Polycomb contraction differentially regulates terminal human hematopoietic differentiation programs

Lorzadeh A.[1], Hammond C.[2,3], Knapp D.J.H.F.[2,3], Cao Q.[1], Wang F[2,4], Heravi-Moussavi A.[5], Wong M[1], Bilenky M.[5], Moksa M.[1], Sharafian Z. [6], Lavoie P.M.[6], Eaves C.J.[2,3,4], Hirst M.[1,5,*]

[1]Department of Microbiology and Immunology, Michael Smith Laboratories, UBC; [2]Terry Fox Laboratory, BC Cancer; [3]Department of Medicine, UBC; [4]Department of Medical Genetics UBC; [5]Canada's Michael Smith Genome Science Centre, BC Cancer; [6]BC Children's Hospital Research Institute; Department of Pediatrics, UBC

*Corresponding author

Summary

Mature blood cells derive from primitive cells through a complex and poorly understood process that involves an interplay of transcription factors and epigenetic modifications. We now show that multiple human hematopoietic progenitor phenotypes display a common repressive H3K27me3 signature (R>0.97) together with their mature lymphoid, but not their differentiated monocyte and erythroid progeny. This signature includes many large organized H3K27me3 domains that we also show are required for the production of differentiating lymphoid cells. These results reveal H3K27me3 contraction to play a previously unknown intermediate step in differentially regulating the activation of terminal differentiation programs in normal human lymphoid and myeloid progenitors.

Introduction

Epigenetic modifications govern local chromatin activity and support gene activation or silencing through the regulation of chromatin structure and DNA accessibility[1,2]. Numerous cell differentiation processes are known to be accompanied by obligatory changes in chromatin structure mediated by proteins that specifically modify histones and thereby establish and maintain defined transcriptional regulatory states[3–7]. However, the details of epigenomic programing changes that accompany lineage restriction, and if, and how regulatory changes may also occur during the activation and execution of terminal differentiation programs in normal tissues remain poorly defined.

Hematopoiesis describes the process through which multiple types of functionally required but short-lived, mature blood cells are produced, and therefore have to be continuously replaced lifelong[8]. A ready accessibility of the more primitive cells responsible for the output of mature blood cells has made this system an attractive one to analyze the dynamics of epigenome programming that underpins this process[9–12]. A prevailing view posits that after birth, a population of self-renewing hematopoietic stem cells (HSCs) residing within a supportive bone marrow niche generate derivative progenitors with reduced proliferative capacity but variable lineage options that then undergo a lineage restriction step prior to activating the final lineage programs that allow the full spectrum of mature blood cell types to be produced[13,14]. The process of lineage restriction is thought to involve the epigenetic demarcation of specific genomic regions that make them accessible to particular transcription factors which, in turn, ultimately activate lineage-specific gene expression programs. However, the molecular details of these steps are not well delineated.

Previous studies using chromatin immunoprecipitation sequencing (ChIP-seq) targeting sites of permissive histone modifications and transposase-accessible chromatin sequencing (ATAC-seq) have provided insight into the dynamics of active regulatory regions in the genomes of variously defined primitive subsets of mouse and human hematopoietic cells[9–11]. Together, these have suggested a model in which certain enhancers are initially "primed" during the earliest stages of hematopoietic cell differentiation by an acquisition of "permissive" histone 3 lysine 4 mono-methylation (H3K4me1) modifications. Then later, these same histones gain additional H3K27ac histone modifications that enable terminal differentiation programs to become activated[9].

However, a comprehensive analysis of the chromatin modifications that take place during the early stages of human hematopoiesis has remained uncharacterized.

To address this gap, we generated detailed genome-wide datasets and measured permissive and repressive histone modifications, DNA methylation and transcriptional states present in a series of phenotypically defined subsets of normal human cord blood (CB) cells. To identify changes associated with different stages in the differentiation process as currently inferred from functional assays[14,15,16] we focused on 8 phenotypically defined subsets that are selectively enriched in cells at different stages of hematopoietic differentiation. Four of these phenotypes identify cells with progenitor activity but distinct lineage options[16]. The other 4 represent different downstream mature or maturing blood cell types. All 8 populations were isolated by fluorescent activated cell sorting (FACS) from CB samples based on their differential expression of historically defined surface marker profiles. The 4 progenitor populations together constitute the bulk of the $CD34^+$ CB population. They consisted of the $CD38^-$ subset, which contains the most primitive cells and all HSCs and 3 phenotypically distinguished subsets of the $CD38^+$ cells historically referred to as common myeloid progenitors (CMPs), granulocyte-macrophage progenitors (GMPs) and megakaryocyte-erythroid progenitors (MEPs). The 4 more differentiated cell types also present in the light-density fraction of CB cells were circulating erythroblasts (hereafter referred to here as "erythroid precursors"), monocytes, mature B cells and mature T cells.

Analysis of repressive chromatin states from these cell types revealed that all 4 of the progenitor populations share a nearly identical H3K27me3 signature (gene promoter Spearman correlation $R>0.97$) including large organized chromatin K27-modification (LOCKs) domains that were not present in the 2 distinct mature myeloid lineages analyzed (i.e., erythroid precursors and monocytes). In contrast, we found the mature lymphoid cells retained the K27 LOCKs common to all 4 progenitor subsets. Analysis of the enhancer landscape revealed that a majority of traditionally defined active enhancers found in differentiated cell types were already evident within the progenitor populations and those unique to differentiated cell types are found almost exclusively within the boundaries of super-enhancers. From these findings, we propose a model in which the chromatin configuration changes that contribute functionally to the normal process of human hematopoietic differentiation include both active and repressive histone modifications that occur after, as well as during, lineage restriction.

Results

Functionally distinct human hematopoietic progenitors share a common polycomb signature

We first undertook a comprehensive mapping of the epigenetic and transcriptional states of historically defined immunophenotypes of human CB cells using our low input ChIP-seq protocol[17,18] to identify H3K4me3, H3K4me1, H3K27me3, H3K27ac, H3K36me3, and H3K9me3 sites genome-wide, plus whole genome bisulfite sequencing and RNA-seq protocols following IHEC standards[19] (**Fig. 1a** and **S1a**). A standardized analytical pipeline was then applied to qualify and analyze the resulting data (see methods).

RNA expression profiles among the progenitor populations analyzed (CD34+CD38- cells, CMPs, GMPs and MEPs) were more highly correlated (Spearman R >0.92) with each other than with any of the 4 more mature CB cell types examined; i.e., erythroid precursors, monocytes and B- and T-cells (average Spearman R = 0.76, 0.83, and 0.84, respectively, Fig 1b and S1b), confirming relationships seen in previously published datasets[3]. Expression signatures derived for each progenitor subset were similarly in agreement with their respective published features (methods and see **Fig. S2a** and **b**). For example, pathway and gene ontology (GO) analysis of uniquely upregulated transcripts in GMPs showed these were enriched in terms related to leukocyte differentiation, inflammation response and regulation of immune response (Benjamini q-value $<10e^{-20}$, **Fig. S2c**) and included the CD135 cell surface marker. Transcripts uniquely up-regulated in MEPs were enriched in terms related to myeloid lineage differentiation (Benjamini q-value <10e-3) (**Fig. S2d**). Genes up-regulated in CD34+CD38- cells and CMPs as compared to the more differentiated cells were enriched in terms related to hematopoiesis regulation and differentiation (Benjamini q-value $<10e^{-3}$) (**Fig. S2e** and **f**).

Examination of the chromatin state of the same cell subsets showed H3K4me3 and H3K36me3 densities correlated with expected transcript levels and cell type-specific signatures (**Fig. 1c, e-g**, **S2g** and **h**). H3K4me3 densities at promoters of genes differentially expressed between different progenitor subsets, or between them and the later cell types, also showed expected relationships (**Fig. S3a**). In contrast, H3K27me3 occupancy was nearly identical across all 4 phenotypically defined CB progenitor subsets examined (**Fig. 1d** and **1h**; promoter Spearman R >0.97), and cell-type specific signatures were apparent only in the mature cell types (**Fig. 1e** and **f**). Amongst the more mature populations, the lymphoid cells showed a significantly higher correlation of

H3K27me3 density at promoters with progenitors as compared to the corresponding data for the erythroid precursors and the monocytes (**Fig. 1d**). This was most readily evident true for the erythroid precursors whose RNA expression, H3K4me3 and H3K27me3 signatures at promoters showed the lowest correlation (average Spearman $R < 0.56$) with the results for the 4 progenitor populations (including MEPs) (**Fig. 1d**). There was no significant difference in H3K27me3 density within promoters of differentially expressed genes in a comparison of GMPs and MEPs (2-sided t-test, $p > 0.01$, **Fig. S3b**). In contrast, genes whose expression appeared down-regulated in monocytes and erythroid precursors compared to GMPs and MEPs, respectively, showed a significant gain of H3K27me3 density at the promoters of these down-regulated genes (2-sided t-test, $p < 2.2 \times 10^{-16}$).

