EBF1 primes B-lymphoid enhancers and limits the myeloid bias in murine multipotent progenitors

Aurelie Lenaerts1,2,3*, Iwo Kucinski3*, Ward Deboutte1*, Marta Derecka3*, Pierre Cauchy3*, Thomas Manke4, Berthold Göttgens4, and Rudolf Grosschedl1

Hematopoietic stem cells (HSCs) and multipotent progenitors (MPPs) generate all cells of the blood system. Despite their multipotency, MPPs display poorly understood lineage bias. Here, we examine whether lineage-specifying transcription factors, such as the B-lineage determinant EBF1, regulate lineage preference in early progenitors. We detect low-level EBF1 expression in myeloid-biased MPP3 and lymphoid-biased MPP4 cells, coinciding with expression of the myeloid determinant C/EBPa. Hematopoietic deletion of Ebf1 results in enhanced myelopoiesis and reduced HSC repopulation capacity. Ebf1-deficient MPP3 and MPP4 cells exhibit an augmented myeloid differentiation potential and a transcriptome with an enriched C/EBPa signature. Correspondingly, EBF1 binds the Cebpα enhancer, and the deficiency and overexpression of Ebf1 in MPP3 and MPP4 cells lead to an up- and downregulation of Cebpα expression, respectively. In addition, EBF1 primes the chromatin of B-lymphoid enhancers specifically in MPP3 cells. Thus, our study implicates EBF1 in regulating myeloid/lymphoid fate bias in MPPs by constraining C/EBPa-driven myelopoiesis and priming the B-lymphoid fate.

Introduction

Hematopoietic stem and progenitor cell (HSPC) differentiation is regulated by an interplay of cell-extrinsic and -intrinsic cues. Hematopoietic stem cells (HSCs) are located at the apex of this process and are characterized by their quiescence and capacity for multi-lineage reconstitution of the entire hematopoietic system (Domen and Weissman, 1999; Orkin and Zon, 2008). HSCs can give rise to multipotent progenitors (MPPs) that are further defined by their lineage potentials. MPP2 differentiates preferentially toward erythroid and megakaryocytic lineages, MPP3 toward myeloid lineages, and MPP4 toward lymphoid lineages (Cabezas-Wallscheid et al., 2014; Pi et al., 2015; Rodriguez-Fraticelli et al., 2018; Weinreb et al., 2020; Sommerkamp et al., 2021). Myeloid-restricted pre-granulocyte-macrophage (pre-GM) progenitors and, further, restricted granulocyte/monocyte progenitors (GMPs) represent the continuum of myeloid differentiation (Pro et al., 2007; Drissen et al., 2016). Common lymphoid progenitors (CLPs), which include all-lymphoid progenitors (ALPs) and B-lymphoid progenitors (BLPs), represent the continuum of early lymphoid differentiation (Inlay et al., 2009; Jensen et al., 2018). In recent years, the hierarchical and stepwise differentiation model of hematopoiesis has been challenged by a model of gradual restriction of lineage potential. Other than the central role of lineage-specifying transcription factors (LS-TFs), the restriction of lineage potential is accompanied by gradual changes in the chromatin landscape and repression of lineage genes (Hu et al., 1997; Miyamoto et al., 2002; Nimmo et al., 2015; Paul et al., 2015; Pali et al., 2019; Ranzoni et al., 2021). Moreover, we have come to appreciate the heterogeneity of the HSC compartment and the plasticity of lineage-biased progenitors that exist along a spectrum of lineage specification rather than as discrete bi-potent populations.

Enhancers associated with lineage-specific genes can be primed in HSCs before their gene expression in committed cells (Heinz et al., 2010; Mercer et al., 2011; Lara-Astiaso et al., 2014). Furthermore, LS-TFs can prime the chromatin landscape and/or initiate lineage-specific gene regulatory networks, thus directing the path of differentiation. This function of LS-TFs is shown by their ability to impose specific cell fates upon forced expression. For example, EBF1 overexpression in LSKs restricts lymphopoiesis to a B-lymphoid output (Zhang et al., 2003), whereas C/EBPa overexpression in fibroblastic cells or B cell progenitors results in transdifferentiation to granulocytic-monocytic cell types (Xie et al., 2004; Fukuchi et al., 2006; Feng et al., 2008; Bussmann et al., 2009; Di Tullio et al., 2011). In MPPs, the relative expression
levels of LS-TFs can determine the resulting cell fate. High levels of PU.1 and low C/EBPa expression are inversely instructive of monocyte versus granulocyte cell fate, whereas low levels of PU.1, achieved through transcriptional repression by Ikaros and Gfi1, are required for B cell specification (DeKoter and Singh, 2000; Dahl et al., 2007; Spooner et al., 2009; Zarnegar and Rothenberg, 2012). In addition to their role in the differentiation of intermediate progenitors, PU.1 and C/EBPs are also required to maintain HSC quiescence and self-renewal capacity (Iwasaki et al., 2005; Ye et al., 2013; Hasemann et al., 2014; Chavez et al., 2021).

EBF1 is at the core of the B cell specification and commitment program, as it is sufficient to rescue the developmental block of PU.1, Ikaros, or E2A deficiency in cultured progenitors (Pongubala et al., 2008; Reynaud et al., 2008). In addition to its role in activating the B cell-specific gene program and suppression of alternative lineage programs in collaboration with Pax5 (Lin and Grosschedl, 1995; Mikkola et al., 2002; Medina et al., 2004; Cobaleda et al., 2007; Treiber et al., 2010; Győry et al., 2012; Nechanitzky et al., 2013; Boller et al., 2016), recent studies have shown that EBF1 can regulate the chromatin landscape by binding to closed chromatin regions and opening these regions via the recruitment of remodeling complexes (Boller et al., 2016; Li et al., 2018; Wang et al., 2020). Furthermore, a knockdown of EBF1 in human cord blood cells was found to result in a reduced frequency of phenotypic HSCs, raising the possibility of additional functions of EBF1 in early hematopoiesis (van Galen et al., 2014).

Here, we find that hematopoietic Ebf1 deletion results in a myeloid-biased HSPC output and decreased HSC quiescence and repopulation capacity. Furthermore, we show that EBF1 confers accessibility upon B-lymphoid enhancers at naive chromatin sites, specifically in myeloid-biased MPP3 progenitors. Additionally, we show that Ebf1 deficiency in MPP3 and MPP4 cells is associated with impaired EBF1-dependent repression of myeloid-determinant C/EBPα, leading to an enhanced myeloid differentiation potential of Ebf1-deficient MPPs.

Results
EBF1 is expressed at low levels in MPP3 and MPP4 cells
To examine whether EBF1 could play a role in HSPCs, we first assessed Ebf1 expression in publicly available bulk RNA-sequencing (RNA-seq) datasets of the following cell populations: LT-HSC (Lin-Sca1+cKit- [LSK] CD34+Flt3-Cd150-Cd48+), MPP1 (LSK CD34+Flt3+ Cd150+CD48+), MPP2 (LSK CD34+Flt3+CD150+CD48+), MPP3 (LSK CD34+Flt3+CD150+CD48+), MPP4 (LSK CD34+ Flt3+CD150+CD48+), CLP (Lin-Sca1+cKit+IL7R/Flt3+), and GMP (Lin-Sca1-cKit+CD41+CD16/32+; Sommerkamp et al., 2020, GEO accession no. GSE162607). As expected, Ebf1 mRNA was detected at high level in the CLPs and was absent in the GMPs (Fig. 1 A). Interestingly, Ebf1 transcripts were detected in multiple HSC populations (defined by the LSK immunophenotype), albeit at much lower levels than in CLPs. To confirm the expression of Ebf1 in HSPCs, we sorted different cell populations by flow cytometry (Fig. S1, A and B) and analyzed their RNA for the presence of Ebf1 transcripts by quantitative reverse transcriptase (qRT)-PCR and by nested RT-PCR. Consistent with bulk RNA-seq results, qRT-PCR analysis of Ebf1 expression, normalized to expression in CLPs, showed low but consistent Ebf1 levels in MPP2, MPP3, and MPP4 cells and no expression in T cells and Ebf1fl/fl Tie2cre (KO) CLPs (Fig. 1 B). Further, recent RT-PCR analysis showed that Ebf1 expression was detectable only in MPP3 and MPP4 cells, albeit at lower levels than in CLPs (Fig. 1 C). In contrast to the qRT-PCR and bulk RNA-seq analysis, Ebf1 transcripts were not detected in MPP2 cells (Fig. 1 C). We also assessed EBF1 protein expression in total cell lysates of 150,000 FACS-sorted HSPCs by immunoblot (Fig. 1 D). EBF1 expression was detected in MPP3 and MPP4 cells, albeit at much lower levels than in CLP and splenic CD19+ B cells. No EBF1 expression was detected in Ebf1fl/fl Tie2cre (KO) pro-B (Fr. A) cells, T cells, and HSCs (LSK Flt3+CD150+CD48+), and was hardly detectable in MPP2 cells (Fig. 1 D). Together, these data suggest that EBF1 is expressed at low but detectable levels in MPP3 and MPP4 cells.

Ebf1 deficiency alters the composition of the HSPC compartment
To elucidate the role of EBF1 in HSPCs, we generated Ebf1wt/wt Tie2cre and Ebf1fl/fl Tie2cre mice, hereafter called Ebf1wt and Ebf1ko mice. Tie2cre drives Cre expression in the endothelium and adult HSCs (Jaffredo et al., 1998; Kisanuki et al., 2001; Zovein et al., 2008). To disentangle the effects of Ebf1 deletion in HSPCs from the effects of B cell deficiency, we analyzed B cell-deficient IghmTm1 mice, also known as muMT- mice, which lack the expression of membrane-bound IgM (Kitamura et al., 1991). The efficiency of Ebf1 and Ighm mutations was confirmed by the absence of CD19+ B cells in the bone marrow (BM; Fig. 1 E). To assess the BM composition, we compared the absolute number of cells within the HSPC compartment. Importantly, no significant differences were observed in the total BM cellularity of Ebf1ko and IghmTm1 mice compared to Ebf1wt and Ighmwt mice (Fig. 1 F). Flow cytometric analysis showed a significant increase in absolute numbers of HSCs, MPP2, and MPP3 cells in the BM of Ebf1ko mice relative to Ebf1wt mice, while no significant differences were observed in IghmTm1 mice (Fig. 1 G). The absolute number of lymphoid-biased MPP4 cells and CLPs showed no significant changes in the BM of Ebf1ko and IghmTm1 mice compared to Ebf1wt mice and Ighmwt mice, respectively (Fig. 1 G and H). CLPs contain ALPs (Lin-Sca1+cKit+IL7R/Flt3+Ly6D-) that give rise to B, T, and natural killer cells, as well as BLPs (Lin-Sca1+cKit+IL7R/Flt3+Ly6D+) that generate B cells (Inlay et al., 2009; Jensen et al., 2018). Therefore, we also examined their relative frequencies in Ebf1ko and IghmTm1 mice and found no significant changes relative to the Ebf1wt and Ighmwt counterparts, indicating that B cell deficiency does not lead to an accumulation of BLPs (Fig. 1 I).

Mesenchymal stromal cell-specific Ebf1 deletion elicits long-term changes in HSC function (Derecka et al., 2020), and EBF1 is expressed in endothelial cells. To examine whether or not a Tie2cre-mediated deletion of Ebf1 in endothelial cells may account for the observed changes in HSPC composition, we performed reverse adaptive transfer (AdT) assays in which lethally irradiated CD45.2 Ebf1wt and Ebf1ko recipients were reconstituted with 5 x 106 BM cells from CD45.1 WT mice (Fig. 1 J). We
Figure 1. **EBF1 expression in MPP3 and MPP4 populations are required for normal HSPC composition.** (A) Normalized read counts of Ebf1 in different hematopoietic cell populations from public RNA-seq datasets (original data from Sommerkamp et al. (2020)) and GEO dataset GSE162662. Data from Lenaerts et al. Journal of Experimental Medicine. doi:10.1084/jem.20212437.
observed a similar overall frequency of WT donor cells in the BM of Ebf1WT and Ebf1KO recipients (Fig. 1 K). Moreover, the frequencies of WT HSCs, MPP2, MPP3, and MPP4 cells, as well as mature B, T, and myeloid cells were similar in Ebf1WT and Ebf1KO recipients (Fig. 1 L and Fig. S1 C). Together, these data show that the Tie2cre-mediated deletion of Ebf1 in the hematopoietic compartment leads to an increase in the numbers of HSCs, MPP2, and myeloid-biased MPP3 cells, while leaving the lymphoid-biased MPP4 cells and CLPs numbers unaffected. Importantly, these effects were neither due to a deletion of Ebf1 in the endothelium of the BM niche nor caused by the lack of B cells, confirming an HSPC-specific role of EBF1.

**EBF1 deficiency leads to reduced HSC quiescence and self-renewal capacity**

Given that HSCs are at the apex of hematopoiesis, we first explored the role of EBF1 for HSCs by extending our flow cytometric analysis. Ebf1WT HSCs showed increased expression of CD41 and increased absolute numbers of CD41+ HSCs in Ebf1KO mice relative to Ebf1WT mice (Fig. 2 A and B; and Fig. S1 D). CD41 expression on HSCs is indicative of a myeloid-biased output (Gekas and Graf, 2013; Yamamoto et al., 2013; Miyawaki et al., 2015) and active HSCs (Wilson et al., 2015). Therefore, we assessed the HSC cell cycle status and detected a significantly lower percentage of Ebf1KO HSCs in G0 phase and a significantly higher percentage of Ebf1KO HSCs in G1 and S/G2/M phase as compared to Ebf1WT HSCs (Fig. 2 C). In contrast, Ig/hmTM HSCs showed no increase in CD41 expression, and no difference in the cell cycle status relative to Ig/hmWT HSCs (Fig. 2 A, C; and Fig. S1 D). We also examined the activation status of HSCs under proliferative stress through chronic 5-fluorouracil (5-FU) treatment. In line with the increased cycling of Ebf1KO HSCs, the median survival of 5-FU–treated Ebf1KO mice was lower than that of 5-FU–treated Ig/hmWT and WT (Ebf1WT and Ig/hmWT) mice (Fig. 2 D). Together, these results indicate that Ebf1KO mice have HSCs that are less quiescent and exhaust faster.

