LETTERS

Comprehensive methylome map of lineage commitment from haematopoietic progenitors

Hong Ji[1]*, Lauren I. R. Ehrlich[2]*†, Jun Seita[2]*, Peter Murakami[1], Akiko Doi[1], Paul Lindau[2], Hwajin Lee[1], Martin J. Aryee[3,4], Rafael A. Irizarry[1,3], Kitai Kim[5], Derrick J. Rossi[2]†, Matthew A. Inlay[2], Thomas Serwold[2]†, Holger Karsunky[2]†, Lena Ho[2], George Q. Daley[5], Irving L. Weissman[2] & Andrew P. Feinberg[1]

Epigenetic modifications must underlie lineage-specific differentiation as terminally differentiated cells express tissue-specific genes, but their DNA sequence is unchanged. Haematopoiesis provides a well-defined model to study epigenetic modifications during cell-fate decisions, as multipotent progenitors (MPPs) differentiate into progressively restricted myeloid or lymphoid progenitors. Although DNA methylation is critical for myeloid versus lymphoid differentiation, as demonstrated by the myeloerythroid bias in *Dnmt1* hypomorphs[1], a comprehensive DNA methylation map of haematopoietic progenitors, or of any multipotent/oligopotent lineage, does not exist. Here we examined 4.6 million CpG sites throughout the genome for MPPs, common lymphoid progenitors (CLPs), common myeloid progenitors (CMPs), granulocyte/macrophage progenitors (GMPs), and thymocyte progenitors (DN1, DN2, DN3). Marked epigenetic plasticity accompanied both lymphoid and myeloid restriction. Myeloid commitment involved less global DNA methylation than lymphoid commitment, supported functionally by myeloid skewing of progenitors following treatment with a DNA methyltransferase inhibitor. Differential DNA methylation correlated with gene expression more strongly at CpG island shores than CpG islands. Many examples of genes and pathways not previously known to be involved in choice between lymphoid/myeloid differentiation have been identified, such as *Arl4c* and *Jdp2*. Several transcription factors, including *Meis1*, were methylated and silenced during differentiation, indicating a role in maintaining an undifferentiated state. Additionally, epigenetic modification of modifiers of the epigenome seems to be important in haematopoietic differentiation. Our results directly demonstrate that modulation of DNA methylation occurs during lineage-specific differentiation and defines a comprehensive map of the methylation and transcriptional changes that accompany myeloid versus lymphoid fate decisions.

Haematopoietic stem cells (HSC) can self renew for life and differentiate into all myeloid and lymphoid blood lineages[2] (Fig. 1a). Recent evidence indicates that DNA methylation has a direct role in regulating both HSC self-renewal and commitment to lymphoid versus myeloid fates[1,3]. Although the frequencies of myeloid progenitors and differentiated cells were normal in *Dnmt1*-hypomorphic mice, lymphoid-restricted CLPs and their downstream thymic T cell progenitors (DN1, DN2 and DN3) were diminished. In the bone marrow of *Dnmt1*-hypomorphs, lymphoid, but not myeloid, transcripts were reduced, and promoters of two myeloerythroid genes were hypomethylated in HSCs. These observations support a critical role for DNA methylation in lymphocyte development, possibly through regulation of gene expression.

Here we have examined genome-wide methylation profiles of the mouse haematopoietic system, because it provides the first opportunity to examine differential methylation of a hierarchical progression of purified cell populations with well-characterized differentiation potentials (Fig. 1a). Eight populations, ranging from uncommitted MPP through oligopotent progenitors specified during myeloid versus lymphoid fate decisions, were fluorescence-activated cell sorting (FACS)-purified and subjected to comprehensive high-throughput array-based relative methylation (CHARM) analysis (Fig. 1a and Supplementary Fig. 1). This approach investigated the methylation status of CpGs throughout the mouse genome using an algorithm favouring regions of higher CpG density (including all CpG islands[4]), but without bias for CpG location relative to genes[5]. Using CHARM, we recently found that differential methylation occurs more frequently in CpG island 'shores' (regions within 2 kilobases of an island) than in CpG islands during multiple cellular differentiation processes[6,7]. Additionally, the mRNA of each population was subjected to microarray and PCR with reverse transcription (RT–PCR) analyses to generate gene expression data. Thus, we were able to compare directly differentially methylated regions (DMRs) throughout the genome with expression levels of nearby genes for all eight populations.