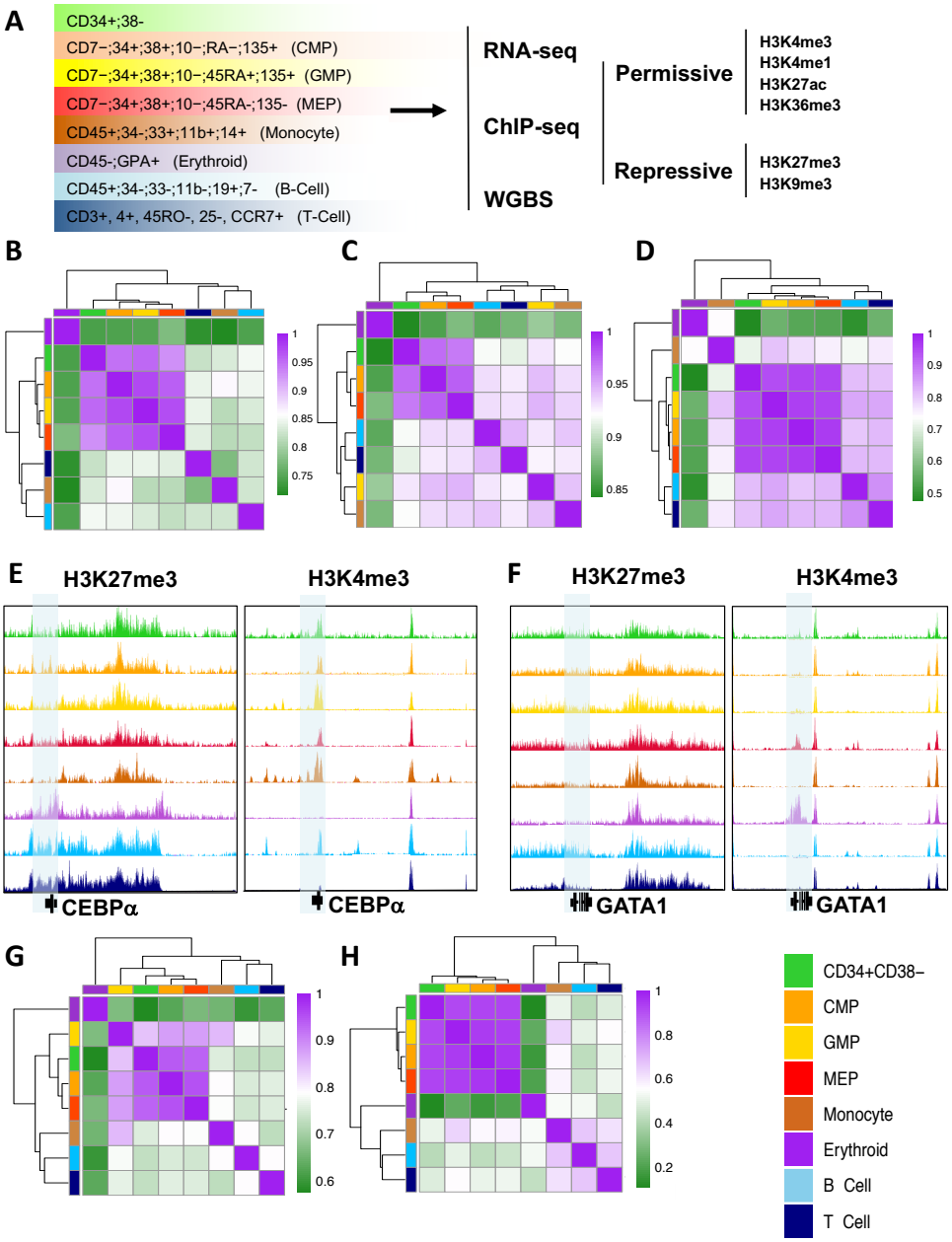

Figure 1. H3K27me3 signatures are shared across functionally distinct hematopoietic progenitor populations A) Schematic of experimental design and colour legend (**also shown in bottom right corner of Figure 1**) for cell types analyzed. Unsupervised hierarchical clustering and heatmap of pairwise Spearman correlations for protein coding gene RPKM values **(B)**, H3K4me3 coding gene promoter (± 2 Kb) density **(C)**, and H3K27me3 coding gene promoter density **(D)** across cell types indicated by colour. Genome browser view of H3K27me3 (left panel) and H3K4me3 (right panel) density tracks across cell populations indicated by track colour at CEBPα **(E)** and GATA1 **(F)** locus. Unsupervised hierarchical clustering and heatmap showing pairwise spearman correlation of H3K4me3 **(G)** and H3K27me3 **(H)** density genomewide.

Terminal differentiation is differentially associated with a genome-wide contraction of H3K27me3 density in myeloid cells

We also identified different patterns of H3K27me3 occupancy in the 4 different mature cell types examined (**Fig. 1d** and **1h**). Both monocytes and erythroid precursors showed significantly significantly fewer (30-52%) of H3K27me3–marked histones genome-wide (**Fig. S3c**) with a pronounced contraction of the broad H3K27me3 domains shared by all 4 of the progenitor populations examined. The remaining H3K27me3 in the erythroid precursors and monocytes took on a more punctate structure reminiscent of that seen in pluripotent cell types (**Fig. 2a-e**). Immunoprecipitated (IP) fragment distributions at promoters (± 2 Kb of transcription start sites (TSS)) were also significantly different in the monocyte and erythroid precursor populations by comparison to the 4 progenitor subsets or to both of the mature lymphoid cell types analyzed (B and T cells, kolmogorov–smirnov test, $p<7x10^{-12}$; **Fig. 2c**). Examination of H3K27me3 promoter distributions in all of these cell phenotypes further revealed an increase in the proportion of H3K27me3-marked histones at promoters in the monocytes and erythroid precursors as compared to the progenitors from which they derive (**Fig. S3d** and **f**), without a measurable change in H3K4me3 promoter occupancy (**Fig. S3e** and **g**). To examine the functional consequence of the global alteration in H3K27me3 occupancy we measured H3K27me3 within gene bodies and related this expression difference across the cell types profiled (**Fig. 2f**). Consistent with genome-wide patterns, we observed an increased number of genes with a loss in H3K27me3 density in monocytes and erythroid precursors (1418 and 1980, respectively) compared to mature lymphoid cells (B cells: 942 and T cells: 873) when compared to $CD34^{+}CD38^{-}$ cells (**Fig 2f**). Loss of H3K27me3 correlated with a greater proportion of up-regulated genes in monocytes and erythroid precursors compared to mature lymphoid cells (**Figure 2g)**. Among the genes which lost H3K27me3 and were upregulated are well-known markers of erythroid and monocyte lineage differentiation including CD14, EPB42 and CD36 (**Fig 2h-j)**.

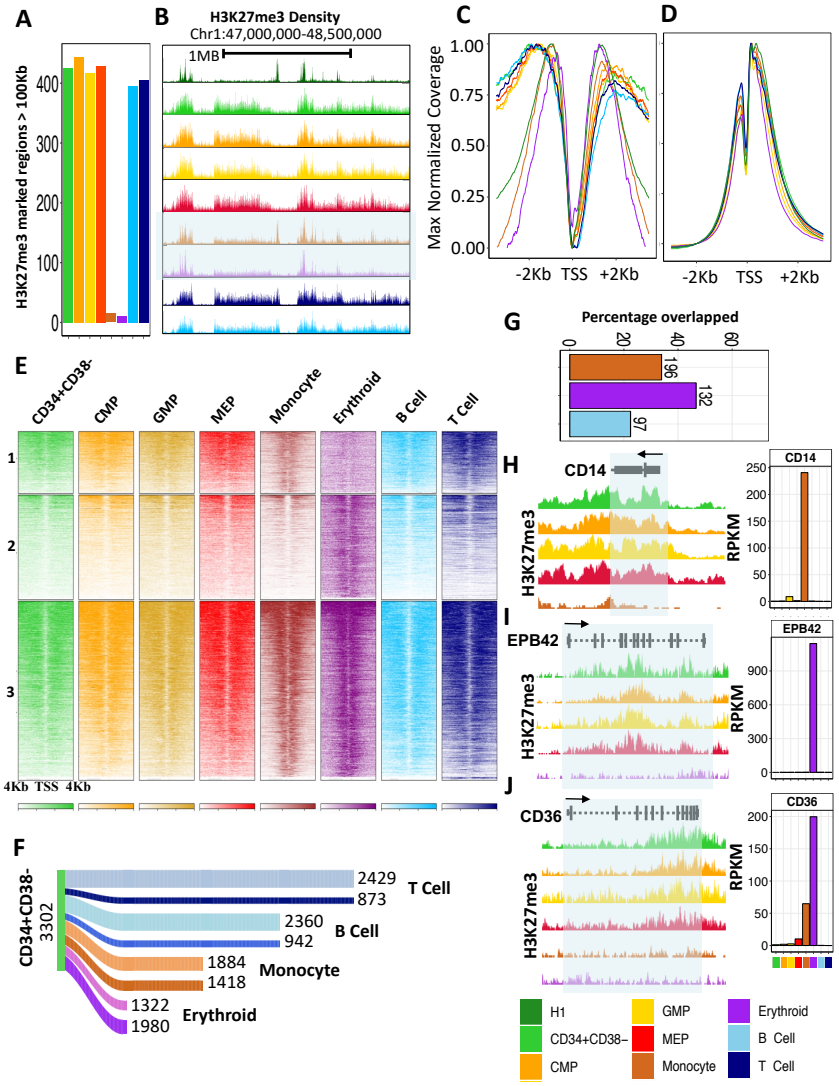

Figure 2. Myeloid differentiation is associated with genome-wide H3K27me3 contraction. A) Number of H3K27me3-marked genomic regions with widths >100 Kb across cell types indicated by the colour legend on the bottom right. **B)** Example genome browser view of H3K27me3 density profiles over a 2Mb window of chromosome 1 highlighting contraction of H3K27me3 in the monocyte (brown) and erythroid precursor (purple) cell types. Maximum value normalized H3K27me3 **(C)** and H3K4me3 **(D)** density at coding gene promoters (± 2 Kb of TSS). **E)** Heatmap of H3K27me3 densities coloured by cell type at gene promoters and flanks (TSS ± 4 Kb) marked by H3K27me3 in monocyte but not erythroid precursor (**panel 1**), in erythroid precursor populations but not monocyte (**panel 2**) or in both monocyte and erythroid precursor (**panel 3**). Shading indicates increased density **F)** Sankey diagram of number of genes retaining or losing H3K27me3 at their gene body in monocyte, erythroid precursor, B cell and T cell with respect to CD34+CD38- population. (Dark shade: loss of H3K27me3, light shade retention of H3K27me3). **G)** Percentage of overlap of genes that lost H3K27me3 within up-regulated gene bodies in monocyte, erythroid precursor and B cells with respect to CD34+CD38- (absolute number of overlap genes is shown on top of each bars). Genome browser view of H3K27me3 density at *CD14* **(H)**, *EPB42* **(I)** and *CD36* **(J)**. Expression (RPKM) of each gene across all profiled cell types are shown in the right panel.

Interestingly, in addition gene specific losses, we noted that H3K27me3 marked domains evident in all progenitor subsets are selectively absent in the differentiated monocytes and erythroid precursors in association with a loss of many large organized chromatin K27me3 domains (LOCKs)[20] (**Fig. 3a-c**). In sharp contrast, these same LOCKs were generally retained in the B and T cells (**Fig. 3a, d** and **e**). A majority of the few remaining LOCKs identified in the monocytes and erythroid precursors (48 and 101, respectively) overlapped with LOCKs present in the 4 progenitor populations (91 and 41, respectively, **Fig. 3b** and **c**). However, even within the LOCKs retained in the monocytes and erythroid precursors, H3K27me3 densities showed evidence of contraction, with an average width of 52 Kb compared to 268 Kb seen in the progenitor subsets, including those (GMPS and MEPs) already largely restricted to the neutrophil/monocyte and the erythroid/megakaryocyte lineages, respectively (**Fig. 3f-h**). These data thus provide evidence of a genome-wide contraction of H3K27me3 density in both gene bodies and within LOCKs during the process by which certain myeloid-restricted CB progenitors transition into their maturing progeny (**Fig. 3i**).

