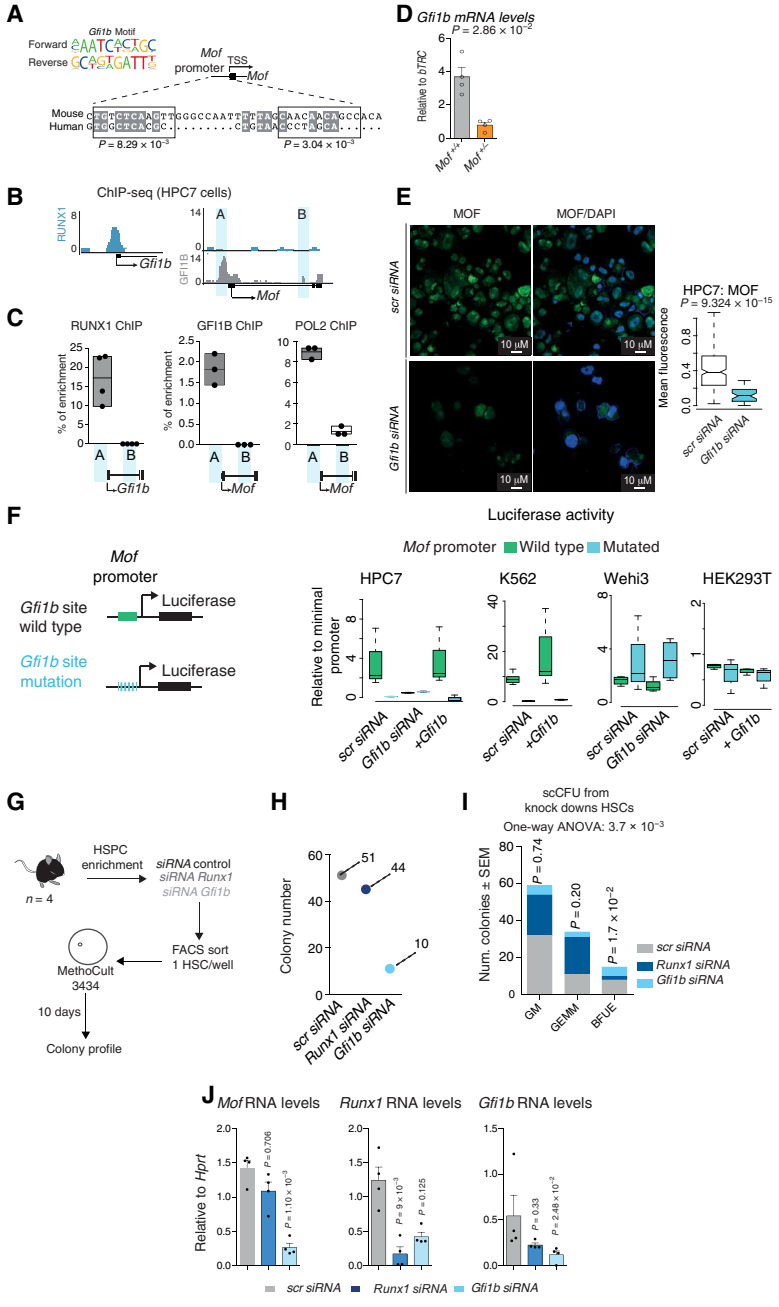

Fig. 5. GFI1B binds the *Mof* promoter and leads to *Mof* activation. (**A**) Find individually motif occurrence analysis showing GFI1B consensus sequence enrichment at the mouse and human *Mof* promoters. (**B**) Genome browser snapshot of RUNX1 and GFI1B ChIP-seq (*40*) at the *Gfi1b* and *Mof* genes in HPC7 cells. (**C**) ChIP-qPCR analysis of RUNX1 at the *Gfi1b* locus and GFI1B at the *Mof* locus in HPC7 cells. Floating plots showing the average of three independent replicates represented by data points. Enrichment was calculated as percentage of input. (**D**) Bar plot showing *Gfi1b* mRNA levels in LSK$^+$ cells from wild-type (gray) and *Mof*$^{+/-}$ (orange) animals. Normality distribution was scored by Shapiro-Wilk normality test, followed by two-tailed Mann-Whitney test for statistical significance. (**E**) Left: Representative immunofluorescence showing MOF and DAPI signal in HPC7 cells in wild-type [*scr* small interfering RNA (siRNA), siRNA control] or upon *Gfi1b* knockdown (KD; *Gfi1b* siRNA). Right: Box plot showing the MOF MFI in HPC7 cells upon *Gfi1b* or *scr* siRNA treatment. (**F**) Left: Scheme showing the wild-type *Mof* promoter construct, in which 790-nt upstream of the *Mof* 5′ untranslated region (5′UTR) was cloned, and the *Mof* mutant promoter construct, in which the GFI1b binding region (39 nt) was scrambled. Right: HPC7, K562, Wehi3, and human embryonic kidney (HEK) 293T cells were transfected with luciferase under the expression of either the wild-type or mutant *Mof* promoter constructs and subjected to either *Gfi1b* siRNA or *Gfi1b* ectopic expression (+*Gfi1b*). Box plot shows the median, maximum, minimum, and the interquartile range for *Mof* luciferase activity. *Mof* promoter activity was normalized over the minimal promoter activity. (**G**) Scheme of the single-cell CFU assay from wild-type (*scr* siRNA), *Runx1* KD, and *Gfi1b* KD in sorted HSCs. (**H**) Dot plot showing the total number of colonies observed from wild-type (*scr* siRNA, gray), *Runx1* (dark blue), or *Gfi1b* (light blue) KD HSCs. (**I**) Stack plot showing the colony types from (H), overall significance is measured by unpaired ANOVA followed by Kruskal-Wallis test, and colony enrichment is calculated by χ^2 test, with the significance threshold set as $P < 0.01$. (**J**) Bar plots showing RT-qPCR for *Mof*, *Runx1*, and *Gfi1b* after 10 days in culture. Expression was normalized to *Hprt* ($n = 3$). Significance is measured by unpaired ANOVA followed by Kruskal-Wallis test, and *P* values are calculated by Dunn's test multiple comparisons test relative to the *scr* siRNA group. Related to fig. S9A.