MPP3 cells can secrete HSC-activating cytokines (Kang et al., 2021) and, therefore, we examined the increase in MPP3 numbers in Ebf1KO BM could contribute to an HSC-activating cytokine environment. To this end, we cultured Cell Trace Yellow (CTY)–labeled WT HSCs in media containing increasing concentrations of Ebf1WT or Ebf1KO BM fluid and evaluated CTY dye dilution as an index of HSC proliferation. The proliferation index of HSCs was unchanged regardless of the BM cytokine environment of WT and Ebf1KO BM (Fig. S1 E). Absolute number of CD19+ B cells in the BM. Ebf1WT and Ebf1KO n = 7, Ig/hmWT and Ig/hmTM n = 11. (F) Total BM cellularity. Ebf1WT n = 24, Ebf1KO n = 14, Ig/hmWT and Ig/hmTM n = 11. (G) Boxplots showing the absolute number of HSC, erythroid-platelet biased MPP2, myeloid-biased MPP3, and lymphoid-biased MPP4 populations in the BM. Ebf1WT n = 19–23, Ebf1KO n = 14–18, Ig/hmWT and Ig/hmTM n = 8. (H) Absolute number of CLP cells in the BM. (I) Proportion of ALP and BLP cells in the CLP population. (H and I) Ebf1WT n = 10, Ebf1KO n = 4, Ig/hmWT and Ig/hmTM n = 4. (J) Schematic of reverse AdT assays with Ebf1WT and Ebf1KO mice. 5 × 106 total BM cells of WT mice (CD45.1) were injected into lethally irradiated Ebf1WT and Ebf1KO recipients (CD45.2). (K) Boxplots showing the frequency of CD45.1 donor cells in the BM of Ebf1WT and Ebf1KO recipients, 16 wk after AdT. (L) Boxplots showing the frequencies of HSCs, MPP2, MPP3, and MPP4 within CD45.1 donor cells in the BM of the Ebf1WT and Ebf1KO recipients 16 wk after AdT. (K and L) Ebf1WT n = 13, Ebf1KO n = 14. (E–L) Statistical significance was determined by Mann–Whitney U test. Source data are from >2 independent experiments. Source data are available for this figure: SourceData F1.

Given that quiescence correlates with repopulation potential (Passegué et al., 2005; Foudi et al., 2009; Lauridsen et al., 2018), we also tested whether Ebf1KO HSCs were functionally impaired. In CFU assays, Ebf1KO, but not Ig/hmTM, HSCs produced significantly fewer colonies relative to WT HSCs (Fig. S1 F). When assessing the type of colonies generated by Ebf1KO HSCs, we find a significant increase in unip- and bi-lineage granulocyte and monocyte colonies, and a significant decrease in multipotent CFU-GEMM colonies relative to Ebf1WT HSCs (Fig. S1 G). These data suggest an impaired self-renewal capacity relative to Ebf1WT HSCs. To assess their self-renewal capacity in vivo, we transferred 150 FACS-sorted HSCs from Ebf1WT, Ebf1KO, Ig/hmWT, and Ig/hmTM mice into primary WT recipients and analyzed their BM chimerism after 16 wk (Fig. 2 E). The total donor contribution in the BM of primary recipients was significantly lower for transplanted Ebf1KO HSCs than for Ebf1WT HSCs, whereas the total donor contribution was similar for Ig/hmWT and Ig/hmTM HSCs (Fig. 2 F). The donor contribution of Ebf1KO HSCs, in the periphery throughout the AdT and in the BM, to T cells and myeloid cells was also reduced, whereas the contribution to HSCs was not significantly changed (Fig. 2 G and Fig. S1 H). As expected, B cell reconstitution was absent for recipients of Ebf1KO and Ig/hmTM HSCs (Fig. 2 G and Fig. S1 H). Secondary AdT assays, performed with 3 × 106 total BM cells from Ebf1WT and Ebf1KO primary recipients, showed a further reduction in BM donor chimerism and an enhanced multilineage reconstitution defect (Fig. 2, H and I). To address whether Ebf1KO HSCs have a decreased repopulating capacity due to a homing defect, we FACSorted 5 × 104 CTY-labeled LSKs from Ebf1WT and Ebf1KO mice as an estimate of HSC homing capacity, injected them into WT recipients, and evaluated the frequency CTY-positive cells in the BM of recipients 14 h after transfer. This analysis revealed that Ebf1WT and Ebf1KO hematopoietic progenitors home to the BM.
Figure 2. Reduced HSC quiescence and self-renewal capacity upon Ebf1 deletion. (A) Representative contour plots showing CD41+ expression on HSCs in Ebf1$^{WT}$ and Ebf1$^{KO}$ mice. (B) Boxplots showing the percentage of CD41+ expressing cells in the HSC population. Ebf1$^{WT}$ n = 11, Ebf1$^{KO}$ n = 9, Ighm$^{WT}$ and Ighm$^{Tm1}$

Lenaerts et al. Journal of Experimental Medicine 5 of 26
EBF1 function in multipotent progenitors

https://doi.org/10.1084/jem.20212437
with similar efficiency (Fig. 2 J). Taken together, these data indicate that Ebf1 deletion in HSPCs leads to an impaired HSC function. Given the lack of EBF1 expression in HSCs and MPP2 cells, the HSC phenotypes are likely indirect and secondary to the EBF1-driven changes in MPP3 and MPP4 cells.

**EBF1 confers de novo accessibility in MPP3 cells**

During B-lymphoid differentiation, EBF1 activates lineage-specific transcriptional networks and induces chromatin landscape changes (Hagman et al., 2011; Boller et al., 2018; Li et al., 2020). Within the HSPC compartment, EBF1 expression was strongest in MPP3 and MPP4 cells (Fig. 1 D). To understand the molecular effects of Ebf1 deletion in HSPCs, we thus analyzed the chromatin and transcriptome of MPP3 and MPP4 cells from Ebf1WT and Ebf1KO mice by bulk assay for transposase-accessible chromatin using sequencing (ATAC-seq) and RNA-seq, respectively. We observed a positive correlation between MPP3/MPP4 differentially accessible (DA) ATAC peaks overlapping with transcription start sites (TSS peaks) and the expression of MPP3/MPP4 differentially expressed (DE) genes (Fig. S2 A). This correlation shows that MPP3 and MPP4 chromatin accessibility profiles reflect the myeloid- and lymphoid-gene expression patterns of MPP3 and MPP4 cells, respectively (Cabezas-Wallscheid et al., 2014; Pietras et al., 2015; Rodriguez-Fraticelli et al., 2018; Weinreb et al., 2020; Sommerkamp et al., 2021). Next, we performed differential peak calling for Ebf1WT versus Ebf1KO MPP3 cells and Ebf1WT versus Ebf1KO MPP4 cells.

The most pronounced effect of the Ebf1 deletion was observed in MPP3 cells, in which 49 sites had gained and 410 sites had reduced accessibility in Ebf1KO MPP3 cells relative to Ebf1WT cells (Fig. 3 A and Table S2). Surprisingly, a minority of KO-reduced MPP3 peaks were detected in Ebf1WT MPP4 cells (91 out of 410 peaks), but these peaks were not DA in Ebf1WT versus Ebf1KO MPP4 cells. A different set of sites showed modest changes in accessibility between Ebf1WT and Ebf1KO MPP4 cells (Fig. 3 A and Table S2). To assess whether the DA sites observed in MPP3 and MPP4 cells were accessible in upstream progenitors, we analyzed the chromatin landscape of Ebf1WT and Ebf1KO LT-HSCs. Notably, the sites that showed reduced accessibility in Ebf1KO MPP3 cells were not accessible in Ebf1WT and Ebf1KO LT-HSCs, whereas the other DA sites were detected in Ebf1WT and Ebf1KO LT-HSCs (Fig. 3 B). This result suggests that the sites of decreased accessibility in Ebf1KO MPP3 cells are not accessible in stem cells.

Motif enrichment analysis of ATAC peaks that were reduced in Ebf1KO MPP3 cells identified the EBF1- and TCF3 (E2A)-binding motifs as top-ranked hits, whereas other motifs, such as Rxrα, Spic, and Stat2, predominated the peaks reduced in Ebf1KO MPP4 cells (Fig. S2 B). More than half of the peaks reduced in Ebf1KO MPP3 cells (230 out of 410 peaks) contained the canonical EBF1 motif and digital footprinting, which determined the cumulative aggregation of Tn5 signals, revealed clear footprints of EBF1 occupancy in Ebf1WT but not in Ebf1KO MPP3 cells, and no footprints of EBF1 occupancy were detected in Ebf1WT MPP4 cells (Fig. 3 C). To assess whether the limited number of sites provided sufficient depth for digital footprinting, we analyzed the Tn5 signal at predicted CTCF sites in random peaks with equal size, which showed similar CTCF occupancy in Ebf1WT and Ebf1KO MPP3 and MPP4 cells (Fig. 3 C). These data suggest a requirement for EBF1 occupancy at peaks reduced in Ebf1KO MPP3 cells. To assess whether forced EBF1 expression can restore accessibility at these sites, we re-expressed EBF1 in Ebf1KO MPP3 cells by retroviral transduction (Fig. S2 C). We find that accessibility is gained specifically at peaks reduced in Ebf1KO MPP3 cells in the EBF1 re-expressed cells, whereas the other DA sites are unchanged by EBF1 re-expression (Fig. 3 D). Together, these data show that EBF1 is necessary and sufficient for de novo accessibility to sites that show reduced accessibility in Ebf1KO MPP3 cells.

**EBF1 primes B-lymphoid accessibility in MPP3 cells**

To assess a possible co-regulation by transcription factors, we performed a co-occurrence analysis of enriched motifs in different DA peaks. Although PU.1 was an enriched motif in DA peaks when using an alternative enrichment analysis (see Materials and methods), we included PU.1 in the co-occurrence analysis given its crucial role in lymphoid and myeloid differentiation of MPPs (DeKoter and Singh, 2000; Pang et al., 2018). We detected co-occurrence of PU.1 motifs with enriched motifs in the different DA peaks, in line with its importance in hematopoietic progenitors (Fig. 3 E and Fig. S2 D). We observed the strongest co-occurrence between EBF1 and TCF3 (E2A) motifs at EBF1-dependent sites detected in MPP3 cells (Fig. 3 E), but not in the other DA peak sets (Fig. S2 D). This co-occurrence may represent functional cis-regulatory modules because EBF1 and E2A synergistically cooperate to establish B-lymphoid gene expression and they often bind chromatin in close proximity to each other (O’Riordan and Grosschedl, 1999; Lin et al., 2010).
Figure 3. **EBF1 primes the B-lymphoid fate in myeloid-biased MPP3 progenitors.** (A and B) Heatmap displaying chromatin accessibility at DA peaks in MPP3 and MPP4 cells (A) and in LT-HSCs (LSK CD34−Flt3−CD150+CD48−; B) from Ebf1WT and Ebf1KO mice. Biological replicates n = 2. (A, B, and D) Peaks are

Lenaerts et al. Journal of Experimental Medicine 7 of 26

EBF1 function in multipotent progenitors

https://doi.org/10.1084/jem.20212437
To inspect EBF1 binding at DA peaks between Ebf1WT and Ebf1KO and MPP3 and MPP4 cells, we analyzed publicly available EBF1 chromatin immunoprecipitation (ChIP)-seq datasets in pro-B cells (Li et al., 2018) and in Hoxb8-FL cells (Kucinski et al., 2020). Hoxb8-FL cells are conditionally immortalized HSPCs that resemble MPPs by maintaining both lymphoid and myeloid potential (Redecke et al., 2013). We find that the EBF1-dependent peaks detected in MPP3 cells are bound by EBF1 in pro-B cells and in Hoxb8-FL cells (Fig. 3 F), suggesting that these EBF1-dependent sites are related to pro-B cells. Consistent with the digital footprinting analysis, no EBF1 binding was detected in Ebf1WT/Ebf1KO DA peaks of MPP4 cells. Additionally, we analyzed the PU.1 ChIP-seq profile in Hoxb8-FL cells (Kucinski et al., 2020) and observed PU.1 binding at peaks reduced in Ebf1KO MPP3 and MPP4 cells and at peaks gained in Ebf1KO MPP4 cells (Fig. 3 F). This observation is in line with the importance of PU.1 in multipotent progenitors.

Gene ontology (GO) analysis of genes annotated to the peaks reduced in Ebf1KO MPP4 cells exhibited enrichment for lymphoid-related molecular signatures (Fig. S2 E). Notably, GO analysis of the genes annotated to peaks reduced in Ebf1KO MPP3 cells showed a specific enrichment for B-lymphoid molecular signatures, such as B cell receptor signaling and B cell differentiation (Fig. 3 G). These genes include B-lymphoid-related genes, such as Cd19 and Igll1, which show EBF1 occupancy in pro-B cells and Hoxb8-FL progenitors at sites accessible in Ebf1WT MPP3 cells (Fig. 3 H and Fig. S2 F). Finally, we assessed whether EBF1 re-expression in Ebf1KO MPP3 and MPP4 cells can rescue their B cell differentiation defect. We observed that re-expression of EBF1 in Ebf1KO MPP3 and MPP4 cells allows their differentiation into CD19+ B cells (Fig. 3 I). Taken together, these data reveal a subset of B-lymphoid-associated regions that display an enhancer signature and EBF1-dependent accessibility in MPP3 cells.

**EBF1-dependent enhancers in MPP3 cells are associated with expression in CLPs**

We turned to bulk RNA-seq of MPP3 and MPP4 progenitors to understand the molecular consequences of Ebf1 deletion in HSPCs (Table S3). Principal component analysis (PCA) of the transcriptomes reveals moderate Ebf1-related changes (aligned with PC2: 8% of variance) compared to the differences between MPP3 and MPP4 populations (PC1: 80% of variance; Fig. S3 A). In Ebf1KO MPP3 cells, we found 113 downregulated and 102 upregulated genes relative to Ebf1WT MPP3 cells, and in Ebf1KO MPP4 cells, we identified 152 downregulated and 153 upregulated genes as compared to Ebf1WT cells (false discovery rate [FDR] <0.1; Fig. S3 B). By an enrichment analysis against various signatures, we find that the Ebf1KO-downregulated genes in MPP3 cells are enriched for B cell–identity signatures, which is in line with the chromatin accessibility analysis, although the number of genes is limited (Fig. 4 A). Interrogation of genes showing reduced ATAC peaks in Ebf1KO MPP3 cells with the transcriptome analysis indicated that roughly half of these genes are weakly expressed in MPP3 cells (Fig. 4 B). Moreover, we found that impaired chromatin accessibility is not associated with DE genes in Ebf1KO versus Ebf1WT MPP3 cells (Fig. 4 B). This poor correlation between altered chromatin accessibility and gene expression was applied to both TSS-associated and DA ATAC peaks (Fig. S3 C). These observations were comparable for MPP4 cells (Fig. 4 B and Fig. S3 D). Thus, the genes associated with EBF1-dependent peaks detected in MPP3 cells are weak and not DE in MPP3 or MPP4 cells.

