

# Inactivation of a Single Copy of *Crebbp* Selectively Alters Pre-mRNA Processing in Mouse Hematopoietic Stem Cells

Madeleine E. Lemieux<sup>1</sup>, Ziming Cheng<sup>2</sup>, Qing Zhou<sup>2</sup>, Ruth White<sup>3</sup>, John Cornell<sup>4</sup>, Andrew L. Kung<sup>1</sup>, Vivienne I. Rebel<sup>2,5\*</sup>

**1** Department of Pediatric Oncology, Dana-Farber Cancer Institute and Harvard Medical School, Boston, Massachusetts, United States of America, **2** Greehey Children's Cancer Research Institute (GCCRI), The University of Texas Health Science Center at San Antonio (UTHSCSA), San Antonio, Texas, United States of America, **3** Department of Cell and Developmental Biology, Oregon Health and Science University, Portland, Oregon, United States of America, **4** Department of Epidemiology and Biostatistics, The University of Texas Health Science Center at San Antonio (UTHSCSA), San Antonio, Texas, United States of America, **5** Department of Cellular and Structural Biology, The University of Texas Health Science Center at San Antonio (UTHSCSA), San Antonio, Texas, United States of America

## Abstract

Global expression analysis of fetal liver hematopoietic stem cells (FL HSCs) revealed the presence of unspliced pre-mRNA for a number of genes in normal FL HSCs. In a subset of these genes, *Crebbp*<sup>+/-</sup> FL HSCs had less unprocessed pre-mRNA without a corresponding reduction in total mRNA levels. Among the genes thus identified were the key regulators of HSC function *Itga4*, *Msi2* and *Tcf4*. A similar but much weaker effect was apparent in *Ep300*<sup>+/-</sup> FL HSCs, indicating that, in this context as in others, the two paralogs are not interchangeable. As a group, the down-regulated intronic probe sets could discriminate adult HSCs from more mature cell types, suggesting that the underlying mechanism is regulated with differentiation stage and is active in both fetal and adult hematopoiesis. Consistent with increased myelopoiesis in *Crebbp* hemizygous mice, targeted reduction of CREBBP abundance by shRNA in the multipotent EML cell line triggered spontaneous myeloid differentiation in the absence of the normally required inductive signals. In addition, differences in protein levels between phenotypically distinct EML subpopulations were better predicted by taking into account not only the total mRNA signal but also the amount of unspliced message present. CREBBP thus appears to selectively influence the timing and degree of pre-mRNA processing of genes essential for HSC regulation and thereby has the potential to alter subsequent cell fate decisions in HSCs.

**Citation:** Lemieux ME, Cheng Z, Zhou Q, White R, Cornell J, et al. (2011) Inactivation of a Single Copy of *Crebbp* Selectively Alters Pre-mRNA Processing in Mouse Hematopoietic Stem Cells. PLoS ONE 6(8): e24153. doi:10.1371/journal.pone.0024153

**Editor:** Kevin D. Bunting, Emory University, United States of America

**Received:** June 2, 2011; **Accepted:** August 1, 2011; **Published:** August 25, 2011

**Copyright:** © 2011 Lemieux et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was supported with funding from the Greehey Children's Cancer Research Institute/University of Texas Health Science Center in San Antonio (UTHSCSA) and a National Cancer Institute-Cancer Center Support Grant (2P30 CA054174-17) for the UTHSCSA Flow Cytometry Core. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: rebel@uthscsa.edu

## Introduction

Cyclic-AMP-responsive element binding protein (CREB) binding protein (CREBBP) – more commonly referred to as CBP – is a multifunctional protein which facilitates gene expression through several mechanisms, including chromatin remodeling, acetylation of associated proteins, and recruitment of the basal transcription machinery to promoters [1]. We have previously shown that CREBBP and its paralog EP300 are essential for proper hematopoietic stem cell regulation but are nevertheless not functionally redundant in this setting: both copies of *Crebbp* are essential for HSC self-renewal while *Ep300*<sup>+/-</sup> HSCs are not compromised in this respect [2]. Inactivation of a single copy of the *Crebbp* gene results in multi-lineage defects in differentiation with a clear excess in myeloid cell production and an age-dependent increase in the incidence of hematologic malignancies [3].

CREBBP also acts as a scaffold in numerous protein-protein interactions [4] so that changes in its levels have the potential to broadly affect cellular processes by altering multiple signaling and

transcriptional pathways. In particular, it has the potential to act as a signal integrator within the hematopoietic system [5] through its interaction with both ubiquitous transcription factors such as SP1 [6,7] and the glucocorticoid receptor (NR3C1) [8,9] and with factors like SFPI1/PU.1 [10] and C/EBPalpha [11] which are essential for HSC function [12,13].

In addition to its activities as coactivator and integrator, confocal microscopy studies have localized CREBBP to nuclear speckles containing splicing proteins [14,15] and it has been shown to regulate 3'-end processing [16]. Both CREBBP and EP300 have furthermore been shown to be concentrated at both 5' and 3' ends of genes with which they associate [17]. It thus appears that CREBBP is involved in pre-mRNA maturation. In addition, experiments in macrophages [18] and T-cells [19] have shown CREBBP/EP300 to be present at the promoters of early response genes, even in the absence of stimulus, and associated with the production of full-length, unspliced transcripts. Other recent studies have reported that a large majority of genes with a paused polymerase produce full-length transcripts, although often at levels

below detection by expression microarrays[20] and RNA-seq studies have documented the presence of low-abundance intronic sequences in B-cell, kidney[21] and embryonic stem cell lines[22]. It has also been noted that HSCs prime multiple lineage programs prior to commitment decisions[23] and that HSCs normally contain unspliced transcripts that disappear as HSCs are driven to proliferate and differentiate[24]. Destabilization of this primed state has been proposed as a first stage of a cascade towards differentiation[25]. The general model that emerges from these findings is that unspliced, full-length transcripts are produced as a means of bookmarking loci and keeping them in an open chromatin state to facilitate subsequent rapid transcriptional up-regulation[18,19].

The presence of unspliced transcripts in HSCs and the links between CREBBP and EP300 with the constitutive production of unspliced RNA and with pre-mRNA processing prompted us to examine more closely an anomaly we had noted in microarray-based gene expression studies but had previously attributed to experimental “noise”. We had noticed that more than half of the probe sets down-regulated in *Crebbp*<sup>+/-</sup> FL HSCs relative to wild-type (WT) mapped entirely within introns, rather than detecting exonic or spliced sequences. We therefore set out to test whether this might be evidence that reduced CREBBP levels selectively alter the generation of full-length, unspliced pre-mRNA. We also asked whether this process might be associated with differentiation since self-renewal and lineage commitment are both responses for which HSCs are primed.

As predicted by our microarray studies, we found that several genes associated with HSC function showed variable ratios of intronic to total RNA signal in *Crebbp*<sup>+/-</sup> FL HSCs relative to WT by quantitative RT-PCR (qRT). In addition to primary FL HSCs, we looked for CREBBP-associated changes in mRNA, intronic message and protein levels in multipotent EML cells which retain lymphoid, myeloid and erythroid potential in the presence of Stem Cell Factor (SCF) but can be induced to undergo differentiation by treatment with appropriate stimuli[26]. Targeted reduction of *Crebbp* by shRNA in EML cells was sufficient to trigger widespread myeloid differentiation of EML cells, bypassing their usual requirement for withdrawal of SCF and treatment with retinoic acid, interleukin-3 (IL-3) and granulocyte-macrophage colony stimulating factor (GM-CSF). A subset of genes tested also showed altered levels of intronic message in subpopulations of EML cells at different phenotypically-defined stages of development which corresponded to changes in protein abundance not predicted by full-length mRNA levels. Furthermore, the differences in intronic levels correlated with differentiation stage-dependent changes in CREBBP levels. Taken together, our data suggest a novel, cell type-specific function for CREBBP in regulating the timing and extent of pre-mRNA splicing of key regulators of HSC maintenance and function.

## Results

### Down-regulation of intronic probe sets without proportional changes in total mRNA levels

We have carried out a global analysis of expression profiles of WT and *Crebbp*<sup>+/-</sup> FL HSCs. Our in-house annotation of the Affymetrix Mouse 430 2.0 microarrays (see Materials and Methods for details) indicated that ~60% of down-regulated probe sets - but none of the up-regulated ones - mapped on the coding strand but entirely within intronic regions of genes (Table 1 and Table S1). This is a significant enrichment over the array background of 15% (6780/45037, excluding Affymetrix control probe sets,  $\chi^2$  test  $p < 1.1 \times 10^{-13}$ ). Of the 83 intronic probe sets

representing 77 distinct transcripts (both protein-coding and non-coding genes), 13 were designed to detect expressed sequence tags (ESTs) in libraries constructed from adult HSCs[24,27,28]. In total, there was evidence in the EST database for intronic message for 50/77 transcripts (Table S1). We will refer to these probe sets and their associated genes as “intronic probe sets” and “intronic targets”, respectively, to distinguish them from “mRNA probe sets” and “exonic probe sets” detecting spliced mRNA or 3'-most exons. Notably, of the 74/77 transcripts with both intronic and mRNA probe sets, we found changes in total mRNA levels for only 2: *Mbnl1* and *Meis1*. In both cases, the mRNA probe set could detect either a putative (non-RefSeq) alternative 3' exon or an intron in a longer isoform.

An attenuated bias was found in an *Ep300*<sup>+/-</sup> FL HSC expression data set we produced concurrently in which 15/35 differentially expressed probe set were intronic ( $\chi^2$  test  $p = 1.3 \times 10^{-5}$ ), again all of them down-regulated. In contrast, probe sets altered in *Crebbp*<sup>+/-</sup> mouse embryonic fibroblasts (MEFs) relative to WT MEFs showed no enrichment for intronic probe sets (Table 1), indicating that the effect is cell type-specific. Like CREBBP, CDKN1A (p21<sup>Cip1/Waf1</sup>) is required for HSC self-renewal and its loss leads to exhaustion of the HSC pool upon serial transplantation[29]. Unlike *Crebbp*<sup>+/-</sup> FL HSCs, however, FL HSCs null for *Cdkn1a* showed both up- and down-regulated intronic probe sets at a frequency close to predicted relative to their WT controls (Table 1). The change in intronic sequence abundance is thus not common to all FL HSCs with compromised self-renewal ability.

### Down-regulated intronic probe sets as a discriminating HSC signature

To determine whether these differentially expressed intronic probe sets were functionally relevant, we first asked whether the intronic probe sets down-regulated in *Crebbp*<sup>+/-</sup> FL HSC could distinguish HSCs from more mature cell types by hierarchical clustering. We took advantage of previously published data sets available for download from the Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/geo/>). In the absence of appropriate data sets from fetal liver, we selected expression series comparing adult bone marrow (BM) HSCs and various populations of early progenitors and more mature hematopoietic cells[30,31,32]. The fetal liver is the primary site of hematopoiesis from day 11 of mouse development, a function taken over by the bone marrow after birth[33]. Although FL HSCs are phenotypically and functionally distinguishable from BM HSCs, FL HSCs are capable of fully reconstituting long-term hematopoiesis in adult animals[34,35] and unspliced transcripts had previously been detected in adult BM[24].

We used either all 83 probe sets for clustering or the single most highly expressed probe set per transcript (77 in total) to avoid overweighting genes detected by multiple intronic probe sets. The cell populations clustered the same way with both approaches. Figure 1A shows the heat map and sample dendrogram for GSE6506[31]. The two adult BM HSC (SP-LSK) samples of this data set are grouped together on a branch (indicated in red) distinct from all the more differentiated cell types. As a control, we generated 10,000 random samples of 83 probe sets annotated as being both intronic and expressed in WT or *Crebbp*<sup>+/-</sup> FL HSCs. Only 0.52% of the random samples yielded such a clean distinction of HSCs from mature cell types, arguing against the separation occurring purely by chance and suggesting that the underlying process may be common to both adult and fetal HSCs.