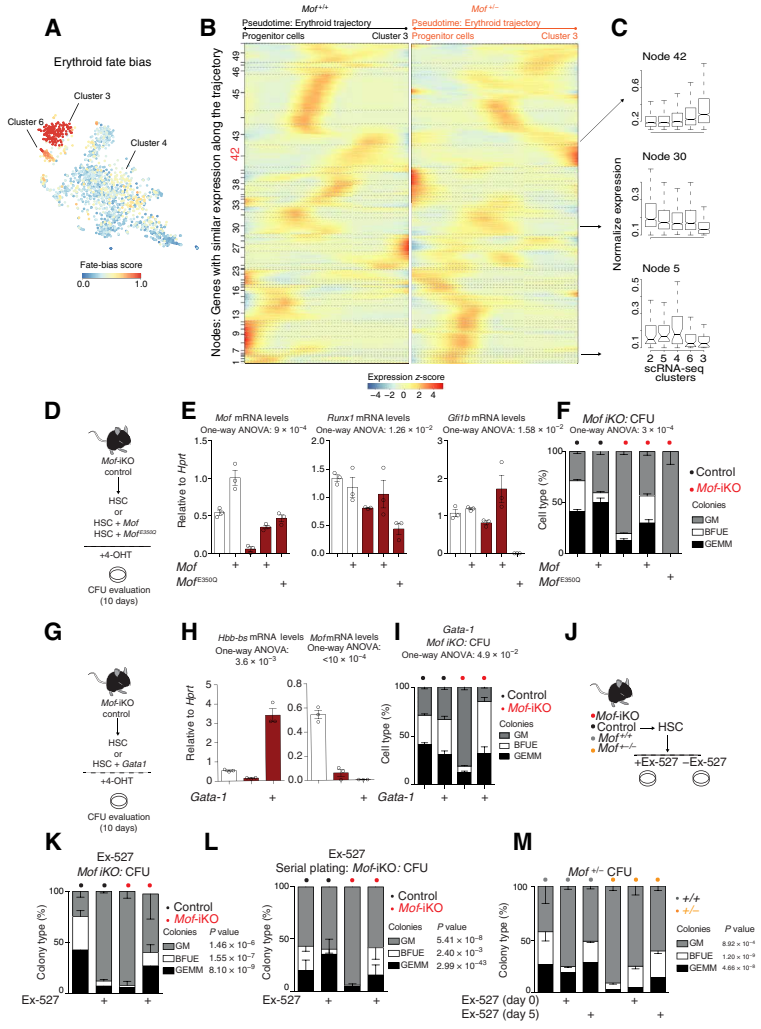

Fig. 6. Erythroid trajectory perturbation can be rescued by *Gata1* expression or by rebalancing acetylation with HDACi. (**A**) t-SNE map showing the erythroid fate bias. Fate bias was calculated per cell (dot in the t-SNE), and colors indicate the power of the bias, 1 being the highest (red) and 0 the lowest (blue) probability of becoming an erythroid cell. (**B**) Self-organizing map (SOM) generated from the erythroid trajectory (principal curve). For SOM ordering, 264 wild-type and 215 *Mof*$^{+/-}$ cells were extracted from the principal curve and fitted to cells with a fate bias of >0.4. The SOM shows the cells in pseudotemporal order on the *x* axis and transcripts that have similar expression profiles along the trajectory aggregated into nodes on the *y* axis (Pearson's correlation coefficient, >0.85), where eventually, every transcript corresponds to a line. Node composition and order are the same for wild-type and *Mof*$^{+/-}$ trajectories (for analysis workflow, see fig. S9B, and the list of genes/nodes and differentially expressed genes along the trajectory is in Supplementary Data). (**C**) Box plot showing the mean expression of all genes belonging to SOM nodes 5, 30, or 42 in the dHSC (cluster 2), aHSC (cluster 5), MPP (cluster 4), MEP (cluster 6), and erythroblast (cluster 3) populations identified in our scRNA-seq dataset (Fig. 2A). (**D**) Graphical scheme showing the experimental design of *Mof* rescue experiments. Control or *Mof*-iKO HSCs (LSK^{+}CD34^{-}Flt3^{-}CD48^{-}CD150^{+}) were sorted, transfected with pTreg3g-*Mof* or pTreg3g-*Mof*^{E350Q} (MOF catalytic mutant), and cultured with 4-OHT in methylcellulose medium for 10 days. After 10 days in culture, colony formation and molecular profiles were evaluated. (**E**) Bar plots showing RT-qPCR analysis of *Mof*, *Runx1*, and *Gfi1b* levels in control (white bars) and *Mof*-iKO (red bars) cells. Levels were determined by dCT (CT-target gene − CT-*Hprt*) and then normalized by the smallest dCT, ddCT. The fold change is 2^{ddCT}. Experimental *P* values were calculated by one-way ANOVA. Overlaid dots indicate the number of animals ($n = 3$). (**F**) Stacked bar plot showing colony profile from transfected HSCs. Error bars represent means ± SEM from $n = 4$ independent biological replicates. Circles indicate the mice genotype: control (black) and *Mof*-iKO (red). Experimental *P* value was determined by one-way ANOVA, and statistical significance was set as $P < 0.05$. For the total number of colonies, see fig. S9D. The "+" signal under the plot indicates the transfection condition. (**G**) Graphical scheme showing the experimental design of *Gata1* rescue experiments. Control or *Mof*-iKO HSCs were sorted, transfected with the pCAG-*Gata1* expression vector to enforce erythroid trajectory commitment, and cultured with 4-OHT in methylcellulose medium for 10 days. After 10 days in culture, colony formation was evaluated. (**H**) Bar plot of mRNA levels *Mof* and *Hbb-bs* in *Mof*-iKO and *Mof*-iKO plus exogenous *Gata1* (indicated by "+") (number of animals, $n = 3$). Relative expression was calculated as above. For the total number of colonies, see fig. S9E. (**I**) Colony quantification is represented by the stacked bar plot. Error bars represent means ± SEM from $n = 4$ independent biological replicates. Experimental *P* values were determined by one-way ANOVA, with statistical significance set as $P < 0.05$. (**J**) Graphical scheme showing the experimental design of HDACi rescue experiments. *Mof*$^{+/+}$, *Mof*$^{+/-}$, *Mof*-iKO, and control HSCs were sorted, treated with 1 mM Ex-527 treatment and cultured with (*Mof*-iKO and control) or without (*Mof*$^{+/+}$ and *Mof*$^{+/-}$) 4-OHT in methylcellulose medium for 10 days. After 10 days in culture, colony formation was evaluated. Mice genotypes are represented by the circles control (black), *Mof*-iKO (red), wild-type (gray), or *Mof*$^{+/-}$ (orange). Ex-527 treatment is indicated by the "+". (**K**) Stacked bar plot representing the colony profile after Ex-527 treatment in primary culture and (**L**), upon serial plating. Error bars represent means ± SEM from $n = 3$ independent biological replicates (for total number of colonies, see fig. S9G). (**M**) As in (K) but in *Mof*$^{+/+}$ and *Mof*$^{+/-}$ HSCs. Cells were treated with Ex-527 either immediately after FACS sorting (day 0) or after 5 days in culture (day 5). Error bars represent means ± SEM from $n = 3$ independent biological replicates. Enrichment was calculated by χ^2 test ($P < 0.05$). A total number of colonies are shown in fig. S9H. Related to fig. S9.

Together, our results suggest that proper dosage and timed expression of *Mof* regulate an intertwined network of TFs such as *Runx1*, *Gfi1b*, and *Gata1* whose expression is regulated by MOF-mediated histone acetylation (fig. S9I). These results establish the critical role of lineage-specific H4K16ac-driven chromatin accessibility in erythroid fate commitment in vivo.

DISCUSSION

Previous studies suggested the existence of a global epigenetic mechanism orchestrating erythropoiesis in a dynamic fashion (*12*). Our data show that MOF, the KAT responsible for H4K16ac, confers this function by regulating chromatin accessibility of HSCs and MEPs. MOF shows remarkably dynamic occupancy during erythropoiesis, which, together with our in vivo and ex vivo models, supports a cell-intrinsic mechanism of differentiation that is exquisitely sensitive to acetylation balance. We also demonstrated that correctly regulated dosage and timing of *Mof* expression are required to enable dynamic chromatin accessibility and TF network rewiring during erythroid commitment.

We show that MOF expression is temporally controlled by a cell type-specific TF network involving the TFs *Runx1* and *Gfi1b*. MOF binds to the *Runx1* promoter in HSCs, which, in turn, regulates *Gfi1b* expression. By directly targeting the *Mof* promoter, GFI1B acts as a positive regulator of *Mof* expression. This network constitutes a "feed-forward" pathway which confers erythroid-specific identity. Thus, appropriate MOF levels in HSCs are essential to prime the naïve chromatin state for transcriptional engagement of the erythroid program.

The existence of these feed-forward pathways might represent a universal principle of KAT function and explain the diverse roles played by other globally expressed epigenetic regulators, for example, the polycomb repressive complex 2 (PRC2) in regulation of the GATA2 to GATA1 switch (*44*). We note that MOF binds not only to TF genes but also to other genes in HSCs, for example, RNA binding proteins and processing factors. It is possible that RNA-based mechanisms such as splicing, translation, or RNA degradation provide an additional global regulatory layer during cellular differentiation. Given that $Mof^{+/-}$ HSCs accumulate at a transient state, which presumably requires deterministic *Mof* activation, one can envision a scenario where the interplay between transcriptional and posttranscriptional processes together enables the very rapid and efficient cellular transitions required during erythropoiesis. Future advances in single-cell technologies beyond expression analysis will allow us to address such questions in the many rare cell types that are critical for proper blood formation in vivo.

In light of our data, it is expected that low levels of MOF are, in turn, linked to acute myeloid leukemia (AML). HDAC inhibitors are currently used as an additional treatment for AML (*45*) but often trigger anemia as a side effect (*46*), which we anticipate could be explained by chromatin acetylation imbalance. Although we did not observe aging features in our mouse models, decreased levels of H4K16ac are associated with premature aging of HSCs, characterized by anemia and myeloid skewing (*47*, *48*). In addition, *Runx1* and *Gfi1b* haploinsufficiencies are related to blood syndromes in humans, which are characterized by platelet and erythroid defects and myelodysplastic syndrome (MDS) (*49*, *50*). Conversely, enhanced dosage of *Runx1* in trisomy 21 was shown to be associated with the development of MDS and poor AML prognosis (*51*), therefore arguing that the proper dosage of this intricate network is required for normal hematopoiesis.

Moreover, it is interesting that the erythropoietic phenotype, but not the myeloid skewing, appears to be regulated through the NSL complex, suggesting that MOF might exert different functions depending on its associated complex and in a cell-specific manner. In this regard, decreased levels of H4K16ac (*52*) and mutations in *Kansl1* (*53*) are recurrent observations in AML. However, the contributions of the MSL complex to these processes are still unknown. Therefore, we envision that the stratification of blood-related disorders between anemia and leukemia based on levels of *Mof* and its related complexes might be an important consideration in personalized medicine.

In addition, we presume that further studies on the contribution of each subunit of the two MOF-associated complexes might shed light on MOF activity during several developmental processes. For instance, we observed differences between *Kansl2* and *Kansl3* mice models during erythroid development, wherein the second one seems to affect mainly terminal erythroid differentiation. Although the underlying mechanism for this discrepancy requires additional investigation, one hypothesis is that MOF retains partial enzymatic activity in the absence of *Kansl3* that enables early erythroid differentiation but affects terminal erythroid differentiation. Another possibility for these differences could be different *Kansl3* isoform usages, which could compensate the effect at early stages but not later in erythroid development.

Moreover, our data revealed that *Mof* itself is regulated in a cell type-specific fashion via GFI1B binding to its promoter. In the future, it will be important to understand which cell type-specific TFs drive *Mof* expression in other organs. Cell type-specific dosage of *Mof* might help explain why tissues are differentially affected by the congenital Koolen–de Vries syndrome caused by whole-body heterozygous *KANSL1* mutation.

Together, we propose that precisely timed expression of *Mof* and the evoked chromatin accessibility changes are a central component to regulation of erythropoiesis. Our observations support the emerging model that controlled changes in mRNA levels and chromatin status are fundamental for directing HSCs toward the erythroid branch (*12*). Since *Mof* haploinsufficiency results in aberrant levels of chromatin accessibility, erythroid trajectory dysregulation, and anemia, we envision that studies carefully dissecting acetylation levels and identification of MOF/H4K16ac-specific activators will be insightful for further understanding of erythropoiesis in healthy and diseased states.

MATERIALS AND METHODS

Ethics statement

Animals were kept on a 14/10-hour light-dark cycle and provided with standard chow food and water ad libitum. Every mouse strain in this study was backcrossed with C57BL/6J mice. All animal procedures are in agreement with and approved by the responsible Animal Welfare Committees (Regierungspräsidium Freiburg, Karlsruhe/Germany, license).

Mouse strains

B6.Cg-Tg(Vav1-cre)A2Kio/J, $Kat8^{tm1Avo}$, and B6.Cg-Tg(CAG-cre/Esr1*)5Amc/J alleles were published elsewhere (*19*). *Vav1*-improved Cre is a recombinase variant with lower chance of silencing in

mammals (the Jackson laboratory, stock number #008610). *Kansl2* $^{fl/fl}$ mice were generated from mouse embryonic stem cells (mESCs) [tm1a(EUCOMM)Wtsi] obtained from the International Knockout Mouse Consortium (*54*). *Kansl3* $^{fl/fl}$ animals were obtained by targeting mESC using a vector obtained from J. Rientjes (Monash, Australia) (*55*). Eight- to 12-week-old mice from both sexes were randomly allocated to experimental groups. Ten-week-old male animals were used for ChIP-seq, RNA-seq, and scRNA-seq.

Cell culture

HPC7 (murine hematopoietic progenitor cell line CVCL_RB19) were cultured in Iscove's modified Dulbecco's medium (IMDM; Gibco, 12440061) with 10% fetal calf serum (FCS; Gibco), penicillin (100 U/ml), and streptomycin (100 μg/ml) (IMDM-10%) and supplemented with stem cell factor (SCF; 10 μg/μl; PeproTech, #250-03). K562 cells (human blood from chronic myelogenous leukemia CVCL_004) and Wehi-3 (*Mus musculus* monocytic leukemia, CVCL_3622) cells were maintained in IMDM-10%. HEK293T cells (CVCL_0063) were kept in Dulbecco's modified Eagle's medium (DMEM; Gibco) with 10% FCS, penicillin (100 U/ml), and streptomycin (100 μg/ml). Cells were cultured in a humidified incubator at 37°C and 5% CO_2. All experiments were performed at early passages (two to four). HPC7 and K562 cell lines were donated by E. Trompouki and Wehi-3 were provided by R. Kemler.