Using datasets of chromatin modifications in HSPC populations (Lara-Astiaso et al., 2014), we found that ATAC peaks reduced or gained in Ebf1KO MPP4 cells were marked by H3K27ac, H3K4me1, and H3K4me3 modifications, suggesting that the peaks coincide with promoter regions (Fig. 4 C). Interestingly, peaks reduced in Ebf1KO MPP3 cells were marked by H3K27ac and H3K4me1 but not by H3K4me3 modifications, suggesting that these chromatin regions may coincide with active enhancers (Fig. 4 C). This observation raised the question of whether EBF1-dependent accessible sites in MPP3 cells are associated with genes that are expressed in further lineage-restricted differentiation intermediates. Therefore, we interrogated the ATAC peak signals in Ebf1WT and Ebf1KO MPP3 cells with changes in gene expression that occur between CLP and GMP progenitors. This analysis indicated a strong association of EBF1-dependent ATAC peaks in MPP3 cells (DA peaks) with a gain of gene expression in CLPs relative to GMPs (Fig. 4 D). A corresponding analysis of ATAC peaks at TSS in MPP3 cells revealed a less pronounced association with the CLP-specific gene expression, consistent with the preferential EBF1-mediated priming of enhancers in MPP3 cells (Fig. 4 D). An association between CLP gene expression and ATAC peak signals in Ebf1WT and Ebf1KO MPP4 cells was not observed (Fig. 4 E). To further assess the association of enhancers in MPP3 and MPP4 cells with downstream gene expression, we generated a myeloid and lymphoid signature using genes that are DE between GMP and CLP progenitors. Then, we selected ATAC peaks that are annotated as enhancers in MPP3 and MPP4 cells. With these means, gene set enrichment analysis (GSEA) analysis showed that
Figure 4. EBF1-dependent chromatin sites in MPP3 cells are associated with expression in CLPs. (A) Selected terms are represented from enrichment analysis of DE genes (FDR < 0.1) between Ebf1 WT and Ebf1 KO MPP3 (left) and MPP4 (right) cells. Enrichr datasets (GO_Biological_Process_2021, Pan-glaoDB_Augmented_2021, Gene_Perturbations_from_GEO_down) were used for enrichment analysis. Biological replicates n = 4. (B) Expression analysis of genes annotated to KO Reduced MPP3 peaks identified in ATAC analysis, in Ebf1 WT and Ebf1 KO MPP3 (upper panel), and MPP4 (lower panel) cells. The x axis represents log10 mean normalized counts of all genes. The y axis represents shrunken |log2 fold change| values of all genes in Ebf1 WT versus Ebf1 KO MPP3 (upper panel), and MPP4 (lower panel) cells, and is capped at 2. Genes in orange indicate the genes annotated to EBF1-dependent accessibility sites in MPP3 cells. (C) Heatmap displaying average ChIP signal for chromatin marks at DA peaks. Original data from Lara-Astiaso et al. (2014). (D and E) Comparison of chromatin
enhancer ATAC peaks in Ebf1<sup>WT</sup> MPP3 cells were specifically enriched for the CLP signature (Fig. 4 F), which was not observed for Ebf1<sup>WT</sup> MPP4 cells (Fig. 4 G). Taken together, these results indicate that EBF1-dependent chromatin accessibility in MPP3 cells is associated with enhancers that confer gene expression in the lymphoid lineage-restricted CLP compartment.

**Ebf1 deficiency enhances the C/EBPa-driven myeloid bias of MPP3 and MPP4 cells**

The impaired lymphoid chromatin landscape of Ebf1<sup>KO</sup> MPP3 cells and the enrichment of myeloid molecular signatures (neutrophil activation) and macrophage cell-identity signatures in the enrichment analysis of the Ebf1<sup>KO</sup> MPP3 and MPP4 transcriptomes raised the question of whether or not these cells have an enhanced myeloid bias. In addition, GSEA analysis found that enhancer ATAC peaks in Ebf1<sup>KO</sup> MPP3 cells were specifically enriched for the GMP signature (Fig. 4 F), although this was not significant for enhancer ATAC peaks in Ebf1<sup>KO</sup> MPP4 cells (Fig. 4 G). To assess whether MPP3 and MPP4 cells have an enhanced myeloid bias upon Ebf1 deletion, we evaluated their myeloid differentiation potential in vitro by culturing them in myeloid differentiating conditions and by evaluating CD11b surface expression after 10 d (Fig. 5 A). In line with their RNA-seq profiles, Ebf1<sup>KO</sup> MPP3 and MPP4 cells both showed higher frequencies of CD11b<sup>+</sup> cells, indicating an increased myeloid potential of these EBF1-deficient progenitors (Fig. 5 A). This increase in myeloid potential was dependent on GM-CSF (data not shown). However, in a CFU assay, the total number of colonies generated by MPP3 and MPP4 cells showed no difference between Ebf1<sup>WT</sup> and Ebf1<sup>KO</sup> cells (Fig. S3 E).

The similar increase in the myeloid bias of Ebf1<sup>KO</sup> MPP3 and MPP4 cells contrasted the MPP3-specific EBF1 dependence of B-lymphoid priming. To explore the common enhanced myeloid bias of MPP3 and MPP4 cells, we found an overlap of 66 DE genes (Fig. S3 F), and we noted that the most significantly enriched term associated with the upregulated genes was a signature previously connected with the downregulation of genes upon Cebpa deletion (Fig. 4 A). Notably, many of the genes that were inversely deregulated in Ebf1<sup>KO</sup> progenitors and Cebpa-deficient cells, including Irf8 and Csf1r, are linked to myeloid differentiation (Fig. 5 B). We also found that more genes belonging to the Cebpa signature were significantly upregulated in Ebf1<sup>KO</sup> MPP3 than in Ebf1<sup>KO</sup> MPP4 cells, although a similar trend was observed in Ebf1<sup>KO</sup> MPP4 cells (Fig. 5 B). To assess whether the upregulated genes in Ebf1<sup>KO</sup> cells are C/EBPa-bound targets, we overlapped these genes with public C/EBPa ChIP-seq datasets of HSPC progenitors (LSKs) and myeloid progenitors (pre-GM and GMP; Hasemann et al., 2014; Pundhir et al., 2018). This analysis showed that almost all of the genes associated with the Cebpa signature in Ebf1<sup>KO</sup> MPP3 and MPP4 cells are bound by C/EBPα (Fig. 5 B). We also overlapped the Ebf1<sup>KO</sup>/Ebf1<sup>WT</sup> DE genes of MPP3 and MPP4 cells with the C/EBPa occupancy in LSK, pre-GMs, and GMPs and found that the DE genes were mainly associated with C/EBPa occupancy in GMPs. Moreover, C/EBPa occupancy was mainly associated with upregulated genes in Ebf1<sup>KO</sup> MPP3 cells and associated with downregulated genes in Ebf1<sup>KO</sup> MPP4 cells (Fig. 5 C and Fig. S3 F), suggesting the enhanced myeloid bias directed by C/EBPα is stronger in Ebf1<sup>KO</sup> MPP3 cells.

C/EBPα is essential for monocyte–neutrophil specification, and Cebpα KO mice display a block in pre-GM progenitors (Zhang et al., 1997; Paul et al., 2015; Pundhir et al., 2018). Analogous to the function of EBF1 in antagonizing alternative lineages in early B cells, C/EBPα represses non-myeloid lineages, whereby its expression increases from pre-GM to GMP progenitors and reaches the highest levels in neutrophils (Paul et al., 2015). To estimate the direction of perturbation upon Ebf1 deletion, we calculated a Direction of Transition (DoT) score, anchored at the lymphoid portion of the landscape marked by the dotted line, for Ebf1<sup>KO</sup>/Ebf1<sup>WT</sup> DE genes of MPP3 and MPP4 cells (Fig. 5 D; Kucinski et al., 2020). Analogous to term-enrichment analysis of DE genes, the DoT score analysis uses a single-cell RNA-seq (scRNA-seq) landscape from a publicly available LK/LSK dataset (Dahlin et al., 2018) as a reference to visualize similarities between cell states. Positive values in red indicate a transition of Ebf1<sup>KO</sup> conditions toward those cell fates, and negative values in blue indicate a transition of Ebf1<sup>KO</sup> conditions away from those cell fates. In line with the enrichment analysis, the DoT score analysis showed that Ebf1<sup>KO</sup> MPP3 and MPP4 cells shift toward a neutrophil, monocyte/dendritic progenitor cell transcriptional state (Fig. 5 D). These data suggest that Ebf1-deficient MPP3 and MPP4 cells upregulate genes that are targets of C/EBPα in GMPs, which may account for the increased myeloid potential of Ebf1<sup>KO</sup> progenitors. To functionally examine whether C/EBPα is responsible for the increased myeloid bias observed in Ebf1<sup>KO</sup> MPP3 cells, we performed knockdown of Cebpa in Ebf1<sup>WT</sup> and Ebf1<sup>KO</sup> MPP3 cells, followed by in vitro myeloid differentiation assay. Upon Cebpa knockdown in Ebf1<sup>KO</sup> MPP3 cells, we observed a decrease in the frequency of CD11b<sup>+</sup> cells relative to control knockdown conditions (Fig. 5 E and Fig. S3 G). No significant effect was observed in Ebf1<sup>WT</sup> MPP3 cells (Fig. 5 E). These data suggest that the increased myeloid bias observed in Ebf1<sup>KO</sup> MPP3 cells may, at least in part, be driven by C/EBPα.

**Impaired B-lymphoid priming and expansion of myeloid-biased progenitors**

To obtain deeper insights into the HSPC cell states affected by the Ebf1 deletion, we performed scRNA-seq of the LK and LSK compartments, comprising myeloid progenitors and HSPCs. We
Figure 5. Increased myeloid bias in Ebf1KO MPP3 and MPP4 cells is enhanced by C/EBPa. (A) Workflow of myeloid differentiation assay with cytokines (IL3, IL6, EPO, GMCSF, SCF) of MPP3 and MPP4 (upper). Frequency of CD11b+ cells after 10 d of differentiation (lower). Red lines indicate the medians. Data obtained from four mice with following colony numbers: Ebf1WT MPP3 n = 99, Ebf1KO MPP3 n = 108, Ebf1WT MPP4 n = 73, Ebf1KO MPP4 n = 69. (B) Heatmap showing gene expression of DE genes associated with Cebpa KO signature identified through Enrichr enrichment analysis in Fig. 4A. Genes are organized into MPP3-specific, common, and MPP4-specific DE genes. Genes with an annotated C/EBPα ChIP peak are labeled in blue. Heatmap scale represents z-scores calculated separately for MPP3 and MPP4 cells. (C) ChIP-seq analysis of C/EBPα occupancy in LSKs, preGM, and GMP. C/EBPα peaks are organized into MPP3-specific, common, and MPP4-specific DE genes. Regions ± 3 kb around C/EBPα summits are shown. Heatmap scale represents read coverage. Original data from Hasemann et al. (2014); Pundhir et al. (2018). (D) Left: Annotated UMAP projection of a scRNA-seq landscape: mouse LK+LSK populations (Dahlin et al., 2018) used as a reference for DoT score analysis; dashed area indicates the point of origin. DoT z-score values for DE genes between MPP3 Ebf1WT versus Ebf1KO (middle) and MPP4 Ebf1WT versus Ebf1KO (right). Red indicates a shift toward that cell fate and blue indicates a shift away from that cell fate. (E) Frequency of CD11b+ cells after 10 d of differentiation following transfection with a control or a Cebpa siRNA pool. Red lines indicate the medians. Data obtained from four mice with following colony numbers: sictrl Ebf1WT MPP3 n = 19, siCebpa Ebf1WT MPP3 n = 25, sictrl Ebf1KO MPP3 n = 36, siCebpa Ebf1KO MPP3 n = 36. (A and E) Statistical significance was determined by Mann-Whitney U test. Data are from >2 independent experiments. Meg, megakaryocyte; Ery, erythrocyte; MC, mast cell; Bas, basophil; Neu, neutrophil; Mono/DC, monocyte/dendritic cell; Ly, lymphocyte.
analyzed 59,505 high-quality cells in four replicates per genotype. We integrated all data and generated a common landscape with clustering and reduced dimensionality representation: Uniform Manifold Approximation and Projection (UMAP; analogously to Dahlin et al., 2018; Kucinski et al., 2020). The landscape, annotated with lineage markers and an HSC score (Hamey and Göttgens, 2019), showed gradual progression from the HSCs to committed myeloid progenitors (Fig. 6 A and Fig. S4, A and B). The Ebf1WT and Ebf1KO cells did not show significant changes in abundance across the major landscape populations (Fig. S4 A). Ebf1 was predominantly expressed in cluster 4, corresponding to lymphoid progenitors within the LSK compartment (Fig. 6 B). By making use of an index-sorted scRNA-seq HSC dataset (Nestorowa et al., 2016), we embedded the MPP3 cells and MPP4/lympho-myeloid primed progenitor (LMPP) cells into our landscape (Fig. 6 C). MPP4/LMPP cells were found to be enriched on the left side of the landscape territory, which shows elevated expression levels of lymphoid genes, including Dntt and Flt3 (Fig. 6 C and Fig. S4 B). MPP3 cells were uniformly distributed, with a few cells belonging to the lymphoid cluster 4 (Fig. 6, A and C). The Nestorowa et al. (2016) dataset had been generated with the Smart-Seq2 technology, therefore, we took advantage of its higher detection rate to observe Ebf1 expression at a higher resolution. Supporting the notion of a continuum of progenitors and the existence of Ebf1-expressing MPP3 and MPP4 cells, we observed Ebf1 expression in the lymphoid-progenitor cluster 4, but also low-level Ebf1 expression scattered in clusters upstream of lineage-restriction, such as clusters 2 and 1 (Fig. 6, A and D).

Next, we examined whether Ebf1 deletion has consequences on gene expression in clusters representing progenitors of lower lineage restriction, in which a putative lymphoid/myeloid choice would occur. In particular, we focused on clusters 0, 1, 2, and 9, in which we observed lower levels of Ebf1 expression than in cluster 4. Joint differential expression analysis identified a limited number of 17 genes, which surprisingly included 11 genes encoding Ig light chain, Ig J chain, and other B-lymphoid markers (Fig. 6 E and Fig. S4 C). All of these genes were found to be downregulated to undetectable or barely detectable levels following Ebf1 deletion (Fig. 6 E). The reduction of Ig gene expression upon Ebf1 deletion was detected before the activation of the main lymphoid program, annotated by markers like Dntt, IL7r, and Ly6d (Fig. S4, B and C), and was also detected in bulk RNA-seq analysis of MPP3 and MPP4 cells (Table S3).

Finally, we investigated changes in gene expression for each cluster between Ebf1WT and Ebf1KO cells. Appropriately, we detected the majority of DE genes in cluster 4, in which Ebf1 was expressed at the highest level (Fig. 6 F). We calculated the DoT score for the DE genes of cluster 4, anchoring the origin point in the same cluster. We observed strong positive score values in the myeloid and erythroid parts of the landscape, indicating a shift toward these states and away from the lymphoid state (Fig. 6 G).

