# Epigenetic characterization of hematopoietic stem cell differentiation using miniChIP and bisulfite sequencing analysis

Joanne L. Attema*†, Peter Papathanasiou*‡, E. Camilla Forsberg*, Jian Xu§, Stephen T. Smale§, and Irving L. Weissman*†

*Institute of Stem Cell Biology and Regenerative Medicine, Departments of Pathology and Developmental Biology, Stanford University School of Medicine, Stanford, CA 94305; and §Howard Hughes Medical Institute, Molecular Biology Institute and Department of Microbiology, Immunology, and Molecular Genetics, University of California, Los Angeles, CA 90095

Contributed by Irving L. Weissman, May 26, 2007 (sent for review March 5, 2007)

**Hematopoietic stem cells (HSC) produce all blood cell lineages by virtue of their capacity to self-renew and differentiate into progenitors with decreasing cellular potential. Recent studies suggest that epigenetic mechanisms play an important role in controlling stem cell potency and cell fate decisions. To investigate this hypothesis in HSC, we have modified the conventional chromatin immunoprecipitation assay allowing for the analysis of 50,000 prospectively purified stem and progenitor cells. Together with bisulfite sequencing analysis, we found that methylated H3K4 and AcH3 and unmethylated CpG dinucleotides colocalize across defined regulatory regions of lineage-affiliated genes in HSC. These active epigenetic histone modifications either accumulated or were replaced by increased DNA methylation and H3K27 trimethylation in committed progenitors consistent with gene expression. We also observed bivalent histone modifications at a lymphoid-affiliated gene in HSC and downstream transit-amplifying progenitors. Together, these data support a model in which epigenetic modifications serve as an important mechanism to control HSC multipotency.**

hematopoiesis | chromatin | gene expression

Hematopoietic stem cells (HSC) are multipotent cells that at the single cell level have the potential to differentiate into all cells of the erythromyeloid and lymphoid lineages, as well as to maintain their numbers by means of controlled self-renewal (1, 2). Progression from HSC to differentiated progeny involves the coordinated regulation of multiple gene expression programs leading to the activation or repression of lineage-specific genes (3–5). It is well established that the gene expression programs governing cellular differentiation involve multiple epigenetic changes beyond the level of transcription factor recruitment (6–8). Epigenetic regulation of gene expression is largely controlled by the posttranslational modification of histones and DNA methylation resulting in the alteration of chromatin structure and function at gene loci (9, 10).

Several observations suggest that epigenetic modifications are present at lineage-specific loci in HSC and that they could be critical for determining cell fate (11). Accordingly, active chromatin structure, as well as partial assembly of critical transcription factors, are present at the regulatory regions of numerous lineage-specifying genes in progenitor cells before high-level expression in mature subsets (12–18). At the single cell level in HSC, low level transcription of lineage-affiliated genes has been observed, a phenomenon known as lineage priming (13, 19, 20). It is possible that a specific chromatin structure exists in HSC at lineage-affiliated genes, which mediates low-level expression for the propagation of transcriptional memory during the differentiation process.

We have investigated the hypothesis that epigenetic modifications exist at lineage-affiliated genes in HSC for lineage commitment using chromatin immunoprecipitation (ChIP) assays and bisulfite sequencing (BS) analysis of genomic DNA. We have overcome the cell number constraints commonly encountered with conventional ChIP through the development of an improved method termed miniChIP, and demonstrate the reproducible analysis of histone and DNA modifications during hematopoietic development at specific loci in 50,000 functionally characterized primary cells. Specifically, we identified histone and DNA modifications associated with transcriptional potentiation at specific regulatory regions of lineage-affiliated genes in HSC and their progeny, multipotent progenitors (MPP), which are then inherited or erased during lineage commitment according to their differential expression. This analysis suggests that the transcriptional-permissive status of lineage-affiliated genes in stem cells is mediated by specific combinations of epigenetic modifications, which may play a key role in HSC multipotency.

## Results

Previous studies suggest that HSC are primed for lineage commitment by means of epigenetic mechanisms that govern low level expression of lineage-affiliated genes (13, 19, 20). To further address this observation, we began by purifying self-renewing long-term HSC, MPP that no longer self-renew (21, 22), common lymphoid progenitors (CLP) (23), common myeloid progenitors (CMP), granulocyte/monocyte progenitors (GMP), megakaryocyte/erythrocyte progenitors (MEP) (24), progenitor T cells, maturing B and T cells, mature granulocyte/macrophage cells, and erythroid progenitors (EP) [supporting information (SI) Fig. 5]. Representative genes expressed at low levels in HSC with well described and functionally important regulatory regions were selected for the epigenetic analyses (Fig. 1*A*). This list included genes affiliated with the erythroid (*Gata1* and *β-globin*), myeloid (*c-fms*), and lymphoid lineages (*Gata3* and *Ptcrα*). We included *Ptcrα* to specifically compare the chromatin structure at a gene not expressed in HSC.

Author contributions: J.L.A. and P.P. contributed equally to this article; J.L.A., P.P., S.T.S., and I.L.W. designed research; J.L.A., P.P., and J.X. performed research; J.L.A. developed and performed miniChIP assays; J.L.A. and J.X. contributed new reagents/analytic tools; J.L.A., P.P., J.X., S.T.S., and I.L.W. analyzed data; and J.L.A., P.P., E.C.F., S.T.S., and I.L.W. wrote the paper.

Conflict of interest statement: I.L.W. has stock in Amgen, Inc., was a member of the Amgen scientific advisory board, and is a cofounder and member of the Board of Directors of Cellerant, Inc., and Stem Cells, Inc.

Freely available online through the PNAS open access option.

Abbreviations: HSC, hematopoietic stem cell; MPP, multipotent progenitor; CLP, common lymphoid progenitor; CMP, common myeloid progenitor; GMP, granulocyte/monocyte progenitor; MEP, megakaryocyte/erythrocyte progenitor; BM, bone marrow; qRT-PCR, quantitative RT-PCR; BS, bisulfite genomic DNA sequencing; EP, erythroid progenitor; HS, hypersensitive; MEL, mouse erythroleukemic; TSA, trichostatin A; TSS, transcriptional start site.

†To whom correspondence may be addressed: joattema@stanford.edu or irv@stanford.edu.

‡Present address: John Curtin School of Medical Research and Australian Phenomics Facility, Australian National University, Canberra, Australia.

This article contains supporting information online at www.pnas.org/cgi/content/full/0704468104/DC1.