Are the intronic probe sets simply surrogate measures of the total mRNA level? For each gene with at least one mRNA probe

**Table 1.** Distribution of Probe Sets Differentially Expressed Relative to Matched Wild-type Cells.

| Probe set target                       | <i>Crebbp+/-</i> HSC |                | <i>Ep300+/-</i> HSC |               | <i>Cdkn1a-/-</i> HSC |               | <i>Crebbp+/-</i> MEF |              |
|--|----------------------|----------------|---------------------|---------------|----------------------|---------------|----------------------|--------------|
|  | Up                   | Down           | Up                  | Down          | Up                   | Down          | Up                   | Down         |
| <b>mRNA</b>                            |                      |                |                     |               |                      |               |                      |              |
| RefSeq                                 | 48                   | 27             | 4                   | 12            | 5                    | 48            | 18                   | 33           |
| non-RefSeq*                            | 6                    | 0              | 1                   | 2             | 0                    | 1             | 0                    | 5            |
| ncRNA                                  | 0                    | 3              | 0                   | 0             | 0                    | 1             | 0                    | 1            |
| putative ncRNA                         | 1                    | 2              | 0                   | 0             | 2                    | 4             | 0                    | 0            |
| putative alt 3'UTR                     | 5                    | 14             | 0                   | 1             | 0                    | 3             | 0                    | 0            |
| <b>Total mRNA</b>                      | 60                   | 46             | 5                   | 15            | 7                    | 57            | 18                   | 39           |
| <b>intronic (expected<sup>†</sup>)</b> | <b>0 (15)</b>        | <b>83 (15)</b> | <b>0 (3)</b>        | <b>15 (3)</b> | <b>3 (7)</b>         | <b>15 (7)</b> | <b>0 (5)</b>         | <b>1 (5)</b> |
| repeat                                 | 0                    | 6              | 0                   | 0             | 0                    | 3             | 0                    | 0            |
| ambiguous/unclear                      | 1                    | 1              | 0                   | 0             | 0                    | 6             | 0                    | 2            |

doi:10.1371/journal.pone.0024153.t001

set (74/77 genes), the one with the largest difference in the *Crebbp+/-* vs WT FL HSC data set was retained. Figure 1B shows the heat map for the same genes as 1A and in the same row order but with the various cell populations clustered by mRNA probe set. In this case, the HSC branch (shown in red) is no longer clearly distinct from mature cell types but is part of a subtree including all lymphoid cell types. Since the two heat maps are row-normalized, it is clear that the relative expression levels of intronic and mRNA probe sets for the same gene are often quite different (for example, compare the relative levels for *Mds1*, marked by a red asterisk in both Figure 1A and 1B). This is inconsistent with intronic transcript levels being a constant proportion of total message produced in all cell types.

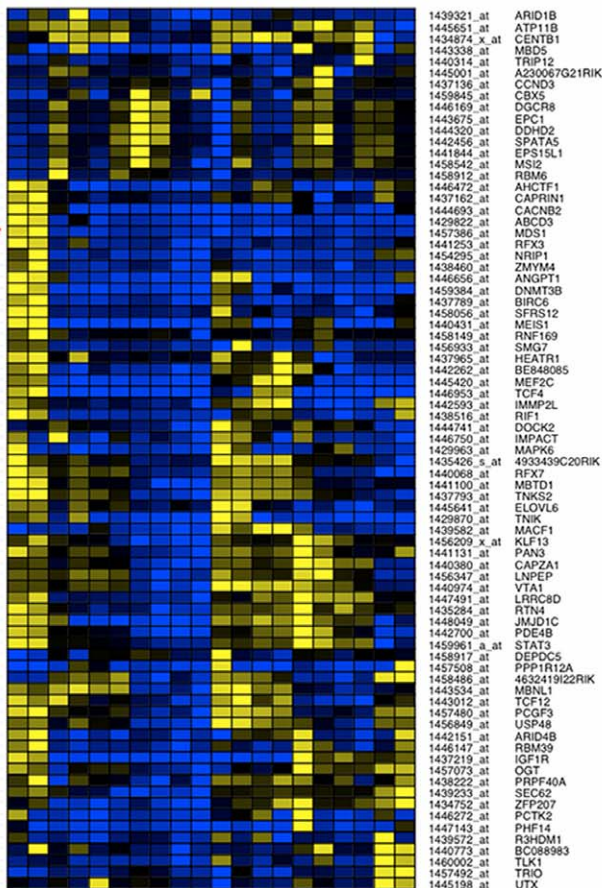
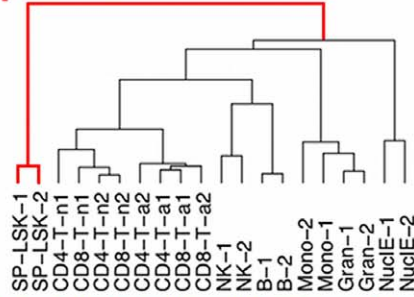
Hierarchical clustering of 2 other published data sets were carried out in the same manner (Figure 1C). In each case, the intronic probe sets (Figure 1C left) segregated the experiment's HSCs (indicated by red bars) from mature cell types. In addition, hierarchical clustering using mRNA probe sets (Figure 1B and 1C right) resulted in somewhat different clustering tree structures than those generated using intronic probe sets (Figure 1A, 1C left), depending on the populations involved. Where intronic and mRNA clustering dendrograms are noticeably different (based on total distance along the dendrogram arms and positions of branch points), the intronic probe sets separate HSCs from early progenitors better than do the mRNA probe sets (shorter overall distance, different branches). The changes in intron levels are therefore not strictly reflections of those in total mRNA differences.

Encouraged by this evidence that the presence of unspliced message was non-random and the finding that our intronic signature was associated with HSC differentiation, we then carried out 3 independent qRT experiments in purified FL HSCs with primers designed to detect spliced or unspliced transcripts (Table 2). Follow-up targets were selected based on previous reports of their involvement in HSC function [36,37,38,39,40,41] or, in the case of *Rbpms*, for its implication in developmental processes [42] and physical association with SMAD4 [43], a factor critical for HSC self-renewal [44]. For exonic amplicons, results from qRT consistently agreed well with the microarray data (Table S2), showing little if any change between WT and *Crebbp+/-* samples. For intronic amplicons, the reproducibility of the results depended on the proximity of the amplicon to poly(A) tracts of sufficient length to potentially prime the oligo(dT) reverse transcription [45] that we used to amplify our starting material

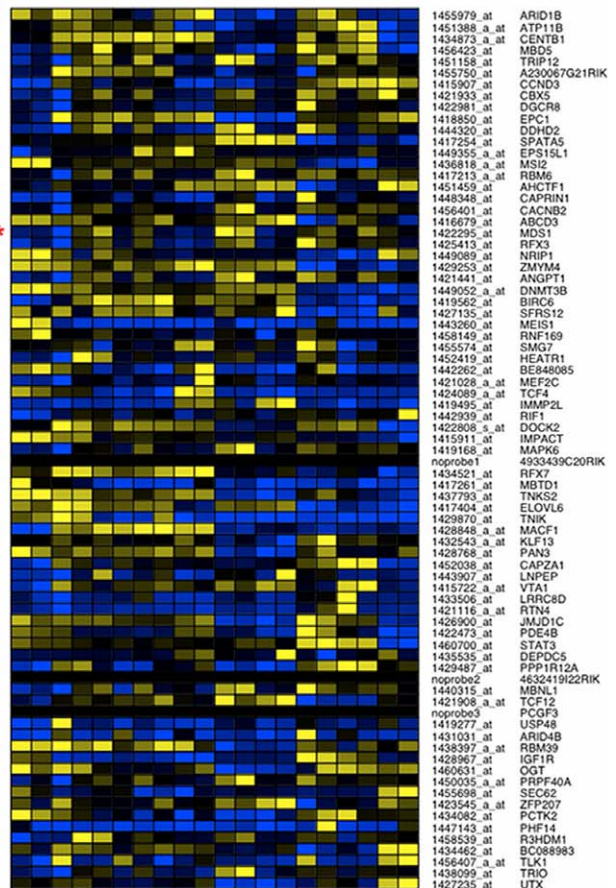
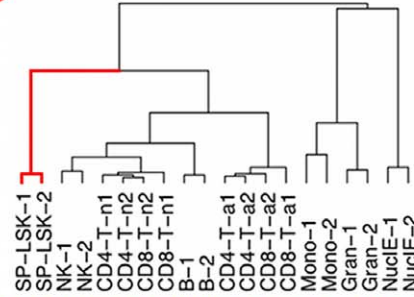
(Table S2). Despite this technical limitation, we found that *Tcf4/E2-2* and *Msi2* intronic amplicons both showed reduced levels in *Crebbp+/-* HSC relative to WT. Two other genes, *Igfa4* and *Rbpms*, which had down-regulated intronic probe sets (although only to a p-value <0.1 significance level in the microarray data) also exhibited greater reduction in intronic than total signal. Figure 2 shows a summary of these results. Plotted are the ratios of total mRNA and intronic signal of WT FL HSCs relative to either *Crebbp+/-* (A) or *Ep300+/-* (B). Genes for which the ratio is similar for total mRNA and unspliced transcripts fall between the vertical and horizontal grey lines marking 1.5-fold differences. *Tcf4*, *Msi2* and *Rbpms*, on the other hand, fall within the vertical lines, indicating similar mRNA levels in WT and *Crebbp+/-* HSCs, but lie well above the horizontal line marking a greater than 1.5-fold excess intronic signal in WT cells. *Igfa4* mRNA levels decreased roughly 2-fold in *Crebbp+/-* HSCs relative to WT but unspliced transcript levels are reduced even more substantially (~10-fold). Consistent with the array data, *Ep300+/-* samples showed attenuated, if any, difference in intronic and total mRNA ratios (Figure 2B).

Many of our target genes were >100 kilobases long and their mRNA was detected at modest levels on the microarrays. Consequently, we wondered whether changes in intronic sequence detection was a widespread phenomenon and whether it correlated with transcript length or expression level. We therefore carried out a further set of qRT reactions using linked primer pairs for which either the forward or the reverse primer was common to both the exonic and intronic amplicons for a given gene (see Table 2 for details). We included genes of varying lengths and predicted expression levels based on the microarray data. Figure 2C shows total and intronic mRNA levels for each gene in WT and *Crebbp+/-* HSCs. As can be seen in the figure, not all transcripts were altered in *Crebbp+/-* HSCs (e.g. *170081L11Rik* indicated as "L11Rik"). In some cases, both intronic and total signal changed proportionately (e.g. *Dhcr7*, *Mllt10* where the slope from the WT square to the *Crebbp+/-* circle is close to 1) while in others the variation was greater in the intronic signal than in total mRNA level (e.g. *Ly6*, *Pole*, *Meis1*). The change in intronic signal relative to total message thus seems gene-specific rather than a function of length or mRNA abundance (compare *Ly6E*, *Igf1R* and *Angpt1*, for example). Overall, the clustering and qRT results indicate that the mechanism underlying the alterations in intronic signal is both cell-type specific and selective in its targets.