Cloning strategy and luciferase activity

The *Gata1* coding sequence was cloned into a pCAG vector with a puromycin resistance cassette. *Mof* or *Mof* ^{E350Q} was cloned into a TRE3G promoter. HSPCs were transfected with this vector by electroporation using the Nucleofector Technology (Lonza) with the P3 Primary Cell Kit (Lonza, #V4SP-3096) and Lonza's "CD34$^+$" default protocol, and HSCs were sorted using an Aria FACS Fusion II with a 100-μm nozzle and maintained in MethoCult (STEMCELL Technologies, #M3434) with IMDM-10% supplemented with puromycin (1 mg/ml; Gibco, #A1113802) upon transfection.

For the Luciferase reporter, 790-bp upstream of the *Mof* 5′ untranslated region (5′UTR) (*Mof* promoter) were cloned into the pGL4.23 [lic/minP] vector (Promega, #E8411). *Gfi1b* (mouse isoform 1) was cloned from a gBlock gene fragment (Integrated DNA Technologies) with PacI and XhoI sites into the mammalian expression vector pcDNATM5/FRT (Invitrogen, #V601020).

Luciferase activity was evaluated by cotransfection of *Mof* promoter–pGL4.23 [lic/minP] vector (wild-type or GFI1B binding site mutated; see Supplementary Data) and pRL-TK in a 100:1 ratio (100 to 1 ng each on a 96-well plate) with a 3:1 lipid [Lipofectamine 2000 (microliters)]–to–DNA (micrograms) ratio. After 6 hours, the media were changed for the respective culture media, and luciferase activity was scored 24 hours after transfection by the Dual-Luciferase Reporter Technology (Promega, #E1910). *Gfi1b*-overexpressing HPC7, K562, and HEK293T cells were generated by transfection with 10 μg of pcDNATM5/FRT-*Gfi1b* (full length) with a 3:1 lipid [Lipofectamine 2000 (microliters)]–to–DNA (micrograms) ratio. Six hours after transfection, cells were pelleted and the media were changed to the respective cell culture media. After 12 hours, the first transfection, pGL4.23 [lic/minP] vector, and pRL-TK vector were cotransfected, as described above. Primers and fragment information can be found in Supplementary Data. For KD experiments, BM-resident HSPCs and HPC7 were transfected with Silencer Select Pre-designed *siRNA*: *Gfi1b* ID#s66608, *Runx1*ID#s201126, or with a scrambled *siRNA* as control.

Fluorescence-activated cell sorting

For flow cytometry, mouse BM cells were isolated from pooled femora and tibia by flushing them with cold phosphate-buffered saline (PBS) containing 2% FCS. Lysis of erythrocytes was performed using ACK Lysing Buffer (Thermo Fisher Scientific). To deplete lineage-positive cells, we used the MojoSort Mouse Hematopoietic Progenitor Cell Isolation Kit (BioLegend, #48003). Total BM was incubated with biotin antibody cocktail, followed by incubation with magnetic streptavidin nanobeads from the kit. HSPCs were sorted into defined progenitor cells (see Supplementary Data) using an Aria FACS Fusion II cytometer (BD). List of antibody panel repertoires are provided in Supplementary Data.

Different collection approaches were applied depending on the experiment. For RT–qPCR/immunoblot and ChIP-seq, sorted cells were collected into ice-cold PBS. For scRNA-seq, cells were sorted into 384-well plates loaded with RNA lysis buffer (see the "Single-cell RNA-seq" section). For single-cell clonal assessment, HSCs were sorted into 96-well plates coated with either MethoCult (STEMCELL Technologies, #M3434) or IMDM-10% supplemented with a cytokine cocktail (see the "Single-cell liquid cultures of mouse BM progenitors" section). Sort purity was >90% in all cases. FACS analyses were conducted by FlowJo 10.v.

For histone immunostaining of primary cells, progenitor cells were fixed and permeabilized using the "Perm/Fix" solution from the eBioscience Foxp3/Transcription Factor Staining Buffer Set (Affymetrix, eBioscience, USA, #00-5523-00), as described in the manufacturer's protocol. Cells were labeled with the H4K16ac-Alexa Fluor 488 (A488) (1:500; Millipore, #07-329-AF488) or with total H4 antibody (1:1000; Active Motif, 39269). Before staining, total H4 antibody was labeled with A488 using the APEX Antibody Labeling Kit (Thermo Fisher Scientific, #A10468). After adding the primary antibody, samples were kept at 4°C for 1 hour. The cells were then washed with wash buffer (Affymetrix, eBioscience, USA, #00-5523-00) and washed twice in FACS buffer [PBS, 0.2% bovine serum albumin (BSA), and 0.1 mM EDTA]. Last, the cells were resuspended in FACS buffer (PBS, 0.2% BSA, and 0.1 mM EDTA) and measured using the Fortessa III flow cytometer (BD).

HPC7 cells were stained with H4K16ac, total H4, H4K8ac (1:500; Active Motif, 61103), H4K12ac (1:500; Active Motif, 39165), total H3 (1:1000; Active Motif, 39763), and H3K27ac (1:500; Active Motif, 39133) antibody. Cells were fixed and permeabilized following the same protocol as for primary cells. After adding the primary antibody, samples were kept at 4°C for 1 hour. The cells were then washed with wash buffer (Affymetrix, eBioscience, USA, #00-5523-00). The cells were resuspended in FACS buffer (PBS, 0.2% BSA, and 0.1 mM EDTA) and incubated with A488-labeled secondary antibody (1:5000) at room temperature for 1 hour. Cells were washed twice with FACS buffer before acquisition.

The following antibodies were used for flow cytometry: Lineage-negative cocktail fluorescein isothiocyanate (FITC; BioLegend, #13:3302; clone: 7A2,RB6-8C5,RA3-6B2,Ter119,M1/70), cKit-APC (BioLegend, #105812; clone: 2B8), Sca-1-BV421 (BioLegend, #108128; clone: D7), Ftl3-PE-CF594 (BD, #562537; clone: A2F10), CD34-Alexa Fluor 700 (BD, #560518; clone:RAM34), CD48-APC-Cy7 (BioLegend, #103431; clone: HM48-1), Slam1-BV510 (BioLegend, #115929; clone: TC15-12F12.2), CD41-PE (BioLegend, #133905; clone: MWReg30), CD105-PERCP-Cy5.5 (BioLegend, #120415; clone: MJ7/18), CD16/32-BV711 (BioLegend, #101337; clone: 93), CD44-APC (BD, #559250; clone: IM7), CD71-FITC (BD, #553266;

clone:C2F2), Ter119-PE (BD, #553673; clone: ter-119), CD11b-PERCP-Cy5.5 (BD, #550993, and BioLegend, #101227; clone: M1/70), F4/80-BV421 (BD, #565411; clone: T45-2342; and Biolegend, #123131; clone: BM8), Ly6G-PE (BD, #551461; clone: 1A8), Ly6C-Alexa700 (BioLegend, #128023; clone: HK1.4), CD45.1-APC (BioLegend, #110713; clone: A20), CD45.2 (BioLegend, #109808; clone: 104), CD45-FITC (BioLegend, #103107; clone: 30-F11), CD19-PE-Cy7 (BD, #561739; clone: 1D3), immunoglobulin D (IgD)–FITC (BD, #562022, and BioLegend, #405703; clone: 11-26c.2a), CD3-APC or PERCP-Cy5.5 (BioLegend, #100235 and #100217; clone: 17A2), and H4K16ac-A488 (Millipore, #07-329-AF488).

Western blots

For whole-cell extracts, cells were lysed in 50 mM NaCl, 1.0% IGEPAL CA-630, and 50 mM tris-HCl (pH 8.0). Buffer was supplemented with cOmplete (Roche, #4693132001) and PhosSTOP (Roche, #4906845001). Protein concentration was determined by Qubit protein assay reagent (Thermo Fisher Scientific, #Q33212). Samples were denatured at 95°C for 5 min in a Roti-Load reducing buffer (Carl Roth, #K929.1) before SDS–polyacrylamide gel electrophoresis (PAGE) using NuPAGE bis-tris gels and then transferred to 0.45 μM polyvinylidene difluoride membranes. Membranes were blocked for 30 min with 5% milk in PBS with 0.3% Tween (PBST). Membranes were washed twice with 0.3% PBST. The membrane was incubated with primary antibodies against MOF (1:1000; Bethyl, #A300-992A), actin (1:10000; Sigma-Aldrich, #A2066), H3 (1:5000; Active Motif, #39763), and H4K16ac (1:2000; Millipore, #07-329) diluted in PBS–5% BSA. The membrane was then incubated with PBS–5% BSA containing horseradish peroxidase–conjugated anti-mouse (GE Healthcare, #NA931-1ML) or anti-rabbit (GE Healthcare, #NA934) (1:10,000).

RNA extraction

After sorting, primary cells were pelleted at 5000*g* and lysed in TRIzol (Invitrogen, #15596026). To isolate RNA, chloroform was added followed by extraction and precipitation of the aqueous phase using 5 μg of ribonuclease (RNase)–free glycogen and 0.25 ml of isopropanol. The supernatant was discarded and the pellet was washed twice in 80% ethanol (EtOH) and lastly dissolved in 10 μl of H_2O.

Reverse transcription quantitative polymerase chain reaction (RT-qPCR)

RNA (200 ng) was reversely transcribed into complementary DNA using Superscript III RT (Thermo Fisher Scientific, #18080-093). RT-qPCR reactions were performed using SYBR Green Master Mix (Roche, #4309155); the mix contained 6.25 μl of SYBR, 0.75 μl of primers (forward and reverse, 300 nM), and 4.5 μl of H_2O.

Tissue histology

Histological tissue preparation was performed as in (*56*), with small variations. Briefly, tibiae were fixed with 4% paraformaldehyde (PFA) in PBS and incubated with PBS–0.5 M EDTA for 1 week at 4°C for demineralization. Every second day, EDTA solution was exchanged. The demineralized tissue was dehydrated by EtOH serial dilution and a xylene substitute Histolemon-Erba (Carlo-Erba, #8028-48-6), followed by paraffin embedding [formaldehyde-fixed paraffin embedding (FFPE)]. Paraffin blocks were sectioned with a RM2155 microtome (Leica) into 5- to 7-μm slices. For downstream analysis by immunofluorescence, FFPE sections were dewaxed and rehydrated stepwise using an EtOH gradient. Antigen retrieval was performed by incubation with 40 mM tris-HCl/64 mM EDTA (pH 9.0) at 95°C. Specimens were blocked for 1 hour with avidin/biotin block (BioLegend, #SIG-31126), followed by an additional 15-min incubation step in PBS–1% BSA. Following five times PBST washes, primary antibodies were added and incubated overnight at 4°C in blocking buffer. Unspecific binding was reduced by serial washes with PBS–0.05% Triton X-100. For downstream analysis by hematoxylin and eosin (H&E) and FFPE, sections were dewaxed and rehydrated stepwise using an EtOH gradient. The specimens were stained with hematoxylin solution [0.1% hematoxylin, 5% $KAl(SO_4)_2$, and 0.02% KIO_3]. Counterstaining was performed by incubating the slides in eosin solution (1% eosin). To analyze iron deposition, spleen sections were stained with Prussian blue (iron kit staining, Sigma-Aldrich, #HT20). Slides were mounted in the Permount Mounting Medium (Thermo Fisher Scientific) and analyzed by brightfield microscopy (Axio Imager Apotome 390, Zeiss).