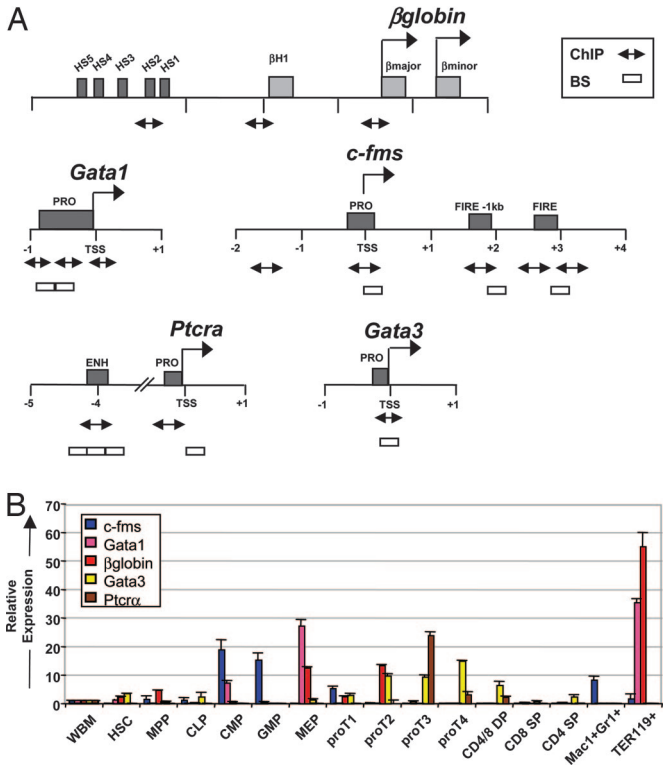


**Fig. 1.** Expression analysis of lineage-affiliated genes in purified hematopoietic cell subsets. (*A*) Locus maps of the *β-globin*, *Gata1*, *c-fms*, *Ptcrα*, *and Gata3* regulatory regions. Numbers indicate the nucleotide position in kilobases relative to the TSS indicated by right-facing arrows. Dark gray boxes represent the regulatory regions, designated names of the elements are shown, and light gray boxes indicate coding regions of the *β-globin* genes. qRT-PCR products for ChIP assays are indicated by double-ended arrows. Open boxes indicate the amplicons for the BS analysis. (*B*) Distribution of lineage-affiliated gene expression. Expression of *c-fms*, *β-globin*, *GATA1*, *GATA3*, and *Ptcrα* in hematopoiesis as measured by qRT-PCR. Relative expression to BM after β-actin normalization is shown on the *y* axis by using the comparative Ct ($2^{-\Delta\Delta Ct}$) method. Error bars represent mean ± SD of three independent sort experiments.

To obtain quantitative changes in differential expression patterns of lineage-affiliated genes in early hematopoiesis, we first performed quantitative real time-PCR (qRT-PCR). The expression of lineage-affiliated genes in the committed progenitors showed an expected distribution, with gene expression increasing in differentiating progenitors, and reaching peak levels in mature progeny (Fig. 1*B*). In addition, these results confirmed the lineage priming model that lymphoid and myeloid-affiliated lineage genes are transcriptionally permissive in HSC before lineage commitment (13, 19, 20).

To analyze chromatin structure at these genes in early hematopoiesis, we investigated the distribution of histone modifications using miniChIP. This assay allowed for the analysis of 50,000 cells, which made the technique amenable to the numbers of HSC and progenitors that can be reasonably obtained to high purity from adult mouse bone marrow (BM). We established the method on 50,000 cells through the systematic adjustment of formaldehyde cross-linking, sonication, preclearing, salt concentration and antibody immunoprecipitation conditions (see *Materials and Methods*).

The feasibility and specificity of the miniChIP method was addressed by examining active histone modifications across functionally important sites controlling *β-globin* expression in erythroid cells (Fig. 1*A*). High-level transcription of *β-globin* genes is mediated by a locus control region that comprises at least five DNaseI hypersensitive (HS) sites (25–27). We used antibodies specific for acetylated histone H3 (AcH3), AcH4, dimethylated lysine 4 of histone H3 (H3K4me2), and rabbit IgG for assessment of background signals. As previously demonstrated, we detected about a 4- to 6-fold increase in AcH3 and AcH4 levels at the HSII and β-major promoter regions relative to the β-H1 embryonic promoter in both $1 \times 10^6$ and 50,000 mouse erythroleukemic (MEL) cells (Fig. 2*A*) (28). Robust levels of histone acetylation were also observed at the active *β-actin* promoter, whereas the silent *albumin* and *α-fetoprotein* promoter regions lacked detectable AcH3 and AcH4 (Fig. 2*A*). We next investigated the effect of trichostatin A (TSA) on histone acetylation across all regions in 50,000 MEL cells. TSA treatment led to large increases in AcH4 levels across the entire locus with the highest levels of enrichment observed at the βH1 embryonic promoter, consistent with previous findings (Fig. 2*B*) (28). These data clearly demonstrated the specificity of miniChIP to detect dynamic changes in histone acetylation. Our following experiment analyzed the levels of AcH3 and AcH4 at these control regions in primary EP cells isolated from BM that expressed high levels of *β-globin* (Fig. 1*B* and SI Fig. 5). Consistent with the miniChIP on MEL cells, increases in AcH3 and AcH4 across the HSII and β-major promoter regions in 50,000 primary EP cells were seen (Fig. 2*C*). As further confirmation of miniChIP specificity, purified splenic $CD4^+$ T cells that lacked *β-globin* expression failed to show AcH3 and AcH4 modifications (Figs. 1*B* and 2*C*).

Previous studies have found that the HS sites of the *β-globin* locus control region associate with active histone modifications in hematopoietic progenitor cells (13, 17). We confirmed these observations by analyzing active histone modifications across the *β-globin* locus in HSC and MPP (SI Fig. 5) (29). H3K4me2 associates with the 5′ regulatory region of genes that are transcriptionally permissive (16, 30–33) and is proposed to be present at the HSII site in HSC before full transcriptional activation occurs in committed erythroid cells (17, 27). Although we found robust enrichment of H3K4me2 and AcH3 at the HSII site in HSC and MPP, little or no levels were detected at the β-major promoter in HSC (Fig. 2*D*). However, the β-major promoter was found to harbor higher levels of these histone modifications in MPP consistent with the observation that these cells expressed higher levels of *β-globin* (Figs. 1*B* and 2*D*). Thus, by using miniChIP, covalent histone modifications were able to be reproducibly examined across numerous loci in highly purified primary hematopoietic stem and progenitor cells.

We then began to analyze epigenetic histone modifications in HSC and MPP compared with CLP and three erythromyeloid precursor populations, CMP, GMP, and MEP (SI Figs. 5 and 6). We first examined histone modifications spanning the *Gata1* promoter in HSC and MPP (Figs. 1*A* and 3*A*). H3K4me2 was detected at all three sites of the promoter in both cell types, whereas H3K4me3 and AcH3 were absent with the exception of a small enrichment of H3K4me3 at the transcriptional start site (TSS) in HSC (Fig. 3*A*). Interestingly, the *Gata1* promoter lacked AcH3 in HSC suggesting that this modification was associated with low-level transcription (Figs. 1*B* and 3*A*). As differentiation progressed to CMP and MEP, H3K4me2 was sustained at all three promoter sites and increased (≈3-fold) at the TSS in MEP. H3K4me3 accumulated at the TSS in CMP and MEP and across the promoter in MEP (Fig. 3*A*). Whereas active histone modifications were found to be absent from the *Gata1* promoter region in CLP and GMP, high levels of H3K27me3, a mark of silent chromatin, were observed in these cell types (Figs. 1*B* and 3*A*).