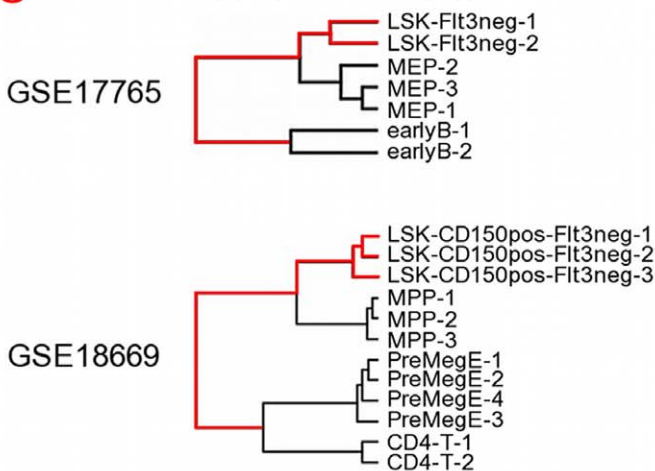
**A** GSE6506 - Intronic Probe Sets



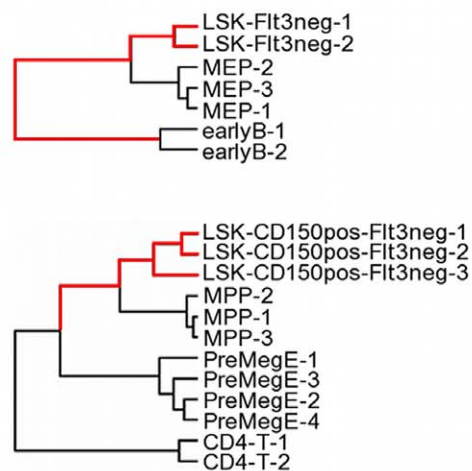
**B** GSE6506 - mRNA Probe Sets



**C** Intronic Probe Sets



mRNA Probe Sets



**Figure 1. Discriminatory power of intronic probe sets in hierarchical clustering of hematopoietic cell populations.** (A) Hierarchical clustering of expression data from Chambers *et al*[31] using intronic probe sets down-regulated in *Crebbp*<sup>+/-</sup> vs WT HSCs. Cell types represented are adult HSCs (SP-LSK, red branch on the dendrogram), activated (a) and naive (n) CD4<sup>+</sup> and CD8<sup>+</sup> lymphocytes, natural killer (NK) cells, B cells, nucleated erythrocytes (NuclE), monocytes (Mono), and granulocytes (Gran). (B) Same data set as in (A) but with cell populations clustered using mRNA probe sets for the same target genes. In cases where multiple probe sets were available, the one with the largest difference in *Crebbp*<sup>+/-</sup> vs WT HSCs was used. Genes are shown in the same order as in (A) with HSC branches shown in red to the nearest junction point with a mature cell type subtree. (C) Clustering as in (A) (left) and (B) (right) of 2 other published data sets[30,32] with different subpopulations. As before, red dendrogram branches highlight the distance from HSCs to the nearest mature cell type junction point. LSK: Lin- SCA1+ KIT+; SP: side population; LK: Lin- KIT+; MEP: megakaryocyte/erythrocyte progenitor, LK SCA1- CD34- Fc-gammaRII/III+; MPP: multi-potential progenitor, LSK CD150- FLT3+; PreMegE: megakaryocyte/erythrocyte progenitors, LK SCA1- CD150+ CD105- CD41-. doi:10.1371/journal.pone.0024153.g001

## Classification of genes hosting differentially expressed intronic probe sets

We found no particular gene ontology term or pathway significantly over-represented among the 77 distinct transcripts hosting these 83 intronic probe sets. The list of functional categories represented by their encoded proteins, however, included kinases as well as genes associated with transcription, cell cycle regulation and cellular division, growth regulation and organismal development, chromatin binding and modification, and RNA binding and processing (Table 3). This last category is notable as it suggests that CREBBP might indirectly modulate intronic signal levels by regulating the expression of proteins that are themselves involved at different stages of RNA metabolism. Furthermore, CREBBP itself has also been localized by immunofluorescence to splicing speckles[14,15] suggesting that it may also play a direct role in the splicing process. Interestingly, among the non-intronic targets summarized in Table 1 and listed in Table S1 are 2 down-regulated probe sets that detect the non-coding RNA, *Malat1*, which is also associated with splicing speckles[46] and regulates alternative splicing of pre-mRNA[47]. CREBBP may thus participate at several levels, both directly and indirectly, in a feed-forward loop regulating pre-mRNA processing.

We do not see a generalized impact on intronic levels with the reduction in *Crebbp* expression, however, indicating that there is some selectivity in the genes exhibiting these changes. We found evidence that the promoters of genes with differentially expressed intronic probe sets are enriched for a different set of transcription factor binding sites (TFBS) proximal to their transcription start sites (TSS, -500bp to +250bp) relative to the exonic target genes (Table 4 and Table S1). TFBS for ARID3A/BRIGHT, C/EBPalpha, GFI1, IRF1, NR3C1 (glucocorticoid receptor), SOX5 or SOX17 where significantly enriched only in the intronic target set. At least one of these TFBS was present in each of the intronic target promoters and 81% of promoters had sites for 3 or more of these transcription factors, all of which are expressed in FL HSCs based on our microarray data. Each of these factors except SOX5 has been implicated in hematopoietic regulation at various stages[13,48,49,50,51,52] and CREBBP has been shown to interact (directly or by homology with EP300) with C/EBPalpha, IRF1, NR3C1 and multiple SOX family members[4].

Taken together, these results indicate that reduction in CREBBP level in HSCs alters pre-mRNA processing in a regulated manner for a subset of genes. It is difficult to obtain sufficient material reproducibly from FL or adult HSC for in-depth molecular analyses so we turned for our subsequent studies to HSC-like EML cells which possess lymphomyeloid differentiation potential[26] and have previously been used as a model system for studying hematopoietic lineage commitment[53,54,55].

## Differentiation-associated down-regulation of CREBBP and spontaneous commitment to myeloid lineages upon knock-down of *Crebbp* expression in EML cells

*Crebbp* expression is widespread but developmentally regulated during embryogenesis[56] and its transcript[57] and protein levels (Figure 3A) are lower in certain committed progenitors and mature cells than in HSCs. Similarly, EML cells driven to myeloid differentiation by growth factor and all-trans retinoic acid (ATRA) stimulation down-regulate CREBBP levels (Figure 3B). Parental EML cells (Figure 3C i) showed a characteristic hand-mirror morphology that was lost as the cells began differentiating upon reduction of SCF and addition of IL-3 and ATRA (Figure 3C ii). As previously reported[53], the shift to medium containing only GM-CSF resulted in some cell death but cultures further stimulated with GM-CSF + ATRA gave rise to mature granulocytes (~70% of cells, Figure 3C iii) mixed with macrophages (Figure 3C iv).

Cells exposed to control shRNA against GFP continued to express CREBBP at WT levels (compare Figure 3D EML and shGFP) and grew in suspension like parental EML cells. In contrast, reduction of CREBBP levels by lentiviral delivery of shRNA led to widespread death or stimulus-independent differentiation of EML cells (Figure 3D i-ii).

Forty *Crebbp* knock-down (shCrebbp) clones were isolated in SCF-containing methylcellulose and then passaged in liquid culture in the presence of SCF (i.e. under non-differentiating conditions). Within 3–4 weeks, 39% of clones were dead or dying, 39% had differentiated and the remainder underwent extensive cell death before subclones emerged, either looking and growing like the parental EML cells or else differentiating. Figure 3E shows growth histories for 37 clones still alive after 40 days. Each point reflects the state of the culture on the indicated day post-cloning: “EML-like” cells grew rapidly in suspension (<5% adherent, blue circles), “mixed” cultures contained less than 50% differentiated (adherent) cells (purple circles), “differentiated” cultures contained mostly adherent cells (red circles) and “dying” cultures contained a majority of dead cells (black circles). Clones with lower levels of *Crebbp* message (as measured by qRT at week 5) were more likely to differentiate and eventually die. Clones that survived for the full length of the experiment (120 days) generally grew in suspension like parental EML cells and seemed to arise as subclones in cultures that had undergone substantial cell death. As controls, we also followed 20 clones receiving control shRNA targeting GFP and 20 that were cloned but not exposed to virus. All grew like parental EML cells for more than a month, indicating that differentiation and death were not predominantly side-effects of lentiviral treatment or cloning.

Down-regulation of CREBBP in EML cells thus mimics the mix of cell death and myeloid differentiation seen in these cells when they are driven to differentiate by ATRA, IL-3 and GM-CSF stimulation[53], suggesting that CREBBP either acts downstream

**Table 2.** Primer Sequences Used in qRT Reactions.

| Gene                 | Amplicon Type   | Forward (5'→3')          | Reverse (5'→3')          |
|----------------------|-----------------|--------------------------|--------------------------|
| <i>Crebbp</i>        | exonic          | AAGTCACCCAGCTCTCTCA      | GGCTGATTGGCCACATACTT     |
| <i>Ep300</i>         | exonic          | GCCTTCTCCACACCATGTTT     | CGAGCTGTGAAAGCATTGAA     |
| <i>Gapdh</i>         | exonic          | ACGTGCCCGCTGGAGAA        | CATGCCTGCTTACCACCTT      |
| <i>Ly6E</i>          | intron-spanning | CTTCCAACATGAGAGCTTCC     | GCAGATAACGTGATACAGTAATGG |
|                      | exon-intron     | GTTGTCAATGGCTGGCTTTT     | GCAGATAACGTGATACAGTAATGG |
| <i>1700081L11Rik</i> | intron-spanning | AAGCGTTTGATTCGGATGTC     | CTGTTGTATGATCTGGGAAAGG   |
|                      | exon-intron     | AAGGGAAGGCCTCTGGTTA      | CTGTTGTATGATCTGGGAAAGG   |
| <i>Dhcr7</i>         | intron-spanning | GATTGTAGCCTGGACCTCA      | GAGAGCTGCACAGGTTGGTA     |
|                      | exon-intron     | GAAGCTGGCGTGACAAGTCT     | GAGAGCTGCACAGGTTGGTA     |
| <i>Pole</i>          | intron-spanning | CTGATGGCCTTCACACTCA      | CTTTTGGAGCAGGCAGTAGG     |
|                      | exon-intron     | CTGATGGCCTTCACACTCA      | TGCTCTGCTCCTGCACTCTA     |
| <i>Mllt10</i>        | intron-spanning | TTCCATGCAGTATCGACATGA    | TGACCAGATGACTGCTGAGG     |
|                      | exon-intron     | AGGGAAGGGATCAAGGAGAA     | TGACCAGATGACTGCTGAGG     |
| <i>Igf1r</i>         | intron-spanning | AGCCCATGTGTGAGAAGACC     | ACGCAGGTTGTTGTCTGT       |
|                      | exon-intron     | AGCCCATGTGTGAGAAGACC     | CGGAGCAGAAAGTGGAAAGTC    |
|                      | intronic        | GTTGTGCCCCAGTGTTTTCT     | TGGTCCTCAGCCAAGAGACT     |
| <i>Angpt1</i>        | intron-spanning | TCTGCCAGCTGTGCTTGTT      | CGTGTGGTTTTGAACAGCAT     |
|                      | exon-intron     | TCTGCCAGCTGTGCTTGTT      | CGTGTGGTTTTGAACAGCAT     |
|                      | exonic          | GCGCTGGCAGTACAATGACAGTTT | CACATTGCCCATGTTGAATCCGGT |
|                      | intronic        | ACAACAGAGGCCACAAGCCTTAGA | CACCAGCCACTGCACAAGACATT  |
| <i>Meis1</i>         | intron-spanning | ACAATTCTGCCACCCGGTAT     | GTTCTCTCTGAACGAGTGGA     |
|                      | exon-intron     | CCCTGTAGCCCTCCTCAGTT     | GTTCTCTCTGAACGAGTGGA     |
|                      | exonic          | TAGGAGGAGGCATGGAAACCCAAA | AGGTGATGGTTGTTGTTGCATGGG |
|                      | intronic        | TGTGGTCTATTGCCACTGCTA    | TGCAACTCCATCTCTGACACCTT  |
| <i>Tcf4</i>          | exonic          | AGGACGAAATGCTGACCCTGAAGT | CGGGCAATGCTGCATGTAATTCCT |
|                      | intronic        | AACTCCCGTGGGAATGAGTGATGT | ATGCCCACTGCCAACATCATGAAC |
| <i>Msi2</i>          | exonic          | TGGTACTTAAGGCACAGCCAGTGT | TTCTGCGCTATCCCGTCTGTGAAT |
|                      | intronic        | TGCAGAAGCCAGCAAGTAGACTGT | AGGACTCGTAGCCAAGATGGTTT  |
| <i>Rbpms</i>         | exonic          | GGCAAACACGAAGATGGCCAAGAA | TATGGCTCCCTGGCAATGAACTGA |
|                      | intronic        | AGGTTCTGCCAATCCTGTCTTTT  | AAACTGTGCTGCAAGGTTTCCAAG |
| <i>Iiga4</i>         | exonic          | AGACCTTTGTACCTTTGCCCTCCA | AGAGAGAGCCTTCCCTGTTTGCTT |
|                      | intronic        | GCAGGACATCAAGTTGCCCTTGT  | AGGAATTCCCACCTGCTACCAACA |