This analysis revealed DMRs in numerous genes known to play a role in lymphoid or myeloid fate specification. For example, *Lck*, the Src family kinase member responsible for initiating signalling downstream of the T cell receptor (TCR)[8], was transcriptionally upregulated from DN1 to DN3, consistent with its role in pre-TCR signal transduction (Fig. 1b). Interestingly, as *Lck* transcription was upregulated, CpGs in exon 1 through intron 2 were progressively demethylated (Fig. 1b). Similarly, myeloid specification from MPP through GMP was accompanied by transcriptional upregulation and progressive hypomethylation of *Mpo*, which encodes an enzyme central to the microbicidal activity of neutrophils[9] (Fig. 1c). Additionally, *Cxcr2*, which encodes a chemokine receptor responsible for neutrophil chemotaxis[10], was upregulated during myeloid commitment from CMP through GMP, while the gene was demethylated (Supplementary Fig. 2a).

[1]Center for Epigenetics and Department of Medicine, Johns Hopkins University School of Medicine, 570 Rangos, 725 N. Wolfe St., Baltimore, Maryland 21205, USA. [2]Institute for Stem Cell Biology and Regenerative Medicine, Stanford University School of Medicine, Stanford, California 94305, USA. [3]Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland 21205, USA. [4]Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins University, Baltimore, Maryland 21231, USA. [5]Stem Cell Transplantation Program, Division of Pediatric Hematology/Oncology, Manton Center for Orphan Disease Research, Howard Hughes Medical Institute, Children's Hospital Boston and Dana Farber Cancer Institute; Division of Hematology, Brigham and Women's Hospital; Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School; Harvard Stem Cell Institute; Boston, Massachusetts 02115, USA. †Present addresses: Institute for Cellular and Molecular Biology, Section of Molecular Genetics and Microbiology, University of Texas at Austin, Austin, Texas 78712, USA (L.I.R.E.); Immune Disease Institute, Harvard Stem Cell Institute Department of Pathology, Harvard Medical School Boston, Massachusetts 02115, USA (D.J.R.); Joslin Diabetes Center, Department of Medicine, Harvard Medical School, Boston, Massachusetts 02215, USA (T.S.); Cellant Therapeutics, San Carlos, California 94070, USA (H.K.).
*These authors contributed equally to this work.