Moreover, the analysis of enriched GO terms among genes upregulated in cluster 4 in Ebf1KO cells showed a strong enrichment for molecular signatures related to proliferation, such as Myc, increased metabolic pathways, E2F targets, and G2-M checkpoint (Fig. S4 D). Accordingly, we detected increased expression of G2-M cell-cycle phase-associated genes, specifically in Ebf1KO cells of cluster 4 (Fig. 6 H and Table S4). We observed increased expression of Myc and Myc target genes, such as Ccct and Eif3b, in cells of the myeloid and erythroid parts of the landscape, but not in the HSC cluster (Fig. 6 I). We also observed an increase in Myc and Myc target gene expression in Ebf1KO but not Ebf1WT cells of cluster 4, whereas a similar expression pattern was observed in the other clusters (Fig. 6 I). The proliferative advantage of myeloid-biased Ebf1KO cells was mirrored by the significant increase in the frequencies and absolute numbers of myeloid cells in the BM of Ebf1KO mice relative to Ebf1WT mice (Fig. 6 J and Fig. S4 E). Although the frequencies of myeloid cells were also increased in the BM of Ighm−/− mice relative to IghmWT mice, the increase was more pronounced in Ebf1KO mice relative to Ebf1WT mice (Fig. 6 J).

**EBF1 binds the Cebpa +37 kb hematopoietic regulatory region**

The +37 kb Cebpa enhancer is responsible for hematopoietic expression of Cebpa (Cooper et al., 2015; Guo et al., 2016). Interestingly, C/EBPα has been implicated in inducing Ebf1 expression (Guo et al., 2018; Barberi et al., 2020). In addition, EBF1 has been shown to bind within a H3K27ac-marked predicted regulatory region of the Cebpa gene in Hoxb8-FL cells (Kucinski et al., 2020). To further investigate the potential regulation of Cebpa by EBF1, we analyzed previous EBF1 ChIP-seq datasets for EBF1 binding at the Cebpa regulatory region in pro-B cells and Ebf1+/− cKit+ progenitors, in which a doxycycline-responsive ectopic Ebf1 gene was induced for 24 or 72 h (Li et al., 2018). In the progenitors with induced EBF1 expression, we detected EBF1 binding within the Cebpa regulatory region, adjacent to the +37 kb Cebpa enhancer, whereas neither EBF1 binding nor H3K27ac marks were found in pro-B cells (Fig. 7 A). Furthermore, Cebpa expression was decreased in pro-B cells relative to progenitors in which EBF1 expression was induced for 24 or 72 h (Fig. S5 A). Although the +37 kb Cebpa enhancer showed similar accessibility in Ebf1WT and Ebf1KO MPP3 and MPP4 cells, the EBF1-bound site was found to reside in inaccessible chromatin, consistent with other EBF1-repressed targets that are transiently bound before silencing (Fig. 7 A; Li et al., 2018). We observed no accumulation of H3K27me3 at the Cebpa regulatory region in pro-B cells, suggesting that the regulatory domain does not become epigenetically silenced (Fig. 7 A). Moreover, the abundance of Cebpa transcripts, but not that of Spiil/PU.1 transcripts, increased in Ebf1KO MPP3 and MPP4 cells relative to Ebf1WT cells (Fig. 7 B and Fig. S5 B). Finally, overexpression of Ebf1 in Ebf1WT MPP3 and MPP4 cells resulted in decreased Cebpa expression compared to empty vector (Fig. 7 C and Fig. S5 C). Together, these data raise the possibility that EBF1 binding at the Cebpa regulatory region results in the downregulation of Cebpa expression.

Accordingly, we observed an increase in Cebpa-expressing Ebf1KO cells in cluster 4 (Fig. 7 D). Corresponding to the C/EBPα-driven myeloid differentiation observed in Ebf1KO MPP3 cells, we observed an increase in the frequencies and absolute numbers of neutrophils in the BM of Ebf1KO mice relative to Ebf1WT mice (Fig. 7, E and F; and Fig. S5 D). Although an increase in the frequencies of neutrophils was also observed in the BM of
Figure 6. *Ebf1* controls lymphoid/myeloid balance in HSPCs, and its loss causes increased myeloid cells in the BM. (A and C) Annotated UMAP projection of scRNA-seq landscape derived from *Ebf1*WT and *Ebf1*KO LK and LSK populations. Biological replicates n = 4. (B and D) Log-normalized *Ebf1* expression in WT and KO conditions.

**F** DE genes upregulated in *Ebf1*KO clusters.

**G** Dot z-score for *Ebf1*KO upregulated genes.

**H** G2/M cell cycle phase expression score for WT and KO.

**I** UMAP projection of expression of *Cyc1*, *Myc*, *Npm1*, and *Elf3b*.

**J** Percentage of myeloid cells in the BM plotted against *Ebf1* genotype.
expression for Ebf1<sup>WT</sup> cells shown in A and C, respectively. (C) Cells from Nestorowa data (Nestorowa et al., 2016, colored) were embedded into the Ebf1<sup>WT</sup>/Ebf1<sup>KO</sup> landscape (gray) with annotated immunophenotypic populations. (E) Log-normalized expression of Ig and lymphoid-associated DE genes in Ebf1<sup>WT</sup> and Ebf1<sup>KO</sup> cells of selected clusters. Each dot represents the mean expression per mouse (four mice in total). (F) Number of DE genes between Ebf1<sup>WT</sup> and Ebf1<sup>KO</sup> cells per cluster. (G) DoT scores calculated using DE genes in cluster 4 between Ebf1<sup>WT</sup> and Ebf1<sup>KO</sup> cells in the context of the LK/LSK landscape. Mean expression in cluster 4 was used as the point of origin. Red indicates a shift toward that cell fate and blue indicates a shift away from that cell fate. (H) Violin plots showing the gene expression score of G2/M cell cycle–associated genes in Ebf1<sup>WT</sup> and Ebf1<sup>KO</sup> cells in selected clusters. (I) UMAP visualization of the expression of Myc and selected Myc target genes in Ebf1<sup>WT</sup> and Ebf1<sup>KO</sup> cells. (J) Boxplots showing the frequency of myeloid cells (CD11b<sup>+</sup>) in the BM. Ebf1<sup>WT</sup> n = 18, Ebf1<sup>KO</sup> n = 16, IgM<sup>WT</sup> and IgM<sup>KO</sup> n = 11. Statistical significance was determined by Mann–Whitney U test. Data are from >2 independent experiments. LK defined as Lin<sup>−</sup> cKit<sup>+</sup> cells. ESLAM, CD45<sup>+</sup> EPCR<sup>−</sup> CD48<sup>+</sup> CD48<sup>−</sup> HSCs; Meg, megakaryocyte; Ery, erythrocyte; Neu, neutrophil; Mono/DC, monocyte/dendritic cell; Ly, lymphocyte.

Ighm<sup>Tm1</sup> mice relative to Ighm<sup>WT</sup> mice, the frequencies of neutrophils in Ebf1<sup>KO</sup> mice were significantly higher than in Ighm<sup>Tm1</sup> mice (Fig. 7, E and F). This suggests an additive effect of B cell depletion and Ebf1 deletion on neutrophil differentiation in Ebf1<sup>KO</sup> mice.

To address whether a similar Ebf1–Cebpα axis could exist in human hematopoietic progenitors, we analyzed a public data set of human fetal liver progenitors (Popescu et al., 2019). Consistently, the HSC/MPP cluster marked by Procr expression also contains mRNA for Ebf1 and Cebpα (Fig. S5, E–G). Altogether, our scRNA-seq and chromatin profiling analysis suggest that Ebf1 takes part in B-lymphoid enhancer priming and in an attenuation of the myeloid fate potential driven by C/EBPa in MPP3 cells.

**Discussion**

The functions of lineage-determining transcription factors, including EBF1, have been well studied at the onset of lineage specification. However, their potential roles in generating lineage bias of stem and progenitor cells are still largely obscure. Here, we show that the B cell determinant EBF1 is expressed at a low level in MPPs, MPP3, and MPP4, that have a myeloid and lymphoid bias, respectively. We find that the deletion of Ebf1 in all hematopoietic cells results in an increased myeloid bias of both MPP3 and MPP4 populations, which may be attributed to a loss of EBF1 binding at the Cebpα enhancer region and the increased expression of the myeloid transcription factor gene Cebpα. In addition, we find that EBF1 is required for the de novo accessibility and priming of B-lymphoid enhancers specifically in MPP3 cells. This result is in line with extensive literature on the importance of EBF1 in establishing the B-lymphoid transcriptional program in pre-pro-B cells (Medina et al., 2004; Pongubala et al., 2008; Hagman et al., 2011; Welinder et al., 2011; Boller et al., 2018). However, it also introduces the additional dimension that EBF1 is responsible for the priming of the B-lymphoid fate in MPPs. Finally, we observe a diminished functionality of HSCs in the BM Ebf1<sup>−</sup>-deficient mice, which may be an indirect consequence of the augmented myeloid potential of MPPs and/or the enhanced myeloiposisis.

Our single-cell analysis of LK/LSK cells from Ebf1<sup>WT</sup> and Ebf1<sup>KO</sup> cells revealed relatively strong Ebf1 expression in lymphoid progenitors (cluster 4) and weak Ebf1 expression scattered among early progenitors of the LSK clusters, in which lymphoid/myeloid lineage bias may occur. Moreover, scRNA-seq data suggested that EBF1 expression increases in the MPP continuum of MPP3 and MPP4 cells. We show through multiple lines of evidence that as EBF1 expression increases, EBF1 plays a role in limiting myeloid differentiation. While ATAC analysis of Ebf1<sup>KO</sup> MPP3 cells points to a loss of B lymphoid priming, we also observed an enhanced myeloid signature of chromatin accessibility that is associated with a GMP pattern of gene expression. This augmented myeloid chromatin landscape is paralleled by the Cebpα-dependent enhanced capacity of Ebf1<sup>KO</sup> MPP3 and MPP4 cells to generate CD11b<sup>+</sup> myeloid cells in vitro. Moreover, our scRNA-seq data indicate that the Ebf1 deficiency in progenitors of lymphoid cluster 4 results in an increase of Cebpα expression, a myeloid fate enrichment, and an increase in proliferation, consistent with the accumulation of mature CD11b<sup>+</sup> myeloid cells and neutrophils in Ebf1<sup>KO</sup> mice. However, we did not observe an increase in the total number of colonies in a CFU assay with MPP3 and MPP4 cells from Ebf1<sup>WT</sup> and Ebf1<sup>KO</sup> mice. The total number of colonies does not take into account the size or type of the colonies generated and therefore, it is possible that other parameters would reflect a myeloid bias in Ebf1<sup>KO</sup> MPP3 and MPP4 cells. A role of Ebf1 in cell proliferation has been previously shown in pro-B cells (Győry et al., 2012), and Ebf1 has been implicated in participating with Myc and Max in distinct transcription modules to regulate DNA replication and cell cycle (Kucinski et al., 2020). Moreover, a hematopoietic enhancer cluster in the Myc locus is bound and regulated by EBF1 (Bahri et al., 2018; Ramamoorthy et al., 2020; Somasundaram et al., 2021).

Insight into the molecular basis for the enhanced myeloid basis was provided by the bulk RNA-seq analysis of MPP3 and MPP4 cells, which showed an increased expression of a large set of genes that have been previously identified as C/EBPa target genes in GMPs (Ye et al., 2015). C/EBPa is a lineage-determining transcription factor critical for the neutrophilic/monocytic cell fate (Zhang et al., 1997; Paul et al., 2015; Pundhir et al., 2018). Notably, the DoT-score analysis of the bulk RNA-seq data, evaluating similarities with cell states among the scRNA-seq landscape of WT LSK cells, indicated that Ebf1<sup>KO</sup> MPP3 and MPP4 cells gain neutrophil–monocytic fate potential. Moreover, the knockdown of Cebpα in Ebf1<sup>KO</sup> MPP3 cells suggests a decreased myeloid differentiation potential relative to the non-targeting control knockdown in Ebf1<sup>KO</sup> MPP3 cells. However, we did not observe an increase in myeloid differentiation of control knockdown Ebf1<sup>KO</sup> MPP3 cells relative to control knockdown Ebf1<sup>WT</sup> MPP3 cells. Thus, further approaches will be needed to address the Cebpα–Ebf1 relationship in MPP3 cells.

The increased expression of C/EBPa target genes, together with the binding of EBF1 to a Cebpα enhancer region upon
Figure 7. EBF1 limits C/EBPa driven myeloid bias in MPP3 and MPP4 cells. (A) Cebpa gene body and regulatory region is displayed, showing EBF1 occupancy 24 and 72 h after EBF1 induction, in pro-B cells and Hoxb8-FL cells; H3K27ac marks in pro-B cells and Hoxb8-FL cells; H3K27me3 marks in pro-B cells.
transient EBF1 induction in Ebf1+/- progenitors and in Hoxb8-FL cells, suggests that EBF1 may antagonize Cebpa expression in hematopoietic progenitors (this study; Kucinski et al., 2020; Li et al., 2018). In particular, transient EBF1 occupancy in Ebf1+/- progenitors has been associated with alternative lineage genes that are silenced in committed pro-B cells (Li et al., 2018). In agreement with a transient binding of EBF1 in HSPCs, Ebf1 has been predicted as a silencer of the +37 kb Cebpa regulatory region in WT pro-B cells. In addition, the over-expression of Ebf1 in MPP3 and MPP4 cells results in a reduction of Cebpa expression. In contrast to our findings, Cebpa expression was not increased upon Ebf1 knock-down in Hoxb8-FL cells, which could be due to a partial depletion of EBF1 and/or due to differences in the chromatin accessibility at the EBF1 binding site (Kucinski et al., 2020). The EBF1 binding site in the Cebpα regulatory region is accessible in Hoxb8-FL cells, but not in Ebf1+/- cKit+ progenitors and MPP3 and MPP4 progenitors (this study; Kucinski et al., 2020; Li et al., 2018). Of note, while C/EBPa is not expressed in B cells, it is expressed in lymphoid progenitors and is required to initiate Ebf1 expression in CLPs (Guo et al., 2018; Barberi et al., 2020), indicating these transcription factors function in a transcriptional regulatory network. In this network, we hypothesize that EBF1 restrains the expression of C/EBPa to regulate the balance between B lymphoid and myeloid potential in MPPs. According to this scheme, changes in the concentrations of either of these countering transcription factors would shift the balance of myeloid versus lymphoid lineage bias. In support of an opposing function of EBF1 and C/EBPa, ectopic expression of C/EBPa in committed B cells leads to their rapid transdifferentiation to macrophages, which includes the downregulation of EBF1 and PAx5 (Xie et al., 2004; Laiosa et al., 2006).

Changes in the function of transcription factors by altered gene dosage have been extensively shown for the myeloid determinant PU.1 which regulates B-lymphoid versus myeloid cell fates in a graded manner (Laslo et al., 2006). High concentrations of PU.1 favor the macrophage developmental program, whereas low concentrations of PU.1 are important for B-lymphoid development (DeKoter and Singh, 2000; Pang et al., 2018). In this context, low levels of PU.1 are achieved by repression of PU.1 by Gfi1 (Laslo et al., 2006). EBF1 functions also in a dose-dependent manner as Ebf1+/- heterozygosity results in a diminished B-lymphoid output (Lukin et al., 2011; Åhsberg et al., 2013). The low-level expression of Ebf1 in MPPs may involve its regulation by Ikaros and PU.1, whereas the high-level expression of Ebf1, observed in CLPs, may be governed by additional IL7Ra signaling and regulation by E2A and FOXO1 (Seet et al., 2004; Dias et al., 2005; Kikuchi et al., 2005; Lin et al., 2010; Mansson et al., 2012). Resembling the dosage-dependent function of EBF1 on B cells, low-level expression of EBF1 in MPPs versus high-level expression in CLPs and pro-B cells may account for the disparate function of EBF1 in multipotent and B lineage-restricted hematopoietic progenitors.