doi:10.1371/journal.pone.0024153.t002

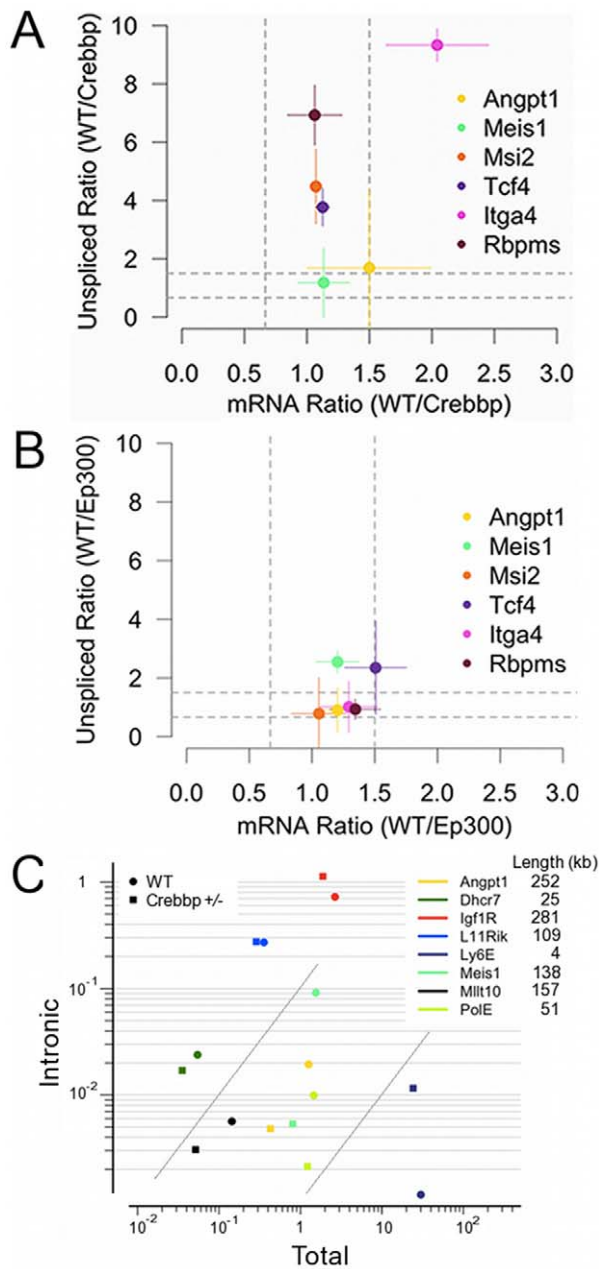
of the growth factor signaling cascade or can trigger differentiation through an independent, parallel pathway. The spontaneous myeloid differentiation of EML cells after *Crebbp* knock-down, reminiscent of the increased myelopoiesis seen in *Crebbp*<sup>+/-</sup> animals[3], and the differentiation-induced down-regulation of CREBBP protein levels, similar to that observed in differentiating hematopoietic cells (Figure 3A and [57]), both indicate that regulated reduction of CREBBP abundance is a key event in early hematopoietic differentiation.

#### Fluctuations in intronic-to-total mRNA ratios and protein levels in HSC-like EML cells depending on differentiation stage and CREBBP levels

We had confirmed by qRT that intron levels for *Iiga4*, *Msi2* and *Tcf4*, each of which has been implicated in hematopoietic development[38,41,58,59], were reduced in *Crebbp*<sup>+/-</sup> FL HSCs relative to WT (Figure 2). We now wanted to know whether a natural reduction in CREBBP protein occurring as a consequence of normal

differentiation processes in EML cells would result in similar intronic changes. More importantly, we wondered whether fluctuations in intronic levels could have a measurable impact on protein levels.

EML cell cultures kept undifferentiated by the presence of SCF are a mixture of CD34+, SCF-responsive, replicating cells and their progeny CD34-, SCF-independent cells that become responsive to interleukin-3[53]. In addition to HSC-like cells (EML-HSC: Lin- SCA1+ CD34+ KIT++), the cultures also contain cells which have cell surface marker profiles similar to early myeloid progenitors (EML-CMP/GMP: Lin- SCA1- CD34+ KIT++) and megakaryocytic-erythroid progenitors (EML-MEP: Lin- SCA1- CD34- KIT++). Down-regulation of CREBBP abundance can be seen in phenotypically distinct subpopulations of EML cells (Figure 4A, inset). In particular, EML-MEP have roughly 50% of EML-HSC CREBBP protein levels, much like primary BM MEPs do relative to HSCs (Figure 3A). In each case, protein levels were lowest in EML-MEPs (Figure 4A). Relative mRNA and intronic signals varied by target and cell subpopulation but, again, changes were most pronounced in EML-MEPs (Figure 4B).



**Figure 2. Changes in the ratio of unspliced message relative to total mRNA ratios in WT vs *Crebbp*<sup>+/-</sup> HSCs.** (A) WT/*Crebbp*<sup>+/-</sup> ratios for total mRNA vs unspliced mRNA qRT signal for the indicated genes. Dashed grey lines mark the 1.5-fold limit used as a cut-off for differential expression. Points lying within the vertical grey lines but above the horizontal ones indicate little change in mRNA levels in WT vs *Crebbp*<sup>+/-</sup> cells but a reduction in the ratio of unspliced message. Plotted are median values with horizontal and vertical lines extending to indicate median absolute deviation for the mRNA and unspliced ratios, respectively. (B) WT/*Ep300*<sup>+/-</sup> ratios for total mRNA vs unspliced mRNA qRT signal for the indicated genes as in (A) showing little if any discrepant change in total and unspliced transcript levels. (C) Total mRNA vs Intronic qRT signal for the indicated genes in WT and *Crebbp*<sup>+/-</sup> HSCs. Grey lines indicate a slope of 1 on the log-log plot and are provided as a visual reference only. Gene lengths in kilobases (kb) are shown at the right. doi:10.1371/journal.pone.0024153.g002

Our model, based on the above results, is that the regulated production of unspliced, full-length transcripts can represent a sufficiently large proportion of the total mRNA pool to have a

detectable impact on protein levels. Figure 4C shows a cartoon of scenarios in which relative protein abundance and total mRNA levels would not correlate well. In each case illustrated (X, Y and Z), the same full-length mRNA signal is present (as detected by qRT, for example, by a 3' exonic probe indicated by the magenta spots). However, varying levels of unspliced transcript detected by an intronic probe (yellow spots) would result in different protein levels since only fully processed mRNA can be translated. In the absence of some correction for the presence of unspliced transcripts, a comparison of full-length message levels in X, Y and Z would thus predict lower than observed protein levels for Y relative to X or, conversely, higher than observed protein levels in Z relative to X.

Our data show instances of each of these cases. Full-length transcript levels of *Itga4* in EML-HSCs are on average lower than in EML-MEPs but protein levels are higher (Figure 4D). Factoring in intronic signal levels yields a better estimate of protein abundance. In the case of *Msi2* (Figure 4E), the combination of both lower total mRNA and a greater proportion of unspliced product in EML-MEPs relative to EML-HSCs results in a more substantial reduction in protein levels than predicted solely by relative mRNA levels.

Taken together, these data indicate that the differences in unspliced pre-mRNA levels we detected originally in FL HSCs are context-sensitive, varying with both developmental and differentiation stage. More importantly, these changes have the potential to alter protein levels and, consequently, to have an impact on downstream HSC functions.

## Discussion

The production of mature, fully spliced mRNA is subject to regulation not only at the initiation stage but also during elongation, 3'-end processing and through splice site selection. Although there is a preferred order in intron removal, excision does not necessarily proceed in a linear or directional manner[60] and, in some cases, splicing of one intron can alter subsequent splice site selection[61,62]. In addition, elongation and splicing are regulated in a cell type- and differentiation stage-specific manner[63]. Recent studies in HSCs[24], macrophages[18] and T-cells[19] have raised the possibility of a further level of control to this already complex system whereby production of full-length but unprocessed transcripts can serve to maintain a locus accessible but non-productive until appropriate signals are received. The role of CREBBP as a context-dependent transcriptional coactivator or inhibitor has been known for some time[64,65]. Less familiar are its function in 3'-end processing[16] and its colocalization with splicing factors and nascent transcripts in nuclear speckles[14,66]. Hargreaves et al. have also recently proposed a role for CREBBP in macrophages in the production of unspliced mRNA in primary response genes with GC-rich promoters (so-called PRG-I genes) [18].

In light of these studies linking CREBBP not only with initiation of transcription but also with pre-mRNA processing, we set out to determine whether fluctuations in intronic signals we detected in microarray studies of *Crebbp*<sup>+/-</sup> FL HSCs were real and might reveal something novel about the underlying biology of HSCs beyond what corresponding mRNA levels could tell us. Although our initial observations came from fetal HSCs, we found that the same probe set signature could distinguish adult HSC from differentiated cell types better than non-intronic probes sets for the same genes. In addition, despite the fact that each independently published data set we examined had slightly different isolation procedures for their HSC populations, our intronic signature

**Table 3.** Functional Annotation of Intronic Target Genes.

| Cell cycle & division            | AHCTF1 | Kinase activity          | CENTB1  |
|----------------------------------|--------|--------------------------|---------|
|                                  | CCND3  |                          | IGF1R   |
|                                  | MACF1  |                          | MAPK6   |
|                                  | MAPK6  |                          | PCTK2   |
|                                  | RIF1   |                          | TLK1    |
|                                  | TLK1   |                          | TNIK    |
|                                  |        |                          | TRIO    |
| Chromatin binding & modification | ARID4B | RNA binding & processing | AHCTF1  |
|                                  | CBX5   |                          | DGCR8   |
|                                  | DNMT3B |                          | MBNL1   |
|                                  | EPC1   |                          | MSI2    |
|                                  | JMJD1C |                          | PAN3    |
|                                  | MBDT1  |                          | PRPF40A |
|                                  | TLK1   |                          | RBM39   |
|                                  | UTX    |                          | SFRS12  |
|                                  |        |                          | SMG7    |
| Development & growth             | AHCTF1 |                          |         |
|                                  | ANGPT1 |                          |         |
|                                  | MBNL1  |                          |         |
|                                  | MEF2C  |                          |         |
|                                  | MEIS1  |                          |         |
|                                  | RTN4   |                          |         |
|                                  | SPATA5 |                          |         |
|                                  | TCF4   |                          |         |
|                                  | TCF12  |                          |         |

doi:10.1371/journal.pone.0024153.t003

consistently separated them from progenitors and mature cells, indicating that it reflects functionality rather than a specific cell surface phenotype.

Reduction in CREBBP levels did not result in a generalized defect in splicing, however. In fact, for most of the transcripts we tested by qRT, the level of intronic signal tracked with changes in total mRNA levels so that the proportion of unspliced message remained constant regardless of expression level. Nevertheless, in both FL HSCs and phenotypically immature EML cell subpopulations, there was a subset of genes that deviated from this linear relationship. Our results in FL HSCs had suggested that reduced levels of CREBBP triggered increased splicing since altered intronic probe sets were invariably down-regulated but results with EML cells indicate that, like other mechanisms regulating pre-mRNA production and maturation, the process (or combination of processes) we are detecting varies by target gene, cell type and differentiation state.