The myeloid-specific *c-fms* locus contains several regulatory sites (Fig. 1*A*). Robust levels of H3K4me2 were observed in HSC, MPP, CMP, and GMP across all regulatory regions, with reduced levels in CLP and MEP (Fig. 3*B*). A striking enrichment of H3K4me3 was found at the promoter in CLP, CMP, and GMP, whereas low enrichment was seen at the FIRE region in HSC and MPP. These data may have reflected a transient role of H3K4me3 at the enhancer for transcriptional potentiation in HSC and MPP before

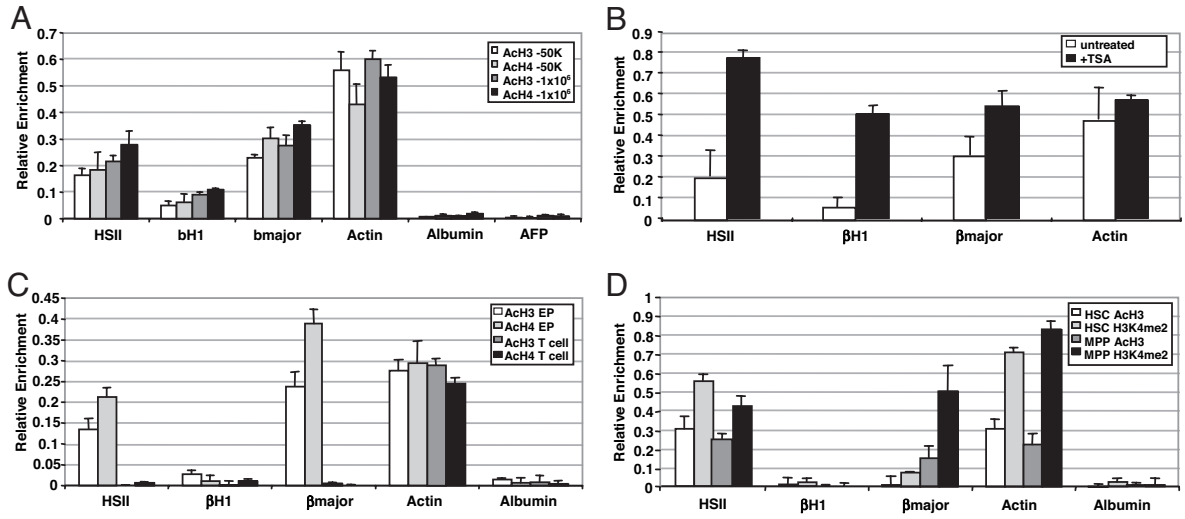


**Fig. 2.** Establishment of the miniChIP assay. (*A*) Comparison of histone modifications detected at the *β-globin* locus in $1 \times 10^6$ and 50,000 MEL cells. Gene-specific enrichment was examined by qRT-PCR. IgG background was subtracted from the histone-specific antibody enrichment values. Error bars represent the mean ± SD of six independent experiments. (*B*) MiniChIP analysis of 50,000 TSA-treated MEL cells using antibodies to AcH4. Open bars represent enrichment values in the untreated control, and black bars indicate AcH4 levels in cells treated with TSA. (*C*) MiniChIP analysis of 50,000 primary EP and splenic $CD4^+$ T cells with antibodies against acetylated histones H3 and H4. (*D*) MiniChIP assays using antibodies specific for AcH3 and H3K4me2 in 50,000 HSC and MPP. Error bars represent the mean ± SD of three independent experiments.

higher expression levels in CMP and GMP. AcH3 levels across the locus were comparable in HSC, MPP, CMP, and GMP. Although histone modifications at the FIRE region in CLP were low or absent, we observed active marks at the *c-fms* promoter potentially reflecting the latent myeloid potential of these cells (Figs. 1*B* and 3*B*) (34, 35). Interestingly, H3K27me3 was robustly detected across the entire locus in MEP, suggesting that K27 methylation may act to permanently silence *c-fms* in cells not destined for a myeloid fate.

Conversely, we identified accumulated levels of active histone modifications at the lymphoid-affiliated genes *Ptcrα* and *Gata3* in CLP, along with increased levels of H3K27me3 in most of the erythromyeloid populations (Fig. 3 *C* and *D*). *Ptcrα* transcripts were undetectable in the six stem and progenitor subsets examined, although differential levels of histone modifications were detected at the enhancer and promoter regions. Robust levels of H3K4me2 were observed at the *Ptcrα* enhancer in HSC, MPP, and CLP, whereas lower levels were seen at the promoter (Fig. 3*C*). We failed to detect robust H3K4me3 levels in HSC but observed very low enrichment of H3K4me3 at the *Ptcrα* promoter in MPP and CLP, with highest AcH3 levels in CLP. Thus, increased levels of activating chromatin modifications were found to correspond with the lymphoid potential of these cells despite *Ptcrα* not being expressed in these cells (Fig. 1*B*). This trend was also reflected by a reduction in H3K27me3 at the promoter region in CLP compared with the five