The enrichment of the myeloid GMP signature was associated specifically with Ebf1-deficient MPP3 enhancers and not with Ebf1-deficient MPP4 enhancers. Moreover, the set of EBF1-dependent sites of chromatin accessibility at B-lymphoid genes was found to be specific to MPP3 cells. Since EBF1 is expressed at similarly low levels in both cell populations, the question arises as to why the EBF1 dependence of B-lymphoid enhancer priming is detected only in MPP3 progenitors. The differential chromatin accessibility sites of Ebf1WT and Ebf1KO MPP4 cells displayed strong PU.1 binding and were independent of EBF1 re-expression, suggesting that other pioneer factors, such as PU.1, could maintain accessibility at these chromatin sites. In addition, MPP3 and MPP4 (LMPP) cells also contain similar levels of Cebpa and SplI (PU.1) transcripts; however, they differ in the abundance of Ikaros transcripts. Ikaros is an important regulator of all lymphoid lineages, and Ikaros deficiency results in impaired lymphoid cell differentiation at the LMPP stage (Winandy et al., 1995; Yoshida et al., 2006). Therefore, it is possible that Ikaros acts independently of EBF1 in regulating the lymphoid potential in MPP4/LMPP cells. However, ectopic expression of EBF1 can rescue the impaired B cell differentiation of Ikaros-deficient mice (Reynaud et al., 2008), suggesting overlapping regulatory functions of both transcription factors.

Finally, we find that the hematopoietic deletion of Ebf1 leads to impaired HSC quiescence and activation of HSCs, along with a diminished HSC re-population capacity. While we observed an increase in myeloid-biased MPP3 cells and CD11b+ cells in the BM of Ebf1KO mice, this increase was not observed in recipients of Ebf1KO HSPCs. This is likely due to the dominant effect of an impaired self-renewal capacity compared to a shift in lineage bias of Ebf1KO HSPCs under homeostatic conditions. Although the effects of the Ebf1 deletion on HSC function are compelling and robust, they are likely secondary to the EBF1-driven changes in MPP3 and MPP4 cells because Ebf1 expression was not detected in HSCs. Recently, acute immune stimulation by LPS has been shown to induce persistent changes in the myeloid potential of exposed HSCs (de Laval et al., 2020). Moreover, the effects on HSC functionality are not due to the absence of B cells as no
obvious changes were observed in Ighm<sup>Tm1</sup> mutant mice, in which a signaling-incompetent Ig<sub>µ</sub> heavy chain results in an early block of B cell differentiation (Kitamura et al., 1991). Therefore, we favor the view that the enhanced myeloid bias and/or expansion of the myeloid compartment in the BM leads to the activation of HSCs. Consistent with this view, inflammatory signaling has been found to result in activation of HSCs (Mirantes et al., 2014; McCabe and MacNamara, 2016; Mitroulis et al., 2018; Hormaechea-Agulla et al., 2020; Bousounis et al., 2021). However, the addition of Ebf<sup>WT</sup> and Ebf<sup>KO</sup> BM fluid to WT HSCs or co-culture with myeloid cells (data not shown) did not influence HSC proliferation, suggesting that the effects of EBF1 deficiency on HSC homeostasis may be independent of changes in soluble cytokines.

In conclusion, our study emphasizes the importance of EBF1 in regulating myeloid/lymphoid fate bias in MPPs by constraining C/EBPα-driven myelopoiesis and priming the B-lymphoid fate. This expression pattern is reflected in human hematopoietic progenitors. Upon aging, EBF1 expression decreases (Lescale et al., 2010; Lescale et al., 2015; Riley, 2013) and hematopoiesis is shifted toward an enhanced myeloid output (Snoeck, 2013; Yamamoto et al., 2013; Lee et al., 2019; Dorshkind et al., 2020; Mejia-Ramirez and Florian, 2020). This altered EBF1 expression may be involved in aging-associated changes of lymphoid and myeloid trajectories.

**Materials and methods**

**Mice**

Ebf<sup>1</sup>-floxed mice were generated as detailed in Győry et al. (2012). Tie2-Cre (Kisanuki et al., 2001) and Ighm<sup>Tm1</sup>, also known as muM<sup>tt+</sup>-/ mice (Kitamura et al., 1991), were purchased from The Jackson Laboratory; stock no. 008863 and 002288, respectively. Only male heterozygous Tie2<sup>Cre</sup>Ebf<sup>1</sup>-<sup>WT</sup> mice were used for breeding, as litters with female Tie2-Cre drivers demonstrated germline Ebf<sup>1</sup> deletion. muM<sup>tt+/-</sup> littermates were used as controls for the Ighm strain and flox/flox or +/Cre littermates were used as controls for the Tie2-Cre driving Ebf<sup>1</sup>-<sup>WT</sup> strain. C57Bl/6<sup>6j</sup> (CD45.1, CD45.2, or CD45.1.2) mice were bred in-house. 8–14-wk-old animals of both sexes were used for experiments. All mice were maintained, bred, and analyzed on the C57Bl/6<sup>6j</sup> background in the animal facility of the Max Planck Institute of Immunobiology and Epigenetics under specific pathogen-free conditions. Animals were housed on a 14-h/10-h light-dark cycle and provided with standard rodent chow and water ad libitum. All animal procedures were performed in compliance and approved by responsible Animal Welfare Committees (Regierungspräsidium Freiburg, Nr. 35-9185.81/G-18/104, Nr. 35-9185.81/G-17/65).

**Cell suspensions and flow cytometry (cell cycle staining)**

Single-cell suspensions were prepared from BM (femora, tibiae, and ilia bones) by crushing them in cold PBS-3%FCS using a pestle and mortar followed by erythrocyte lysis using RBC lysis buffer (BioLegend). For cell sorting, BM cells were depleted using a magnet, enriching for the lineage-negative fraction. Briefly, BM cells were incubated for 15 min with a biotin antibody cocktail (2 ml/mouse). Labeled cells were then incubated for 15 min with 50 µl of magnetic MojoSort Streptavidin Nanobeads (BioLegend). In the case of HSPC collection for Western blot, cells were prepared in cold PBS-1%BSA, and labeled cells were incubated with 400 µl/mouse of magnetic washed anti rat IgG-coated Dynabeads provided in the Dynabeads Untouched Mouse CD4 Cells Kit (Invitrogen). Lineage-negative cells were subsequently incubated for 30 min with the appropriate antibody cocktail to purify HSPCs. Following cell surface staining, intracellular staining of Ki67 and Hoechst was performed by fixing and permeabilizing with the eBioscience Foxp3/Transcription Factor Staining Buffer Set (eBioscience). For the homing assay and BM fluid culture assay, cells were labeled with 5 µM CTY (Invitrogen) for 20 min at room temperature. Surface antibodies were diluted in PBS-3%FCS, intracellular antibodies were diluted in permeabilization buffer (eBioscience). Between each step, cells were washed twice with PBS-3%FCS or permeabilization buffer and centrifuged at 400 g at 4°C. Flow cytometry analysis was performed on an LSRFortessa instrument and cell sorting on FACSAria III, FACSAria Fusion, or FACSymphony instruments (BD Biosciences). Data were analyzed with FlowJo Software V.10 (TreeStar). Antibodies were purchased from BD Biosciences, BioLegend, or Invitrogen; a list of antibodies used and population definitions can be found in Table S1, gating strategy for HSPCs in Fig. S1.

**BM AdT**

Primary HSC AdT experiments were performed by injecting 150 donor HSCs (CD45.2) mixed with 500,000 supportive total BM cells (CD45.1) into the tail vein of recipients (CD45.1.2). Secondary HSC AdT experiments were performed by injecting 3 × 10<sup>6</sup> total BM cells from primary recipients at 16 wk after AdT into the tail vein of secondary recipients (CD45.1.2).

Reverse AdT experiments were performed by injecting 5 × 10<sup>5</sup> total BM cells from WT (CD45.1) donors into irradiated Ebf<sup>WT</sup> and Ebf<sup>KO</sup> (CD45.2) recipient mice.

Recipient mice (8–10-wk-old) were lethally irradiated with 9.5 Gy using a γ irradiator with a 137Cs source (Biobeam GM 8000) and injected within 24 h of irradiation. One donor was injected into two to three recipient mice, and for reverse AdT experiments, a single WT (CD45.1) donor was used per experiment. Donor hematopoietic reconstitution was monitored every 4 wk in peripheral blood by flow cytometry and sacrificed at 16–24 wk after AdT.

Homing assays were performed by injecting 50,000 sorted LSKs from Ebf<sup>WT</sup> and Ebf<sup>KO</sup> (CD45.2) mice, labeled with CTY (5 µM; Invitrogen) into recipients (CD45.1). Recipients were sacrificed 14–16 h after injection. Spleen, peripheral blood, and BM were analyzed for CTY donor contribution.

**5-FU treatment**

Mice were injected intraperitoneally with 5-FU (Sigma-Aldrich) at 150 mg/kg every 10 d over the course of 100 days. 5-FU is dissolved in DMSO, in 10% of the final volume, then immediately 5-FU is dissolved in DMSO, in 10% of the final volume, then immediately and centrifuged at 400 g at 4°C. After 5-FU is dissolved, it is filtered through 0.2 µm and resuspended to the appropriate final volume of sterile PBS. Control mice were injected with 100 µl PBS per 10 g of mouse.
Western blots
Cells were sorted directly into cold PBS-1%BSA and the cell pellet was frozen in liquid nitrogen before further processing. For whole cell extracts, cells were incubated for 15 min on ice in 150 mM NaCl, 1% NP-40, 50 mM Tris–HCl, pH 7.4, 1 mM EDTA, 0.25% deoxycholic acid, supplemented with complete protease inhibitor cocktail (Roche). After lysis, samples were sonicated in a water bath with 12 strokes and 50% output and incubated for another 15 min on ice. Cell debris was then removed by centrifugation at 16,000 g for 10 min at 4°C. Samples were denatured at 95°C for 5 min in NuPAGE LDS Sample Buffer (Invitrogen) before SDS-PAGE using NuPAGE 4–12% Tris-Glycine gels (Invitrogen) and then transferred to 0.45 μM polyvinylidene difluoride membranes (Bio-Rad). Membranes were blocked for 1 h at room temperature with 5% milk in PBS with 0.1% Tween (PBST). The following antibodies were used: anti-H3 (Abcam) and anti-EBF1 (in-house), which were raised against the first 14 amino acids of EBF1, as detailed in Roessler et al. (2007). Primary antibody incubation was performed overnight at 4°C, and secondary incubation with HRP conjugated anti-rabbit antibody (Invitrogen) was performed the next day at room temperature. Membranes were washed with PBST thrice between each antibody incubation. Antibodies were diluted in PBST-5% milk.

Bulk CFU assays
150 LT-HSCs, HSCs, MPP3, or MPP4 cells were cultured in Methylcellulose Complete Media (R&D systems) in triplicate. CFU colonies were counted after 10–12 d in culture. Cultures were maintained at 37°C at 5% CO₂.

BM fluid cultures
300 HSCs labeled with CTY (5 μM; Invitrogen) were sorted into 96-well U-bottom plates. HSCs were cultured for 48 h in Stempro-34 SFM (Gibco) supplemented with cytokines stem cell factor (SCF; 50 ng/ml; Peprotech) and Thrombopoietin (TPO; 10 ng/ml; Peprotech) and the indicated amount of BM fluid. BM fluid was harvested by crushing clean bones in 400 μl cold PBS and cells were removed by centrifugation at 400 g for 5 min. Supernatant was collected with Microtainer SST tubes (BD Biosciences) at 12,000 g for 1 min and then stored at −80°C. HSC division was analyzed on day 2 and day 5; see Table S1 for antibody panel.

Myeloid differentiation assay and siRNA knockdown of Cebpa
10 MPP3 and MPP4 cells were sorted into 96-well U-bottom plates. Cells were cultured for 10 d in Stempro-34 SFM (Gibco) supplemented with 10% StemPro Nutrient Supplement and cytokines SCF (50 ng/ml; Peprotech), IL-3 (5 ng/ml; Peprotech), Erythropoietin (EPO; 2 μU/ml; BioLegend), IL-6 (10 ng/ml; Peprotech), and GM-CSF (10 ng/ml; Peprotech). Media with cytokines were refreshed on day 6. Differentiation was assessed by flow cytometry on days 10–12 (see Table S1 for antibody panel). Cultures were maintained at 37°C at 5% CO₂. Lineage-negative cells from Ebf1WT and Ebf1KO mice were transfected with DharmaFECT1 (Dharmacon), with 25 nM control non-targeting or Cebpa siRNA (D-001810-10-05, L-040561-00-0005; Dharmacon) together with 25 nM siGLO Green (Dharmacon). After 48–72 h, 10 MPP3 and MPP4 siGLO Green-positive cells were sorted into 96-well U-bottom plates as described above.

Retroviral expression of EBF1
Retroviral transduction was performed as previously described (Treiber et al., 2010). Briefly, retrovirus was produced by transient transfection of Plat-E cells with retroviral constructs pMYs-Ebf1-IRES-EGFP or pMYs-IRES-EGFP. Lineage-negative cells from Ebf1WT and Ebf1KO mice were resuspended in retroviral supernatant at 1-2 × 10⁶ cells/ml and centrifuged at 2600 rpm for 2–3 h in 6-well plates (2 ml/well). After spin infection, 70% of the viral supernatant was removed and cells were cultured with Stempro-34 SFM (Gibco) supplemented with cytokines SCF (50 ng/ml; Peprotech), TPO (10 ng/ml; Peprotech), and Flt3L (10 ng/ml; Peprotech). After 40 h, GFP-positive MPP3 and MPP4 cells were sorted for B cell differentiation assay or ATAC-seq.

B cell differentiation assay
OP9-feeder cells were grown in MEM Alpha (Sigma-Aldrich) supplemented with 20% FCS, 1% sodium pyruvate, and 1% penicillin/streptomycin, L-glutamine. Feeder cells were grown until 80–90% confluent and treated with 10 μg/ml mitomycin C (Sigma-Aldrich) for 2 h. Cells were washed four times with PBS, detached with trypsin–EDTA (Gibco) at 37°C for 5 min, and seeded at ~3,500 cells per well of a 48-well plate, 48 h before the experiment. 50 MPP3 and MPP4 cells were FACs sorted into the OP9-seeded 48-well plates. Cells were cultured in Opti-MEM medium (Gibco) supplemented with 4% FCS, 1% penicillin/streptomycin, L-glutamine, 50 μM β-mercaptoethanol, SCF (10 ng/ml; Peprotech), Flt3L (10 ng/ml; Peprotech), and IL-7 (5 ng/ml; Peprotech). Media was refreshed on day 6. Differentiation was assessed by FACs on days 6, 9, and 12.