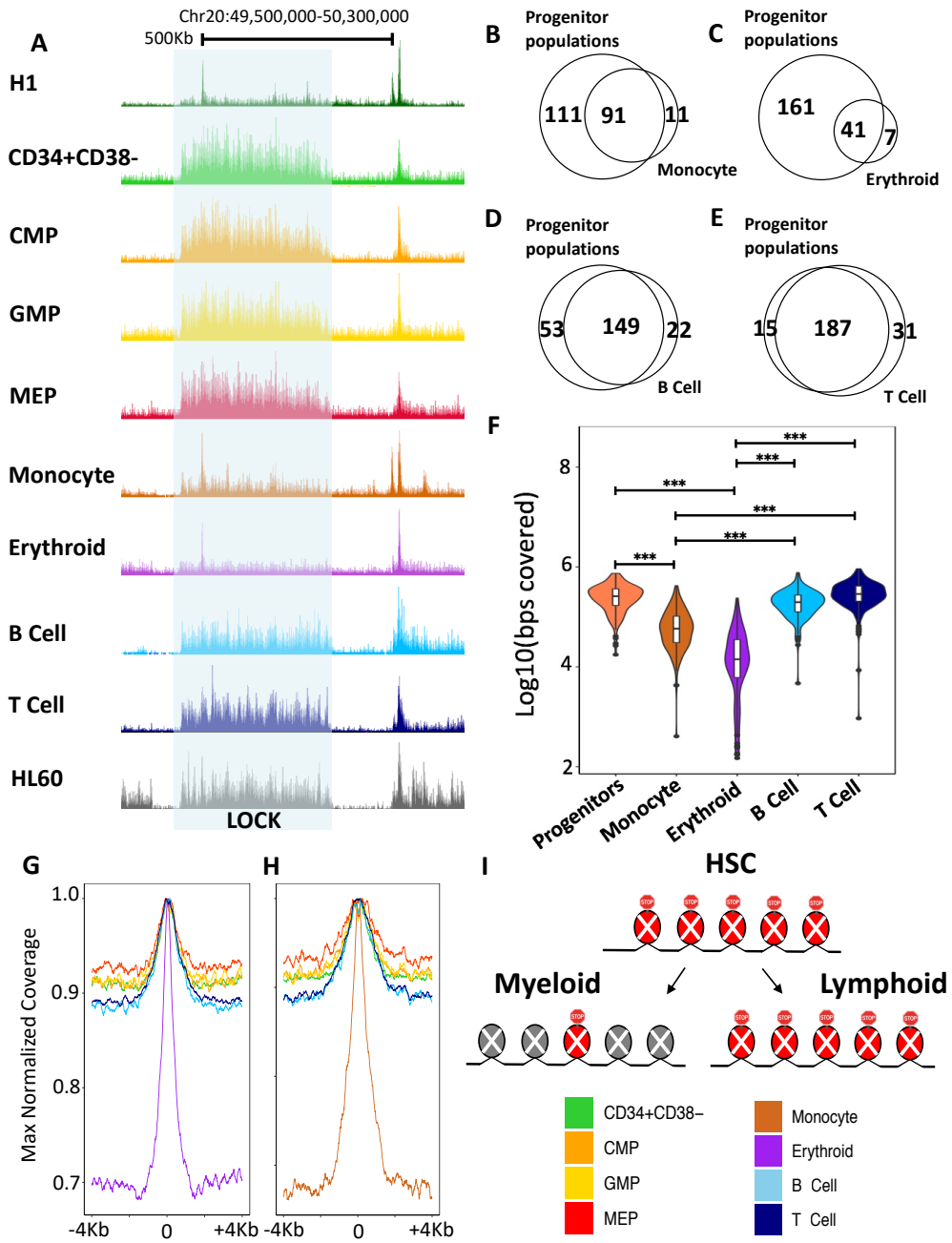

Figure 3. H3K27me3 contraction during myeloid differentiation coincides with a loss of progenitor LOCKs. A) Genome browser view of H3K27me3 density on chromosome 20 across cell types as indicated by the colour legend on the bottom right. A H3K27me3 LOCK (FDR <0.05) present in progenitor cells but absent in monocyte and erythroid precursors is indicated by the shaded box. Venn diagram of LOCKS common to progenitors (CD34+CD38-; MEP; CMP; GMP) compared to those called in monocytes **(B)**, erythroid precursors **(C)**, B cells **(D)**, and T cells **(E)**. **F)** Violin plot of base pairs marked by H3K27me3 within LOCKs across cell types as indicated (*** P < 0.001). Maximum value normalized H3K27me3 density-enriched within LOCKs called in erythroid precursors **(G)** and monocytes **(H)**. **I)** Cartoon illustrating expansion and contraction of H3K27me3 marked nucleosomes (red) during myeloid differentiation.

H3K27me3 LOCKs are enriched in lamina-associated domains (LADs)

To further characterize the LOCKs that are lost during the terminal differentiation of erythroid precursor cells and monocytes, we next examined the co-occurrence of other histone modifications and DNA methylation in these same regions. For this we applied ChromHMM[21] to generate an 18-state model based on H3K4me1, H3K4me3, H3K27me3, H3K27ac, H3K36me3, H3K9me3 occupancy across all of the 8 different hematopoietic subsets examined (**Fig. 4a**). Interestingly, H3K9me3 was found to be the most highly enriched mark within H3K27me3-defined LOCKS (**Fig. 4b** and **c**). Polycomb-repressed regions with or without H3K9me3 were found to be enriched in lamina associated domains (LADs) and in intergenic regions in both the progenitor subsets and in the B and T cells but, again, not in monocytes or erythroid precursors (**Fig. 4a, d, e** and **S4a**). In contrast, regions enriched in H3K9me3 alone were strongly associated with LADs in all cell types (**Fig. 4a** and **S4a**). Monocytes and erythroid precursors showed reduced DNA methylation at LOCKs compared to the 4 progenitor subsets (2-sided t-test $p<5.5 \times 10^{-8}$; **Fig. 4f** and **S4b**). Erythroid precursors particularly, but also the monocytes showed reduced DNA methylation more broadly in chromatin states identified as polycomb-repressed and/or heterochromatin-repressed in the ChromHMM model (**Fig. S4c-f**). These findings are consistent with the previously reported reduced global level of DNA methylation of monocytes and neutrophils relative to different progenitor populations[12]. They also corroborate previous relationship between DNA methylation and H3K9me3 densities[22–24] and, in the present context, suggest that $CD34^+$ hematopoietic progenitors share a higher order chromatin structure that is associated strongly with LADs and enriched in sites of H3K27me3 and H3K9me3.

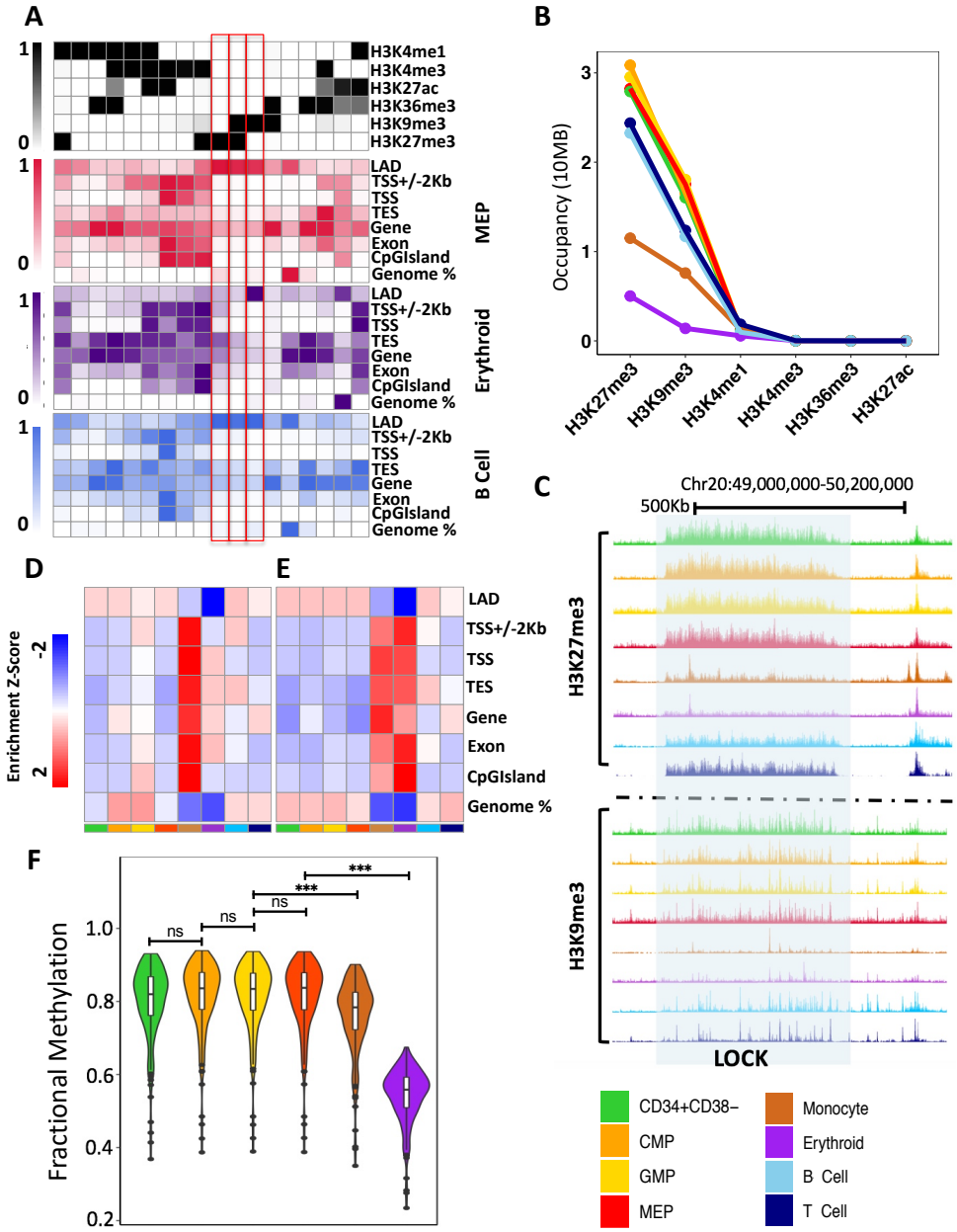

Figure 4. H3K27me3 LOCKs lost during myeloid differentiation are associated with LADs in progenitor cells. A) Emission probabilities by histone mark for 18 states ChromHMM chromatin state model across all cell types (upper panel). Enrichment of chromatin states within genomic features for MEPs (red), erythroid precursors (purple) and B cells (blue). **B)** Cumulative number of bases marked by indicated histone modification within progenitor population's LOCKS as indicated by colour legend on bottom right. **C)** Genome browser view of H3K27me3 and H3K9me3 density on chromosome 20 across cell types as indicated by colour legend on bottom left. A H3K27me3 LOCK (FDR < 0.05) present in progenitor cells but lost in monocyte and erythroid precursor is indicated by the shaded box. **D)** Heatmap of z-score enrichment of ChromHMM identified polycomb repressed and dually repressed chromatin states at genomic features. **F)** Violin plot of fractional CpG methylation levels within progenitor's LOCKs (*** P < 0.001).