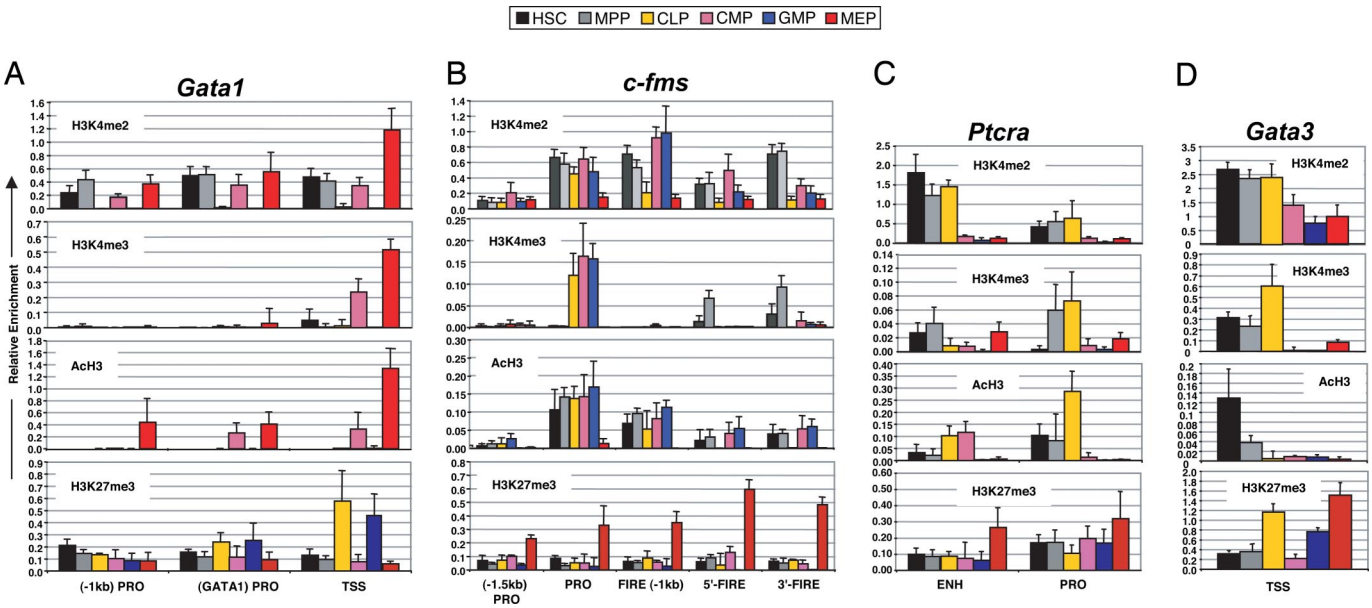


**Fig. 3.** Active histone modifications associate with lineage-affiliated genes during early hematopoiesis. MiniChIP analysis of H3K4me2, H3K4me3, AcH3, and H3K27me3 levels at the regulatory regions of *Gata1* (*A*), *c-fms* (*B*), *Ptcrα* (*C*), and *Gata3* (*D*) genes in HSC (black bars), MPP (gray bars), CLP (yellow bars), CMP (pink bars), GMP (blue bars), and MEP (red bars) using qRT-PCR. Error bars represent the mean ± SD of three independent miniChIP experiments.

A

| Location | | HSC | | MPP | | CLP | | CMP | | GMP | | MEP | | Liver | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | % | Ratio | % | Ratio | % | Ratio | % | Ratio | % | Ratio | % | Ratio | % | Ratio |
| *Gata1* PRO | -762 | 92 | 11/12 | 75 | 9/12 | 100 | 12/12 | 50 | 6/12 | 90 | 9/10 | 20 | 2/10 | 87 | 13/15 |
| | -753 | 92 | 11/12 | 92 | 11/12 | 100 | 12/12 | 50 | 6/12 | 100 | 10/10 | 20 | 2/10 | 100 | 15/15 |
| | -743 | 83 | 10/12 | 83 | 10/12 | 92 | 11/12 | 50 | 6/12 | 100 | 10/10 | 20 | 2/10 | 100 | 15/15 |
| | -670 | 50 | 6/12 | 83 | 10/12 | 100 | 12/12 | 17 | 2/12 | 50 | 5/10 | 0 | 0/10 | 64 | 9/14 |
| | -651 | 45 | 5/12 | 45 | 5/12 | 75 | 9/12 | 8 | 1/12 | 70 | 7/10 | 0 | 0/10 | 7 | 1/14 |
| | -600 | 58 | 7/12 | 100 | 12/12 | 100 | 12/12 | 17 | 2/12 | 90 | 9/10 | 20 | 2/10 | 92 | 12/13 |
| | -585 | 50 | 6/12 | 100 | 12/12 | 92 | 11/12 | 0 | 0/12 | 90 | 9/10 | 20 | 2/10 | 85 | 11/13 |
| | -580 | 75 | 9/12 | 92 | 11/12 | 100 | 12/12 | 17 | 2/12 | 70 | 7/10 | 20 | 2/10 | 100 | 13/13 |

B

| Location | | HSC | | MPP | | CLP | | CMP | | GMP | | MEP | | Liver | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | % | Ratio | % | Ratio | % | Ratio | % | Ratio | % | Ratio | % | Ratio | % | Ratio |
| *c-fms* PRO | +189 | 0 | 0/8 | 0 | 0/8 | 0 | 0/8 | 0 | 0/7 | 0 | 0/10 | 0 | 0/11 | 60 | 6/10 |
| | +205 | 13 | 1/8 | 0 | 0/8 | 0 | 0/8 | 14 | 1/7 | 0 | 0/10 | 0 | 0/11 | 50 | 5/10 |
| | +235 | 0 | 0/8 | 0 | 0/8 | 0 | 0/8 | 0 | 0/7 | 0 | 0/10 | 0 | 0/11 | 50 | 5/10 |
| | +245 | 0 | 0/8 | 0 | 0/8 | 0 | 0/8 | 0 | 0/7 | 0 | 0/10 | 27 | 3/11 | 60 | 6/10 |
| | +263 | 75 | 6/8 | 13 | 1/8 | 0 | 0/8 | 14 | 1/7 | 0 | 0/10 | 45 | 5/11 | 60 | 6/10 |
| | +270 | 75 | 6/8 | 38 | 3/8 | 13 | 1/8 | 43 | 3/7 | 0 | 0/10 | 64 | 7/11 | 60 | 6/10 |
| | +388 | 88 | 7/8 | 100 | 8/8 | 100 | 8/8 | 100 | 4/4 | 50 | 5/10 | 90 | 9/10 | 70 | 7/10 |
| *c-fms* FIRE -1 Kb | +1985 | 100 | 8/8 | 100 | 8/8 | 75 | 6/8 | 92 | 11/12 | 83 | 10/12 | 50 | 6/12 | 73 | 8/11 |
| | +2028 | 75 | 6/8 | 13 | 1/8 | 88 | 7/8 | 17 | 2/12 | 0 | 0/12 | 67 | 8/12 | 91 | 10/11 |
| | +2041 | 63 | 5/8 | 88 | 7/8 | 88 | 7/8 | 33 | 4/12 | 8 | 1/12 | 33 | 4/12 | 91 | 10/11 |
| | +2056 | 25 | 1/8 | 75 | 6/8 | 88 | 7/8 | 8 | 1/12 | 8 | 1/12 | 50 | 6/12 | 91 | 10/11 |
| | +2076 | 13 | 1/8 | 63 | 5/8 | 88 | 7/8 | 17 | 2/12 | 8 | 1/12 | 58 | 7/12 | 91 | 10/11 |
| | +2083 | 13 | 1/8 | 13 | 1/8 | 38 | 3/8 | 0 | 0/12 | 8 | 1/12 | 0 | 0/12 | 91 | 10/11 |
| | +2092 | 38 | 3/8 | 0 | 0/8 | 38 | 3/8 | 0 | 0/12 | 8 | 1/12 | 67 | 8/12 | 91 | 10/11 |
| | +2159 | 0 | 0/8 | 0 | 0/8 | 0 | 0/8 | 0 | 0/12 | 0 | 0/12 | 25 | 3/12 | 100 | 11/11 |
| | +2217 | 0 | 0/8 | 0 | 0/8 | 0 | 0/8 | 0 | 0/12 | 10 | 1/10 | 17 | 2/12 | 100 | 11/11 |
| *c-fms* FIRE | +2966 | 0 | 0/8 | 0 | 0/7 | 0 | 0/8 | 0 | 0/11 | 0 | 0/8 | 40 | 4/10 | 0 | 0/8 |
| | +2985 | 0 | 0/8 | 0 | 0/7 | 13 | 1/8 | 0 | 0/11 | 0 | 0/8 | 10 | 1/10 | 0 | 0/8 |
| | +3002 | 0 | 0/8 | 0 | 0/7 | 0 | 0/8 | 0 | 0/11 | 0 | 0/8 | 0 | 0/10 | 0 | 0/8 |
| | +3023 | 0 | 0/8 | 0 | 0/7 | 0 | 0/8 | 0 | 0/11 | 0 | 0/8 | 30 | 3/10 | 13 | 1/8 |
| | +3036 | 0 | 0/8 | 0 | 0/7 | 0 | 0/8 | 0 | 0/11 | 0 | 0/8 | 0 | 0/10 | 0 | 0/8 |
| | +3042 | 0 | 0/8 | 0 | 0/7 | 0 | 0/8 | 0 | 0/11 | 0 | 0/8 | 0 | 0/10 | 0 | 0/8 |
| | +3137 | 0 | 0/8 | 0 | 0/7 | 0 | 0/8 | 0 | 0/11 | 0 | 0/7 | 0 | 0/9 | 25 | 2/8 |
| | +3183 | 0 | 0/8 | 0 | 0/7 | 0 | 0/7 | 0 | 0/8 | 0 | 0/7 | 43 | 3/7 | 38 | 3/8 |
| | +3247 | 0 | 0/7 | 0 | 0/7 | 0 | 0/7 | 0 | 0/8 | 0 | 0/7 | 83 | 5/6 | 50 | 4/8 |