Blood smear

Smear from freshly taken blood collected into EDTA-treated tubes was fixed with 100% methanol for 30 s, followed by H&E staining. Images were acquired with an Axio Imager Apotome 390 (Zeiss).

Immunofluorescence

Whole-mount ear staining was performed by isolating dermis, after tissue was kept overnight in 30% sucrose at 4°C. Then, tissue was washed three times with PBST (wash buffer) and blocked for 1 hour in PBS–2% BSA. After blocking, tissue was washed three times with wash buffer, and primary antibodies were added. Dermis were stained with collagen IV (1:500; Abcam, #ab19808) and Ly6G (1:100; BioLegend, #127607) antibodies.

Images were acquired by confocal microscopy (LSM780, Zeiss) and processed with ImageJ and Zen Black analysis software. Images show the orthogonal projection from multiple stacks.

Chromatin accessibility with visualization technology (ATAC-see)

For ATAC-see, we followed the published protocol in (*26*). Briefly, labeled oligonucleotides (Tn5ME-A-ATTO488 or A555, Tn5ME-B-ATTO488 or A555, and Tn5MErev) were denatured at 95°C and slowly annealed to room temperature in a thermocycler. The hyperactive Tn5 (Nextera Illumina, #1502824) was assembled according to (*26*).

For immunofluorescence, sorted MEPs ($Lin^-Kit^{high}Sca\text{-}1^-CD34^-FcγRII/III^-$) were fixed with 1% methanol-free formaldehyde (Sigma-Aldrich) for 10 min at room temperature and centrifuged onto the glass slide with cytospin, 1000 rpm for 5 min, 10,000 to 30,000 cells per slide. After fixation, the cells were permeabilized with lysis buffer [10 mM tris-HCl (pH 7.4), 10 mM NaCl, 3 mM $MgCl_2$, and 0.01% IGEPAL CA-630] for 10 min. Slides were washed with PBS and blocked with blocking solution (1% BSA, 0.1% Tween 20 in PBS, and PBST) for 1 hour at room temperature in constant movement on a shaker, followed by three PBST washes for 10 min each. Primary antibodies diluted in PBS (1:200, MOF, 1:300; Bethyl, #A300-992A) were incubated overnight at 4°C. PBST washes were followed by 40-min incubation with secondary antibody (1:10,000; goat anti-rabbit A594, Thermo Fisher Scientific, #A-11037). Last, the slides were incubated with transposase mixture solution (100 nM Tn5-ATTO-488N in a total volume of 50 μl of 1× TD buffer L) at 37°C in a humidified chamber for 30 min and washed with PBST three times

for 10 min each. Coverslips were mounted using Fluoromount-G with 4′,6-diamidino-2-phenylindole (DAPI; Thermo Fisher Scientific, #F4680). Slides were imaged with Zeiss Observer.Z1 with the CSU-X1 spinning disk head (Yokogawa) and the Axiocam camera (Zeiss), and images were analyzed with CellProfiler software.

For FACS analysis, BM was processed as described in the "Fluorescence-activated cell sorting" section and stained with the following fluorescence-labeled antibodies: anti-CD11b, anti-Ter119, anti-CD3, anti-B220, anti-NK1.1 and anti-Ly6G cocktail, anti-CD117, anti–Sca-1, anti-CD34, and anti-CD150 (see Supplementary Data) and incubated at 4°C for 40 min. After washes with ice-cold PBS, the cells were fixed with 1% methanol-free formaldehyde (Pierce, Thermo Fisher Scientific, #28906) for 10 min. Excess of formaldehyde was removed, and cells were permeabilized with lysis buffer [10 mM tris-HCl (pH 7.4), 10 mM NaCl, 3 mM $MgCl_2$, and 0.01% IGEPAL CA-630] for 10 min at room temperature. BM-enriched HSPC were then incubated with the transposed mixture or only with the oligo mixture (negative control) for 30 min at 37°C. After the transposase reaction, cells were centrifuged and resuspended in PBS and analyzed using a FACS-Fortessa II (BD Biosciences). FACS analyses were conducted by FlowJo 10.v.

Chromatin accessibility analysis from human hematopoietic cells (ATAC-seq)

Published dataset (*27*) was analyzed following the original paper documentation. The quantile-normalized fragments were plotted using the graphic package ggplot2.

Single-cell liquid cultures of mouse BM progenitors

Freshly collected mouse BM was labeled according to panels exemplified in Supplementary Data. Single HSCs ($LSK^+CD34^-Flt3^-CD150^+CD48^-$) were sorted into 96-well plates, using an Aria FACS Fusion II with a 100-μm nozzle. Cells were cultured for 10 days in IMDM-10% (see the "Cell culture" section), supplemented with SCF (50 ng/ml; recombinant murine SCF, PeproTech, #250-03), interleukin-3 (IL-3; 10 ng/ml; recombinant murine IL-3, PeproTech, #213-13), IL-6 (10 ng/ml; recombinant murine IL-6, PeproTech, #216-16), erythropoietin (EPO; 2 U/ml; BioLegend, #587602), IL-11 (50 ng/ml; recombinant murine IL-11, PeproTech, #220-11), IL-5 (10 ng/ml; recombinant murine IL-5, PeproTech, #215-15), thrombopoietin (TPO; 50 ng/ml; recombinant murine TPO, PeproTech, #315-14), IL-4 (10 ng/ml; recombinant murine IL-4, PeproTech, #214-14), granulocyte-macrophage colony-stimulating factor (GM-CSF; 15 ng/ml; recombinant murine GM-CSF, PeproTech, #315-03), and IL-7 (10 ng/ml; recombinant murine IL-7, PeproTech, #217-17). Growth factors/cytokines were refreshed at day 5. The clones/colonies in each well were labeled on day 10 with anti-CD117, anti-Ter119, anti-CD11b, and anti-IgD antibody. Clones were analyzed by flow cytometry using the BD Fortessa II (BD Biosciences) and processed with FlowJo 10.v.

Single-cell RNA-seq

Single-cell sorting

Age-matched C57BL/6J littermates were used in this analysis, and BM cells were isolated as described in the "Fluorescence-activated cell sorting" section. The following gating strategy was used to sort the populations present in the scRNA-seq analysis: megakaryocyte progenitor ($Lin^-Sca\text{-}1^-c\text{-}Kit^+CD150^+CD41^+$, 48 cells were sorted), GMP ($Lin^-Sca\text{-}1^-c\text{-}Kit^+CD41^-FcγRII/III^+$, 24 cells were sorted), pre-GMP ($Lin^-Sca\text{-}1^-c\text{-}Kit^+CD41^-FcγRII/III^-CD150^-CD105^-$, 24 cells were sorted), pre-MEP ($Lin^-Sca\text{-}1^-c\text{-}Kit^+CD41^-FcγRII/III^-CD150^+CD105^-$, 48 cells were sorted), pre–CFU-E ($Lin^-Sca\text{-}1^-c\text{-}Kit^+CD41^-FcγRII/III^-CD150^+CD105^+$, 48 cells were sorted) (*57*), MEP ($Lin^-Sca\text{-}1^-c\text{-}Kit^+CD150^+CD41^-FcγRII/III^-CD105^-$, 48 cells were sorted), pro-erythroid ($CD71^+Ter119^{low}$, 48 cells were sorted), myeloid progenitor ($c\text{-}Kit^{high}Sca\text{-}1^-$, 768 cells were sorted), early progenitor ($LSK^+Flt3^-CD34^-$, 576 cells were sorted), LT-HSC ($LSK^+CD150^+CD48^-$, 792 cells were sorted), LSK^{high} (864 cells were sorted), and LSK^{low} (576 cells were sorted). Only cells negative for Zombie Dye (BioLegend, #423212) were included to ensure the sort of living cells. Cells were sorted into 384-well plates loaded with RNA lysis buffer.

RNA and library preparation from single cells

For CEL-Seq2 single-cell library preparation (*58*), a nanoliter pipetting robot (mosquito HTS, TTP Labtech) was used to reduce the CEL-Seq2 protocol original volumes by fivefold. The 384-well plates (Corning, PCR-384-RGD-C) were prepared with 240 nl of lysis buffer in every well in these final ratios: 20 nl of 10 mM deoxynucleotide triphosphate (1:12), 40 nl of 1:100,000 ERCC spike mix 1 or 2 (2:12), 140 nl of water with 0.35% Triton X-100 (7:12), and 40 nl of 25 ng/μl of uniquely barcoded polydT primers with unique molecular identifier (UMI; 2:12). Hydrophobic encapsulation barrier (1.2 μl; Vapor-Lock, QIAGEN, 981611) was lastly added to every well.

Single-cell data analysis

For transcript quantification, we followed the analysis described in (*25*). Briefly, paired-end reads were aligned to the transcriptome by bwa (version 0.6.2-r126), with default parameters. The transcriptome contained all gene models based on the mouse ENCODE VM9 release downloaded from the University of California, Santa Cruz (UCSC) genome browser comprising 57,207 isoforms derived from 57,207 gene loci with 57,114 isoforms mapping to fully annotated chromosomes (1 to 19, X, Y, and mitochondrial). All isoforms of the same gene were merged to a single gene locus, and gene loci were merged to larger gene groups, if loci overlapped by >75%. This procedure resulted in 34,111 gene groups. The right mate of each read pair was mapped to all gene groups and to the set of 92 ERCC spike-ins in sense direction. Multimapper reads were discarded. The left mate contains the barcode information: The first six bases corresponded to the cell-specific barcode, followed by six bases representing the UMI. The remainder of the left read contains the polyT stretch and the adjacent gene sequence. For each cell barcode and gene locus, the number of UMIs was aggregated and, on the basis of binomial statistics (*25*), converted into transcript counts.