Hematopoietic progenitor subsets share lineage-specific enhancers marked by H3K27ac

We next sought to identify and compare the enhancer states of the 4 progenitor populations and the 4 more mature cell types examined. Accordingly, we identified the H3K27ac and H3K4me1 marked regions in each population and used the results to create a catalogue of active (H3K27ac and H3K4me1) and primed (H3K4me1) enhancers. Primed enhancers were relatively consistent across all 4 progenitor populations (Spearman, average R >0.8; **Fig. 5a**) with active enhancers showing a higher degree of progenitor specificity (Spearman, average R >0.52; **Fig. 5b and S5c**). The progenitor populations also showed a consistently higher number of total enhancers, as measured by the sum of both the active and primed enhancer states, by comparison to the 4 more mature cell types (**Fig. S5a**). In contrast, the number of active enhancers was higher in 3 of the 4 more mature cell types (2-sided t-test, p =0.026), the exception being the erythroid precursors that had a lower number of active enhancers compared with all other cell types (**Fig. S5b**).

We next traced the gain or loss of H3K27ac and H3K4me1 from the most primitive $CD34^{+}CD38^{-}$ cells to each differentiated cell type according to predicted trajectories. This exercise confirmed a directional loss of total enhancers during progenitor differentiation with the erythroid precursors showing the greatest overall loss of enhancers (**Fig. 5c**). However, the directional loss was largely restricted to primed enhancers with active enhancers staying either the same or increasing during differentiation (**Fig. 5c).** A majority of primed enhancers (>90%) and active enhancers (>72%) found in mature cells were already evident in the progenitor populations (**Fig. 5d, e and S5d**). A significant fraction (>80%) of the active enhancers in the differentiated cell types were also already primed across the progenitor subsets, as noted earlier for similar subsets of mouse hematopoietic cells[9] (**Fig. S5e and f**). Interestingly, genes associated with enhancers found to be active in monocytes and erythroid precursors, but already active in progenitors also produced significantly higher levels of transcripts in the mature cells compared to their inferred parental progenitors (pairwise Wilcoxon signed-rank test Bonferroni corrected $p < 2x10^{-5}$, **Fig. 5f**). They are also enriched in terms related to specific hematopoietic differentiation programs (Benjamini corrected $p < 10e^{-30}$, **Fig. S6a and b**). For example, *SPI1* and *EPB41* are associated with an active enhancer in progenitor and in lineage-restricted cells, with their expression increasing significantly during the full course of the differentiation process (**Fig. 5g and h**).

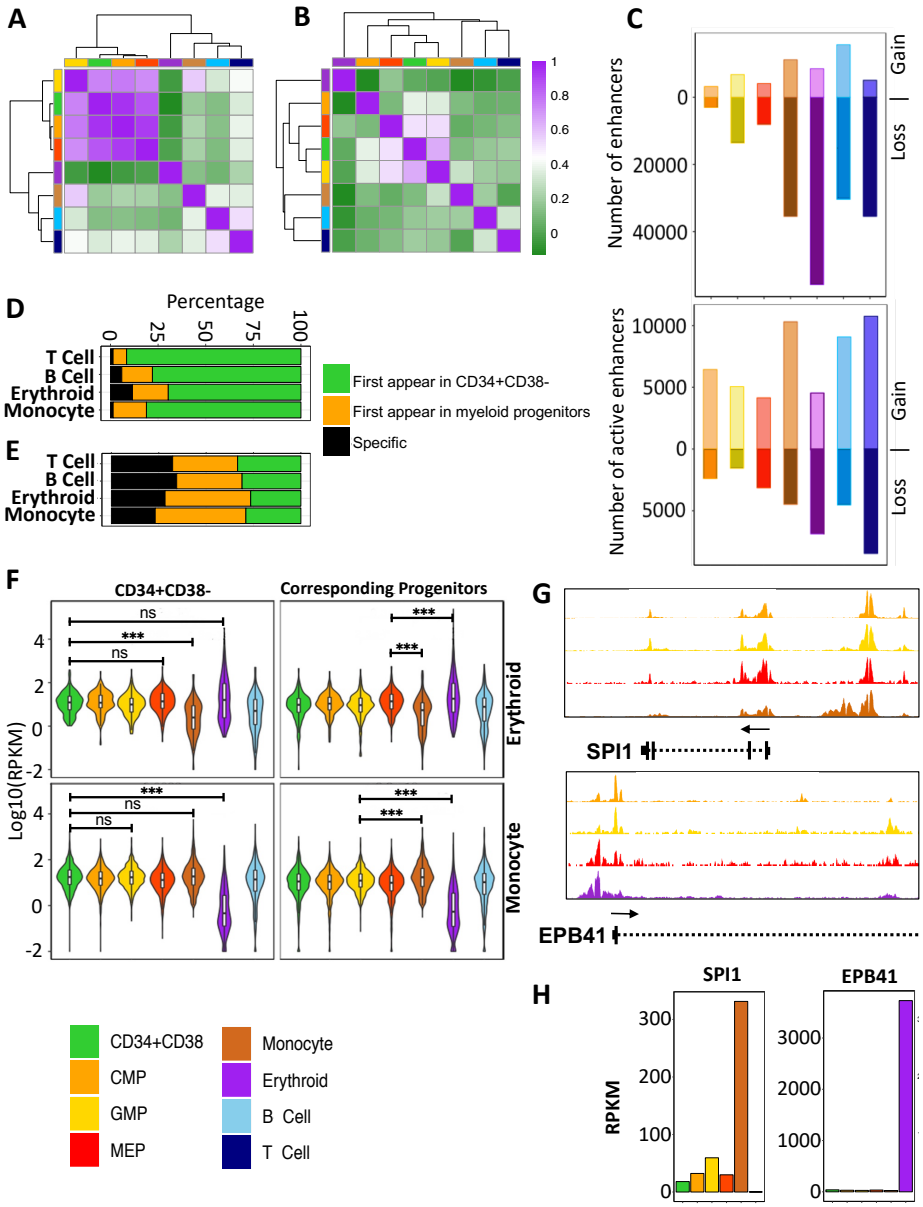

Figure 5. A majority of active enhancers in myeloid cells are marked by H3K27ac in progenitor cell populations. Unsupervised hierarchical clustering and heatmap of pairwise Spearman correlations of H3K4me1 **(A)** and H3K27ac **(B)** densities across cell types as indicated by the colour legend on the bottom left. **C)** Number of total enhancers (top panel) and active enhancers (bottom panel) that are gained or lost in comparison to CD34+CD38- cells in each cell type. Percentage of enhancers **(D)** and active enhancers **(E)** present in differentiated cells that are first detected in CD34+CD38- cells, myeloid progenitor cells, or uniquely in monocytes, erythroid precursors, or lymphoid cells as indicated. **F)** Expression of associated genes with active enhancers in monocytes and erythroid precursors that first appear in CD34+CD38- cells or their corresponding progenitor populations. **G)** Genome browser view of H3K27ac density at SPI1 (top panel) and EPB41 (bottom panel) locus. **L)** Expression (RPKM) of SPI1 (left panel) and EPB4 (right panel). *** P <0.001.

To further understand how active enhancer states differ between progenitors and terminally differentiated cells, we leveraged self-organizing maps (SOM) and used ranked normalized H3K27ac signals across all hematopoietic enhancers as input to identify progenitor, monocyte, erythroid, B cell and T cell enhancer clusters (methods; **Fig. S6c**). Relative to the $CD34^{+}CD38^{-}$ population, MEPs and GMPs showed progression towards, and enrichment in, enhancers belonging to erythroid precursor and monocyte enhancer clusters, respectively (**Fig. S6d**). Enrichment of differentiated enhancer clusters in MEPs and GMPs is indicative of presence of a pre-existing H3K27ac signature at cell type-specific enhancers in progenitors. Similarly, the appearance of monocyte- and erythroid-specific enhancers in GMPs and MEPs, respectively, support this view. Interestingly, CMPs show enrichment in both lymphoid and monocyte enhancer clusters consistent with data suggesting that CMPs are quite heterogeneous and could thus appear to share a greater similarity in enhancer landscape with lymphoid cells rather than with MEPs and GMPs in which cells with lymphoid potential appears largely eliminated[15,16].

We next used the ROSE RANK algorithm[25] to identify the high amplitude active enhancers and clusters of H3K27ac (super-enhancers) in each of the 8 cell types analyzed. We observed a greater than 2-fold increase in the number of super-enhancers in the mature cell types compared to the numbers of these in the progenitor populations (**Fig. 6a**). Surprisingly, we also found that a majority (>70%) of the active enhancers seen only in the mature cells were located within the boundaries of super enhancers, including the human counterparts of enhancers previously identified in differentiating mouse hematopoietic cells (**Fig. 6b and c**). Pathway analysis of genes associated with monocyte and erythroid precursor super-enhancers were enriched in terms related to leukocyte activity and erythroid differentiation, respectively (**Fig. S6e and f**). Taken together these data suggests a model in which a majority of active enhancers are initially marked in progenitor populations and then further reinforced during the final processes of terminal differentiation.