C

| Location | | HSC | | MPP | | CLP | | CMP | | GMP | | MEP | | proT3 | | proT4 | | Thym. | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | % | Ratio | % | Ratio | % | Ratio | % | Ratio | % | Ratio | % | Ratio | % | Ratio | % | Ratio | % | Ratio |
| *Ptcra* ENH | -4362 | 100 | 12/12 | 100 | 12/12 | 100 | 12/12 | 100 | 12/12 | 100 | 12/12 | 100 | 12/12 | 17 | 2/12 | 31 | 4/13 | 0 | 0/3 |
| | -4318 | 50 | 6/12 | 75 | 9/12 | 75 | 9/12 | 75 | 9/12 | 58 | 7/12 | 67 | 8/12 | 33 | 4/12 | 23 | 3/13 | 0 | 0/3 |
| | -4292 | 58 | 7/12 | 67 | 8/12 | 50 | 6/12 | 50 | 6/12 | 33 | 4/12 | 50 | 6/12 | 25 | 3/12 | 31 | 4/13 | 0 | 0/3 |
| | -4220 | 25 | 3/12 | 42 | 5/12 | 33 | 4/12 | 25 | 3/12 | 58 | 7/12 | 58 | 7/12 | 17 | 2/12 | 23 | 3/13 | 0 | 0/3 |
| | -4130 | 25 | 3/12 | 50 | 6/12 | 36 | 4/11 | 25 | 3/12 | 58 | 7/12 | 17 | 2/12 | 0 | 0/13 | 15 | 2/13 | 0 | 0/3 |
| | -4080 | 0 | 0/12 | 0 | 0/12 | 0 | 0/11 | 0 | 0/12 | 0 | 0/12 | 0 | 0/12 | 0 | 0/13 | 0 | 0/13 | 0 | 0/3 |
| | -4042 | 50 | 6/12 | 42 | 5/12 | 73 | 8/11 | 67 | 8/12 | 50 | 6/12 | 33 | 4/12 | 0 | 0/13 | 0 | 0/13 | 0 | 0/3 |
| | -3997 | 42 | 5/12 | 17 | 2/12 | 42 | 5/12 | 8 | 1/12 | 0 | 0/12 | 17 | 2/12 | 0 | 0/8 | 0 | 0/8 | 0 | 0/6 |
| | -3965 | 100 | 12/12 | 100 | 12/12 | 100 | 12/12 | 83 | 10/12 | 50 | 6/12 | 92 | 11/12 | 0 | 0/8 | 0 | 0/8 | 0 | 0/6 |
| | -3947 | 42 | 5/12 | 58 | 7/12 | 83 | 10/12 | 67 | 8/12 | 8 | 1/12 | 92 | 11/12 | 0 | 0/8 | 0 | 0/8 | 0 | 0/6 |
| | -3900 | 25 | 3/12 | 75 | 9/12 | 100 | 12/12 | 58 | 7/12 | 0 | 0/12 | 75 | 9/12 | 13 | 1/8 | 0 | 0/8 | 0 | 0/6 |
| | -3818 | 25 | 3/12 | 42 | 5/12 | 8 | 1/12 | 36 | 4/11 | 0 | 0/12 | 58 | 7/11 | 0 | 0/8 | 0 | 0/8 | 0 | 0/6 |
| | -3806 | 92 | 11/12 | 100 | 12/12 | 83 | 10/12 | 55 | 6/11 | 25 | 3/12 | 100 | 12/12 | 38 | 3/8 | 25 | 2/8 | 17 | 1/6 |
| | -3800 | 75 | 9/12 | 58 | 7/12 | 67 | 8/12 | 55 | 6/11 | 17 | 2/12 | 67 | 8/12 | 13 | 1/8 | 13 | 1/8 | 0 | 0/6 |
| | -3722 | 90 | 9/10 | 90 | 9/10 | 100 | 10/10 | 90 | 9/10 | 50 | 5/10 | 60 | 6/10 | 13 | 1/8 | 13 | 1/8 | 14 | 1/7 |
| | -3693 | 100 | 10/10 | 100 | 10/10 | 90 | 9/10 | 100 | 10/10 | 70 | 7/10 | 100 | 10/10 | 38 | 3/8 | 13 | 1/8 | 29 | 1/7 |
| *Ptcra* PRO | +111 | 92 | 11/12 | 83 | 10/12 | 50 | 3/6 | 71 | 5/7 | 88 | 7/8 | 88 | 7/8 | 0 | 0/10 | 0 | 0/10 | 0 | 0/4 |
| | +152 | 92 | 11/12 | 92 | 11/12 | 100 | 6/6 | 100 | 7/7 | 100 | 8/8 | 100 | 8/8 | 0 | 0/10 | 0 | 0/10 | 0 | 0/4 |


**Fig. 4.** HSC and committed progenitors exhibit distinct DNA methylation patterns. BS analysis across the regulatory regions of *Gata1* (*A*), *c-fms* (*B*), and *Ptcrα* (*C*) genes in HSC, MPP, CLP, CMP, GMP, MEP, proT3, proT4, and liver cells. Results obtained for the *Ptcrα* gene are also described in ref. 49. The percent methylation of each CpG dinucleotide was determined by the number of methylated clones in the total number of clones sequenced as indicated in the heat map ranging from 100% methylated (red) to 0% methylated (green).

other subpopulations. The TSS of *Gata3* revealed H3K4me2 in all cell types although levels were lowest in the myeloid-committed subsets (Fig. 3*D*). H3K4me3 and AcH3 levels were highest in CLP and present at lower levels in HSC and MPP. Analysis of H3K27me3 revealed high levels at the TSS in GMP and MEP, consistent with the qRT-PCR data (Figs. 1*B* and 3*D*). High levels of H3K27me3 were also observed in CLP, with modest signals detected in HSC and MPP. The presence of H3K27me3 was unexpected based on the gene expression profiles across these subsets and suggested that the *Gata3* TSS is marked by bivalent domains harboring both activating (H3K4me3) and silencing (H3K27me3) histone modifications (36, 37). Taken together, these results revealed that there were gene-specific combinations of active and silent chromatin domains present at gene regulatory regions that were either expressed at low levels or not expressed (*Ptcrα*) in HSC. These particular histone modifications may execute important roles in lineage commitment.