Rare cell-type identification

Bioinformatics analysis was conducted in R environment (for package version, see Supplementary Data, R settings). For identification of rare and abundant cell types, we applied the *RaceID3* algorithm (*25*). Sorted cells with high expression of *Kcnq1ot1* (>2% of all transcripts) were removed to ensure scRNA quality data, since this gene was previously identified marker of low-quality cells. Then, all reads mapping to ERCC spike-ins were discarded, and the data were then processed using *RaceID3*. Cells with less than 2000 transcripts were removed, and genes with more than three transcripts (*minexpr* input) in at least one cell (*minnumber* input) were kept. The cutoff for identification of outlier genes was set to five transcript of a gene in a cluster (*outminc* input), with a probability threshold for outlier calling at 10^{-4} (*probthr* input). To avoid clustering being dominated by nonspecific processes, such as technical batch effects or cell proliferation, we filtered out *Pcna*, *Mki67*, *Ptma*, *Hsp90ab1*, *Actb*, *Jun*,

Fos, *Gnas*, and *Hspa8* and correlated genes expression (*CGenes*) for cell-type inference. This procedure excludes ubiquitous similarities originating from cell proliferation (*Pcna*, *Mki67*, and *Ptma*) or stress-associated mechanisms (*Jun*, *Fos*, *Hspa8*, *Hspa90ab1*, and *Gnas*) (*25*, *59*). In addition, we used *Malat1* and *Igkc* as *FGenes*, which are also discarded before distance matrix computation. *FGenes* were removed because they were exceedingly highly expressed (hundreds of copies for cells) and thus dominated the clustering, obscuring the information of the entire transcriptome. There is no risk of overfitting since excluded genes were based on associated functions and transcriptome signatures (cell cycle and stress response), and the number of excluded genes is very small. Normalization was then performed by dividing transcript counts in each cell by the total number of transcripts in this cell, followed by multiplication with the median of the total number of transcripts across cells (median normalization). After the normalization a pseudocount of 0.1 was added to the expression data. Clustering of the processed data was performed using the *k*-medoids method [*clustexp()*], the maximum number of cluster for the computation was set to 30 (*clustnr* input), the number of bootstrapping runs to 50 (*bootnr* input), cluster distances were computed with metric = "pearson," that is, Pearson's correlation was applied, and on the basis of the saturation behavior of the within-cluster dispersion, a cluster number of 13 was chosen (*cln* input). Outliers of the initial *k*-medoids clusters were identified on the basis an internally computed background model of the expected gene expression variability. This assumes that transcript counts follow a negative binomial distribution defined by the mean and the variance of the expression of each gene per cluster [*findoutliers()*]. t-SNE maps for cluster visualization were generated using *comptsne()*. The differentially express genes were identified by *clustdiffgenes()* and $P < 0.01$. The *RaceID3* algorithm is available at GitHub (https://github.com/dgrun/RaceID3_StemID2).

Gene expression visualization

The normalized data (pseudocount) was used for heatmap and boxplot representations in this study. The *StemID2* algorithm is available at GitHub (https://github.com/dgrun/RaceID3_StemID2).

Prediction of stem identity

For inference of trajectories and stem status, the *StemID2* algorithm was used using the default settings. This algorithm infers links between clusters generated by *RaceID3*, which are more populated with intermediate single-cell transcriptomes than expected by chance. The number of links to other clusters and maximum transcriptome entropy were used to predict their stem status.

Fate-bias analysis and pseudotime inspection

Fate-bias probability was calculated using FateID (*25*). The *FateID* algorithm applies an iterative random forest classification to quantify fate bias in naïve progenitors using cells classified in *RaceID3* and *StemID2* as training set. The expression data and cluster partitions filtered by *RaceID3* were used as input data for the fate-bias analysis. The end points of the differentiation trajectories determined by *RaceID3* clusters, i.e., erythroid (e.g., $Gypa^{+}Gata1^{+}Klf1^{+}$), myeloid bias (e.g., $Itgam^{+}Mpo^{+}Elane^{+}$), and lymphoid ($Rag1^{+}$, $Ebf1^{+}$) were used as end points (*tar* input). The fate bias toward target clusters was calculated by *fatebias()* from the *FateID* algorithm using default settings. Afterward, dimensional reduction representations were computed on the basis of t-SNE maps using *compdr()*. The results from *fatebias()* and *compdr()* were used to generate the principal curve using *prcurve()*. The fraction threshold of random forest votes required for the inclusion of a given cell in the principal curve was set to 0.4.

Gene expression changes along the erythroid pseudotime

Cells with fate bias toward the erythroid lineage (264 wild-type cells and 215 $Mof^{+/-}$) were extracted from the principal curve. To eliminate lowly expressed genes on the trajectory, the RaceID-normalized transcript expression values were filtered with *filterset()*. The following parameters were applied: ≧0.5 normalized expression counts per gene (*minexpr* input), genes expressed at least in one cell (*minnumber* input), and a vector containing the cells names extracted from the principal curve (*n*, input).

The self-organizing map (SOM) of the pseudotemporal order was generated using *getsom()* with the following inputs: the filtered gene list, $n = 1000$ maximum nodes (default value) and $\alpha = 0.5$. The *getsom()* returns a data frame with smoothed and normalized expression profiles and *z*-score–transformed pseudotemporal expression profiles. The SOM was further processed by *prsom()*. The *prsom()* groups the nodes generated by *getsom()* into larger nodes, in which genes having higher than 0.85 correlation of the SOM *z*-scores are aggregated into the same node. The minimal number of genes per node was set to 3. The processed SOM was plotted with *plotheatmap()*.

Differential gene expression analysis

Differentially expressed genes between two cell subgroups were identified similar to a previously published method. First, we inferred a negative binomial distribution reflecting the gene expression variability within each subgroup. We used a background model for the expected transcript count variability computed by the same strategy as in *RaceID3* (*25*). Using these distributions, a *P* value was computed for the observed difference in transcript counts between the two subgroups as described in DESeq (*60*). These *P* values were corrected for multiple testing by the Benjamini-Hochberg method. We used *diffexpnb()* to perform this calculation (*25*).

Differential gene expression analysis along the erythroid pseudotime

Determination of differentially expressed genes along the erythroid trajectory was conducted using Monocle (*61*). *RaceID3* normalized counts were used as expression input matrix for Monocle. The estimate size and dispersion were evaluated by *estimateSizeFactors()* and *estimateDispersions()* functions. The genes extracted from the *filterset()* (see the "Fate-bias analysis and pseudotime inspection" section) were subsetted from the expression matrix and defined as "*marked_genes*." Last, *differentialGeneTest()* was applied for the marked genes and genotype comparison.

Bulk CFUs assay

One hundred HSCs ($LSK^{+}CD34^{-}Flt3^{-}CD150^{+}CD48^{-}$) from $CreERT2^{+/tg}Mof^{fl/fl}$ and wild-type 8- to 10-week-old mice were sorted by Aria FACS Fusion II (BD) and cultured in MethoCult (STEMCELL Technologies, #M3434) in technical duplicates. To deplete *Mof* in vitro, 200 nM 4-OHT (Sigma-Aldrich, #H7904) was added to the culture media. After 10 days in culture, CFU capacity was evaluated by visual inspection. Then, we conducted serial plating in which 10,000 cells originated from the primary CFU assay were transferred to a new 24-well plate and cultured in MethoCult (STEMCELL Technologies, #03434). After 10 days, colonies were quantified. For lineage-committed cells, 100 sorted MEPs ($Lin^{-}Sca\text{-}1^{-}cKit^{high}CD34^{-}IL7R^{-}FcR\gamma II/III^{-}$) from $CreERT2^{+/tg}Mof^{fl/fl}$ or $CreERT2^{+/tg}Mof^{+/l+}$ mice were plated into MethoCult (STEMCELL Technologies, #M3434) in technical duplicates and treated with 4-OHT. After 10 days in culture, colony unit formation (CFU) capacity was evaluated by visual inspection.

BM adoptive transplantations

Two hundred HSCs ($LSK^+CD34^-Flt3^-CD150^+CD48^-$) from *Mof*$^{+/-}$ or wild-type 8-week-old mice (CD45.2) were FACS sorted and transplanted into irradiated (1× 350 rad) mice (CD45.1). Contribution of CD45.2-donor cells was monitored weekly in peripheral blood over the course of 8 weeks after transplantation in primary recipients. Chimerism and blood cell populations were investigated by flow cytometry using the following antibodies: anti–CD451/2-FITC, anti–CD45.1-APC, anti–CD45.2-PE, CD11b-PERCP-Cy5.5, Ter119-PE, CD71-FITC, and CD44-APC (see Supplementary Data). Two months after transplantation, animals were euthanized following the ethics end point. BM-resident cells and spleen were analyzed by flow cytometry. Total CD45.2 BM was sorted and 1×10^6 CD45.2 cells were transplanted into lethally irradiated mice (1× 450 and 1× 500 rad). Animals were monitored weekly for contribution of CD45.2-donor cells in peripheral blood.

ChIP sequencing

Femora, pelvis, and tibiae were isolated from five male mice (10 weeks old) from either wild-type or *Mof*$^{+/-}$. BM cells were collected by flushing and crushing with a mortar and pestle in cold PBS containing 2% FCS. Lysis of erythrocytes was performed using the ACK Lysing Buffer (Thermo Fisher Scientific). Cells were enriched with a Mojo HSPC enrichment kit according to the manufacturer's guidelines. Resulting cells were stained with anti-CD11b, anti-Ter119, anti-CD3, anti-B220, anti-NK1.1, and anti-Ly6G cocktail (clone 145-2C11, RB6-8C5, M1/70, RA3-6B2, Ter119, and PE), anti-CD117, anti–Sca-1, anti-CD34, anti-CD16/32, Ftl3, and anti–IL-7R for 40 min at 4°C. $LSK^+CD34^-Flt3^-$ (HSC profile, 15,000 cells) and Lin^-Kit^{high}Sca-1-CD34-FcγR-II/III (MEP profile, 30,000 cells) were sorted in the FACS Fusion II cell sorter (BD Biosciences). Before FACS sorting, cells were fixed with 1% PFA for 10 min at room temperature and quenched with 0.125 M glycine.

For ChIP-seq in the HPC7 cell line, 500,000 living cells (0.4% trypan-blue counter staining) were counted into a chamber and fixed either with 0.1% PFA (H4K16ac profile) or 1% PFA (MOF profile) for 10 min at room temperature and quenched with 0.125 M glycine for 4 min. Cells were pelleted by centrifugation at 1000 rpm for 10 min, resuspended in ChIP lysis buffer (5 mM Pipes, 85 mM KCl, 0.5% and IGEPAL CA-630) and 10-min incubation on ice, followed by a 10-min spin at 1000 rpm. The nuclear pellet was then resuspended in nuclei lysis buffer [50 mM tris-HCl (pH 8.0), 10 mM EDTA, 0.5% SDS, cOmplete, and phosphoStop] and incubated for 15 min on ice. Lysates were transferred into 0.5-ml Bioruptor Microtubes (Diagenode C30010015, Liège, Belgium), and chromatin was sheared using an NGS Bioruptor Sonicator (Diagenode) at default intensity and cycles of 30″ ON/30″ OFF. Sonication time was optimized for the different cell types: 10 min for MEPs and 15 min for HPC7/HSCs.