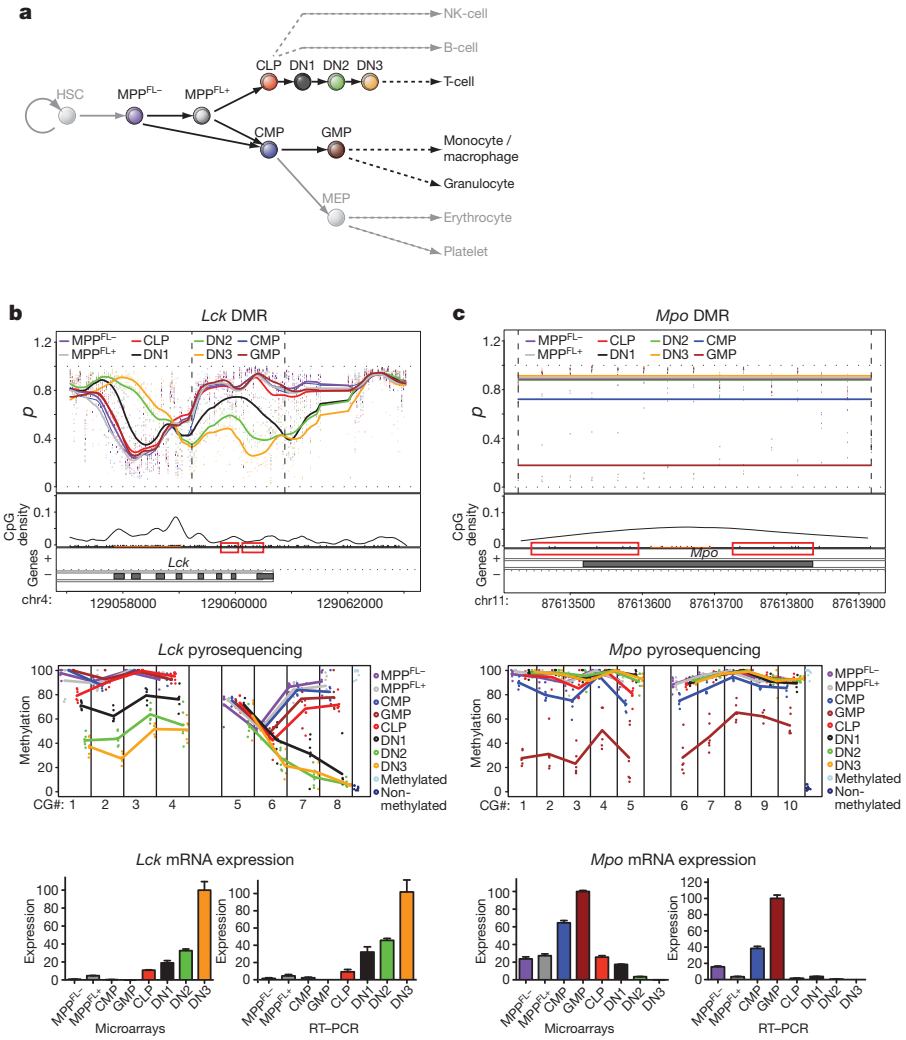

Figure 1 | Examples of known lineage-related genes showing differential DNA methylation between lymphoid and myeloid progenitors. **a**, Haematopoietic progenitors included in this study. Dashed-arrow indicates existence of intermediate progenitors. DMR in **b**, *Lck* and **c**, *Mpo*. Top panels: top half: CpG methylation (*p*); lower half: CpG dinucleotides (black tick marks), CpG density (curve), CpG islands (orange lines) and the gene annotation (see online Methods). Middle panels: methylation of individual CpGs (in the red boxes), mean values connected by lines. Bottom panels: mRNA expression levels, normalized to the highest expression among the populations (mean ± s.d., $n = 3$; $n = 5$ for MPP^{FL-} for microarrays).

Furthermore, *Gadd45α* (also known as *Gadd45a*), which is implicated in myeloid development[11], was found to be concomitantly upregulated and demethylated in the CMP to GMP transition (Supplementary Fig. 2b). *Gadd45α* can actively demethylate DNA in different model systems[12,13]; thus, hypomethylation of *Gadd45α* during myelopoiesis may promote further hypomethylation of genes regulating myeloid commitment. However, the role of *Gadd45α* in promoting demethylation is still controversial[14]. Taken together, these data indicate that CHARM analysis correctly identifies DMRs in known lymphoid and myeloid specifying genes, each confirmed by pyrosequencing and gene expression analysis, making it a valuable tool for identifying candidate genes important for lymphoid or myeloid fate specification.

Viewed globally, CHARM analysis revealed striking epigenetic plasticity, resulting in increased overall methylation upon lymphoid relative to myeloid commitment (Table 1). Most DMRs distinguishing MPP^{FL+} cells from CLP lost methylation during this step of early lymphoid commitment, but upon the subsequent transition to DN1, 15-fold more DMRs showed gain, as opposed to loss, of methylation. Similarly in the earliest step of myeloid commitment from MPP^{FL+} to CMP there were substantially more hypermethylated than hypomethylated DMRs, but nearly all DMRs showed loss of methylation on transition from CMP to GMP. Comparing DN1 to GMP, two populations similarly differentiated towards lymphoid and myeloid fates, respectively, there were eightfold more DMRs with higher-level methylation in DN1 cells, suggesting a skewing towards greater methylation in lymphoid compared to myeloid haematopoiesis. These observations might explain why *Dnmt1*-hypomorphic mice, which are unable to maintain CpG methylation properly, have normal myeloid, but diminished lymphoid development[1,3].