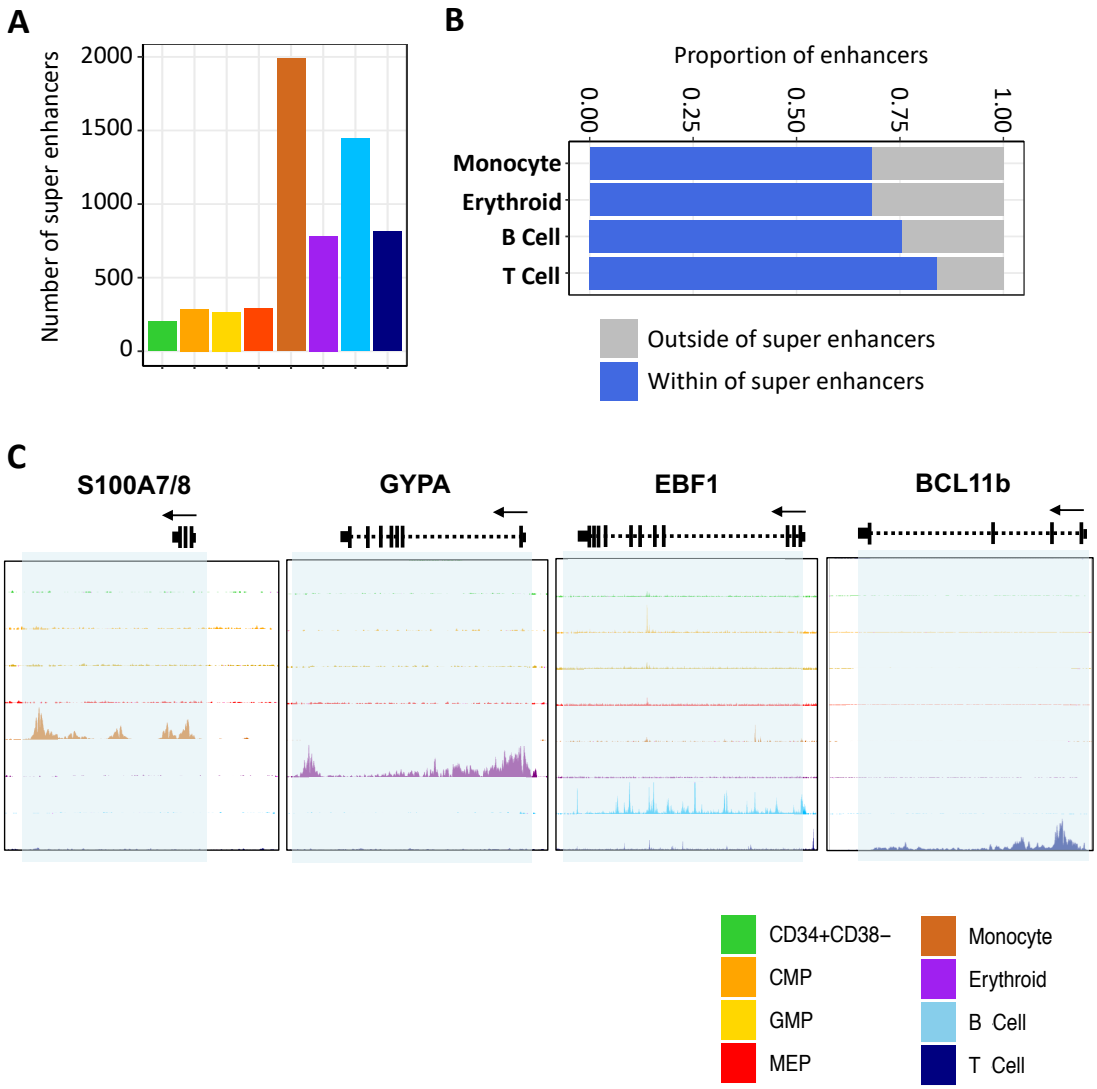

Figure 6. The majority of active enhancers first appearing in differentiated cell populations contribute to the establishment of super enhancers. A) Total number of super enhancers across cell types as indicated by colour legend on the bottom right. **B)** Fraction of cell type-specific active enhancers that reside within super enhancers. **C)** Genome browser view of H3K27ac density across cell types as indicated by colour at lineage-specific super enhancers indicated by the shaded box.

EZH2 inhibition differentially alters the production of different hematopoietic lineages consistent with their acquired epigenomic features.

The broad H3K27me3 domains shared by CB progenitors that are no longer present in monocyte and erythroid precursors but persist in mature lymphoid cells (**Fig. 3**) led us to hypothesize that these differences may have functional roles in the processes that allow these different lineages to be produced. To test this possibility, we first compared the ability of 2 inhibitors that target complexes implicated in the maintenance of repressive chromatin; i.e., EPZ-6438, an inhibitor of EZH2, a major component of the polycomb repressive complex-2 (PRC2) and GSK-J4, a H3K27me3 demethylase inhibitor to mimic the ability of ATRA to induce the granulopoietic differentiation potential of HL60 cells[26]. After 3 days of treatment with EPZ-6438, HL60 cells showed the same growth-arrest obtained by treatment with ATRA (**Fig. 7a and b**), no evidence in either of decreased viability which then appeared at a similar rate in both (**Fig. 7b**). In contrast, treatment of HL60 cells with GSK-J4 exhibited no significant cell growth arrest or loss of viability for 9 days with results that were indistinguishable from the DMSO-treated controls. Interestingly, neither inhibitor when present together with ATRA produced any alteration to the response to ATRA alone at the concentrations tested. Thus, the inferred EPZ-6438-mediated H3K27me3 removal in HL60 cells appeared to drive the same phenotypic alterations as ATRA in HL60 cells, consistent with the importance of their removal to enable granulopoietic differentiation to proceed.

We then asked how these 2 inhibitors affect the granulopoietic and/or B-lymphopoietic differentiation potential of primary human hematopoietic progenitors maintained *in vitro* under conditions optimized to support the production of both of these lineages[16]. Accordingly, we initiated a series of cultures from a FACS-purified $CD38^{mid}CD71^{-}CD10^{-}$ subset of $CD34^{+}$ CB cells, that we have previously shown is highly enriched specifically in cells with one or both of these differentiation potentialities[16]. Assessment of the mature cell types present 3 weeks later showed a significant (2-sided t-test $p<0.014$) and selective decrease in cells expressing B cell markers in the exclusive presence of EPZ-6438, with no effects of either inhibitor on the output of cells expressing phenotypic markers of monocyte/macrophages or neutrophils (**Fig. 7c and d**). These latter results suggest that continued repression of H3K27me3 sites by PRC2 complexes is required for progenitors with lymphoid potential to differentiate into mature lymphoid cells.

Together the effects of EPZ-6438 in both cell systems are consistent with a model in which the loss of H3K27me3 domains promotes the differentiation of mature myeloid cells from progenitors with that potential, whereas maintenance of H3K27m3 domains is required for the production of mature lymphoid cells.

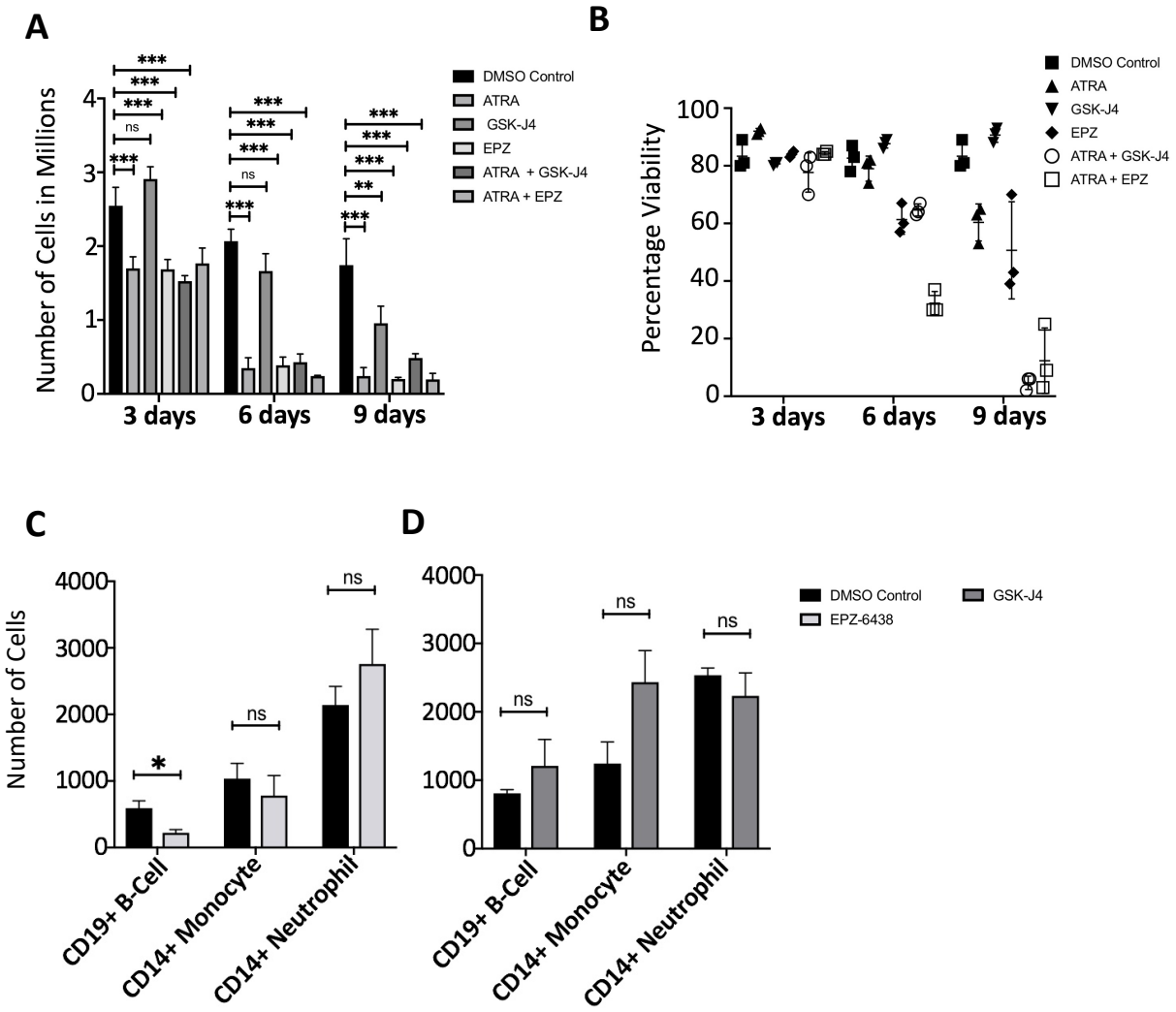

Figure 7. EZH2 inhibition arrests HL60 cell growth and lymphoid production *in vitro*. A) Total number of HL60 cells in culture treated with ATRA, GSK-J4 (K27me demethylase inhibitor) and EPZ (EZH2 inhibitor) every 3 days for 9 days. **B)** Percentage viability of HL60 cells in **(A)**. Total cell number of CD19+ B cells, CD14+ monocytes, and CD15+ neutrophil cells after 3 weeks of culturing CB-derived CD45^{high};34^{high};38^{mid};71-;10- cells with lymphoid and non-lymphoid lineage-specific cytokines in presence or absence of EPZ **(C)** and GSK-J4 **(D).** * P <0.05, ** P <0.01 and *** P <0.001.