Sonicated chromatin was diluted 1:5 with sonication equilibration buffer (10 mM tris-HCl, 140 mM NaCl, 0.1% sodium deoxycholate, 1% Triton X-100, 1 mM EDTA, 1× cOmplete, and phosphoStop) (*62*). Chromatin was precleared by adding 50 μl of protein A/agarose beads. The samples were kept under constant rotation for 2 hours at 4°C, followed by a centrifugation step of 3000 rpm for 5 min at 4°C. Supernatants were transferred to fresh 1.5-ml microcentrifuge tubes and 2 μg of anti-MOF (Bethyl, #A300-992A), or anti-H4K16ac antibody (Millipore, #07-329) was added to the diluted and cleared chromatin extracts and incubated for 20 hours at 4°C. After that, 20 μl of protein A-dynabeads (Thermo Fisher Scientific), previously blocked with 1% BSA, was added to the samples and kept under constant rotation for 2 hours at 4°C.

The beads were washed twice with high-salt buffer [50 mM Hepes (pH 7.9), 500 mM NaCl, 1 mM EDTA, 0.1% SDS, 1% Triton X-100, and 0.1% deoxycholate], once with LiCl buffer (10 mM Tris-EDTA (TE), 250 mM LiCl, 0.5% IGEPAL CA-630, and 0.5% deoxycholate) and once with TE [10 mM tris-HCl (pH 8.0) and 1 mM EDTA]. Samples were resuspended in 10 μl of elution buffer [QIAGEN; 10 mM tris-HCl (pH 8.5)] and treated with 2 μl of RNase A (QIAGEN, #19101) for 30 min at 37°C, followed by 2.5 μl of proteinase K (Thermo Fisher Scientific, #AM2546) treatment for 2 hours at 55°C. Afterward, the temperature was increased to 65°C for 8 hours to revert formaldehyde cross-linking. Chromatin was quantified by Qubit Fluorometric Quantitation DNA high-sensitivity kit (Thermo Fisher Scientific, #Q32854), and the quality was evaluated by fragment analyzer.

Libraries were prepared with the NEBNext Ultra DNA Library Prep Kit for Illumina according to the manufacturer's instructions (New England Biolabs, #E7645) with 15 amplification cycles. ChIP libraries were sequenced using an Illumina HiSeq 3000 sequencing 2 × 75 base paired-end reads.

Processing of ChIP-seq datasets

ChIP-seq datasets were mapped to GRCm38/Mm10 with Bowtie2 (Galaxy version 2.3.0.1) (*63*) with default Bowtie2 parameters. Peaks were called with MACS2 (Galaxy version 2.1.1.20160309.3) (*63*, *64*), and bandwidth was set to 200, lower mfold bound to 5, upper mfold bound to 500, and the *P* value to 0.01 (for HSC and MEP peak list, see Supplementary Data). Peaks from each cell type were merged using cat, BEDtools sort (Galaxy version 2.27.0.0) and piped to BEDtools merge with a distance of 1000 bp. For TRAP, we used TRAP (*33*). Transfac_2010.1 vertebrates was selected as the matrix file and "mouse promoter" as background model. For significant enrichment analysis, motifs were subjected into the Benjamini-Hochberg multiple test corrections. Motif enrichment was based on adjusted $P < 0.05$. Coverage files were generated with deepTools2 (*65*) *BamCompare* (Galaxy version 2.5.1.0.0) with a bin size of 50 bp. The data were normalized as log2 fold change over input. For plotting genome browser snapshots, 1× sequencing depth–normalized coverage files were generated using deepTools2 *BamCoverage* and visualized with Integrative Genomics Viewer (IGV). Heatmaps were generated using deepTools2 *compute matrix* and *plotHeatmap* function. Enrichment scores were calculated using deepTools2 multiBigwigSummary. The Bioconductor packages ChIP-Seeker (*66*) was used to retrieve the nearest genes around the peak, annotate the genomic region of the peak, and peak features annotation. GO term analysis was performed using the Metascape platform and the ShinyGo v0.50 database (http://bioinformatics.sdstate.edu/go/).

ATAC-seq and MOF–ChIP-seq integration

Encode-published available ATAC-seq peaks from HSCs and MEPs were downloaded into galaxy environment and subjected to the BEDtools *FisherBED*, together with the list of MACS2 MOF peaks from HSCs, MEPs, or random peaks and the list of total genomic regions from the UCSC. Significant enrichment was determined by the Fisher's exact test *P* value.

scATAC-seq pseudotime and MOF–ChIP-seq integration

Published available scATAC-seq dataset (*38*) was subjected into the STREAM (*39*) pipeline. The list of *k*-mers and their *P* value for each

leaf branch were extracted as a data frame named *leaf_genes_all*. For pseudotime analysis, we follow the original pipeline and plotted as subway map. HSCs ("S2") were used as roots. In parallel, MOF peaks associated to TSS were selected and *k*-mers, 7-bp upstream, and downstream flanking regions (−b), extracted using the BEDtools function *FlankBed* (v.2.27.0.0), and then subjected into the *GetFastaBed* tool. Then, MOF-associated *k*-mers were selected from the *leaf_genes_all* list and plotted into the subway map or as bar plots. Jupyter Notebook used for STREAM can be found at https://github.com/pinellolab/stream.

GFI1B, RUNX1, and MOF ChIP-qPCR

HPC7 cells were collected by centrifugation and cross-linked with 1% formaldehyde for 10 min at room temperature. The chromatin fragmentation and immunoprecipitation was performed as described in the "ChIP sequencing" section with 2 μg of anti-GFI1B (Thermo Fisher Scientific, #PA5-49692), 2 μg of anti-RUNX1 (Thermo Fisher Scientific,#OSR00271W), or 2 μg of anti-MOF antibody (Bethyl, #A300-992A).

Dihydroethidium staining for ROS detection

Freshly isolated and enriched lineage-negative cells (HSPCs) were incubated for 15 min with 5 μM dihydroethidium (Thermo Fisher Scientific, #D11347) in PBS. ROS measurements were conducted in the BD Fortessa II cytometer and analyzed in FlowJo 10.v.

Cell cycle analysis by FACS

BM cells were retrieved, as described in the "Fluorescence-activated cell sorting" section, and labeled with anti-lineage, anti–c-Kit, and anti–Sca-1 antibody (for fluorochromes and dilution, see Supplementary Data). Cells were then fixed in suspension with 70% ultrapure cold EtOH and kept at 4°C for 20 min. Cells were then washed with FACS buffer (PBS, 0.2% BSA, and 0.1 mM EDTA) and propidium iodide (1 μM) added for nuclei staining. Cells were then acquired in Fortessa II cytometer (BD Biosciences). Downstream analysis was performed in FlowJo (10.v) applying the "Cell cycle" plug-in with the Dean-Jett-Fox pragmatic model.

Quantification and statistical analysis

Data are presented as means ± SEM or SD as indicated in the figure legends. Box plots were generated using the R function *boxplot()* and displayed the median, maximum, minimum, and the interquartile range for each experiment. A minimum of three biological replicates was performed, and the exact replicate numbers *n* are mentioned in the respective figure legends. All statistical analyses were performed using Prism 6 software (GraphPad) or R. Details of statistical tests and *P* values are reported in the figure legends and/or figures.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/6/21/eaaz4815/DC1