A further indication that the changes in intron levels we found in FL HSCs are functionally relevant is the fact that the transcription factor binding sites enriched in the proximal promoter regions of intronic genes were in part different from those for genes in which total mRNA abundance changes. Binding sites for SOX17, a factor essential for fetal but not adult hematopoiesis[52], were over-represented only in intronic targets. Also selectively enriched in intronic target promoters were binding sites for interferon-responsive factor 1 (IRF1) and the glucocorticoid receptor (NR3C1). Interferon alpha has been shown to

trigger HSC cycling[67] and IRF1 and NR3C1 exhibit synergistic or inhibitory crosstalk in signal transduction depending on cellular context[68,69]. One further tantalizing connection is that mice haploinsufficient for *Crebbp* expression develop myelodysplastic disease within their first year of life (manuscript submitted) and IRF1 has also been implicated in the pathophysiology of immune-mediated marrow failure in myelodysplastic syndrome[70] and is frequently deleted in 5q- myelodysplasia[71].

Interestingly, binding sites for the Ets family factors SFPI1/PU.1, ETS1, GABPA and ELK1 were enriched in down-regulated intronic targets and in the up-regulated mRNA set but not in down-regulated mRNA targets. This shared regulation by up-regulated mRNAs and transcripts with reduced intronic levels would make sense if, as our data suggests, a reduction in intronic signal results in greater mRNA splicing leading to increased protein production. Shared promoter elements would thus allow coordinated expression by directly up-regulating transcription for some targets while increasing the processing of unspliced transcripts for others. This partitioning of transcription factor binding site clusters between intronic and mRNA targets is worthy of further study as it points to specific regulatory modules participating in CREBBP-mediated control of mRNA processing.

It is still unclear whether the link between altered intronic signal and differentiation is causative but what is clear from our data is that knocking down *Crebbp* expression pushes multipotent EML cells to commit to myelopoiesis. The HSC-like EML cell line was established by introducing a dominant-negative retinoic acid receptor (RAR $\alpha$ 403) into mouse BM. Due to an RAR $\alpha$ 403 block, EML cells require treatment with ATRA to restore their ability to produce granulocyte-macrophage progenitors (CFU-GM) in response to IL-3 and GM-CSF[26]. Targeted reduction of *Crebbp* expression levels, however, allowed EML cells to bypass this block and differentiate into macrophages even in the presence of SCF doses normally able to maintain the cells in an undifferentiated state. This spontaneous myeloid differentiation of EML cells after *Crebbp* knock-down is reminiscent of the increased myelopoiesis seen in *Crebbp*<sup>+/-</sup> animals[3] and the differentiation-induced down-regulation of CREBBP protein levels is similar to that we observed in differentiating hematopoietic cells.

We have shown here that both splicing changes and spontaneous myeloid differentiation are sensitive to CREBBP dosage and selective alterations in abundance of intronic sequences correlate with differentiation in EML cells and changes in protein levels. Splicing changes varied by cell type and gene but nevertheless appear regulated as opposed to being a surrogate measure of total mRNA. Our data are consistent with a bookmarking model of HSC lineage priming with CREBBP involved in multiple steps of the lineage decision-making process as coactivator of transcription and coregulator of pre-mRNA processing. They also fit well with the concept that key regulators such as CREBBP can alter the landscape of HSC fate decisions[25] by altering the probability of adopting a given path towards self-renewal, death or differentiation.

## Materials and Methods

### Animals

Mice were bred and maintained under pathogen-free conditions at the animal facility of the GCCRI. All animal procedures were approved by the University Health Science Center Institutional Animal Care and Use Committee (protocol numbers 06030 and 06059). *Crebbp*<sup>+/-</sup> and *Ep300*<sup>+/-</sup> mice (kindly provided by Dr. D. Livingston, Dana-Farber Cancer Institute, Boston, MA) are fully back-crossed on a C57BL/6 background. WT littermates



**Table 4.** Binding Sites for Expressed Transcription Factors Enriched in Proximal Promoters of Target Genes.

| Jaspar<br>TFBS Matrix | Factor    | Binds CREBBP?<br>† | Intronic (all down)   |       |       | Down mRNA             |       |       | Up mRNA               |       |       |
|-----------------------|-----------|--------------------|-----------------------|-------|-------|-----------------------|-------|-------|-----------------------|-------|-------|
|                       |           |                    | Score *               | TSS   | Chr19 | Score                 | TSS   | Chr19 | Score                 | TSS   | Chr19 |
|                       |           |                    | N = 85, 61% (G+C)     |       |       | n = 48, 59% (G+C)     |       |       | n = 55, 60% (G+C)     |       |       |
|                       |           |                    | Enrichment<br>P-value |       |       | Enrichment<br>P-value |       |       | Enrichment<br>P-value |       |       |
| MA0079.2              | SP1       | Y                  | 110                   | 0.005 | 0     |                       |       |       |                       |       |       |
| MA0080.2              | SFP1/PU.1 | Y                  | 56.4                  | 0     | 0     |                       |       |       | 37.7                  | 0     | 0     |
| MA0152.1              | NFATC2    | Y                  | 44.5                  | 0.008 | 0.002 |                       |       |       | 35.1                  | 0     | 0     |
| MA0098.1              | ETS1      | Y                  | 25.9                  | 0     | 0     |                       |       |       | 19.6                  | 0     | 0     |
| MA0050.1              | IRF1      | Y                  | 20.2                  | 0     | 0     |                       |       |       |                       |       |       |
| MA0113.1              | NR3C1/GR  | Y                  | 20.2                  | 0     | 0     |                       |       |       |                       |       |       |
| GFI1_Q6 #             | GFI1      |                    | 11.2                  | 0     | 0     |                       |       |       |                       |       |       |
| MA0151.1              | ARID3A    |                    | 8.88                  | 0     | 0     |                       |       |       |                       |       |       |
| MA0028.1              | ELK1      | Y                  | 8.82                  | 0     | 0     | 2.42                  | 0.006 | 0.001 | 13.6                  | 0     | 0     |
| MA0062.2              | GABPA     | Y                  | 6.81                  | 0     | 0     |                       |       |       | 15.7                  | 0     | 0     |
| MA0078.1              | SOX17     |                    | 6.39                  | 0     | 0.001 |                       |       |       |                       |       |       |
| MA0108.2              | TBP       | Y                  | 5.63                  | 0.002 | 0     |                       |       |       |                       |       |       |
| MA0087.1              | SOX5      |                    | 5.61                  | 0.001 | 0     |                       |       |       |                       |       |       |
| MA0102.2              | CEBPA     | Y                  | 4.73                  | 0.001 | 0     |                       |       |       |                       |       |       |
| MA0060.1              | NFYA      | Y                  |                       |       |       | 10.5                  | 0.001 | 0     | 6.4                   | 0.004 | 0     |
| MA0018.2              | CREB1     | Y                  |                       |       |       | 1.93                  | 0.001 | 0     |                       |       |       |

†Bedford et al. (2010) and references therein.

\*Scores comparable within but not between groups.

#From Transfac; Matys et al. (2003).

doi:10.1371/journal.pone.0024153.t004

served as controls. B6;129S2-*Cdkn1a*<sup>tm1Tvj</sup>/J mice and B6129SF2 WT controls were purchased from Jackson Laboratories (Bar Harbor, ME).

### Isolation of FL HSCs and MEFs

Mated females were checked daily for the presence of vaginal copulation plugs, which defined day 0.5 of gestation. Plugged females were sacrificed on day 14.5 of gestation and their embryos harvested. The average cell count of 21 WT fetal livers (14 litters) was  $21.3 \pm 4.1 \times 10^6$  cells, indicating that the embryos were at a similar stage of development when sacrificed [72].

**FL HSCs.** Single cell suspensions were prepared from each FL and viably frozen. Once enough embryos of each genotype were collected, cells were thawed and prepared for HSC isolation. Lineage negative (Lin-: Ter119-, CD5- or CD4-CD8-, B220-, GR1-) SCA1+ AA4.1+ KIT+[73] cells were purified by fluorescence-activated cell sorting on a FACSaria (BD Biosciences, San Jose, CA).

**MEFs.** Day 14.5 post-coitus embryo bodies were disaggregated in 0.25% trypsin and plated in MEF medium (DMEM, 15% FBS, 2mM glutamine, 100  $\mu$ M non-essential amino acids,  $8.9 \times 10^{-5}$ M  $\beta$ -mercaptoethanol, penicillin/streptomycin; Invitrogen, Carlsbad, CA).

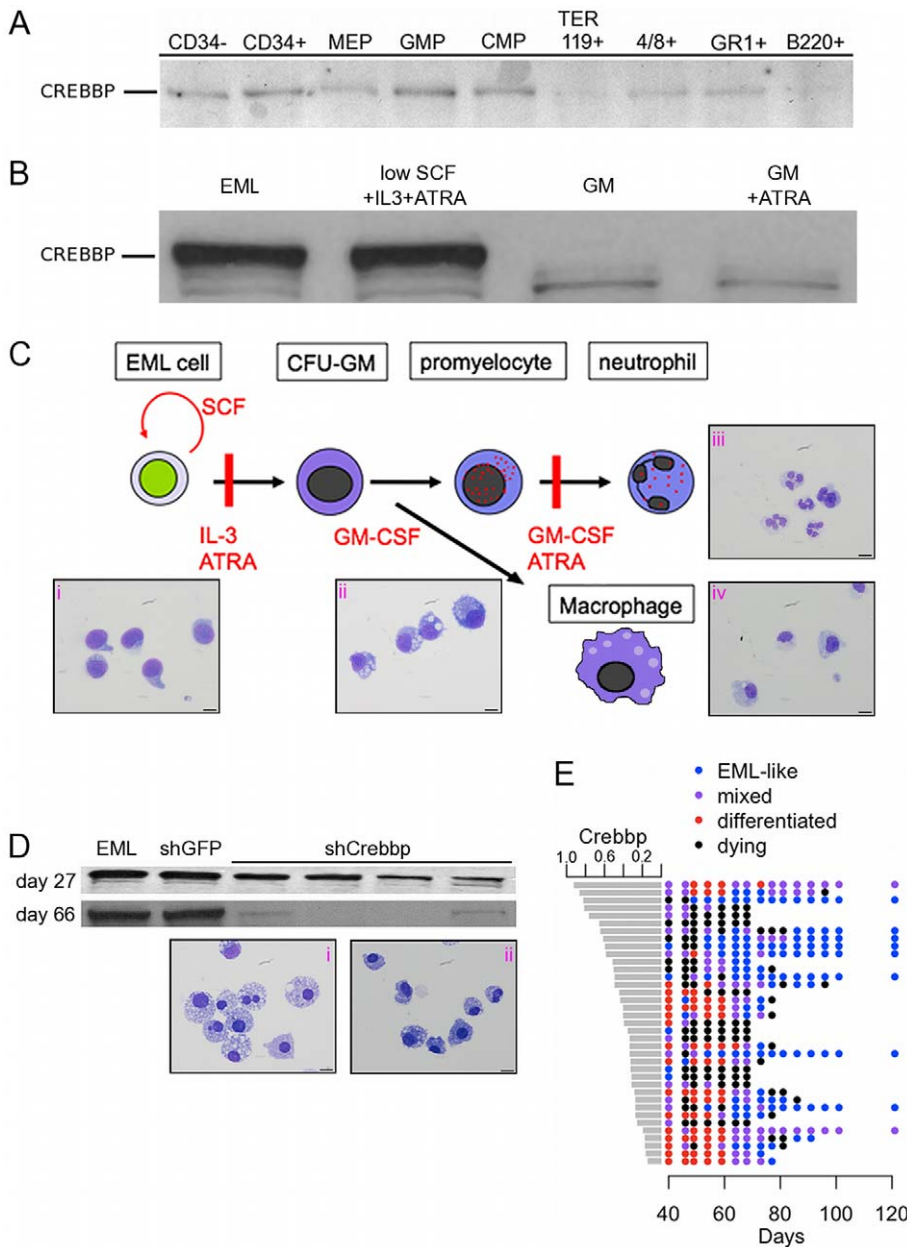
### Microarray analysis

All data is MIAME-compliant and the arrays are available for download from the Gene Expression Omnibus (GSE27987). For each HSC sample, 12 ng of RNA were amplified using the Ovation RNA amplification kit (NuGen Technologies, Inc., San