ATAC-seq
Generation
ATAC-seq was performed as described in Corces et al. (2017) with some variations. Briefly, 5,000 FACs-sorted cells were sorted into 100 μl PBS-1%FCS and spun down at 500 g for 10 min at 4°C and the supernatant was removed. The cell pellet was resuspended in a lysis/transposition buffer (1 μl Tagmentase, 10 μl 2× TD buffer [Diagenode], 0.01% digitonin [Promega], and 0.3% PBS) and incubated for 30 min at 37°C at 700 rpm. Reactions were cleaned up with Zymo DNA Clean and Concentrator –5 columns (Zymo Research). The remainder of the ATAC-seq library preparation was performed as described in Buenrostro et al. (2013); Buenrostro et al. (2015) using Nextera DNA UD Indexes (IDT). Libraries were cleaned up using AmpureXP beads (Beckman Coulter), and double-sized selection was performed by removing large fragments that precipitate with 0.5× volume and purifying fragments that precipitate with 1.8× volume ratio. Two biological replicates were used for ATAC-seq from LT-HSCs, MPP3, and MPP4 cells from Ebf1WT and Ebf1KO mice. Four biological replicates were used for ATAC-seq from MPP3 cells from Ebf1KO mice and transduced with empty or EBF1 overexpression vector. ATAC-seq
paired-end 100 bp reads were generated using the Illumina NovaSeq 6000 system. Samples were first sequenced at shallow depth (~5 million reads) to assess the quality before sequencing at a depth of 50 million reads. Reads from both flowcells were merged at the fastq level.

**Analysis**

The Nextera transposase adapters were trimmed from the raw FASTQ using Cutadapt (version 2.10 [Martin, 2011]) using flags -A CTGTCTCTTATA -a CTGTCTCTTATA. Alignments against the mm10 reference genome (ensembl release 91) were created using the DNA-mapping module implemented in snakePipes (version 2.5.1 [Bhardwaj et al., 2019]) using the --dedup --mapq 5 flags. Data quality was assessed using FastQC (version 0.11.9 [Andrews, 2010], MultiQC (version 1.8 [Ewels et al., 2016]), and deepTools (version 3.3.2 [Ramírez et al., 2016]).

Analysis downstream was performed using a custom pipeline (https://github.com/maxplanck-ie/ATACofthesnake) implemented in snakemake (version 6.0.0 [Koster and Rahmann, 2012]). Briefly, fragments mapping to read attracting regions (Amemiya et al., 2019) or exceeding 150 bp (to capture nucleosome-free regions) were filtered out using the alignmentSieve module in deepTools. Subsequently, peaks were called using MACS2 (version 2.2.7.1 [Zhang et al., 2008]) with the settings --nomodel, --shift -75, --extsize 150 -g 1.87e9 -q 0.01 and --keep-dup all, and a union of peaks per cell type (MP3P and MPF4) was created using bedtools merge (version 2.30.0 [Quinlan and Hall, 2010]). Genes were annotated to the closest gene using Uropa, allowing a 20-kb distance upstream and 10-kb distance downstream. Count matrices were created using the multiBamSummary module in deepTools, and differential accessible regions were called using the glmQLFit and glmQLFT functions (with design: --genotype) from the edgeR package (version 3.32.0 [Robinson et al., 2010]) implemented in R (version 4.0.3 [R Core Team, 2020]). Heatmaps were visualized using the plotHeatmap function from deepTools after normalizing the alignments with the reads per kilobase per million mapped reads (RPKM) method implemented in the bamCoverage function from deepTools. Motif enrichment for DA peaks was performed with the same method (MEME-suite, version 5.0.2 [McLean and Bailey, 2010]), using the non-redundant vertebrate motif database from JASPAR [Fornes et al., 2019]. As control sequences, the reciprocal differential peak sets were used. For the alternative enrichment analysis, a shuffled DA peak set was used by using the --control --shuffle-- flag. Motif clustering was done using the ClusterMotifs module implemented in TOBIAS (version 0.12.11 [Bentsen et al., 2020]), using flags -t 0.4, --dist_method sec gor --clust_method complete. Co-motif enrichment was assessed by counting overlapping consensus motif occurrences in either the differential peak set or 1,000 background sets of equal size containing a random subsample of all peaks. Consensus motif occurrences were considered overlapping if they were not further than 50 bp apart. Aggregation plots were created with the TOBIAS workflow, using the ATACorrect and PlotAggregate functions with default parameters. GSEA was performed using the prerank module implemented in GSEApy (version 0.10.5 [Subramanian et al., 2005]). A ranked list was created by subsetting ATAC peaks overlapping with predicted enhancer regions (see below) and by omitting peaks without any annotated gene. A rank value was calculated per ATAC peak by multiplying the log2 fold change with the -log10 FDR. The final list was deduplicated by retaining the gene symbol associated with the highest log2 fold change. The CLP and GMP gene sets were created by extracting 300 significant genes with the lowest FDR value (FDR < 0.05 and log2 fold change >2 or -2). GSEA was performed using 1,000 permutations. GO was performed using clusterProfiler (version 4.0.5 [Wu et al., 2021]) implemented in R on genes annotated to DA peak sets (KO Reduced and KO Gained, for MP3P and MPF4 cells) with FDR < 0.1. Read coverages for track visualization were obtained via deepTools bamCoverage using RPKM normalization (binsize 5bp). For peak calling and visualization, replicates were merged per condition at the bam level.

**Public data integration**

Publicly available data were collected for EBF1 ChIP experiments in 24 h, 72 h, and pro-B cells samples (GSE107242 [Li et al., 2018]), and Hoxb8-FL cells (GSE146128 [Kucinski et al., 2020]). Publicly available data were collected for H3K27Ac ChIP experiments in pro-B cells (GSE107242 [Li et al., 2018]) and in Hoxb8-FL cells (GSE146128 [Kucinski et al., 2020]), and for H3K4Me1, H3K4Me3, and H3K27Ac ChIP experiments in HSPC populations (GSE59636 [Lara-Altisao et al., 2014]). Publicly available data were collected for C/EBPα ChIP experiments in LSK, pre-GMs, and GMP cells (GSE89767, GSE43007 [Hasemann et al., 2014; Pundhir et al., 2018]). Alignments against the mm10 reference genome (ensembl release 91) were performed using the DNA-mapping mode implemented in snakePipes (version 2.5.1), using the flags --trim, --dedup, --mapq 3, and --fastqc. Subsequently, the ChIP-seq mode from snakePipes was invoked between 0 and 1 per ChIP over all the regions. Enhancer prediction was performed on replicate-merged bam (from histone modification ChIPs) files using the Condition-specific Regulatory Units Prediction (CRUP) algorithm (version 1.0 [Ramisch et al., 2019]), implemented in R (version 3.6.3) using default parameters. Only regions with a probability >0.8 were included for downstream analysis. Read coverages for track visualization were obtained via deepTools bamCoverage using RPKM normalization (binsize 5bp) and subsequently plotted using pyGenomeTracks (3.3 [Ramírez et al., 2018]). C/EBPα ChIP was analyzed further by unifying peaks over replicates. Genes were annotated to the closest gene using Uropa, allowing a 20-kb distance upstream and 10-kb distance downstream. Heatmaps were visualized using the plotHeatmap function from deepTools after normalizing the alignments with the RPKM method.
implemented in the bamCoverage function from deepTools (binsize 5 bp).

**RNA extraction and RT-qPCR**

Total RNA was prepared using the RNA Clean & Concentrator ~5 RNA isolation kit (Zymo Research) according to the manufacturer’s instructions and treated with DNase (Qiagen). cDNA was synthesized with Oligo(dT)18 primers using the SuperScript II Reverse Transcriptase Kit (Invitrogen). RT-PCR was performed using the GoTaq Green master mix (Promega). The external primers used were Ebf1F: 5'-CCAGCCCGTGAGATGAAGATGAGC GAG-3', Ebf1R: 5'-TCTTTCTACATGGGACCAATCAT-3' and Ash2lF: 5'-CACCCTTGAATAGAAGCTGTC-3', Ash2lR: 5'-TCC GCCAGTGACCTGGCTGTC-3', yielding products ~1,000–1,500 bp. The first round of PCR was carried out as follows: initial denaturation at 95°C for 2 min, followed by 10 cycles of denaturation at 95°C for 30 s, annealing at 62°C for 20 s, and extension at 72°C for 1.5 min, followed by a final extension at 72°C for 5 min. The second round of amplification was carried out using 2 μl of the first PCR product as a template and using the primers Ebf1R: 5'-GATGAGGCGCACATAGAAATCCTGT-3' and Ash2lR: 5'-CTGGTCCCAGCCCTTAGGTAACC-3', which amplify a 150–200 bp product. The same conditions and amplification program were followed as for the first round, except the extension time was 20 s, with an increasing number of cycles from 25 to 40 cycles. PCR products were resolved on 2% agarose gels.

qRT-PCR was performed with Taqman Gene Expression Assays (FAM; Thermo Fisher Scientific) with Taqman Fast Advanced Master Mix (Thermo Fisher Scientific) using target probes against Ebf1 (Mm01288947_g1) or Cebpa (Mm01265914_s1) normalized to Act2b (Mm00607939_s1 Actb). Reactions were performed in duplicate and run on a StepOne Real-Time PCR system (Applied Biosystems).

**Bulk RNA-seq**

MPP3 and MPP4 cells were sorted into 50 μl of extraction buffer. Total RNA was prepared using the Arcturus PicoPure RNA isolation kit (Applied Biosystems) according to the manufacturer’s instructions and treated with DNase (Qiagen). Four biological replicates were used for RNA-seq from MPP3 and MPP4 cells from Ebf1WT and Ebf1KO mice. RNA-seq libraries were prepared by the deep-sequencing facility (Max Planck Institute-Freiburg) using NEBNext SingleCell/Low Input RNA library prep kit (NEB). Bulk RNA-seq paired-end 100 bp reads were generated using the Illumina NovaSeq 6000 system at a depth of 50 million reads.

**Analysis**

Raw FASTQ files were mapped against the mm10 reference genome using the mRNA-seq module implemented in snakePipes (version 2.5.1 [Bhardwaj et al., 2019]), using the --trim flag. Briefly, reads were aligned using STAR (version 2.7.3a [Dobin et al., 2013]), followed by expression count quantification with featureCounts (version 2.0.0 [Liao et al., 2014]). Data quality was assessed using Deeptools QC (version 3.3.2 [Ramirez et al., 2016]). Genes with an average expression higher than 100 counts in any condition were selected for further analysis.
3.1.0) with default settings. Cells with <1,200 detected genes were excluded from further analysis. Count data were logarithmized and each cell was analyzed for the expression of Xist and Y-chromosome genes. Cells with Xist expression but no Y-chromosome gene expression were assigned female and vice versa. The small minority of cells expressing both Xist and Y-chromosome genes were excluded from the analysis as suspected doublets, and cells lacking expression of Xist and Y-chromosome genes were excluded. Data for the remaining 59,505 cells were used to generate a common scRNA-seq landscape as follows. 7,000 top variable genes (excluding Ebf1, Xist, Y-chromosome genes, and cell-cycle associated genes [as in Dahlin et al., 2018]) were used to compute 50 PCs. Batch effects were removed using the Harmony method (Korsunsky et al., 2018). The corrected components were used to identify eight nearest neighbors, which served as a basis for clustering (leiden algorithm) and UMAP (Traag et al., 2019; McNmes et al., 2020; https://github.com/TomKellyGenetics/leiden). Clusters were manually annotated using commonly used marker genes (see Dahlin et al., 2018) and the integration of Nestorowa HSPC scRNA-Seq data (Nestorowa et al., 2016; see below). Cell cycle scores were assigned using the score genes_cell_cycle function in the Scanypy framework and gene sets from Macosko et al. (2015). Differential expression was performed using the pseudobulk method approach (Lun and Marioni, 2017) treating each cluster as a cell state (Crowell et al., 2020). Pseudobulk profiles were compared between KO and WT cells using edgeR (Robinson et al., 2010) quasi-likelihood F-test test, including covariates for sex and experimental batch. DE genes were called using the thresholds: FDR of 0.1 and |log2(Fold Change)| of 0.2. To find genes DE in the wider population of undifferentiated progenitors, we performed analogous analysis for combined clusters 0, 1, 2, and 9. Enrichment analysis was performed on DE genes separately for up- and down-regulated subsets using the Enrichr framework and GSEApy interface (Moehoa et al., 2003; Subramanian et al., 2005; Chen et al., 2013; Kuleshov et al., 2016). DoT score was calculated as previously described (Kucinski et al., 2020) using point of origin indicated in the respective figures.

The human fetal liver data were obtained from ArrayExpress (E-MTAB-7407) and processed as previously described (Kucinski et al., 2020). Briefly, cells expressing <1,000 genes and with 10% counts aligned to the mitochondrial genes were filtered out. Based on the original annotation provided by the authors, only cells belonging to blood and endothelial populations were considered. 7,984 highly variable genes were used to compute 50 PCs, and batches with <400 cells were excluded. The remaining cells were integrated with batch-balanced k-nearest neighbor method (Polaniski et al., 2019).

Smart-Seq2 data integration

We integrated immunophenotypically defined populations from the Nestorowa Smart-Seq2 dataset (Nestorowa et al., 2016; projected data) into our 10x Genomics scRNA-seq data (reference) using the Cellproject (https://github.com/Iwo-K/cellproject) framework. Briefly, Smart-Seq2 data were log-normalized analogously to the 10x data and corrected to match the expression space of the reference data using the Seurat batch correction method (Stuart et al., 2019). Subsequently, the reference PCA rotation matrix was used to compute PCs for the projected data, and these were used to identify 15 nearest neighbors between Smart-Seq2 and 10x data. Finally, we used the nearest-neighbor regression to predict PC coordinates in the Harmony-corrected PCA space (basis for our integrated landscape) and fitted Smart-Seq2 data into the reference UMAP coordinates.

Statistical analysis

Data are presented as boxplots or medians, as indicated in the figure legends. Details of statistical tests and the exact replicate numbers are reported in the figure legends and/or figures. Except for sequencing analysis, all statistical analyses were performed using Prism 8 software (GraphPad).

Code availability

All code used for data analysis in this manuscript are available at https://github.com/AurelieLen/EBF1_MPPs. ATAC analysis available at https://github.com/maxplanck-ie/ATACofthesnake. DoT score available at https://github.com/Iwo-K/dotscore. Cellproject available at https://github.com/Iwo-K/cellproject.

Online supplemental material

Fig. S1 shows data related to Fig. 1 and Fig. 2. Fig. S2 shows data related to Fig. 3. Fig. S3 shows data related to Fig. 4 and Fig. 5. Fig. S4 shows data related to Fig. 6. Fig. S5 shows data related to Fig. 7. Table S1 contains a list of antibodies, reagents, and flow cytometry staining panels used in the study. Table S2 contains a list of the DA peak sets identified from bulk ATAC-seq of MPP3 and MPP4 cells from Ebf1WT and Ebf1KO mice. Table S3 contains a list of differentially regulated genes and normalized read counts identified from bulk RNA-seq of MPP3 and MPP4 cells from Ebf1WT and Ebf1KO mice, as well as the DEseq2 results table of the comparison of MPP3 and MPP4 cells from Ebf1WT mice. Table S4 contains a list of genes used for the gene sets G2-M and S cell cycle phase, related to Fig. 6 H.