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES

1. E. Dzierzak, S. Philipsen, Erythropoiesis: Development and differentiation. *Cold Spring Harb. Perspect. Med.* **3**, a011601 (2013).
2. H. Shi, S. Yamamoto, M. Sheng, J. Bai, P. Zhang, R. Chen, S. Chen, L. Shi, O. Abdel-Wahab, M. Xu, Y. Zhou, F.-C. Yang, ASXL1 plays an important role in erythropoiesis. *Sci. Rep.* **6**, 28789 (2016).
3. J.-B. Micol, A. Pastore, D. Inoue, N. Duployez, E. Kim, S. C.-W. Lee, B. H. Durham, Y. R. Chung, H. Cho, X. J. Zhang, A. Yoshimi, A. Krivtsov, R. Koche, E. Solary, A. Sinha, C. Preudhomme, O. Abdel-Wahab, ASXL2 is essential for haematopoiesis and acts as a haploinsufficient tumour suppressor in leukemia. *Nat. Commun.* **8**, 15429 (2017).
4. J. Palis, Primitive and definitive erythropoiesis in mammals. *Front. Physiol.* **5**, 3 (2014).
5. E. Dzierzak, N. A. Speck, Of lineage and legacy: The development of mammalian hematopoietic stem cells. *Nat. Immunol.* **9**, 129–136 (2008).
6. A. E. Rodriguez-Fraticelli, S. L. Wolock, C. S. Weinreb, R. Panero, S. H. Patel, M. Jankovic, J. Sun, R. A. Calogero, A. M. Klein, F. D. Camargo, Clonal analysis of lineage fate in native haematopoiesis. *Nature* **553**, 212–216 (2018).
7. P. Wong, S. M. Hattangadi, A. W. Cheng, G. M. Frampton, R. A. Young, H. F. Lodish, Gene induction and repression during terminal erythropoiesis are mediated by distinct epigenetic changes. *Blood* **118**, e128–e138 (2011).
8. C. R. Vakoc, S. A. Mandat, B. A. Olenchock, G. A. Blobel, Histone H3 lysine 9 methylation and HP1gamma are associated with transcription elongation through mammalian chromatin. *Mol. Cell* **19**, 381–391 (2005).
9. G. Fromm, C. de Vries, R. Byron, J. Fields, S. Fiering, M. Groudine, M. A. Bender, J. Palis, M. Bulger, Histone hyperacetylation within the β-globin locus is context-dependent and precedes high-level gene expression. *Blood* **114**, 3479–3488 (2009).
10. F. Paul, Y. Arkin, A. Giladi, D. A. Jaitin, E. Kenigsberg, H. Keren-Shaul, D. Winter, D. Lara-Astiaso, M. Gury, A. Weiner, E. David, N. Cohen, F. K. B. Lauridsen, S. Haas, A. Schlitzer, A. Mildner, F. Ginhoux, S. Jung, A. Trumpp, B. T. Porse, A. Tanay, I. Amit, Transcriptional heterogeneity and lineage commitment in myeloid progenitors. *Cell* **163**, 1663–1677 (2015).
11. N. K. Wilson, D. G. Kent, F. Buettner, M. Shehata, I. C. Macaulay, F. J. Calero-Nieto, M. Sánchez Castillo, C. A. Oedekoven, E. Diamanti, R. Schulte, C. P. Ponting, T. Voet, C. Caldas, J. Stingl, A. R. Green, F. J. Theis, B. Göttgens, Combined single-cell functional and gene expression analysis resolves heterogeneity within stem cell populations. *Cell Stem Cell* **16**, 712–724 (2015).
12. B. K. Tusi, S. L. Wolock, C. Weinreb, Y. Hwang, D. Hidalgo, R. Zilionis, A. Waisman, J. R. Huh, A. M. Klein, M. Socolovsky, Population snapshots predict early haematopoietic and erythroid hierarchies. *Nature* **555**, 54–60 (2018).
13. X. Yu, C. Wu, D. Bhavanasi, H. Wang, B. D. Gregory, J. Huang, Chromatin dynamics during the differentiation of long-term hematopoietic stem cells to multipotent progenitors. *Blood Adv.* **1**, 887–898 (2017).
14. A. Grigoryan, N. Guidi, K. Senger, T. Liehr, K. Soller, G. Marka, A. Vollmer, Y. Markaki, H. Leonhardt, C. Buske, D. B. Lipka, C. Plass, Y. Zheng, M. A. Mulaw, H. Geiger, M. C. Florian, LaminA/C regulates epigenetic and chromatin architecture changes upon aging of hematopoietic stem cells. *Genome Biol.* **19**, 189 (2018).
15. M. Shogren-Knaak, H. Ishii, J.-M. Sun, M. J. Pazin, J. R. Davie, C. L. Peterson, Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* **311**, 844–847 (2006).
16. K. K. Lee, J. L. Workman, Histone acetyltransferase complexes: One size doesn't fit all. *Nat. Rev. Mol. Cell Biol.* **8**, 284–295 (2007).
17. E. Seto, M. Yoshida, Erasers of histone acetylation: The histone deacetylase enzymes. *Cold Spring Harb. Perspect. Biol.* **6**, a018713 (2014).
18. J. Seita, D. Sahoo, D. J. Rossi, D. Bhattacharya, T. Serwold, M. A. Inlay, L. I. R. Ehrlich, J. W. Fathman, D. L. Dill, I. L. Weissman, Gene expression commons: An open platform for absolute gene expression profiling. *PLoS One* **7**, e40321 (2012).
19. T. Thomas, M. P. Dixon, A. J. Kueh, A. K. Voss, Mof (MYST1 or KAT8) is essential for progression of embryonic development past the blastocyst stage and required for normal chromatin architecture. *Mol. Cell. Biol.* **28**, 5093–5105 (2008).
20. C. Joseph, J. M. Quach, C. R. Walkley, S. W. Lane, C. Lo Celso, L. E. Purton, Deciphering hematopoietic stem cells in their niches: A critical appraisal of genetic models, lineage tracing, and imaging strategies. *Cell Stem Cell* **13**, 520–533 (2013).
21. M. F. Basilicata, A.-L. Bruel, G. Semplicio, C. I. K. Valsecchi, T. Aktaş, Y. Duffourd, T. Rumpf, J. Morton, I. Bache, W. G. Szymanski, C. Gilissen, O. Vanakker, K. Õunap, G. Mittler, I. van der Burgt, S. El Chehadeh, M. T. Cho, R. Pfundt, T. Y. Tan, M. Kirchhoff, B. Menten, S. Vergult, K. Lindstrom, A. Reis, D. S. Johnson, A. Fryer, V. McKay; DDD Study, R. B. Fisher, C. Thauvin-Robinet, D. Francis, T. Roscioli, S. Pajusalu, K. Radtke, J. Ganesh, H. G. Brunner, M. Wilson, L. Faivre, V. M. Kalscheuer, J. Thevenon, A. Akhtar, De novo mutations in MSL3 cause an X-linked syndrome marked by impaired histone H4 lysine 16 acetylation. *Nat. Genet.* **50**, 1442–1451 (2018).
22. C. I. K. Valsecchi, M. F. Basilicata, G. Semplicio, P. Georgiev, N. M. Gutierrez, A. Akhtar, Facultative dosage compensation of developmental genes on autosomes in Drosophila and mouse embryonic stem cells. *Nat. Commun.* **9**, 3626 (2018).
23. B. N. Sheikh, A. Akhtar, The many lives of KATs — Detectors, integrators and modulators of the cellular environment. *Nat. Rev. Genet.* **20**, 7–23 (2019).
24. S.-I. Tsukumo, K. Yasutomo, Notch governing mature T cell differentiation. *J. Immunol.* **173**, 7109–7113 (2004).
25. J. S. Herman, Sagar, D. Grün, FateID infers cell fate bias in multipotent progenitors from single-cell RNA-seq data. *Nat. Methods.* **15**, 379–386 (2018).