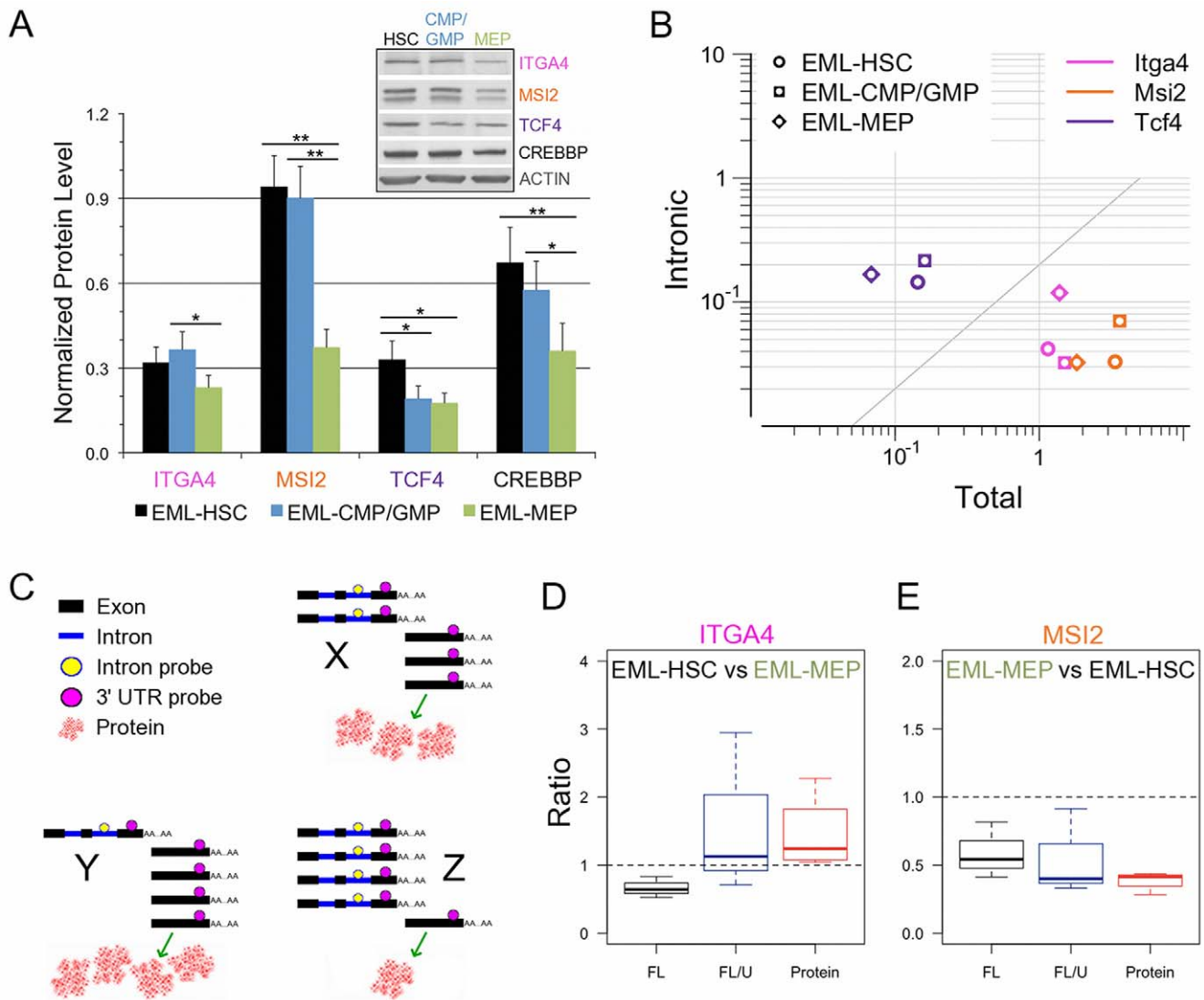
Carlos, CA). MEF RNA was isolated using TRIzol (Invitrogen) and used without amplification. After biotin labeling, samples were hybridized to Affymetrix Gene Chip Mouse Genome 430 2.0 microarrays. Each microarray was checked for large-scale defects, RNA degradation and the proportion of probe sets called "Present" or "Absent" using MAS5 (~50% in all cases). All arrays were comparable and satisfactory by these metrics. Arrays for each group were normalized and corrected for background using the Bioconductor [74] *gcrma* package. We excluded from further consideration probe sets for which all samples in a group were called "Absent", for which all samples had intensities below  $\log_2(100)$  or for which the interquartile range was  $<0.5$ . In the case of *Crebbp*<sup>+/-</sup>, *Ep300*<sup>+/-</sup> and their WT control FL HSCs, the values of the technical replicates were averaged prior to filtering based on minimum intensity levels and variation. Paired Student T-tests with p-value  $<0.05$  and a fold-change  $>1.5$  were used as the cut-off for calling significant change.

### Annotation

Probe sets on the arrays used in these studies typically comprise 11 25-mers designed to interrogate GenBank mRNA and EST sequences. We used Blast [75] to map the individual 25-mers for each probe set to mouse RefSeq mRNA sequences (May 2009). Probe sets that did not match reference mRNA sequences were then aligned to genomic sequences corresponding to RefSeq genes (all bases between the 5'-most transcription start site and the 3'-most end position) based on Mouse genome Build 37 (Mm9, July 2007). In both cases, a probe set was called a match if at least 8/11 25-mers aligned perfectly.



**Figure 3. Knock-down of CREBBP levels triggers stimulus-independent myeloid differentiation of EML cells.** (A) Western blots for CREBBP in lysates from 15,000 sorted hematopoietic cells per lane for a single experiment. Sort phenotypes: CD34- = LT HSCs (Lin- SCA1+ KIT++ CD34-), CD34+ = ST-HSC (Lin- SCA1+ KIT++ CD34+), MEP = (Lin- SCA1- KIT+ CD34- CD16/32-), GMP = (Lin- SCA1- KIT+ CD34+ CD16/32+), CMP = (Lin- SCA1- KIT+ CD34+ CD16/32-), others as indicated. (B) Western blots for CREBBP in 12.5 µg of nuclear extracts of undifferentiated EML cells (EML) and EML cells induced to differentiate by reducing the dose of SCF and addition of IL3 and ATRA (low SCF+IL3+ATRA), switching to GM-CSF only (GM) and then adding ATRA (GM+ATRA). Shown is a representative of at least 3 experiments. (C) Depiction of factors required to either maintain undifferentiated EML cell cultures (SCF) or to trigger myeloid differentiation (IL-3, ATRA, GM-CSF). Vertical red lines indicate RARα403 blocks to differentiation and the factors used to overcome them. (i-iv) Representative cytopins stained with Wright-Giemsa. (i) Undifferentiated EML cells with characteristic hand-mirror morphology. (ii) EML cells in low SCF + IL-3 + ATRA. (iii, iv) Fully differentiated EML cells after culture in GM-CSF alone followed by GM-CSF + ATRA. (D) CREBBP protein levels in  $3 \times 10^5$  cloned cells, either control (EML, shGFP) or in representative clones expressing shRNA against *Crebbp* measured at the indicated times post-cloning. (i, ii) Representative cytopins of cultures treated with shCrebbp. Scale bar in cytopins = 10 µm at 60X magnification. (E) Summary of outcomes for clones expressing shRNA against *Crebbp*. Each point represents the state of the culture at the indicated time. Relative *Crebbp* levels measured at 5 weeks post-cloning are indicated by the histogram on the left. Growth characteristics are indicated by color: EML-like (rapid growth with <5% adherent cells), mixed (<50% adherent), differentiated (>50% adherent), dying (>50% cells dead). Clone histories were recorded for 120 days or ended when the clone could no longer be passaged.  
doi:10.1371/journal.pone.0024153.g003



**Figure 4. Regulated changes in protein levels and intronic sequence abundance as a function of differentiation stage.** (A) Quantification of expression levels for the indicated proteins in phenotypically HSC-like cells (EML-HSC: Lin- SCA1+ CD34+ KIT++), early myeloid progenitors (EML-CMP/GMP: Lin- SCA1- CD34+ KIT++) and megakaryocyte/erythrocyte progenitors (EML-MEP: Lin- SCA1- CD34- KIT++). Bars indicate means + SEM of 4 independent experiments normalized to beta-actin levels. Significant differences in protein levels between cell types indicated by \* for ANOVA p-values <0.05 and \*\* for <0.01. Inset: representative Western blots. (B) Full-length vs intronic qRT signals for the indicated genes in the EML subpopulations described in A. Shown are median values from 3 independent experiments. (C) Model of the potential impact of varying intronic levels for a constant level of full-length mRNA. Exonic sequences are represented by black rectangles and introns by blue lines. Probes for intronic and full-length sequences are indicated by yellow and magenta circles, respectively. For each of X, Y and Z, the full-length level is the same (5 magenta circles) but the proportion of unspliced transcript varies (X = 2 probes, Y = 1 and Z = 4). The corresponding protein levels (pink blobs) reflect the relative proportion of spliced and unspliced transcript since only the latter can be translated. (D) Ratio of full-length mRNA (black box, F), full-length/unsliced (blue box, F/U) and protein (red box, P) in EML-HSC relative to EML-MEP for *Itga4* (E) As in (D) but for *Msi2* in EML-MEP vs EML-HSC. Whisker plots in (D) and (E) show median (line inside box), interquartile range (limits of the box) and data range (limits of whiskers). doi:10.1371/journal.pone.0024153.g004

### Hierarchical clustering of published data sets

Data sets of hematopoietic populations using the same microarray platform as our studies were downloaded from the Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/geo/>) and either reprocessed from image files (GSE17765[30], GSE6506[31]) or the normalized data used as published (GSE18669[32]). Two sets of probe sets were used for clustering subpopulations: (a) intronic probe sets differentially expressed in *Crebbp*<sup>+/−</sup> vs WT HSCs and (b) the most variable mRNA probe set for each gene corresponding to an intronic probe set in (a). GenePattern[76] was used for hierarchical clustering and visualization. For randomization studies, the *hclust* and

*cutree* functions from the R *stats* package[77] were used to generate dendrograms of Pearson correlations and determine the separation of HSCs from other cell types for 10,000 random samples of expressed intronic probe sets.

### Gene set overlap and TFBS analysis

Gene lists were submitted to web-based MSigDB[78] and DAVID[79,80] applications and related or overlapping terms collapsed for simplicity of presentation. Vertebrate position-specific weight matrices were downloaded from JASPAR[81] and supplemented with Transfac[82] matrices for GFI1 and LEF1. These were

used with Clover[83] to quantify over-representation of TFBS in target proximal promoters extending 500bp upstream and 250bp downstream of the transcription start site (TSS). We used two backgrounds to correct for compositional biases: one comprising 5Kb regions centered on all RefSeq (May 2008 download) and another consisting of mouse chromosome 9. Background sequences were derived from mouse NCBI Build 37 (Mm9). Results for intronic vs exonic targets were compared to identify TFBS uniquely enriched in intronic targets.

### EML cell culture and differentiation

EML cells (EML clone 1, CRL-11691) were obtained from ATCC (Manassas, VA). Cells were cultured in IMDM medium supplemented with 20% FBS (06952; StemCell Technologies, Vancouver, BC, Canada) and rmSCF (100ng/ml, R&D Systems, Minneapolis, MN) to maintain them in an undifferentiated state (non-differentiation medium, NDM). Myeloid differentiation was induced by a 3-step protocol adapted from a previous study[26]. Briefly, on day 1 cells were switched from NDM to IMDM plus 20% FBS with a lower dose of rmSCF (50 ng/ml), 25 ng/ml IL-3 (R&D Systems) and 10  $\mu$ M all-trans retinoic acid (ATRA; R2625, Sigma, St Louis, MI). On day 4, the medium was replaced with fresh medium containing only GM-CSF (25 ng/ml) and 7 days later 10  $\mu$ M ATRA was added.