Discussion

Epigenetic mechanisms play a central role in hematopoietic cell fate decisions[27–32]. Of particular interest here is the repressive H3K27me3 chromatin modification that has been implicated in the regulation of hematopoiesis *in vivo* and *in vitro*[33–37] and how manipulation of H3K27me3 may impact the steps progenitors undergo to produce mature myeloid and lymphoid cell types[35,38]. Both overexpression and inactivation of PRC2 components, responsible for the methylation of H3K27 have been reported in hematopoietic malignancies, further reinforcing the importance of H3K27me3 in regulating normal hematopoiesis[39,40]. To further our understanding of the role of epigenetic states in these aspects of human hematopoiesis, we analyzed the epigenomic states of immunophenotypically defined, FACS-purified subpopulations of $CD34^+$ and mature cells isolated from normal CB samples. Our analysis revealed a broad and stable repressive H3K27me3 landscape across all progenitor subsets analyzed; i.e., those defined phenotypically as $CD34^+CD38^-$, CMPs, GMPs and MEPs, which contrasted markedly with the dynamic and cell-type specific landscape for active histone modifications. We also identified a striking and lineage-selective genome-wide H3K27me3 signature evident in 2 mature myeloid cell types (monocytes and erythroid precursors) and not seen in 2 mature lymphoid cell types (B cells and T cells). Intriguingly, the H3K27me3 signature common to the monocytes and erythroid precursors displayed a punctuated structure reminiscent of the H3K27me3 signature associated with embryonic stem cells (ESCs)[41].

In the classical Waddington[42] view, the chromatin of very primitive cells is envisaged to exist in a privileged state allowing for the dynamic remodeling required for lineage restriction and later activation of a differentiation program. One early finding in support of such privileged chromatin in ESCs is their bivalent state[41,43–45] in which the nucleosomes in the promoters of developmentally important genes are marked by both permissive and repressive histone modifications (H3K4me3 and H3K27me3, respectively) that subsequently resolve to an homogenously active or repressed chromatin state as differentiation occurs[41,43]. However, subsequent epigenomic studies across a broad range of primary tissues have revealed that bivalent promoters are not unique to ESC chromatin and can be found in many cell types including fully functional, terminally differentiated cell types[3,18]. Another prevailing model of H3K27me3 occupancy during ESC differentiation posits that, upon differentiation, H3K27me3 genomic occupancy spreads outwards from discrete focal regions to occupy large genomic

regions[45], an observation which has correlated with the differentiation capacity of these cells. Here, we demonstrate that H3K27me3 reverts back to a punctuated structure in cells that have differentiated into certain types of mature blood cells. This unexpected observation suggests that at some point in the differentiation of these cells there is an overall loss of repressive chromatin rather than a further compaction as might have been predicted from studies of ESC differentiation[45]. Why B and T cells retain a broader H3K27me3 landscape also raises interesting questions. One possibility is that this could be related to their known ability to activate a "stem-like" state upon antigenic stimulation and then execute a large expansion of their progeny[46].

We also found that this surprising H3K27me3 contraction evident in erythroid precursors and monocytes includes a specific loss of H3K27me3 marked LOCKs that in the progenitors and mature lymphoid cells show co-occupancy with another suppressive histone modification, H3K9me3. The lineage-specific restructuring of H3K27me3 specifically during myeloid differentiation highlights the importance of higher order chromatin structure in differentiation and establishment of cellular identity in the hematopoietic system. H3K27me3 and H3K27me3/K9me3 enriched regions common to progenitor and lymphoid cells and lost in different mature myeloid cell types are also strongly enriched in LADs, thus reinforcing the concept that myeloid cells can also be distinguished from lymphoid and progenitors based on lamin distribution and nucleus rigidity[47]. Taken together these observations provide molecular support of the observation that manipulation of lamin expression specifically modulates myeloid cell differentiation[47].

In contrast to H3K27me3, we found the active enhancer mark H3K27ac is dynamic across all progenitor populations analyzed and the majority of active enhancers identified in the mature cells were first detected in their traditionally defined progenitor populations despite recent evidence of the considerable heterogeneity in differentiation potential and other molecular features they display[16]. Key lineage-specific regulators were identified among the genes associated with active enhancers in the progenitor populations. During differentiation, these active enhancers increase in width and amplitude and expression of associated genes increases. Our results thus suggest that priming of key regulatory regions during hematopoietic differentiation occurs in the context of traditionally described active enhancers whose activity is reinforced and increases during lineage restriction. This finding refines the currently proposed primed (H3K4me1) to an active enhancer

model[9] and suggests that additional features beyond the presence of H3K27ac, for example formation of lineage specific phase separated condensates[48], may be required for full enhancer activity.

The appearance of lineage-specific active enhancers in myeloid progenitor cells prior to the acquisition by their mature progeny of a global H3K27me3 contraction is particularly intriguing. This finding suggests that active enhancers already present in progenitor cells are constrained by H3K27me3 and fully activate their target genes only after large scale chromatin structure reorganization. Thus the broad H3K27me3 domains observed in both progenitors and mature lymphoid cells may act as a general block for the generation of multiple mature myeloid cell types whose differentiation is normally completed within 3-5 cell divisions. These results reveal a previously unknown relationship of H3K27 modification in myeloid differentiation and further reinforce the importance of chromatin structure in the regulation of normal hematopoietic differentiation processes.

Experimental Procedures

Preparation of human cord blood cells

Anonymized consented anticoagulated samples of normal CB cells were obtained with informed consent according to University of British Columbia Research Ethics Board approved protocols. CD34+ cells were enriched by EasySep™ from the light-density fraction of Lymphoprep™ or RosetteSep™-depleted CD11b+;CD3+;CD19+ cells (STEMCELL Technologies) and then used after cryopreservation in DMSO and fetal bovine serum (FBS, STEMCELL Technologies).

Isolation of human CB populations

Frozen CB cells were thawed by drop-wise addition to Iscove's Modified Dulbecco's Medium (IMDM) (STEMCELL Technologies) supplemented with 10% FBS and 10 µg/mL DNase (Sigma-Aldrich). Cells were suspended in Hanks Balanced Salt Solution (HBSS) supplemented with 5% human serum and 1.5 µg/mL anti-human CD32 antibody (Clone IV.3; STEMCELL Technologies) and then stained with designated antibodies for 1-2 hours on ice prior to sorting on a Becton

Dickinson FACSAria™ Fusion or FACSAria™ III sorter. Cells from the following populations were sorted directly into DNA LoBind tubes (Eppendorf) containing HBSS + 2% FBS: CD34+CD38-, CD34+CD38+CD10-CD7-CD135-CD45RA- (MEP), CD34+CD38+CD10-CD7-CD135+CD45RA- (CMP), CD34+CD38+CD10-CD7-CD135+CD45RA+ (GMP), CD45-CD34-GPA+ (Erythroid Precursors), CD45+CD34-CD11b+CD33+CD14+ (Monocytes), CD45+CD34-CD11b-CD33-CD19+CD7- (B cells). Naïve CD4 T cells were isolated from CB mononuclear cells (CBMCs) obtained using Lymphoprep (StemCell Technologies Inc., Canada), followed by a first step of negative selection by magnetic bead separation using EasySep Human Naïve CD4 T cell isolation kit (STEMCELL Technologies Inc., Canada) and FACS on a BD FACSAriaTM II using the following markers/staining: anti-CD3 PE (clone UCHT1; BD Bioscience), anti-CD4 BV605 (clone OKT4; BioLegend,), anti-CD25 PE-Cy7 (clone M-A251; BD Bioscience), anti-CD45RO Alexa Fluor700 (clone UCHL1; BioLegend), anti-CCR7 Alexa Fluor 647 (clone 3D12; BD Bioscience, ON, Canada), and anti-CD235 eFluor 450 (clone 6A7M eBioscience). Isolated cells were then centrifuged (500rcf, 6min) and had their supernatant removed prior to rapid freezing using dry ice or liquid nitrogen then stored at -80°C.

RNA-seq

Total RNA was rRNA depleted using NEBNext rRNA Depletion Kit (New England BioLabs, E6310L). 1st strand cDNA was generated using Maxima H minus First Strand cDNA Synthesis Kit (Thermo Scientific, K1652) with the addition of 1ug of Actinomycin D (Sigma, A9415). The product was purified using in-house prepared 20% PEG, 1M NaCL Sera-Mag bead solution at 1.8X ratio and eluted in 35uL of Qiagen EB buffer. Second Strand cDNA was synthesized in a 50uL volume using SuperScript Choice System for cDNA Synthesis (Life Technologies, 18090-019) with 12.5mM GeneAmp dNTP Blend with dUTP. Double stranded cDNA was then purified with 20% PEG, 1M NaCL Sera-Mag bead solution at 1.8X ratio, eluted in 40uL of Qiagen EB buffer, and fragmented using Covaris E220 (55 seconds, 20% duty factor, 200 cycles per burst). Sheared cDNA was End Repaired/Phosphorylated, single A-tailed, and Adapter Ligated using custom reagent formulations (New England BioLabs, E6000B-10) and in-house prepared Illumina forked small adapter. 20% PEG, 1M NaCl Sera-Mag bead solution was used to purify the template in-between each of the enzymatic steps. To complete the process of generating strand directionality, adapter-ligated template was digested with 5U of AmpErase Uracil N-Glycosylase

(Life Technologies, N8080096). Libraries were then PCR amplified and indexed using Phusion Hot Start II High Fidelity Polymerase (Thermo Scientific, F 549-L). An equal molar pool of each library was sequenced on HiSeq2000 (Illumina) PE75. Libraries were sequenced on a HiSeq 2500 platform following the manufacture's protocols (Illumina, Hayward CA.). Sequence reads were aligned to a transcriptome reference generated by JAGuaR (version 2.0.2)[49] using the GRCh37-lite reference genome supplemented by read-length-specific exon–exon junction sequences (Ensemble v75 gene annotations) and bam files were generated using Sambamba (version 0.5.5)[50]. The resulting bam files were repositioned to GRCh37-lite by JAGuaR. To quantify exon and gene expression reads per kilobase per million mapped reads (RPKM) metrics was calculated. A standardized analytical pipeline introduced by Canadian Epigenetics, Environment, and Health Research Consortium [CEEHRC] and International Human Epigenomic Consortium (IHEC) (http://ihec-epigenomes.org/) was applied to qualify the resulting data. Pairwise differential expression analysis was carried by an inhouse MATLAB script, DEFine, on RPKM values that are transcripts GC biased corrected. MetaScape version 2.0[51] (http://metascape.org) was used for genome ontology analysis of differentially expressed genes. All figures were generated by R statistical software[52].