Data availability

Data sets generated in this study are available as a superseries in the GEO database under accession code GSE189078. Individual series can be obtained as follows: GSE188884 (RNA-seq), GSE189049 (ATAC-seq), and GSE189051 (scRNA-seq). All other data supporting the findings of this study are available at DOI: 10.5281/zenodo.6962483.

Acknowledgments

We are grateful to E. Trompouki, N. Cabezas-Wallscheid, and C. Pessoa Rodrigues for critical feedback on the manuscript, and to M. Rott for the assistance with the preparation of figures. The authors thank I. Falk, D. Komandin, and H.-J. Schwarz for technical assistance and members of Grosschedl, Trompouki, and Cabezas-Wallscheid laboratories for discussions. We thank the Max Planck Institute of Immunobiology and Epigenetics core facilities for Deep Sequencing, Mouse, and Flow Cytometry (especially K. Schuldes) for their invaluable assistance.
I. Kucinski and B. Göttgen were supported by Wellcome (206328/Z/17/Z and 203151/Z/16/Z), Blood Cancer UK (18002), Cancer Research UK (C11623/A21762) and UKRI Medical Research Council (MC_PC_17230). Work in the Grosschedl laboratory is supported by funds from the Max Planck Society. Open Access is supported by funds from the Max Planck Society. Work in the Grosschedl Laboratory is supported by funding provided by the Max Planck Society.

Author contributions: A. Lenaerts designed and performed experiments. M. Derecka performed CFU assay colony identification. A. Lenaerts, I. Kucinski, P. Cauchy and W. Deboutte performed bioinformatics analysis. A. Lenaerts analyzed bulk RNA-seq experiments with input from I. Kucinski, P. Cauchy, and W. Deboutte. W. Deboutte analyzed ATAC-seq experiments. I. Kucinski performed single-cell RNA-seq and DoT score analysis. Conceptualization: A. Lenaerts, M. Derecka and R. Grosschedl. Writing: A. Lenaerts and R. Grosschedl with input of I. Kucinski and W. Deboutte. Supervision: M. Derecka, T. Manke, B. Göttgens and R. Grosschedl.

Disclosures: The authors declare no competing interests exist.

Submitted: 7 December 2021
Revised: 23 June 2022
Accepted: 3 August 2022

References

Åhsberg, J., J. Ungerbäck, T.Strid, E.Welinder, J.Stjernberg, M.Larsson, H.Qian, and M. Sigvardsson. 2013. Early B-cell factor 1 regulates the expansion of B-cell progenitors in a dose-dependent manner. *J. Biol. Chem.* 288:33449–33461. https://doi.org/10.1074/jbc.M113.506261

Amemiya, H.M., A. Kundaje, and A.P. Boyle. 2019. The ENCODE blacklist: Identification of problematic regions of the genome. *Sci. Rep.* 9:9354. https://doi.org/10.1038/s41598-019-45839-2

Andrews, S. 2010. FastQC: A quality control tool for high throughput sequence data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Bahr, C., L. von Paleske, V.V. Uslu, S. Remeseiro, N. Takayama, S.W. Ng, A. Ahmed, R.L. Geissler, Z. Ke, B.M. Idler, et al. 2021. PU.1 enforces quiescence and limits hematopoietic stem cell expansion during inflammatory stress. *J. Exp. Med.* 21:202001619. https://doi.org/10.1084/jem.20212437

Buessmann, L.H., A. Schubert, T.P. Vu Manh, L. De Andreis, S.C. Desbordes, M. Parra, T. Zimmermann, F. Rapino, J. Rodriguez-Ubreva, E. Ballestar, and T. Graf. 2009. A robust and highly efficient immune cell reprogramming system. *Cell Stem Cell.* 5:554–566. https://doi.org/10.1016/j.stem.2009.10.004

Cabezás-Wallscheid, N., D. Klimmek, J. Hansson, D.B. Lipka, A. Reyes, Q. Wang, D. Weichenhan, A. Lier, V. von Paleiske, S. Renders, et al. 2014. Identification of regulatory networks in HSCs and their immediate progeny via integrated proteome, transcriptome, and DNA methylome analysis. *Cell Stem Cell.* 15:507–522. https://doi.org/10.1016/j.stem.2014.07.005

Chavez, J.S., J.L. Rabe, D. Loeffler, K.C. Higa, G. Hernandez, T.S. Mills, N. Ahmed, R.L. Geissler, Z. Ke, B.M. Idler, et al. 2021. PU.1 enforces quiescence and limits hematopoietic stem cell expansion during inflammatory stress. *J. Exp. Med.* 21:202001619. https://doi.org/10.1084/jem.20212437

Chen, E.Y., C.M. Tan, Y. Kou, Q. Duan, Z. Wang, G.V. Meirelles, N.R. Clark, and A. Ma‘ayan. 2013. Enrichr: Interactive and collaborative HTML5 gene list enrichment analysis tool. *Mol. Bioinformatics.* 14:128. https://doi.org/10.2147/molbio.s48041-14-128

Cobaleda, C., W. Jochum, and M. Busslinger. 2007. Conversion of mature B cells into T cells by dedifferentiation to uncommitted progenitors. *Nature* 449:473–477. https://doi.org/10.1038/nature06159

Cooper, S., H. Guo, and A.D. Friedman. 2015. The +37 kb Cebpa enhancer is critical for Cebpa myeloid gene expression and contains functional sites that bind SCL, GATA2, C/EBPα, PU.1, and additional ets factors. *PLoS One.* 10:e0126385. https://doi.org/10.1371/journal.pone.0126385

Corces, M.R., A.E. Trevino, E.G. Hamilton, P.G. Greenside, N.A. Sinnamon-Armstrong, S. Vesuna, A.T. Sapathy, A.J. Rubio, K.S. Montine, B. Wu, et al. 2017. An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat. Methods* 14:959–962. https://doi.org/10.1038/nmeth.4396

Crowell, H.L., C. Soneson, P.L. Germain, D. Calini, L. Collin, C. Raposo, D. Malhotra, and M.D. Robinson. 2020. Muscat detects subpopulation-specific state transitions from multi-sample multi-condition single-cell transcriptomics data. *Nat. Commun.* 11:6077. https://doi.org/10.1038/s41467-020-19894-4

Dahl, R., S.R. Iyer, K.S. Owens, D.D. Cuyler, and M.C. Simon. 2007. The transcriptional repressor GFI-1 antagonizes PU.1 activity through protein–protein interaction. *J. Biol. Chem.* 282:6473–6483. https://doi.org/10.1074/jbc.M607613200

Dahlin, J.S., F.K. Hamey, B. Pijuan-Sala, M. Shepherd, W.W.Y. Lau, S. Nistorowa, C. Weinreb, S. Wolock, R. Hannah, E. Diamantzi, et al. 2018. A single-cell hematopoietic landscape resolves 8 lineage trajectories and defects in Kit mutant mice. *Blood.* 131:e1–e11. https://doi.org/10.1182/blood-2017-12-821413

de Lavall, B., J. Maurizio, P.K. Kandalla, G. Brisou, L. Simonnet, C. Huber, G. Gimenez, O. Matcovitch-Natan, S. Reinhardt, E. David, et al. 2020. C/EBPβ-Dependent epigenetic memory induces trained immunity in hematopoietic stem cells. *Cell Stem Cell.* 26:657–674.e8. https://doi.org/10.1016/j.stem.2020.01.017

DeKoter, R.P., and H. Singh. 2000. Regulation of B lymphocyte and macrophage development by graded expression of PU.1. *Science.* 288:1439–1441. https://doi.org/10.1126/science.288.5470.1439

Derecka, M., J.S. Herman, P. Cauchy, S. Ramamoorthy, E. Lupar, D. Grun, and R. Grosschedl. 2020. EBF1-deficient bone marrow stroma elicits persistent changes in HSC potential. *Nat. Immunol.* 21:261–273. https://doi.org/10.1038/s41590-020-0595-7

Di Tullio, A., T.P. Vu Manh, A. Schubert, G. Castellano, R. Mansson, and T. Graf. 2011. CCAT1/ enhancer binding protein alpha (C/EBPalpha)-induced transdifferentiation of pre-B cells into macrophages involves no overt retrodifferentiation. *Proc. Natl. Acad. Sci. USA.* 108:17016–17021. https://doi.org/10.1073/pnas.112169108

Dias, S., H. Silva Jr, A. Cumano, and P. Vieira. 2005. Interleukin-7 is necessary to maintain the B cell potential in common lymphoid progenitors. *J. Exp. Med.* 201:971–979. https://doi.org/10.1084/jem.20042393
Reynaud, D., I.A. Demarco, K.L. Reddy, H. Schjerven, E. Bertolino, Z. Chen, Repele, A., S. Krueger, T. Bhattacharyya, M.Y. Tuineau, and Manu. 2019. The R Core Team. 2020. R: A language and environment for statistical computing.

Pronk, C.J.H., D.J. Rossi, R. Mansson, J.L. Attema, G.L. Norddahl, C.K.F. Chan, Pundhir, S., F.K. Bratt Lauridsen, M.B. Schuster, J.S. Jakobsen, Y. Ge, E.M. Ranzoni, A.M., A. Tangherloni, I. Berest, S.G. Riva, B. Myers, P.M. Strzelecka, Popescu, D.-M., R.A. Botting, E. Stephenson, K. Green, S. Webb, L. Jardine, EBF1 function in multipotent progenitors https://doi.org/10.1084/jem.20212437

Riley, R.L. 2013. Impaired B lymphopoiesis in old age: A role for inflammatory signaling, Myc expression, and folate metabolism.

Robinson, M.D., D.J. McCarthy, and G.K. Smyth. 2010. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 26:139–140. https://doi.org/10.1093/bioinformatics/btp616

Rodriguez-Fraticelli, A.E., S.L. Wolock, C.S. Weinreb, R. Panero, S.H. Patel, M. Jankovic, J. Sun, R.A. Calgero, A.M. Klein, and F.D. Camargo. 2018. Clonal analysis of lineage fate in native haematopoiesis. Nature. 553: 212–216. https://doi.org/10.1038/nature25168

Roesler, S., I. Gory, S. Imhof, M. Spivakov, R.R. Williams, M. Busslinger, A.G. Fisher, and R. Grosschedl. 2007. Distinct promoters mediate the regulation of Ebf1 gene expression by interleukin-7 and Pax5. Mol. Cell Biol. 27:579–594. https://doi.org/10.1099/mcb.0.1192-06

Seet, C.S., R.L. Brumbaugh, and B.L. Kee. 2004. Early B cell factor promotes B lymphopoiesis with reduced interleukin 7 responsiveness in the absence of E2A. J. Exp. Med. 199:1689–1700. https://doi.org/10.1084/jem.20032202

Snoeck, H.-W. 2013. Aging of the hematopoietic system. Curr. Opin. Hematol. 20:355–361. https://doi.org/10.1097/MOH.0b013e3283623c77

Somasundaram, R., C.T. Jensen, J. Tingvall-Gustafsson, J. Ahberg, K. Okuyama, M. Prasad, J.R. Hagman, X. Wang, S. Sonej, T. Strid, et al. 2021. EBF1 and PAX5 control pro-B cell expansion via opposing regulation of the Myc gene. Blood. 137:3037–3049. https://doi.org/10.1182/blood.2020009564

Sommerkamp, P., S. Altamura, S. Renders, A. Nall, L. Ladel, P. Zeisberger, P.L. Elben, M. Kawaz, M.A. Rieger, N. Cabezas-Wallscheid, and A. Trumpol. 2020. Differential alternative polyadenylation landscapes mediate hematopoietic stem cell activation and regulate glutamine metabolism. Cell Stem Cell 26:722–738.e7. https://doi.org/10.1016/j.stem.2020.03.003

Sommerkamp, P., M.C. Romero-Mulerio, A. Nall, L. Lustin, K. Schönberger, S. Renders, S. Altamura, P. Zeisberger, K. Jacklein, et al. 2021. Mouse multipotent progenitor 5 cells are located at the interface between hematopoietic stem and progenitor cells. Blood. 137:3218–3224.

Spooner, C.J., J.X. Cheng, E. Pujadas, P. Laslo, and H. Singh. 2009. A recurrent network involving the transcription factors PU.1 and Gfi1 orchestrates innate and adaptive immune cell fates. Immunity. 31:576–586. https://doi.org/10.1016/j.immuni.2009.07.011

Stephens, M. 2017. False discovery rates: A new deal. Biostatistics. 18:275–294. https://doi.org/10.1093/biostats/kbx041

Stuart, T., A. Butler, P. Hoffman, C. Hafemeister, E. Papalexi, W.M. Mauck 3rd, Y. Hao, M. Stoeckius, P. Smibert, and R. Satija. 2019. Comprehensive integration of single-cell data. Cell. 177:1888–1902.e21. https://doi.org/10.1016/j.cell.2019.05.031

Subramanian, A., P. Tamayo, V.K. Mootha, S. Mukherjee, B.L. Ebert, M.A. Gillette, P. Paulovich, S.L. Pomeroy, T.R. Golub, E.S. Lander, and J.P. Mesirov. 2005. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression datasets. Proc. Natl. Acad. Sci. USA. 102:15545–15550. https://doi.org/10.1073/pnas.0506580102

Traag, V.A., L. Waltman, and N.J. van Eck. 2019. From louvain to leiden: Guaranteeing well-connected communities. Sci. Rep. 9:5233. https://doi.org/10.1038/s41598-019-41699-2

Treiber, T., E.M. Mandel, S. Pott, I. Györ, S. Färn, E.T. Liu, and R. Grosschedl. 2010. Early B cell factor 1 regulates B cell gene networks by activation, repression, and transcription-independent poised of chromatin. Immunity. 32:714–725. https://doi.org/10.1016/j.immuni.2010.04.013

van Galen, P.A., K. Kreso, E. Wienholds, E. Laurenti, K. Eppert, E.R. Lechman, N. Mhong, K. Hermans, S. Dobson, C. April, et al. 2014. Reduced lymphoid lineage priming promotes human hematopoietic stem cell expansion. Cell Stem Cell 14:94–106. https://doi.org/10.1016/j.stem.2013.11.021

Wang, Y., N. Zolotarev, C.Y. Yang, A. Rambold, G. Mittler, and R. Grosschedl. 2020. A prion-like domain in transcription factor EBF promotes phase separation and enables programming of progenitor chromatin. Immunity. 53:1151–1167.e6. https://doi.org/10.1016/j.immuni.2020.10.009

Weinreb, C., A. Rodriguez-Fraticelli, F.D.C. Camargo, and A.M. Klein. 2020. Lineage tracing on transcriptional landscapes links state to fate during differentiation. Science. 367:eaa83381. https://doi.org/10.1126/science.aaw3381