It has been suggested that histone modifications and methylation states of CpG dinucleotides cooperate to control embryonic stem (ES) cell pluripotency (37–40). To address whether active histone modifications colocalize with altered DNA methylation densities, we next used BS analysis. As expected, numerous CpGs located within the *albumin* enhancer were consistently methylated in hematopoietic populations, whereas they were unmethylated in liver cells expressing *albumin* (SI Fig. 7). The *Gata1* promoter contains eight CpG sites located ≈700 bp upstream from the TSS (Fig. 4*A*). Partial methylation of this region observed in HSC was increasingly lost in CMP and MEP, consistent with the miniChIP and qRT-PCR data (Figs. 1*B*, 3*A*, and 4*A*). By contrast, these CpGs were methylated in CLP and GMP. The CpG site located at −651bp is adjacent to a functionally critical *Gata1* binding motif (41). This CpG, and to a lesser extent adjacent CpGs, were partially methylated in HSC and MPP, raising the possibility that *Gata1* binding to this site primed the HSC toward an erythroid cell fate. This site was also unmethylated in liver, which may have be a reflection of transcription factor occupancy including *Gata1* complexes in hepatocytes (Fig. 4*A*).

We next examined DNA methylation across the *c-fms* locus. Consistent with its location across a CpG island, CpGs within the FIRE region were mostly unmethylated in all cell types (Fig. 4*B*). By contrast, CpGs within the −1 kb FIRE site were completely methylated in liver, with very low methylation signals observed in all hematopoietic populations. The only exception we found was the partial methylation in CLP, which was a result consistent with previous findings (42). We found lowest methylation densities in the CMP and GMP subpopulations, with slightly higher levels in MEP. A similar pattern was observed for the CpGs within the promoter, which were mostly unmethylated in all six hematopoietic populations, with lowest methylation levels observed in GMP (Fig. 4*B*). By contrast, methylation was observed at all promoter CpGs in liver. Reduced methylation levels at these control regions in CMP and GMP were in agreement with the *c-fms* expression profiles. Fur-

thermore, the unmethylated CpGs at the promoter and −1 kb FIRE sites were found to coincide with the distribution of active histone modification regions (Figs. 3*B* and 4*B*).

Finally, we investigated DNA methylation at the *Ptcrα* enhancer and promoter located between two constitutively expressed genes (*Rik23* and *Tnrc5*) that contain completely unmethylated CpG island promoters (Fig. 4*C* and SI Fig. 8). DNA methylation analysis of the enhancer revealed an unmethylated CpG site (−4080) in all cell types examined including cell samples from spleen and liver (Fig. 4*C*). This CpG coincided with a functionally critical Myb binding site (43). We observed varying degrees of methylation at other CpGs in all subsets, which were overall less methylated in proT3 and proT4 cells (Fig. 4*C*). BS analysis of the *Ptcrα* promoter revealed one partially methylated CpG site in CLP, which may indicate increased transcriptional potentiation before high-level expression in proT3 and proT4 cells (Fig. 4*C*). Accordingly, both enhancer and promoter regions were unmethylated in proT3 and proT4 cells expressing the highest levels of *Ptcrα* (Figs. 1*B* and 4*C*). BS analysis of CpGs within the *Gata3* promoter revealed a complete lack of methylation in HSC, the downstream transit-amplifying progenitors, and in liver which was likely due to its location within a CpG island (SI Fig. 9).

Together, these data demonstrate that many of the unmethylated CpGs in HSC colocalize with well defined *cis*-elements functionally required for transcriptional activity. Most regulatory regions comprising these selectively undermethylated CpGs were also marked by active histone modifications, and, in one instance, bivalent domains revealing a close association between histone and DNA epigenetic modifications.

## Discussion

Here, we have examined chromatin structure in highly purified primary hematopoietic stem and progenitor cells using miniChIP and BS analysis. The analysis of histone modifications in HSC and daughter progenitors was possible by the optimization of conventional ChIP assays to 50,000 cells. The miniChIP method described here allows for the quantitative analysis of fixed chromatin under optimal sonication and immunoprecipitation conditions. When used in conjunction with the recently developed carrier ChIP (44) and $Q^2$ChIP (45) assays, which were described during the preparation of this paper, epigenetic analysis should be possible on extremely rare cells (i.e., 100) contained within primary tissues. Therefore, these newly described ChIP methods will serve as invaluable tools for the epigenetic analysis of primary cells in numerous developmental systems and neoplastic cell populations, including cancer stem cells (46), where cell numbers are limiting.

The epigenetic analysis of early hematopoiesis raises the exciting possibility that chromatin structure plays a critical role for the establishment and maintenance of multipotency. We hypothesize that chromatin structure at these regulatory regions ensures inheritance of epigenetic states to maintain lineage-specific expression or repression during the process of lineage commitment. Although the data presented here are a limited data set, clear trends are discernable. First, the patterns of H3K4me2 were strikingly similar between the two multipotent populations examined, HSC and MPP. This observation is consistent with the notion that although the genes investigated were expressed at low levels or not at all, these genes are "primed" for expression later in development. Second, H3K4me2 enrichment was most prominent at the enhancer, not the promoter, in preexpressing cells. For example, HSII of the β-globin locus and the *Ptcrα* enhancer carried high levels of H3K4me2 in HSC and MPP as well as in HSC, MPP, and CLP, respectively. Many recent locus and genome scale analyses have found that genes in a transcriptionally poised nonexpressing state are enriched in H3K4me2 (16, 30–33, 47, 48). It is possible that H3K4me2 at regulatory regions, particularly at enhancers, maintains these genes in a transcriptionally permissive state to prime the HSC for lineage commitment. Third, H3K4me3, a marker of actively transcribing genes (30), was observed at the promoter or TSS of genes in cells showing highest expression (*c-fms* and *Gata1*) or developmentally nearest to cells expressing the gene (*Gata3* and *Ptcrα*). Thus, enhancers could be important for priming genes for expression later in hematopoiesis whereas promoter modifications may serve as true indicators of active transcription. A similar trend was observed with DNA methylation of the *Ptcrα* locus. Whereas the enhancer displayed varying degrees of methylation, promoter methylation strictly correlated with *Ptcrα* expression. Fourth, H3K27me3, which typically associates with transcriptionally silent genes, was highest at promoter regions in the cells most developmentally distant from those expressing the gene (*Gata1*, *c-fms*, and *Ptcrα*). Bivalent domains were detected at the TSS of *Gata3* in HSC and MPP, with highest enrichment observed in CLP, thus perhaps reflecting the lymphoid potential across these subsets for high level *Ptrcα* expression later in T cell development (2, 23, 35). Interestingly, the *Gata3* TSS is located within a CpG island, which has been shown to strongly coincide with bivalent domains (37). It will be important to determine whether bivalent histone methylation is indeed present at the *Gata3* promoter and not just a reflection of distinct subsets within the CLP population. We are currently investigating the extent to which bivalent histone modifications mark additional lineage-affiliated genes in the defined progenitor populations.