26. X. Chen, Y. Shen, W. Draper, J. D. Buenrostro, U. Litzenburger, S. W. Cho, A. T. Satpathy, A. C. Carter, R. P. Ghosh, A. East-Seletsky, J. A. Doudna, W. J. Greenleaf, J. T. Liphardt, H. Y. Chang, ATAC-see reveals the accessible genome by transposase-mediated imaging and sequencing. *Nat. Methods* **13**, 1013–1020 (2016).
27. M. R. Corces, J. D. Buenrostro, B. Wu, P. G. Greenside, S. M. Chan, J. L. Koenig, M. P. Snyder, J. K. Pritchard, A. Kundaje, W. J. Greenleaf, R. Majeti, H. Y. Chang, Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. *Nat. Genet.* **48**, 1193–1203 (2016).
28. D. G. Valerio, H. Xu, M. E. Eisold, C. M. Woolthuis, T. K. Pandita, S. A. Armstrong, Histone acetyltransferase activity of MOF is required for adult but not early fetal hematopoiesis in mice. *Blood* **129**, 48–59 (2017).
29. T. Chelmicki, F. Dündar, M. J. Turley, T. Khanam, T. Aktas, F. Ramírez, A.-V. Gendrel, P. R. Wright, P. Videm, R. Backofen, E. Heard, T. Manke, A. Akhtar, MOF-associated complexes ensure stem cell identity and Xist repression. *eLife* **3**, e02024 (2014).
30. G. C. A. Taylor, R. Eskeland, B. Hekimoglu-Balkan, M. M. Pradeepa, W. A. Bickmore, H4K16 acetylation marks active genes and enhancers of embryonic stem cells, but does not alter chromatin compaction. *Genome Res.* **23**, 2053–2065 (2013).
31. B. Aranda-Orgilles, R. Saldaña-Meyer, E. Wang, E. Trompouki, A. Fassl, S. Lau, J. Mullenders, P. P. Rocha, R. Raviram, M. Guillamot, M. Sánchez-Díaz, K. Wang, C. Kayembe, N. Zhang, L. Amoasii, A. Choudhuri, J. A. Skok, M. Schober, D. Reinberg, P. Sicinski, H. Schrewe, A. Tsirigos, L. I. Zon, I. Aifantis, MED12 regulates HSC-specific enhancers independently of mediator kinase activity to control hematopoiesis. *Cell Stem Cell* **19**, 784–799 (2016).
32. S. M. Chambers, N. C. Boles, K.-Y. K. Lin, M. P. Tierney, T. V. Bowman, S. B. Bradfute, A. J. Chen, A. A. Merchant, O. Sirin, D. C. Weksberg, M. G. Merchant, C. J. Fisk, C. A. Shaw, M. A. Goodell, Hematopoietic fingerprints: An expression database of stem cells and their progeny. *Cell Stem Cell* **1**, 578–591 (2007).
33. M. Thomas-Chollier, A. Hufton, M. Heinig, S. O'Keeffe, N. E. Masri, H. G. Roider, T. Manke, M. Vingron, Transcription factor binding predictions using TRAP for the analysis of ChIP-seq data and regulatory SNPs. *Nat. Protoc.* **6**, 1860–1869 (2011).
34. T. E. North, T. Stacy, C. J. Matheny, N. A. Speck, M. F. T. R. de Bruijn, Runx1 is expressed in adult mouse hematopoietic stem cells and differentiating myeloid and lymphoid cells, but not in maturing erythroid cells. *Stem Cells* **22**, 158–168 (2004).
35. Y. Fujiwara, C. P. Browne, K. Cunniff, S. C. Goff, S. H. Orkin, Arrested development of embryonic red cell precursors in mouse embryos lacking transcription factor GATA-1. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 12355–12358 (1996).
36. M. L. Ratliff, T. D. Templeton, J. M. Ward, C. F. Webb, The bright side of hematopoiesis: Regulatory roles of ARID3a/bright in human and mouse hematopoiesis. *Front. Immunol.* **5**, 113 (2014).
37. ENCODE Project Consortium, An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).
38. J. D. Buenrostro, M. R. Corces, C. A. Lareau, B. Wu, A. N. Schep, M. J. Aryee, R. Majeti, H. Y. Chang, W. J. Greenleaf, Integrated single-cell analysis maps the continuous regulatory landscape of human hematopoietic differentiation. *Cell* **173**, 1535–1548.e16 (2018).
39. H. Chen, L. Albergante, J. Y. Hsu, C. A. Lareau, G. L. Bosco, J. Guan, S. Zhou, A. N. Gorban, D. E. Bauer, M. J. Aryee, D. M. Langenau, A. Zinovyev, J. D. Buenrostro, G.-C. Yuan, L. Pinello, Single-cell trajectories reconstruction, exploration and mapping of omics data with STREAM. *Nat. Commun.* **10**, 1903 (2019).
40. N. K. Wilson, S. D. Foster, X. Wang, K. Knezevic, J. Schütte, P. Kaimakis, P. M. Chilarska, S. Kinston, W. H. Ouwehand, E. Dzierzak, J. E. Pimanda, M. F. T. R. de Bruijn, B. Göttgens, Combinatorial transcriptional control in blood stem/progenitor cells: Genome-wide analysis of ten major transcriptional regulators. *Cell Stem Cell* **7**, 532–544 (2010).
41. R. Thambyrajah, R. Patel, M. Mazan, M. Lie-A-Ling, A. Lilly, A. Eliades, S. Menegatti, E. Garcia-Alegria, M. Florkowska, K. Batta, V. Kouskoff, G. Lacaud, New insights into the regulation by RUNX1 and GFI1(s) proteins of the endothelial to hematopoietic transition generating primordial hematopoietic cells. *Cell Cycle* **15**, 2108–2114 (2016).
42. J. Füllgrabe, M. A. Lynch-Day, N. Heldring, W. Li, R. B. Struijk, Q. Ma, O. Hermanson, M. G. Rosenfeld, D. J. Klionsky, B. Joseph, The histone H4 lysine 16 acetyltransferase hMOF regulates the outcome of autophagy. *Nature* **500**, 468–471 (2013).
43. D. Saidi, M. Cheray, A. M. Osman, V. Stratoulias, O. R. Lindberg, X. Shen, K. Blomgren, B. Joseph, Glioma-induced SIRT1-dependent activation of hMOF histone H4 lysine 16 acetyltransferase in microglia promotes a tumor supporting phenotype. *Oncoimmunology* **7**, e1382790 (2018).
44. J. Xu, Z. Shao, D. Li, H. Xie, W. Kim, J. Huang, J. E. Taylor, L. Pinello, K. Glass, J. D. Jaffe, G.-C. Yuan, S. H. Orkin, Developmental control of polycomb subunit composition by GATA factors mediates a switch to non-canonical functions. *Mol. Cell* **57**, 304–316 (2015).
45. K. Wahaib, A. E. Beggs, H. Campbell, L. Kodali, P. D. Ford, Panobinostat: A histone deacetylase inhibitor for the treatment of relapsed or refractory multiple myeloma. *Am. J. Health Syst. Pharm.* **73**, 441–450 (2016).
46. C. V. Andreu-Vieyra, J. R. Berenson, The potential of panobinostat as a treatment option in patients with relapsed and refractory multiple myeloma. *Ther. Adv. Hematol.* **5**, 197–210 (2014).
47. M. C. Florian, K. J. Nattamai, K. Dörr, G. Marka, B. Uberle, V. Vas, C. Eckl, I. Andrä, M. Schiemann, R. A. J. Oostendorp, K. Scharffetter-Kochanek, H. A. Kestler, Y. Zheng, H. Geiger, A canonical to non-canonical Wnt signalling switch in haematopoietic stem-cell ageing. *Nature* **503**, 392–396 (2013).
48. M. C. Florian, K. Dörr, A. Niebel, D. Daria, H. Schrezenmeier, M. Rojewski, M.-D. Filippi, A. Hasenberg, M. Gunzer, K. Scharffetter-Kochanek, Y. Zheng, H. Geiger, Cdc42 activity regulates hematopoietic stem cell aging and rejuvenation. *Cell Stem Cell* **10**, 520–530 (2012).
49. E. Anguita, F. J. Candel, A. Chaparro, J. J. Roldán-Etcheverry, Transcription factor GFI1B in health and disease. *Front. Oncol.* **7**, 54 (2017).
50. R. Sood, Y. Kamikubo, P. Liu, Role of RUNX1 in hematological malignancies. *Blood* **129**, 2070–2082 (2017).
51. A. Tefferi, D. Idossa, T. L. Lasho, M. Mudireddy, C. Finke, S. Shah, M. Nicolosi, M. M. Patnaik, A. Pardanani, N. Gangat, C. A. Hanson, R. P. Ketterling, Mutations and karyotype in myelodysplastic syndromes: TP53 clusters with monosomal karyotype, *RUNX1* with trisomy 21, and *SF3B1* with inv(3)(q21q26.2) and del(11q). *Blood Cancer J.* **7**, 658 (2017).
52. M. F. Fraga, E. Ballestar, A. Villar-Garea, M. Boix-Chornet, J. Espada, G. Schotta, T. Bonaldi, C. Haydon, S. Ropero, K. Petrie, N. G. Iyer, A. Pérez-Rosado, E. Calvo, J. A. Lopez, A. Cano, M. J. Calasanz, D. Colomer, M. A. Piris, N. Ahn, A. Imhof, C. Caldas, T. Jenuwein, M. Esteller, Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nat. Genet.* **37**, 391–400 (2005).
53. E. Jaskula, A. Lange, M. Sedzimirska, J. Chrustowicz, A. Tarnowska, J. Lange, Microdeletion in chromosome 17 cytoband q21.31 identified as located within the *KANSL1* gene prevailed in *Flt3-ITD* negative normal karyotype AML patients as opposed to their AML counterparts and influenced the phenotype of the disease. *Blood* **130**, 1420–1420 (2017).
54. R. H. Friedel, C. Seisenberger, C. Kaloff, W. Wurst, EUCOMM—The European conditional mouse mutagenesis program. *Brief. Funct. Genomic. Proteomic.* **6**, 180–185 (2007).
55. A. Karoutas, W. Szymanski, T. Rausch, S. Guhathakurta, E. A. Rog-Zielinska, R. Peyronnet, J. Seyfferth, H.-R. Chen, R. de Leeuw, B. Herquel, H. Kimura, G. Mittler, P. Kohl, O. Medalia, J. O. Korbel, A. Akhtar, The NSL complex maintains nuclear architecture stability via lamin A/C acetylation. *Nat. Cell Biol.* **21**, 1248–1260 (2019).
56. M. F. Basilicata, M. Frank, D. Solter, T. Brabletz, M. P. Stemmler, Inappropriate cadherin switching in the mouse epiblast compromises proper signaling between the epiblast and the extraembryonic ectoderm during gastrulation. *Sci. Rep.* **6**, 26562 (2016).
57. R. Drissen, N. Buza-Vidas, P. Woll, S. Thongjuea, A. Gambardella, A. Giustacchini, E. Mancini, A. Zriwil, M. Lutteropp, A. Grover, A. Mead, E. Sitnicka, S. E. W. Jacobsen, C. Nerlov, Distinct myeloid progenitor-differentiation pathways identified through single-cell RNA sequencing. *Nat. Immunol.* **17**, 666–676 (2016).
58. T. Hashimshony, N. Senderovich, G. Avital, A. Klochendler, Y. de Leeuw, L. Anavy, D. Gennert, S. Li, K. J. Livak, O. Rozenblatt-Rosen, Y. Dor, A. Regev, I. Yanai, CEL-Seq2: Sensitive highly-multiplexed single-cell RNA-Seq. *Genome Biol.* **17**, 77 (2016).
59. T. Ilicic, J. K. Kim, A. A. Kolodziejczyk, F. O. Bagger, D. J. McCarthy, J. C. Marioni, S. A. Teichmann, Classification of low quality cells from single-cell RNA-seq data. *Genome Biol.* **17**, 29 (2016).
60. S. Anders, W. Huber, Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).
61. C. Trapnell, D. Cacchiarelli, J. Grimsby, P. Pokharel, S. Li, M. Morse, N. J. Lennon, K. J. Livak, T. S. Mikkelsen, J. L. Rinn, The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381–386 (2014).
62. D. Lara-Astiaso, A. Weiner, E. Lorenzo-Vivas, I. Zaretsky, D. A. Jaitin, E. David, H. Keren-Shaul, A. Mildner, D. Winter, S. Jung, N. Friedman, I. Amit, Chromatin state dynamics during blood formation. *Science* **345**, 943–949 (2014).
63. E. Afgan, D. Baker, M. van den Beek, D. Blankenberg, D. Bouvier, M. Čech, J. Chilton, D. Clements, N. Coraor, C. Eberhard, B. Grüning, A. Guerler, J. Hillman-Jackson, G. Von Kuster, E. Rasche, N. Soranzo, N. Turaga, J. Taylor, A. Nekrutenko, J. Goecks, The galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Res.* **44**, W3–W10 (2016).
64. Y. Zhang, T. Liu, C. A. Meyer, J. Eeckhoute, D. S. Johnson, B. E. Bernstein, C. Nusbaum, R. M. Myers, M. Brown, W. Li, X. S. Liu, Model-based analysis of ChIP-seq (MACS). *Genome Biol.* **9**, R137 (2008).
65. F. Ramírez, D. P. Ryan, B. Grüning, V. Bhardwaj, F. Kilpert, A. S. Richter, S. Heyne, F. Dündar, T. Manke, deepTools2: A next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).
66. G. Yu, L.-G. Wang, Q.-Y. He, ChIPseeker: An R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **31**, 2382–2383 (2015).
67. M. Samata, A. Akhtar, Dosage compensation of the X chromosome: A complex epigenetic assignment involving chromatin regulators and long noncoding RNAs. *Annu. Rev. Biochem.* **87**, 323–350 (2018).

68. N. Cabezas-Wallscheid, F. Buettner, P. Sommerkamp, D. Klimmeck, L. Ladel, F. B. Thalheimer, D. Pastor-Flores, L. P. Roma, S. Renders, P. Zeisberger, A. Przybylla, K. Schönberger, R. Scognamiglio, S. Altamura, C. M. Florian, M. Fawaz, D. Vonficht, M. Tesio, P. Collier, D. Pavlinic, H. Geiger, T. Schroeder, V. Benes, T. P. Dick, M. A. Rieger, O. Stegle, A. Trumpp, Vitamin A-retinoic acid signaling regulates hematopoietic stem cell dormancy. *Cell* **169**, 807–823.e19 (2017).

Acknowledgments: We are grateful to M. F. Basilicata for valuable support in writing/reviewing the manuscript and guidance. E. Trompouki, N. Cabezas-Wallscheid, T. Aktas, I. A. Ilik, M. Samata, G. V. Renschler, A. A. M. Lenaerts, B. Sheikh, and M. Shvedunova for critical reading of the draft manuscript. We are very thankful to T. Lammerman for providing the *Vav1-iCre* animals. M. F. Basilicata, T. Aktas, and A. Karoutas for help in various experiments. G Semplicio for sharing bioinformatic knowledge. The MPI-IE core facilities for FACS, deep sequencing, mouse, and imaging, which have been invaluable for this project. **Funding:** This work was supported by the Deutsche Forschungsgemeinschaft (DFG) GR4980/3-1 awarded to D.G. and by the German Research CRC992 (A02), and CRC1381 (B3) awarded to A.A. **Author contributions:** Conceptualization, C.P.R. and A.A. Methodology, C.P.R. Investigation, C.P.R. Generated *Vav1*-iCre *Kansl3*$^{fl/fl}$ and *Kansl2*$^{fl/fl}$ animals, B.H. Animal breeding and handling, C.P.R. and T.S. Luciferase assay conceptualization, C.I.K.V. scRNA-seq supervision, D.G. and J.S.H.; ChIP-seq supervision, C.I.K.V.; visualization, C.P.R.; writing (original draft), C.P.R. and C.I.K.V. Writing (review and editing), C.P.R., C.I.K.V., A.A., and D.G. Funding acquisition, A.A. Supervision , A.A. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors. The accession number for the scRNA-seq and HSC MOF ChIP-seq reported in this paper is GEO: GSE107154.

Submitted 12 September 2019
Accepted 9 March 2020
Published 20 May 2020
10.1126/sciadv.aaz4815

Citation: C. Pessoa Rodrigues, J. S. Herman, B. Herquel, C. I. K. Valsecchi, T. Stehle, D. Grün, A. Akhtar, Temporal expression of MOF acetyltransferase primes transcription factor networks for erythroid fate. *Sci. Adv.* **6**, eaaz4815 (2020).