Weinreb, C., E. R. Mansson, E.M. Mercier, B. Mjöberg, M. Sigvardsson, and C. Murre. 2011. The transcription factors E2A and HER play an act in concert to induce the expression of FOXP1 in the common lymphoid progenitor. Proc. Natl. Acad. Sci. USA. 108:17402–17407. https://doi.org/10.1073/pnas.1117661018

Wickham, H. 2016. ggplot2: Elegant Graphics for Data Analysis. 2nd ed. Springer International Publishing (Use R!), Cham. https://doi.org/10.1007/978-3-319-24277-4

Wilson, N.K., D.G. Kent, F. Buettrner, M. Shetha, I.C. Macaulay, F.J. Calero-Nieto, M. Sanchez Castillo, C.A. Odekoens, E. Diamanti, R. Schulte, et al. 2015. Combined single-cell functional and gene expression analysis resolves heterogeneity within stem cell populations. Cell Stem Cell 16:712–724. https://doi.org/10.1016/j.stem.2015.04.004

Wu, T., E. Hu, S. Xu, M. Chen, P. Guo, Z. Dai, T. Feng, L. Zhou, W. Tang, L. Zhan, et al. 2021. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innovation. 2:100141. https://doi.org/10.1016/j.xinn.2021.100141
Xie, H., M. Ye, R. Feng, and T. Graf. 2004. Stepwise reprogramming of B cells into macrophages. Cell. 117:663–676. https://doi.org/10.1016/S0092-8674(04)00419-2

Yamamoto, R., Y. Morita, J. Ooebara, S. Hamanaka, M. Onodera, K.L. Rudolph, H. Ema, and H. Nakauchi. 2013. Clonal analysis unveils self-renewing lineage-restricted progenitors generated directly from hematopoietic stem cells. Cell. 154:1112–1126. https://doi.org/10.1016/j.cell.2013.08.007

Ye, M., H. Zhang, G. Amabile, H. Yang, P.B. Staber, P. Zhang, E. Levantini, M. Alberich-Jorda, J. Zhang, A. Kawasaki, and D.G. Tenen. 2013. C/EBPa controls acquisition and maintenance of adult haematopoietic stem cell quiescence. Nat. Cell Biol. 15:385–394. https://doi.org/10.1038/ncb2698

Ye, M., H. Zhang, H. Yang, R. Koche, P.B. Staber, M. Cusan, E. Levantini, R.S. Welner, C.S. Bach, J. Zhang, et al. 2015. Hematopoietic differentiation is required for initiation of acute myeloid leukemia. Cell Stem Cell. 17:611–623. https://doi.org/10.1016/j.stem.2015.08.011

Yoshida, T., S.Y. Ng, J.C. Zuniga-Pflucker, and K. Georgopoulos. 2006. Early hematopoietic lineage restrictions directed by Ikaros. Nat. Immunol. 7:382–391. https://doi.org/10.1038/ni1314

Zarnegar, M.A., and E.V. Rothenberg. 2012. Ikaros represses and activates PU.1 cell-type-specifically through the multifunctional Sfp1 URE and a myeloid specific enhancer. Oncogene. 31:4647–4654. https://doi.org/10.1038/onc.2011.597

Zerbino, D.R., N. Johnson, T. Juittemann, S.P. Wilder, and P. Flicek. 2014. WiggleTools: Parallel processing of large collections of genome-wide datasets for visualization and statistical analysis. Bioinformatics. 30:1008–1009. https://doi.org/10.1093/bioinformatics/btt737

Zhang, D.-E., P. Zhang, N.D. Wang, C.J. Hetherington, G.J. Darlington, and D.G. Tenen. 1997. Absence of granulocyte colony-stimulating factor signaling and neutrophil development in CCAAT enhancer binding protein -deficient mice. Proc. Natl. Acad. Sci. USA. 94:569–574. https://doi.org/10.1073/pnas.94.2.569

Zhang, Z., C.V. Cotta, R.P. Stephan, C.G. deGuzman, and C.A. Klug. 2003. Enforced expression of EBF in hematopoietic stem cells restricts lymphopoiesis to the B cell lineage. EMBO J. 22:4759–4769. https://doi.org/10.1093/emboj/cdg464

Zhang, Y., T. Liu, C.A. Meyer, J. Eeckhoute, D.S. Johnson, B.E. Bernstein, C. Nusbaum, R.M. Myers, M. Brown, W. Li, and X.S. Liu. 2008. Model-based analysis of ChIP-seq (MACS). Genome Biol. 9:R137. https://doi.org/10.1186/gb-2008-9-9-r137

Zovein, A.C., J.J. Hofmann, M. Lynch, W.J. French, R.A. Turlo, Y. Yang, M.S. Becker, L. Zanetta, E. Dejana, J.C. Gasson, et al. 2008. Fate tracing reveals the endothelial origin of hematopoietic stem cells. Cell Stem Cell. 3:625–636. https://doi.org/10.1016/j.stem.2008.09.018
Extended characterization of HSC functionality upon Ebf1 deficiency. Related to Figs. 1 and 2. (A) Gating strategies for HSC, MPP2, MPP3, and MPP4 cells. (B) Gating strategies for CLP, ALP, and BLP cells. Representative plots are shown for Ebf1WT and Ebf1KO mice. (C) Frequencies of myeloid cells (CD11b+), T cells (CD4+ CD8+), and B cells (CD19+) within CD45.1 donor cells in the BM of the Ebf1WT and Ebf1KO recipients 16 wk after AdT. Ebf1WT n = 3, Ebf1KO n = 4. (D) Absolute number of CD41+ HSCs in the BM. Ebf1WT n = 11, Ebf1KO n = 9, IghmWT n = 11, and IghmTm1 n = 14. (E) 300 WT CTV-labeled HSCs were cultured in TPO and SCF, with increased concentrations of BM fluid collected from Ebf1WT and Ebf1KO mice. Division index (CTY mean fluorescence intensity [MFI] from day n/CTY MFI from day 0) at day 2 and day 5, of CTV-labeled HSCs cultured with different BM fluid concentrations from Ebf1WT and Ebf1KO mice, n = 6–8. (F) CFU assay performed with 150 sorted HSCs plated in triplicate in methocult. Total number of colonies formed after 10–12 d of plating. Biological replicates Ebf1WT and Ebf1KO n = 2, IghmWT and IghmTm1 n = 3. (G) CFU assay performed with 150 sorted LT-HSCs (LSK CD34− Flt3− CD150+ CD48−) plated in triplicate in methocult. Colony identification after 10–12 d of plating. Biological replicates Ebf1WT and Ebf1KO n = 5. (H) Frequency of CD45.2 donor-derived cells in the peripheral blood within T cells (CD4+ CD8+), ****, P < 0.0001; ***, P < 0.0004; **, P = 0.0013), myeloid cells (CD11b+), ***, P = 0.0002, ***, P < 0.0001; **, P = 0.0010; ***, P < 0.0024) and B cells (CD19+), ***, P = 0.0060; ****, P < 0.0001), during the AdT. Data are represented as mean ± SEM. Ebf1WT n = 33 and Ebf1KO n = 40, IghmWT n = 17, and IghmTm1 n = 19. (C–H) Statistical significance was determined by Mann-Whitney U test. Data are from >2 independent experiments.
Figure S2. Chromatin accessibility changes in Eb1-deficient MPP3 and MPP4 cells. Related to Fig. 3. (A) Comparison of chromatin accessibility changes to gene expression changes in MPP3 versus MPP4 cells, in Eb1WT conditions. The y axis represents log2 fold changes of gene expression, and the x axis represents log2 fold changes of peak read counts in MPP3 versus MPP4 cells in Eb1WT conditions. Left: Dots represent ATAC peaks that overlap their TSS (TSS peaks). Right: Dots represent DA peaks between MPP3 versus MPP4 cells in Eb1WT conditions. Biological replicates for RNA-seq n = 4. Biological replicates for ATAC-seq n = 2. (B) Sequence logos of top ranked enriched motifs underlying the DA peaks in MPP3 and MPP4 cells from Eb1WT and Eb1KO mice. P values are depicted in red, percentages in blue reflect the fraction of hits found in the peak set, percentages in black reflect the fraction of hits found in the background set. (C) Immunoblot analysis of EBF1 protein levels in total cell extracts of 150,000 FACS-sorted cells of lineage-negative GFP+ cells transduced with empty vector or EBF1 vector, biological replicates for ATAC-seq used in Fig. 3 D (n = 4). Eb1WT MPP3 cells and CLPs (Lin−cKitintSca1intIL7R+Flt3+) were used as a positive control. T (CD4+CD8+) cells and Eb1KO Fr. A cells were used as a negative control. Histone H3 was used as a loading control. (D) Heatmaps of co-occurrence counts in KO Gained MPP3 peaks and KO Reduced and KO Gained MPP4 peaks of enriched motifs in these DA peak sets versus co-occurrence counts in all MPP3 and MPP4 peaks (1,000 replicates), respectively. Co-occurrence counts with z-scores > 5 are displayed. (E) Enrichment analysis of genes associated with KO Reduced MPP4 peaks. (F) Genome tracks showing ATAC signal and EBF1 ChIP signal at B-lymphoid related KO Reduced MPP3 peaks annotated to the Cd19 gene. Original data from Li et al. (2018); Kucinski et al. (2020). Source data are available for this figure: SourceData FS2.
Figure S3. Role of Ebf1 deficiency in myeloid-biased transcriptome. Related to Figs. 4 and 5. (A) PCA of bulk RNA-seq data from MPP3 and MPP4 cells of Ebf1 WT and Ebf1 KO conditions; based on DE genes. n = 4 biological replicates. (B) MA plots of MPP3 and MPP4 cells comparing Ebf1 WT and Ebf1 KO transcriptomes. Dots in blue represent genes with an FDR <0.1. The y axis represents shrunken log2 fold change (FC), capped at |log2 FC| 2, and the x axis represents log10 basemean. (C and D) Comparison of chromatin accessibility changes to gene expression changes in MPP3 (C) and MPP4 (D) cells upon Ebf1 deletion. The y axis represents shrunken log2 fold changes of gene expression and the x axis represents log2 fold changes of peak read counts in Ebf1 WT versus Ebf1 KO MPP4 cells. Left: Dots represent DA peaks in MPP3 (C) and MPP4 (D) cells. Right: Dots represent DE genes and ATAC peaks that overlap their TSS (TSS peaks). (E) CFU assay performed with 150 sorted MPP3 and MPP4 cells plated in triplicate in methocult. Total number of colonies formed after 10–12 d of plating. Biological replicates Ebf1 WT and Ebf1 KO n = 6. Statistical significance was determined by unpaired t test. (F) Heatmap showing gene expression of DE genes in MPP3 and MPP4 cells, between Ebf1 WT and Ebf1 KO conditions. Genes are organised into MPP3-specific, common, and MPP4-specific DE genes. Genes with an annotated C/EBPα ChIP peak (original data from Hasemann et al. [2014]; Pundhir et al. [2018]) are labeled in blue. Heatmap scale represents z-scores calculated separately for MPP3 and MPP4 cells. (G) qPCR analysis of Cebpα mRNA expression in Ebf1 WT and Ebf1 KO MPP3 cells transfected with a control or a Cebpα siRNA pool. Cebpα mRNA expression relative to Actb was normalized to the control transfected samples. Data are represented as mean ± SEM. Ebf1 WT and Ebf1 KO n = 4. (E and G) Data are from >2 independent experiments.
Figure S4. Characterization of the scRNA-seq landscape in Ebf1WT and Ebf1KO LSK and LK cells. Related to Fig. 6. (A) scRNA-seq landscape derived from Ebf1WT and Ebf1KO LK and LSK populations, annotated for LSK and LK cells (top) and Ebf1WT and Ebf1KO cells (bottom). (B) UMAP visualisation of the HSC score and the expression of lineage marker genes used for the identification of cell clusters (Procr, Dntt, Elane, Pf4, Flt3, Ly6d). (C) UMAP visualisation of the expression of Il7r and Ig genes in Ebf1WT and Ebf1KO cells. (D) GO enrichment analysis of the upregulated genes in cluster 4, between Ebf1WT and Ebf1KO cells. (E) Boxplots showing the absolute number of myeloid cells (CD11b+) in the BM. Ebf1WT n = 18, Ebf1KO n = 16, IghmWT and IghmTm1 n = 11. Statistical significance was determined by Mann-Whitney U test. Data are from ≥2 independent experiments.
Figure S5  Extended analysis of the Ebf1-Cebpa regulatory relationship.  Related to Fig. 7.  (A) Normalized read counts of Ebf1 (left) and Cebpa (right) in Ebf1$^{-/-}$ pre-pro-B cells 24 and 72 h after EBF1 induction, and pro-B cells (original data from Li et al. [2018]).  (B) Normalized read counts of Sp1 (PU.1) in MPP3 and MPP4 cells in Ebf1WT and Ebf1KO populations.  (C) qRT-PCR analysis of Ebf1 expression in Ebf1WT MPP3 and MPP4 cells transduced with an empty or an EBF1 vector.  Ebf1 mRNA expression relative to Actb was normalized to the empty vector transduced samples.  Data are represented as mean ± SEM.  Ebf1WT and Ebf1KO n = 3–4.  (D) Boxplots showing the absolute number of neutrophils (Ly6G$^+$) in the BM.  Ebf1WT n = 13, Ebf1KO n = 14, IghmWT and IghmTm1 n = 11.  Statistical significance was determined by Mann-Whitney U test.  (E and D) Data are from >2 independent experiments.  (E) Annotated UMAP projection of scRNA-seq landscape derived from human fetal liver progenitors.  (F) UMAP visualization of the expression of Procr in human fetal liver progenitors.  Box indicates selected clusters (HSC/MPP, Ly/B progenitors, and Neu/Mono/DC progenitors) shown in G.  (G) UMAP visualization of the expression of Ebf1 and Cebpa in selected clusters of human fetal liver progenitors.  (E–G) Original data from Popescu et al. (2019).  Meg, megakaryocyte; Ery, erythrocyte; Bas, basophil; Neu, neutrophil; Mono, monocyte; DC, dendritic cell; Ly, lymphocyte; prog, progenitors; pDC, plasmacytoid DC.
Provided online are four tables. Table S1 contains a list of antibodies, reagents, and flow cytometry staining panels used in the study. Table S2 contains a list of the DA peak sets identified from bulk ATAC-seq of MPP3 and MPP4 cells from $Ebf1^{WT}$ and $Ebf1^{KO}$ mice. Table S3 contains a list of differentially regulated genes and normalized read counts identified from bulk RNA-seq of MPP3 and MPP4 cells from $Ebf1^{WT}$ and $Ebf1^{KO}$ mice, as well as the DEseq2 results table of the comparison of MPP3 and MPP4 cells from $Ebf1^{WT}$ mice. Table S4 contains a list of genes used for the gene sets G2-M and S- cell cycle phase, related to Fig. 6 H.