Low Input Native ChIP-seq

ChIP-seq was performed as previously described[17]. In brief, cells were lysed in 0.1 % Triton X-100 and Deoxycholate with protease inhibitor cocktail. Chromatin were digested by Microccocal nuclease (MNase, New England BioLabs, M0247S) at room temperature for 5 minutes and 5ul of 0.25mM EDTA was used to stop the reaction. Antibodies against epitopes of interest, H3K4me1 (Diagenode: Catalogue# pAb-037-050, lot# A1657D), H3K27me3 (Diagenode: Catalogue# pAb-069-050, lot# A1811-001P), H3K9me3 (Diagenode: Catalogue# pAb-056-050, lot# A1675-001P), H3K4me3 (Cell Signaling Technology: Catalogue# 9751S, lot# 8), H3K27ac (PMID: 18227620), and H3K36me3 (PMID: 18227620) and digested chromatin were incubated were incubated with anti-IgA magnetic bead (Dynabeads,Themo Finsher, 10001D) for 2 hours. Pre-cleared chromatin was incubated with antibody-bead complex overnight in IP buffer (20 mM Tris-HCl pH 7.5, 2 mM EDTA, 150 mM NaCl, 0.1 % Triton X-100, 0.1 % Deoxycholate) at 4 °C. IPs were washed 2 times by Low Salt (20 mM Tris-HCl pH 8.0, 2 mM EDTA, 150 mM NaCl, 1 % Triton X-100, 0.1 % SDS) and High Salt (20 mM Tris-HCl pH 8.0, 2 mM EDTA, 500 mM NaCl, 1 % Triton X-100,

0.1 % SDS) wash buffers. IPs were eluted in elution buffer (1 % SDS, 100 mM Sodium Bicarbonate) for 1.5 hours at 65 °C. Histones were digested by Protease (Qiagen 19155) for 30min at 50 C and DNA fragments were purified using Sera Mag magnetic beads in 30% PEG. Illumina sequencing libraries were generated by end repair, 3' A-addition, and Illumina sequencing adaptor ligation (New England BioLabs, E6000B-10). Libraries were then indexed and PCR amplified (10 cycles) and sequenced on Illumina HiSeq 2500 sequencing platform following the manufacture's protocols (Illumina, Hayward CA.).Sequence reads were aligned to GRCh37-lite using Burrows-Wheeler Aligner (BWA) 0.5.7.[53] and converted to bam format by Sambamba (version 0.5.5). Sequence reads with BWA mapping quality scores <5 are discarded and reads that aligned to the same genomic coordinate were counted only once in the profile generation. A standardized analytical pipeline developed by the International Human Epigenomic Consortium (IHEC) (http://ihec-epigenomes.org/) was applied to qualify the resulting data.

ChIP-seq Analysis

Genome browser tracks were generated by converting bam files to wiggle files using a custom script (http://www.epigenomes.ca/tools-and-software). Wiggle files then converted to bigwig for display on UCSC genome browser by UCSC tool, wig2Bigwig script. MACS2[54] was employed to identify enriched regions with a false discovery rate (FDR) value of ≤ 0.01 for H3K4me3, H3K4me1 and H3K27ac peaks in ChIP-seq data. Finder2.0 (http://www.epigenomes.ca/tools-and-software/finder) was used to identify enriched regions for H3K27me3 and H3K9me3 ChIP-seq data. Enriched regions overlapping with ENCODE blacklist regions were eliminated. ChIP-seq signal was calculated as tag density generated using HOMER v4.10[55] and normalized to total number of tags within enriched regions. Heat map of fragment distribution at promoters were generated by deeptools[56]. All other figures were generated by R statistical software[52]. ChromHMM[57] was used to identify 18 chromatin states based on MACS2[54] identified enriched regions in each cell type as previously described[21]. To identify active enhancers, we first found MACS2 identified enriched H3K27ac regions in all populations (CD34+ CD38-, CMP, GMP, MEP, Monocyte, Erythroid, B cell and T cell) and created an enhancer catalogue of hematopoietic cells. Any enriched region that overlap with +/-2Kb of TSS defined by Ensemble v75 human gene annotation was discarded. All H3K27ac libraries were subsampled to the depth of the library with lowest depth. Next, enhancers with H3K27ac tags > 40 in at least one cell type were selected. Then

normalized signal H3K27ac tag density was calculated across all regions in our enhancer catalogue for each cell type and enhancers with signal > 25 were marked as active in the corresponding cell type. Same strategy was used to identify poised enhancers with H3K4me1 enriched regions with the exception of signal threshold of 40. SOM analysis was carried by oposSOM[58] R package on ranked normalized signal of H3K27ac at union of H3K27ac enriched regions identified across all cell types. Enhancer gene association was done by bedtools[59] closestBed command with default parameters using Ensemble v75 human gene annotation. Promoters (+/2kb of TSS of protein coding) with distance of greater than 50Kb to enhancer regions were eliminated. We utilize GREAT, an online ontology analysis tool, for gene ontology analysis of regulatory regions.
To identify H3K27me3 LOCKs, MACS2 broad mode was used to call peaks for H3K27me3 with FDR cut off of 0.05. The resulting peaks were then entered into CREAM[60] with WScutoff of 1.5, MinLength of 1000 and peakNumMin of 2 to identify H3K27me3 LOCKs. H3K27me3 marked genes in $CD34^{+}CD38^{-}$, monocyte, erythroid precursor, B and T cells were annotated by overlapping FindER identified H3K27me3 enriched regions with hg19v75 ensemble gene annotations. We then selected for genes with >=20% of their genome covered with H3K27me3.

Whole Genome Bisulfite Sequencing

Whole genome bisulfite libraries were constructed as previously described[61]. In brief genomic DNA was sheered and subjected to bisulfite conversion using the MethylEdge Bisulfite Conversion kit (Promega, N1301) using a bead-based automated protocol. Bisulfite-converted DNA was mixed with 180µl of MethylEdge Binding Buffer and 1.8µl of 20 mg/ml decontaminated MagSi-DNA all round silica beads (MagnaMedics, MD02018) and left at room temperature for 15 minutes and washed twice with 220µl of 80% ethanol for 30 seconds. 60µl of MethylEdge desulfonation buffer was added to the beads and incubated at room temperature for 15 minutes, then washed twice with 100µl of 80% ethanol and dried for 1 minute. To elute the DNA, 20µl 10 mM Tris-HCL, pH 8.5 (Qiagen, 19086) was added to DNA-bead mix and incubated in a Thermomixer C (Eppendorf, 5382000015) at 56°C while being centrifuged at 2,000 rpm for 15 minutes. The Bisulfite converted single stranded DNA was converted to double stranded DNA through 1 cycle of PCR with random hexamer as previously described[61] followed by standard illumine library construction. Libraries were aligned using Novoalign V3.02.10 (www.novocraft.com) to human genome assembly GRCh37 (hg19). Duplicate reads were marke by Picard V1.31 (http://picard.sourceforge.net). and discarded (http://picard.sourceforge.net).

Methylation calls were generated using Novomethyl V1.01 (www.novocraft.com). All figures were generated by R statistical software[52].

HL60 treatment with ATRA, EPZ-6438 and GSK-J4

and 1X GlutaMAX (Gibco, Life Technologies, Fisher Thermo. 35050061) in the presence of 1µM ATRA (Sigma, **R2625**), 20 µM EPZ-6438 (Selleckchem, S7128) and 1µM GSK-J4 (Sigma, **420205**) in 96-well V-shape bottom plates for 9 days. Cells were treated with each compound every 3 days and reset to original seeding density. Concentration and viability were assessed every 3 days using a Countess II automated cell counter (Thermo Finsher, AMQAX100).

Stroma-cell containing cultures of CB cells

CB cells pooled from 3 donors were thawed, stained with antibodies. 50 of a P-NML ($CD45^{high}$;34^{high};38^{mid};71-;10-) phenotype were deposited into each well preloaded with 100,000 MS-5 stromal cells and alpha-MEM medium with 2 mM glutamine, 7.5% FBS and 10^{-4} M β-mercaptoethanol (Sigma) plus 50 ng/ml SCF (Novartis), 10 ng/ml FLT3L (Immunex), 1 ng/ml IL-3 (Novartis) , and 3 units/mL EPO (STEMCELL Technologies) in the presence of 1µM GSK-J4 (Sigma, 420205), 0.4 µM EPZ-6438 (Selleckchem, S7128) , 0.4 µM A-395[62]. The same dose of inhibitor was added to each half medium change performed after 4 and 7 days, respectively. Cells were cultured with 50 ng/ml SCF, 10 ng/mL FLT3L, 10 ng/ml IL-7 (R&D Systems), 3 units/ml EPO for the second week and 10 ng/ml IL-7 and 3 units/ml EPO only for the third week. After 3 weeks, all cells were harvested, stained with antibodies, and assessed by FACS to detect cultures of >10 cells belonging to one or more of the following cell populations: monocytes (CD45+;33+;14+), neutrophils (CD45+;33+;15+) and B cells (CD45+;33-;14+;15-;19+).

Acknowledgements

The authors thank the production and technical staff at Canada's Michael Smith Genome Sciences Centre for generating of human CB T cell reference epigenomes and their technical expertise in sequence generation and analysis. The authors also thank Darcy Wilkinson and Margaret Hale in the Eaves lab for collecting and processing CB cells,respectively, and to Dr. Cheryl Arrowsmith (University of Toronto, Ontario, Canada) for providing the A-395 compound. A.L., and C.H. were

supported by Frederick Banting and Charles Best Canada Graduate Scholarship Doctoral Awards (CGS-D), and F.W. by a University of BC Graduate Fellowship. This work was supported by the Canadian Institutes of Health Research (EP1-120589 & CEE-151619) and Genome Canada (C41EMT &C32EMT) under the Canadian Epigenetics, Environment and Health Research Consortium (to M.H.) and by the Terry Fox Research Institute (#1039 to MH, #1074 & TFF-122869 to M.H. and C.J.E.)

Contributions

A.L. and M.H. conceptualized the study design. A.L. performed the analysis. M.H. C.H., M.B., A.H.M and F.W. contributed to data analysis and interpretation. A.L., C.H., D.J.H.F.K., F.W., C.Q, M.W., M.M. P.L. and S.Z. performed experimental work. A.L., M.H. and C.J.E. wrote the manuscript.

Accession Numbers

The sequencing data reported in this study is accessible through the European Genome-Phenome Archive (EGA) under accession number EGAD00001004950.

References

1. Li, L., Lorzadeh, A. & Hirst, M. Regulatory variation: An emerging vantage point for cancer biology. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **6**, (2014).