### shRNA-mediated knock-down of CREBBP in EML cells

A lentivirus-encoded shRNA targeting the sequence 5'-CAAGCACTGGGAATTCCTCT-3' in mouse *Crebbp* (shCrebbp) was created by cloning oligonucleotides into the FSIPPW vector as previously described[84]. A control lentiviral shRNA targeting enhanced green fluorescent protein (shGFP, 5'-AAGAACGGCATCAAGGTGAACTT-3') was generated similarly. Both were packaged as previously described[85]. Co-transfection of 293TD cells was performed using Lipofectamine 2000 according to manufacturer's instructions (Invitrogen). Undifferentiated, early passage EML cells were split one day prior to infection. Virus-containing supernatant supplemented with 8  $\mu$ g/ml protamine was added to the cells for ~16 hours after which the cells were placed in fresh NDM. A second round of infection was performed ~8 hours later using a flow-through infection protocol[86]. Culture medium was replaced with fresh NDM supplemented with puromycin (3  $\mu$ g/ml) the following day to select for transduced cells. Surviving EML cells were cloned in NDM methylcellulose plus puromycin (M3234, StemCell Technologies) for 8 days, after which individual clones were picked and expanded in liquid NDM.

### EML subpopulation isolation

Phenotypically primitive EML subpopulations[53] were purified by first removing B220+, CD19+ and Ter119+ cells by magnetic separation (Miltenyi, Auburn, CA) and then incubating with SCA1-biotin, followed by streptavidin-APC-Cy7 and CD34-FITC (BD Biosciences). No anti-KIT antibody was necessary since all B220- CD19- Ter119- cells were KIT+. 7-Aminoactinomycin D (BD Biosciences) was used to exclude dead cells. Purification was done on a FACSAria (BD Biosciences).

### Quantitative RT-PCR

Total RNA from cloned EML cells was isolated by TRIzol extraction and purified with the Qiagen RNeasy kit (Qiagen, Valencia, CA). Total RNA from all other samples was extracted using the Qiagen RNeasy (Plus) Micro Kit as directed by the manufacturer. RNA was then reverse transcribed using the High Capacity cDNA Reverse Transcription Kit (Applied Biosciences, Foster City, CA). FL HSC was amplified using the Ovation RNA

amplification kit. PCR was performed using either the SYBR Green PCR kit (Qiagen) or the GoTaq qPCR Master Mix (Promega, Madison, WI) on a 7500 Real-Time PCR System (Applied Biosciences). Data were analyzed using the  $2^{-\Delta\Delta Ct}$  relative quantification method[87] and normalized to *Gapdh*. Primers were designed using the Primer3 web interface[88] (<http://frodo.wi.mit.edu/primer3/>). See Table 2 for primer sequences.

### Western blots

Anti-CREBBP antibody AC26[89] was kindly provided by Dr D. Livingston and anti-EP300 antibody RW128 was purchased from Upstate Biotechnology, Lake Placid, NY. Antibodies against ITGA4 (Ab65984), MSI2 (Ab76148), TCF4 (Ab72586) and ACTB (Ab6276) were purchased from Abcam (Cambridge, MA). Visualization of proteins was done by staining with a goat anti-mouse or goat anti-rabbit HRP (Bio-Rad Laboratories, Hercules, CA), as appropriate, and Pierce ECL Western Blotting Substrate (Thermo Scientific, Rockford, IL) or Amersham ECL Plus (GE Healthcare, Piscataway, NJ).

**Hematopoietic cells.** Whole cell lysates from 15,000 cells of each phenotype were suspended in RIPA buffer containing protease inhibitors (Roche Diagnostics, Indianapolis, IN) and PMSF (Sigma-Aldrich, St. Louis, MO) then loaded in each well after heat denaturation (10 min @ 70°C).

**EML cells.** Whole cell lysates from  $4-5 \times 10^5$  cells suspended in RIPA buffer with protease inhibitors and PMSF or 12.5  $\mu$ g of nuclear extracts were loaded per well, as indicated.

**Image quantification and statistics.** In all cases, ImageJ software (<http://rsb.info.nih.gov/ij>) was used to quantify band intensities. Significance levels for the differences between cell types were determined by correlated-samples one-way ANOVA and Tukey HSD test using the web-based calculator at VassarStats: Website for Statistical Computation (<http://faculty.vassar.edu/lowry/anova1u.html>).

### Supporting Information

**Table S1 Annotation of Differentially Expressed Probe Sets in *Crebbp*<sup>+/-</sup> vs WT HSCs.** Detailed Excel table of differentially expressed probe sets including EST support for intronic sequences and expression levels for mutant and WT HSCs. (XLS)

**Table S2 Quantitative RT-PCR Validation of Microarray Results.** Detailed Excel table of locations of primer sequences relative to microarray probe sets and qRT validation results for exonic and intronic targets. For intronic probe sets, the table also includes the length of potential poly(A) RT priming sites and predicted presence of intronic poly(A) signals. (XLS)

### Acknowledgments

The authors thank the GCCRI Laboratory Animal Resources for care of experimental animals and Karla Moncada of the UTHSCSA Flow Cytometry Core and Jennifer Rebeles of the GCCRI Shared Instrumentation Facility for assistance with FACS sorting and analysis.

### Author Contributions

Conceived and designed the experiments: MEL VIR. Performed the experiments: ZC QZ RW. Analyzed the data: MEL JC ALK VIR. Contributed reagents/materials/analysis tools: MEL ALK VIR. Wrote the paper: MEL VIR.

## References

- Kalkhoven E (2004) CBP and p300: HATs for different occasions. *Biochem Pharmacol* 68: 1145–1155.
- Rebel VI, Kung AL, Tanner EA, Yang H, Bronson RT, et al. (2002) Distinct roles for CREB-binding protein and p300 in hematopoietic stem cell self-renewal. *Proc Natl Acad Sci U S A* 99: 14789–14794.
- Kung AL, Rebel VI, Bronson RT, Ch'ng LE, Sief CA, et al. (2000) Gene dose-dependent control of hematopoiesis and hematologic tumor suppression by CBP. *Genes Dev* 14: 272–277.
- Bedford DC, Kasper LH, Fukuyama T, Brindle PK (2010) Target gene context influences the transcriptional requirement for the KAT3 family of CBP and p300 histone acetyltransferases. *Epigenetics* 5: 9–15.
- Blobel GA (2000) CREB-binding protein and p300: molecular integrators of hematopoietic transcription. *Blood* 95: 745–755.
- Liu Z, Simpson ER (1999) Molecular mechanism for cooperation between Sp1 and steroidogenic factor-1 (SF-1) to regulate bovine CYP11A gene expression. *Mol Cell Endocrinol* 153: 183–196.
- Hausen M, Tonjes RR, Grez M (1998) The transcription factor Sp1 regulates the myeloid-specific expression of the human hematopoietic cell kinase (HCK) gene through binding to two adjacent GC boxes within the HCK promoter-proximal region. *J Biol Chem* 273: 31844–31852.
- Ahmlof T, Wallberg AE, Gustafsson JA, Wright AP (1998) Role of important hydrophobic amino acids in the interaction between the glucocorticoid receptor tau 1-core activation domain and target factors. *Biochemistry* 37: 9586–9594.
- von Lindern M, Zauner W, Mellitzer G, Steinlein P, Fritsch G, et al. (1999) The glucocorticoid receptor cooperates with the erythropoietin receptor and c-Kit to enhance and sustain proliferation of erythroid progenitors in vitro. *Blood* 94: 550–559.
- Yamamoto H, Kihara-Negishi F, Yamada T, Hashimoto Y, Oikawa T (1999) Physical and functional interactions between the transcription factor PU.1 and the coactivator CBP. *Oncogene* 18: 1495–1501.
- Kovacs KA, Steinmann M, Magistretti PJ, Halfon O, Cardinaux JR (2003) CCAAT/enhancer-binding protein family members recruit the coactivator CREB-binding protein and trigger its phosphorylation. *J Biol Chem* 278: 36959–36965.
- Iwasaki H, Somoza C, Shigematsu H, Duprez EA, Iwasaki-Arai J, et al. (2005) Distinctive and indispensable roles of PU.1 in maintenance of hematopoietic stem cells and their differentiation. *Blood* 106: 1590–1600.
- Zhang P, Iwasaki-Arai J, Iwasaki H, Fenyus ML, Dayaram T, et al. (2004) Enhancement of hematopoietic stem cell repopulating capacity and self-renewal in the absence of the transcription factor C/EBP alpha. *Immunity* 21: 853–863.
- Bex F, Yin MJ, Burny A, Gaynor RB (1998) Differential transcriptional activation by human T-cell leukemia virus type 1 Tax mutants is mediated by distinct interactions with CREB binding protein and p300. *Mol Cell Biol* 18: 2392–2405.
- McManus KJ, Stephens DA, Adams NM, Islam SA, Freemont PS, et al. (2006) The transcriptional regulator CBP has defined spatial associations within interphase nuclei. *PLoS Comput Biol* 2: e139.
- Shimazu T, Horinouchi S, Yoshida M (2007) Multiple histone deacetylases and the CREB-binding protein regulate pre-mRNA 3'-end processing. *J Biol Chem* 282: 4470–4478.
- Ramos YF, Hestand MS, Verlaan M, Krabbendam E, Ariyurek Y, et al. (2006) Genome-wide assessment of differential roles for p300 and CBP in transcription regulation. *Nucleic Acids Res* 34: 5396–5408.
- Hargreaves DC, Horng T, Medzhitov R (2009) Control of inducible gene expression by signal-dependent transcriptional elongation. *Cell* 138: 129–145.
- Byun JS, Wong MM, Cui W, Idelman G, Li Q, et al. (2009) Dynamic bookmarking of primary response genes by p300 and RNA polymerase II complexes. *Proc Natl Acad Sci U S A* 106: 19286–19291.
- Core LJ, Waterfall JJ, Lis JT (2008) Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* 322: 1845–1848.
- Sultan M, Schulz MH, Richard H, Magen A, Klingenhoff A, et al. (2008) A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. *Science* 321: 956–960.
- Cloonan N, Forrest AR, Kolle G, Gardiner BB, Faulkner GJ, et al. (2008) Stem cell transcriptome profiling via massive-scale mRNA sequencing. *Nat Methods* 5: 613–619.
- Hu M, Krause D, Greaves M, Sharkis S, Dexter M, et al. (1997) Multilineage gene expression precedes commitment in the hemopoietic system. *Genes Dev* 11: 774–785.
- Bowman TV, McCooney AJ, Merchant AA, Ramos CA, Fonseca P, et al. (2006) Differential mRNA processing in hematopoietic stem cells. *Stem Cells* 24: 662–670.
- Enver T, Pera M, Peterson C, Andrews PW (2009) Stem cell states, fates, and the rules of attraction. *Cell Stem Cell* 4: 387–397.
- Tsai S, Bartelmez S, Sitnicka E, Collins S (1994) Lymphohematopoietic progenitors immortalized by a retroviral vector harboring a dominant-negative retinoic acid receptor can recapitulate lymphoid, myeloid, and erythroid development. *Genes Dev* 8: 2831–2841.
- Piao Y, Ko NT, Lim MK, Ko MS (2001) Construction of long-transcript enriched cDNA libraries from submicrogram amounts of total RNAs by a universal PCR amplification method. *Genome Res* 11: 1553–1558.
- Pritsker M, Doniger TT, Kramer LC, Westcot SE, Lemischka IR (2005) Diversification of stem cell molecular repertoire by alternative splicing. *Proc Natl Acad Sci U S A* 102: 14290–14295.
- Cheng T, Rodrigues N, Shen H, Yang Y, Dombkowski D, et al. (2000) Hematopoietic stem cell quiescence maintained by p21cip1/waf1. *Science* 287: 1804–1808.
- Broske AM, Vockentanz L, Kharazi S, Huska MR, Mancini E, et al. (2009) DNA methylation protects hematopoietic stem cell multipotency from myeloerythroid restriction. *Nat Genet* 41: 1207–1215.
- Chambers SM, Boles NC, Lin KY, Tierney MP, Bowman TV, et al. (2007) Hematopoietic fingerprints: an expression database of stem cells and their progeny. *Cell Stem Cell* 1: 578–591.
- Weishaupt H, Sigvardsson M, Attema JL Epigenetic chromatin states uniquely define the developmental plasticity of murine hematopoietic stem cells. *Blood* 115: 247–256.
- Orkin SH, Zon LI (2008) Hematopoiesis: an evolving paradigm for stem cell biology. *Cell* 132: 631–644.
- Rebel VI, Miller CL, Eaves CJ, Lansdorp PM (1996) The repopulation potential of fetal liver hematopoietic stem cells in mice exceeds that of their liver adult bone marrow counterparts. *Blood* 87: 3500–3507.
- Rebel VI, Miller CL, Thornbury GR, Dragowska WH, Eaves CJ, et al. (1996) A comparison of long-term repopulating hematopoietic stem cells in fetal liver and adult bone marrow from the mouse. *Exp Hematol* 24: 638–648.
- Hisa T, Spence SE, Rachel RA, Fujita M, Nakamura T, et al. (2004) Hematopoietic, angiogenic and eye defects in Meis1 mutant animals. *EMBO J* 23: 450–459.
- Arai F, Hirao A, Ohmura M, Sato H, Matsuoka S, et al. (2004) Tie2/angiopoietin-1 signaling regulates hematopoietic stem cell quiescence in the bone marrow niche. *Cell* 118: 149–161.
- Arroyo AG, Yang JT, Rayburn H, Hynes RO (1999) Alpha4 integrins regulate the proliferation/differentiation balance of multilineage hematopoietic progenitors in vivo. *Immunity* 11: 555–566.
- Hope KJ, Cellot S, Ting SB, MacRae T, Mayotte N, et al. (2010) An RNAi screen identifies Msi2 and Prox1 as having opposite roles in the regulation of hematopoietic stem cell activity. *Cell Stem Cell* 7: 101–113.
- Kharas MG, Lengner CJ, Al-Shahrour F, Bullinger L, Ball B, et al. (2010) Musashi-2 regulates normal hematopoiesis and promotes aggressive myeloid leukemia. *Nat Med* 16: 903–908.
- Murre C (2005) Helix-loop-helix proteins and lymphocyte development. *Nat Immunol* 6: 1079–1086.
- Song HW, Cauffman K, Chan AP, Zhou Y, King ML, et al. (2007) Hermes RNA-binding protein targets RNAs-encoding proteins involved in meiotic maturation, early cleavage, and germline development. *Differentiation* 75: 519–528.
- Sun Y, Ding L, Zhang H, Han J, Yang X, et al. (2006) Potentiation of Smad-mediated transcriptional activation by the RNA-binding protein RBPMS. *Nucleic Acids Res* 34: 6314–6326.
- Karlsson G, Blank U, Moody JL, Ehinger M, Singbrant S, et al. (2007) Smad4 is critical for self-renewal of hematopoietic stem cells. *J Exp Med* 204: 467–474.
- Nam DK, Lee S, Zhou G, Cao X, Wang C, et al. (2002) Oligo(dT) primer generates a high frequency of truncated cDNAs through internal poly(A) priming during reverse transcription. *Proc Natl Acad Sci U S A* 99: 6152–6156.
- Hutchinson JN, Ensminger AW, Clemson CM, Lynch CR, Lawrence JB, et al. (2007) A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains. *BMC Genomics* 8: 39.
- Tripathi V, Ellis JD, Shen Z, Song DY, Pan Q, et al. (2007) The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol Cell* 39: 925–938.
- Webb CF, Smith EA, Medina KL, Buchanan KL, Smithson G, et al. (1998) Expression of bright at two distinct stages of B lymphocyte development. *J Immunol* 160: 4747–4754.
- Zeng H, Yucel R, Kosan C, Klein-Hitpass L, Moroy T (2004) Transcription factor Gfi1 regulates self-renewal and engraftment of hematopoietic stem cells. *EMBO J* 23: 4116–4125.
- Testa U, Stellacci E, Pelosi E, Sestili P, Venditti M, et al. (2004) Impaired myelopoiesis in mice devoid of interferon regulatory factor 1. *Leukemia* 18: 1864–1871.
- Wessely O, Deiner EM, Beug H, von Lindern M (1997) The glucocorticoid receptor is a key regulator of the decision between self-renewal and differentiation in erythroid progenitors. *EMBO J* 16: 267–280.
- Kim I, Saunders TL, Morrison SJ (2007) Sox17 dependence distinguishes the transcriptional regulation of fetal from adult hematopoietic stem cells. *Cell* 130: 470–483.
- Ye ZJ, Kluger Y, Lian Z, Weissman SM (2005) Two types of precursor cells in a multipotential hematopoietic cell line. *Proc Natl Acad Sci U S A* 102: 18461–18466.
- Finstad SL, Rosenberg N, Levy LS (2007) Diminished potential for B-lymphoid differentiation after murine leukemia virus infection in vivo and in EML hematopoietic progenitor cells. *J Virol* 81: 7274–7279.
- Suh HC, Gooya J, Renn K, Friedman AD, Johnson PF, et al. (2006) C/EBPalpha determines hematopoietic cell fate in multipotential progenitor cells