A central question that remains to be resolved is precisely how and when epigenetic marks are established during development. A parallel study has revealed that unmethylated CpG residues at lineage-directing genes in HSC are also unmethylated in ES cells at similar levels (49). Furthermore, the unmethylated windows in ES cells are occupied by sequence-specific transcription factors. These results strongly suggest that control regions for hematopoietic-specific genes are transcriptionally primed in pluripotent cells long before the establishment of the definitive, multipotent HSC. In this case, the transition from pluripotent to multipotent developmental potential may be mediated by selective gene silencing rather than activation. It will be of high interest to identify the specific recruitment of positively and negatively acting transcriptional regulators that contribute to the establishment of epigenetic marks at control regions in HSC and functionally assess these features in lineage priming.

In summary, we extend the proposition that histone and DNA modifications represent important characteristics of HSC development and differentiation. A deeper understanding of these epigenetic features in HSC and their daughter progenitors may illuminate the molecular mechanisms that govern multipotency and lineage commitment, both in normal and pathological hematopoiesis.

## Materials and Methods

**Purification of Cells.** HSC and MPP were sorted as described (21). CLP were sorted as described with the addition of PE-conjugated anti-Flk2 (eBioscience) (ref. 23; H. Karsunky and I.L.W., unpublished results). Erythromyeloid progenitors (24) and T cells (50) were sorted as described (24). Cells were double sorted for high purity by using FACSAria and FACSVantage (Becton Dickinson, San Jose, CA).

**RNA Purification and qRT-PCR.** RNA isolation and qRT-PCR were performed as described (29). Primer sequences were designed by using Primer3 and are available on request. Fold expression relative to whole BM was calculated after β-actin normalization.

**ChIP Assays.** ChIP assays using $1 \times 10^6$ cells per reaction were performed as described (Upstate Biotechnology). Antibodies included anti-acetyl histone H3 (Upstate; 06-599), anti-acetyl histone H4 (Upstate; 06-866), anti-dimethyl histone H3K4 (Upstate; 07-030) and anti-trimethyl histone H3K4 (Upstate, 07-473; Abcam, 8580), anti-trimethyl histone H3K27 (Upstate; 07-449), and rabbit

IgG (Upstate; 12-370). Quantitation of ChIP DNA (relative enrichment) used the ABI 7000 Sequence Detection System. DNA concentrations were measured with a ND-1000 spectrophotometer (NanoDrop) and analyzed in triplicate qRT-PCRs.

**MiniChIP Assays.** Prospectively purified hematopoietic cells ranging from 50,000 to 150,000 were sorted into 1.5-ml Eppendorf tubes containing 1 ml of DMEM (Gibco) and 10% FBS and cross-linked in 0.1% formaldehyde for 10 min at room temperature with gentle rocking. Cells were washed in ice-cold HBSS (Gibco) containing protease inhibitor mixture (PIC) (Roche), and each 50,000-cell equivalent was lysed in 100 $\mu$l of lysis buffer (Upstate; 17-295) containing PIC and incubated on ice for 5 min. After lysis, 300 $\mu$l of HBSS containing PIC was added per 50,000-cell equivalent, and 400-$\mu$l aliquots were placed into new Eppendorf tubes. Each 400-$\mu$l aliquot was sonicated by using a Branson Sonifier (Branson Ultrasonics, Danbury, CT), which was empirically determined to give rise to genomic fragments ranging from 200 to 800 bp. Soluble chromatin was collected by 4°C ultracentrifugation (10,000 $\times$ *g* for 10 min), pooled into a new Eppendorf tube, and precleared by using 5 $\mu$l of salmon sperm DNA/protein A agarose beads (Upstate; 16-157) per 50,000-cell equivalent for 30 min at 4°C with rotation. After 4°C ultracentrifugation (1,000 $\times$ *g* relative centrifugal force for 2 min), the supernatant was diluted 2-fold with ChIP dilution buffer (Upstate; 17-295), input was removed, and 800 $\mu$l of chromatin was distributed to new Eppendorf tubes. Antibodies to modified histones were added, ranging from 2 $\mu$g to 10 $\mu$g as empirically tested at control promoters, and a reaction containing an equivalent amount of rabbit IgG was included as the background control. Immunoprecipitations were performed for 8–12 h at 4°C with rotation, and antibody:protein:DNA complexes were then collected with 30 $\mu$l of salmon sperm DNA/protein A agarose beads for 1 h of rotation. The beads were washed as described (Upstate; 17-295) by using 500 $\mu$l of each buffer, then eluted twice from the beads with 200 $\mu$l of fresh elution buffer (1% SDS and 0.1 M $NaHCO_3$). Eluates were pooled, 20 $\mu$l of 5 M NaCl was added, and reactions were incubated at 65°C for 6 h to reverse the protein:DNA cross-links. After incubation, Proteinase K, 1 M Tris (pH 6.0), and 0.5 M EDTA (pH 8.0) were added and incubated at 45°C for 1 h. Genomic DNA was recovered by using phenol chloroform extraction and ethanol precipitation. Pellets were washed in 70% ethanol, briefly air-dried, and resuspended in TE buffer.

**Isolation of Genomic DNA and Bisulfite Sequencing Analysis.** Genomic DNA isolation and bisulfite treatment were performed as described (49).

We thank L. Jerabek for excellent laboratory management, C. Richter for antibody production, and members of the I.L.W. laboratory for valuable discussions. This work was supported in part by National Institutes of Health Grants 5P01DK053074 (to I.L.W.) and R01 DK43726 (to S.T.S and J.X.), by a California Institute of Regenerative Medicine Fellowship (to J.L.A), and by a National Health and Medical Research Council C. J. Martin Fellowship (to P.P.). E.C.F. is supported by National Institutes of Health Grant 1F32 DK72620-01A1.