2. Stricker, S. H., Köferle, A. & Beck, S. From profiles to function in epigenomics. *Nature Reviews Genetics* (2016). doi:10.1038/nrg.2016.138

3. Roadmap Epigenomics Consortium *et al.* Integrative analysis of 111 reference human epigenomes. *Nature* (2015). doi:10.1038/nature14248

4. Dunham, I. *et al.* An integrated encyclopedia of DNA elements in the human genome. *Nature* (2012). doi:10.1038/nature11247

5. Pott, S. & Lieb, J. D. What are super-enhancers? *Nat. Genet.* **47**, 8 (2014).

6. Local, A. *et al.* Identification of H3K4me1-associated proteins at mammalian enhancers. *Nat. Genet.* (2018). doi:10.1038/s41588-017-0015-6

7. Jaenisch, R. *et al.* Histone H3K27ac separates active from poised enhancers and predicts

developmental state. *Proc. Natl. Acad. Sci.* (2010). doi:10.1073/pnas.1016071107

8. Orkin, S. H. & Zon, L. I. Hematopoiesis: An Evolving Paradigm for Stem Cell Biology. *Cell* (2008). doi:10.1016/j.cell.2008.01.025

9. Lara-Astiaso, D. *et al.* Chromatin state dynamics during blood formation HHS Public Access. *Science (80-.).* (2014). doi:10.1126/science.1256271

10. Corces, M. R. *et al.* Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. *Nat. Genet.* (2016). doi:10.1038/ng.3646

11. Buenrostro, J. D. *et al.* Integrated Single-Cell Analysis Maps the Continuous Regulatory Landscape of Human Hematopoietic Differentiation. *Cell* (2018). doi:10.1016/j.cell.2018.03.074

12. Farlik, M. *et al.* DNA Methylation Dynamics of Human Hematopoietic Stem Cell Differentiation. *Cell Stem Cell* (2016). doi:10.1016/j.stem.2016.10.019

13. Eaves, C. J. Hematopoietic stem cells: concepts, definitions, and the new reality. *Blood* **125**, 2605 LP-2613 (2015).

14. Laurenti, E. & Göttgens, B. From haematopoietic stem cells to complex differentiation landscapes. *Nature* (2018). doi:10.1038/nature25022

15. Notta, F. *et al.* Distinct routes of lineage development reshape the human blood hierarchy across ontogeny. *Science (80-.).* (2016). doi:10.1126/science.aab2116

16. Knapp, D. J. H. F. *et al.* A topological view of human CD34+ cell state trajectories from integrated single-cell output and proteomic data. *Blood* blood-2018-10-878025 (2019). doi:10.1182/blood-2018-10-878025

17. Lorzadeh, A., Gutierrez, R. L., Jackson, L., Moksa, M. & Hirst, M. Generation of native chromatin immunoprecipitation sequencing libraries for nucleosome density analysis. *J. Vis. Exp.* **2017**, (2017).

18. Lorzadeh, A. *et al.* Nucleosome Density ChIP-Seq Identifies Distinct Chromatin Modification Signatures Associated with MNase Accessibility. *Cell Rep.* **17**, 2112–2124 (2016).

19. Stunnenberg, H. G. *et al.* The International Human Epigenome Consortium: A Blueprint for Scientific Collaboration and Discovery. *Cell* (2016). doi:10.1016/j.cell.2016.11.007

20. Wen, B., Wu, H., Shinkai, Y., Irizarry, R. A. & Feinberg, A. P. Large organized chromatin K9-modifications (LOCKs) distinguish differentiated from embryonic stem cells. *Nat. Genet.* (2009). doi:10.1038/ng.297.Large

21. Ernst, J. & Kellis, M. Chromatin-state discovery and genome annotation with ChromHMM. *Nat. Protoc.* (2017). doi:10.1038/nprot.2017.124

22. Dong, K. B. *et al.* DNA methylation in ES cells requires the lysine methyltransferase G9a

but not its catalytic activity. *EMBO J.* (2008). doi:10.1038/emboj.2008.193

23. Lehnertz, B. *et al.* Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr. Biol.* (2003). doi:10.1016/S0960-9822(03)00432-9

24. Freitag, M. & Selker, E. U. Controlling DNA methylation: Many roads to one modification. *Current Opinion in Genetics and Development* (2005). doi:10.1016/j.gde.2005.02.003

25. Hnisz, D. *et al.* XSuper-enhancers in the control of cell identity and disease. *Cell* (2013). doi:10.1016/j.cell.2013.09.053

26. Breitman, T. R., Selonick, S. E. & Collins, S. J. Induction of differentiation of the human promyelocytic leukemia cell line (HL-60) by retinoic acid. *Proc. Natl. Acad. Sci.* (2006). doi:10.1073/pnas.77.5.2936

27. Álvarez-Errico, D., Vento-Tormo, R., Sieweke, M. & Ballestar, E. Epigenetic control of myeloid cell differentiation, identity and function. *Nature Reviews Immunology* (2015). doi:10.1038/nri3777

28. Deaton, A. M. *et al.* Cell type-specific DNA methylation at intragenic CpG islands in the immune system. *Genome Res.* (2011). doi:10.1101/gr.118703.110

29. Stewart, M. H. *et al.* The histone demethylase Jarid1b is required for hematopoietic stem cell self-renewal in mice. *Blood* (2015). doi:10.1182/blood-2014-08-596734

30. Bröske, A. M. *et al.* DNA methylation protects hematopoietic stem cell multipotency from myeloerythroid restriction. *Nat. Genet.* (2009). doi:10.1038/ng.463

31. Bock, C. *et al.* DNA Methylation Dynamics during In Vivo Differentiation of Blood and Skin Stem Cells. *Mol. Cell* (2012). doi:10.1016/j.molcel.2012.06.019

32. Hodges, E. *et al.* Directional DNA methylation changes and complex intermediate states accompany lineage specificity in the adult hematopoietic compartment. *Mol. Cell* (2011). doi:10.1016/j.molcel.2011.08.026

33. Xie, H. *et al.* Polycomb repressive complex 2 regulates normal hematopoietic stem cell function in a developmental-stage-specific manner. *Cell Stem Cell* (2014). doi:10.1016/j.stem.2013.10.001

34. Yu, W. *et al.* Depletion of polycomb repressive complex 2 core component EED impairs fetal hematopoiesis. *Cell Death Dis.* (2017). doi:10.1038/cddis.2017.163

35. Petruk, S. *et al.* Structure of Nascent Chromatin Is Essential for Hematopoietic Lineage Specification. *Cell Rep.* (2017). doi:10.1016/j.celrep.2017.03.035

36. Kamminga, L. M. *et al.* The Polycomb group gene Ezh2 prevents hematopoietic stem cell exhaustion. *Blood* (2006). doi:10.1182/blood-2005-09-3585

37. Xu, J. *et al.* Developmental control of polycomb subunit composition by GATA factors mediates a switch to non-canonical functions. *Mol. Cell* (2015). doi:10.1016/j.molcel.2014.12.009

38. Mochizuki-Kashio, M. *et al.* Dependency on the polycomb gene Ezh2 distinguishes fetal from adult hematopoietic stem cells. *Blood* (2011). doi:10.1182/blood-2011-03-340554

39. Ernst, T. *et al.* Inactivating mutations of the histone methyltransferase gene EZH2 in myeloid disorders. *Nat. Genet.* (2010). doi:10.1038/ng.621

40. Ganji, G. *et al.* Mutation of A677 in histone methyltransferase EZH2 in human B-cell lymphoma promotes hypertrimethylation of histone H3 on lysine 27 (H3K27). *Proc. Natl. Acad. Sci.* (2012). doi:10.1073/pnas.1116418109

41. Zhu, J. *et al.* Genome-wide chromatin state transitions associated with developmental and environmental cues. *Cell* (2013). doi:10.1016/j.cell.2012.12.033

42. Waddington, C. H. Genetic Assimilation of the Bithorax Phenotype. *Evolution (N. Y).* (2006). doi:10.2307/2406091

43. Bernstein, B. E. *et al.* A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells. *Cell* (2006). doi:10.1016/j.cell.2006.02.041

44. Mikkelsen, T. S. *et al.* Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* (2007). doi:10.1038/nature06008

45. Hawkins, R. D. *et al.* Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. *Cell Stem Cell* (2010). doi:10.1016/j.stem.2010.03.018

46. Goldrath, A. W. *et al.* Memory T and memory B cells share a transcriptional program of self-renewal with long-term hematopoietic stem cells. *Proc. Natl. Acad. Sci.* (2006). doi:10.1073/pnas.0511137103

47. Swift, J. *et al.* Lamins regulate cell trafficking and lineage maturation of adult human hematopoietic cells. *Proc. Natl. Acad. Sci.* (2013). doi:10.1073/pnas.1304996110

48. Sabari, B. R. *et al.* Coactivator condensation at super-enhancers links phase separation and gene control. *Science (80-.).* (2018). doi:10.1126/science.aap9195

49. Butterfield, Y. S. *et al.* JAGuaR: Junction alignments to genome for RNA-seq reads. *PLoS One* (2014). doi:10.1371/journal.pone.0102398

50. Tarasov, A., Vilella, A. J., Cuppen, E., Nijman, I. J. & Prins, P. Sambamba: Fast processing of NGS alignment formats. *Bioinformatics* (2015). doi:10.1093/bioinformatics/btv098

51. Tripathi, S. *et al.* Meta- and Orthogonal Integration of Influenza 'oMICs' Data Defines a Role for UBR4 in Virus Budding. *Cell Host Microbe* (2015). doi:10.1016/j.chom.2015.11.002

52. Team, R. D. C. & R Development Core Team, R. R: A Language and Environment for Statistical Computing. *R Found. Stat. Comput.* (2016). doi:10.1007/978-3-540-74686-7

53. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* (2010). doi:10.1093/bioinformatics/btp698

54. Feng, J., Liu, T., Qin, B., Zhang, Y. & Liu, X. S. Identifying ChIP-seq enrichment using MACS. *Nat. Protoc.* **7**, 1728–1740 (2012).

55. Heinz, S. *et al.* Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Mol. Cell* **38**, 576–589 (2010).

56. Ramírez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* (2016). doi:10.1093/nar/gkw257

57. Ernst, J. & Kellis, M. ChromHMM: automating chromatin-state discovery and characterization. *Nat. Methods* (2012). doi:10.1038/nmeth.1906

58. Löffler-Wirth, H., Kalcher, M. & Binder, H. OposSOM: R-package for high-dimensional portraying of genome-wide expression landscapes on bioconductor. *Bioinformatics* (2015). doi:10.1093/bioinformatics/btv342

59. Quinlan, A. R. & Hall, I. M. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* (2010). doi:10.1093/bioinformatics/btq033

60. Madani Tonekaboni, S. A., Mazrooei, P., Kofia, V., Haibe-Kains, B. & Lupien, M. CREAM: Clustering of genomic REgions Analysis Method. *bioRxiv* 222562 (2018). doi:10.1101/222562

61. Hui, T. *et al.* High-Resolution Single-Cell DNA Methylation Measurements Reveal Epigenetically Distinct Hematopoietic Stem Cell Subpopulations. *Stem Cell Reports* (2018). doi:10.1016/j.stemcr.2018.07.003

62. He, Y. *et al.* The EED protein-protein interaction inhibitor A-395 inactivates the PRC2 complex. *Nat. Chem. Biol.* (2017). doi:10.1038/nchembio.2306