- by inhibiting erythroid differentiation and inducing myeloid differentiation. *Blood* 107: 4308–4316.
56. Partanen A, Motoyama J, Hui CC (1999) Developmentally regulated expression of the transcriptional cofactors/histone acetyltransferases CBP and p300 during mouse embryogenesis. *Int J Dev Biol* 43: 487–494.
  57. Tersikh AV, Miyamoto T, Chang C, Diatchenko L, Weissman IL (2003) Gene expression analysis of purified hematopoietic stem cells and committed progenitors. *Blood* 102: 94–101.
  58. Kharas MG, Lengner CJ, Al-Shahrour F, Bullinger L, Ball B, et al. (2010) Musashi-2 regulates normal hematopoiesis and promotes aggressive myeloid leukemia. *Nat Med* 16: 903–908.
  59. Hope KJ, Cellot S, Ting SB, MacRae T, Mayotte N, et al. (2010) An RNAi screen identifies Msi2 and Prox1 as having opposite roles in the regulation of hematopoietic stem cell activity. *Cell Stem Cell* 7: 101–113.
  60. Lang KM, Spritz RA (1987) In vitro splicing pathways of pre-mRNAs containing multiple intervening sequences? *Mol Cell Biol* 7: 3428–3437.
  61. Schwarze U, Starman BJ, Byers PH (1999) Redefinition of exon 7 in the COL1A1 gene of type I collagen by an intron 8 splice-donor-site mutation in a form of osteogenesis imperfecta: influence of intron splice order on outcome of splice-site mutation. *Am J Hum Genet* 65: 336–344.
  62. Schor IE, Rascovan N, Pelisch F, Allo M, Kornbliht AR (2009) Neuronal cell depolarization induces intragenic chromatin modifications affecting NCAM alternative splicing. *Proc Natl Acad Sci U S A* 106: 4325–4330.
  63. Uptain SM, Kane CM, Chamberlin MJ (1997) Basic mechanisms of transcript elongation and its regulation. *Annu Rev Biochem* 66: 117–172.
  64. Giordano A, Avantaggiati ML (1999) p300 and CBP: partners for life and death. *J Cell Physiol* 181: 218–230.
  65. Goodman RH, Smolik S (2000) CBP/p300 in cell growth, transformation, and development. *Genes Dev* 14: 1553–1577.
  66. von Mikecz A, Zhang S, Montminy M, Tan EM, Hemmerich P (2000) CREB-binding protein (CBP)/p300 and RNA polymerase II colocalize in transcriptionally active domains in the nucleus. *J Cell Biol* 150: 265–273.
  67. Essers MA, Offner S, Blanco-Bose WE, Waibler Z, Kalinke U, et al. (2009) IFN $\alpha$  activates dormant haematopoietic stem cells in vivo. *Nature* 458: 904–908.
  68. Tliba O, Damera G, Banerjee A, Gu S, Baidouri H, et al. (2008) Cytokines induce an early steroid resistance in airway smooth muscle cells: novel role of interferon regulatory factor-1. *Am J Respir Cell Mol Biol* 38: 463–472.
  69. Jiang X, Norman M, Roth L, Li X (2004) Protein-DNA array-based identification of transcription factor activities regulated by interaction with the glucocorticoid receptor. *J Biol Chem* 279: 38480–38485.
  70. Voulgarelis M, Giannouli S, Ritis K, Tzioufas AG (2004) Myelodysplasia-associated autoimmunity: clinical and pathophysiologic concepts. *Eur J Clin Invest* 34: 690–700.
  71. Willman CL, Sever CE, Pallavicini MG, Harada H, Tanaka N, et al. (1993) Deletion of IRF-1, mapping to chromosome 5q31.1, in human leukemia and preleukemic myelodysplasia. *Science* 259: 968–971.
  72. Morrison SJ, Hemmati HD, Wandycz AM, Weissman IL (1995) The purification and characterization of fetal liver hematopoietic stem cells. *Proc Natl Acad Sci U S A* 92: 10302–10306.
  73. Ivanova NB, Dimos JT, Schaniel C, Hackney JA, Moore KA, et al. (2002) A stem cell molecular signature. *Science* 298: 601–604.
  74. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, et al. (2004) Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* 5: R80.
  75. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215: 403–410.
  76. Reich M, Liefeld T, Gould J, Lerner J, Tamayo P, et al. (2006) GenePattern 2.0. *Nat Genet* 38: 500–501.
  77. R Development Core Team (2009) R: A language and environment for statistical computing. Vienna/Austria: R Foundation for Statistical Computing.
  78. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, et al. (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 102: 15545–15550.
  79. Huang W, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4: 44–57.
  80. Dennis G, Jr., Sherman BT, Hosack DA, Yang J, Gao W, et al. (2003) DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol* 4: P3.
  81. Brync JC, Valen E, Tang MH, Marstrand T, Winther O, et al. (2008) JASPAR, the open access database of transcription factor-binding profiles: new content and tools in the 2008 update. *Nucleic Acids Res* 36: D102–106.
  82. Matys V, Fricke E, Geffers R, Gossling E, Haubrock M, et al. (2003) TRANSFAC: transcriptional regulation, from patterns to profiles. *Nucleic Acids Res* 31: 374–378.
  83. Frith MC, Fu Y, Yu L, Chen JF, Hansen U, et al. (2004) Detection of functional DNA motifs via statistical over-representation. *Nucleic Acids Res* 32: 1372–1381.
  84. Kanellopoulou C, Muljo SA, Kung AL, Ganesan S, Drapkin R, et al. (2005) Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. *Genes Dev* 19: 489–501.
  85. Lois C, Hong EJ, Pease S, Brown EJ, Baltimore D (2002) Germline transmission and tissue-specific expression of transgenes delivered by lentiviral vectors. *Science* 295: 868–872.
  86. Chuck AS, Palsson BO (1996) Consistent and high rates of gene transfer can be obtained using flow-through transduction over a wide range of retroviral titers. *Hum Gene Ther* 7: 743–750.
  87. Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>-Delta Delta C(T)</sup> Method. *Methods* 25: 402–408.
  88. Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* 132: 365–386.
  89. Yao TP, Oh SP, Fuchs M, Zhou ND, Ch'ng LE, et al. (1998) Gene dosage-dependent embryonic development and proliferation defects in mice lacking the transcriptional integrator p300. *Cell* 93: 361–372.