1. Shizuru JA, Negrin RS, Weissman IL (2005) *Annu Rev Med* 56:509–538.
2. Kondo M, Wagers AJ, Manz MG, Prohaska SS, Scherer DC, Beilhack GF, Shizuru JA, Weissman IL (2003) *Annu Rev Immunol* 21:759–806.
3. de Haan G, Gerrits A, Bystrykh L (2006) *Curr Opin Hematol* 13:249–253.
4. Loose M, Patient R (2006) *Curr Opin Hematol* 13:229–236.
5. Ivanova NB, Dimos JT, Schaniel C, Hackney JA, Moore KA, Lemischka IR (2002) *Science* 298:601–604.
6. Turner BM (2002) *Cell* 111:285–291.
7. Rosenfeld MG, Lunyak VV, Glass CK (2006) *Genes Dev* 20:1405–1428.
8. Hsieh J, Gage FH (2005) *Curr Opin Cell Biol* 17:664–671.
9. Bannister AJ, Kouzarides T (2005) *Nature* 436:1103–1106.
10. Hake SB, Allis CD (2006) *Proc Natl Acad Sci USA* 103:6428–6435.
11. Akashi K (2005) *Ann NY Acad Sci* 1044:125–131.
12. Jimenez G, Griffiths SD, Ford AM, Greaves MF, Enver T (1992) *Proc Natl Acad Sci USA* 89:10618–10622.
13. Hu M, Krause D, Greaves M, Sharkis S, Dexter M, Heyworth C, Enver T (1997) *Genes Dev* 11:774–785.
14. Tagoh H, Himes R, Clarke D, Leenen PJ, Riggs AD, Hume D, Bonifer C (2002) *Genes Dev* 16:1721–1737.
15. Tagoh H, Schebesta A, Lefevre P, Wilson N, Hume D, Busslinger M, Bonifer C (2004) *EMBO J* 23:4275–4285.
16. Szutorisz H, Canzonetta C, Georgiou A, Chow CM, Tora L, Dillon N (2005) *Mol Cell Biol* 25:1804–1820.
17. Bottardi S, Aumont A, Grosveld F, Milot E (2003) *Blood* 102:3989–3997.
18. Medina KL, Pongubala JM, Reddy KL, Lancki DW, Dekoter R, Kieslinger M, Grosschedl R, Singh H (2004) *Dev Cell* 7:607–617.
19. Miyamoto T, Iwasaki H, Reizis B, Ye M, Graf T, Weissman IL, Akashi K (2002) *Dev Cell* 3:137–147.
20. Månsson R, Hultquist A, Luc S, Yang L, Anderson K, Kharazi S, Al-Hashmi S, Liuba K, Thorén L, Adolfsson J, *et al*. (2007) *Immunity* 26:407–419.
21. Christensen JL, Weissman IL (2001) *Proc Natl Acad Sci USA* 98:14541–14546.
22. Forsberg EC, Prohaska SS, Katzman S, Heffner GC, Stuart JM, Weissman IL (2005) *PLoS Genet* 1:e28.
23. Kondo M, Weissman IL, Akashi K (1997) *Cell* 91:661–672.
24. Akashi K, Traver D, Miyamoto T, Weissman IL (2000) *Nature* 404:193–197.
25. Grosveld F (1999) *Curr Opin Genet Dev* 9:152–157.
26. Tolhuis B, Palstra RJ, Splinter E, Grosveld F, de Laat W (2002) *Mol Cell* 10:1453–1465.
27. Vieira KF, Levings PP, Hill MA, Crusselle VJ, Kang SH, Engel JD, Bungert J (2004) *J Biol Chem* 279:50350–50357.
28. Forsberg EC, Downs KM, Christensen HM, Im H, Nuzzi PA, Bresnick EH (2000) *Proc Natl Acad Sci USA* 97:14494–14499.
29. Forsberg EC, Serwold T, Kogan S, Weissman IL, Passegué E (2006) *Cell* 126:415–426.
30. Wysocka J, Swigut T, Xiao H, Milne TA, Kwon SY, Landry J, Kauer M, Tackett AJ, Chait BT, Badenhorst P, *et al*. (2006) *Nature* 442:86–90.
31. Santos-Rosa H, Schneider R, Bannister AJ, Sherriff J, Bernstein BE, Emre NC, Schreiber SL, Mellor J, Kouzarides T (2002) *Nature* 419:407–411.
32. Adachi S, Rothenberg EV (2005) *Nucleic Acids Res* 33:3200–3210.
33. Wysocka J, Swigut T, Milne TA, Dou Y, Zhang X, Burlingame AL, Roeder RG, Brivanlou AH, Allis CD (2005) *Cell* 121:859–872.
34. Kondo M, Scherer DC, Miyamoto T, King AG, Akashi K, Sugamura K, Weissman IL (2000) *Nature* 407:383–386.
35. Laiosa CV, Stadtfeld M, Graf T (2006) *Annu Rev Immunol* 24:705–738.
36. Azuara V, Perry P, Sauer S, Spivakov M, Jorgensen HF, John RM, Gouti M, Casanova M, Warnes G, Merkenschlager M, Fisher AG (2006) *Nat Cell Biol* 8:532–538.
37. Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, *et al*. (2006) *Cell* 125:315–326.
38. Ayton PM, Chen EH, Cleary ML (2004) *Mol Cell Biol* 24:10470–10478.
39. Roh TY, Cuddapah S, Zhao K (2005) *Genes Dev* 19:542–552.
40. Terranova R, Agherbi H, Boned A, Meresse S, Djabali M (2006) *Proc Natl Acad Sci USA* 103:6629–6634.
41. Nicolis S, Bertini C, Ronchi A, Crotta S, Lanfranco L, Moroni E, Giglioni B, Ottolenghi S (1991) *Nucleic Acids Res* 19:5285–5291.
42. Tagoh H, Melnik S, Lefevre P, Chong S, Riggs AD, Bonifer C (2004) *Blood* 103:2950–2955.
43. Reizis B, Leder P (2001) *J Exp Med* 194:979–990.
44. O'Neill LP, Vermilyea MD, Turner BM (2006) *Nat Genet* 38:835–841.
45. Arne Dahl J, Collas P (2007) *Stem Cells* 25:1037–1046.
46. Reya T, Morrison SJ, Clarke MF, Weissman IL (2001) *Nature* 414:105–111.
47. Schneider R, Bannister AJ, Myers FA, Thorne AW, Crane-Robinson C, Kouzarides T (2004) *Nat Cell Biol* 6:73–77.
48. Bernstein BE, Kamal M, Lindblad-Toh K, Bekiranov S, Bailey DK, Huebert DJ, McMahon S, Karlsson EK, Kulbokas EJ, III, Gingeras TR, *et al*. (2005) *Cell* 120:169–181.
49. Xu J, Pope SD, Jazirehi AR, Attema JL, Papathanasiou P, Watts JA, Zaret KS, Weissman IL, Smale ST (2007) *Proc Natl Acad Sci USA* 104:12377–12382.
50. Godfrey DI, Kennedy J, Suda T, Zlotnik A (1993) *J Immunol* 150:4244–4252.