To test the hypothesis that reduced methylation preferentially promotes myeloid as opposed to lymphoid differentiation, we turned to an *in vitro* assay system that promotes both myeloid and lymphoid development[15,16]. In the presence of 5-aza-2′-deoxycytidine, the percentage of myeloid progeny increased at the expense of lymphoid progeny for MPP^{FL+}, CLP, DN1 and DN2, but not DN3, which remained lymphoid committed (Supplementary Fig. 3a, b). This myeloid skewing was most pronounced in DN1 cells, perhaps indicating that the large number of methylated DMRs in DN1 compared to CLP is critical for lymphoid specification (Table 1). We conclude that inhibiting DNA methylation promotes myeloid versus lymphoid specification, providing a mechanism for the myeloid skewing observed in *Dnmt1* hypomorphs[1].

Consistent with our previous studies[6,7], most DMRs were in CpG island shores (Table 1). The exceptions were for MPP^{FL-} versus MPP^{FL+}, and MPP^{FL+} versus CLP, in which most DMRs were in CpG islands: interestingly, both of these transitions are involved in early differentiation. Differential DNA methylation and gene expression showed a statistically significant inverse relationship particularly at CpG island shores (Fig. 2 and Supplementary Fig. 4). As the CHARM array design was targeted towards CpG density but not gene architecture *per se*, we also created a new array that included all promoters, and hybridized DNA from three of the groups studied earlier. Analysis showed a similar statistically significant inverse relationship between differential DNA methylation and gene expression, again particularly at CpG island shores (Supplementary Fig. 4g, h). Thus, CpG island shores are the regions with the most variability in DNA methylation between haematopoietic populations, and this variability correlates best with changes in gene expression. However, not

Table 1 | Numbers of differentially methylated regions (DMRs) identified in myeloid versus lymphoid commitment

Comparisons (Group 1 versus Group 2)	Numbers of DMRs*		DMRs within 2 kb of a transcriptional start site	Locations of DMRs relative to CpG islands (%)		
	Group 1 > Group 2	Group 1 < Group 2		Islands	Shores†	>2 kb‡
MPP^{FL-} vs MPP^{FL+}	18	112	69	56.1	35.4	8.5
MPP^{FL+} vs CLP	302	46	182	52.4	39.0	8.6
CLP vs DN1	203	3017	2433	17.8	80.7	1.5
DN1 vs DN2	3923	174	3125	18.9	80.3	0.7
DN2 vs DN3	47	12	13	5.0	78.3	16.7
MPP^{FL+} vs CMP	149	1412	1129	17.3	81.2	1.5
CMP vs GMP	768	11	554	16.8	81.1	2.1
DN1 vs GMP	1011	122	773	16.8	78.2	5.0

* Area cutoff of 2.0 was used to calculate the number of DMRs (see Methods).
† Shores were defined for this table as overlapping islands by <50% and/or extending up to 2 kb from an island.
‡ Regions more than 2 kb from an island.

all DNA methylation changes correlated with changes in gene expression: for example, *Tha1* is demethylated during lymphoid specification (see CHARM plots on http://charm.jhmi.edu/hsc), but is expressed at high levels from MPP through DN3. In converse, and as expected since there are multiple mechanisms for epigenetic regulation, we also identified lineage-specifying genes with changes in expression levels, but not in DNA methylation, such as *Gata3* and *Hes1* (see microarrays deposited in GEO).

Many novel genes with the potential to contribute to myeloid/lymphoid fate specification were revealed by comparing CHARM-identified DMRs with gene expression data. For example, *Arl4c*, a member of the ADP-ribosylation factor family of GTP-binding proteins, was upregulated and hypomethylated in DN1–3 thymocytes (Fig. 3a). *Arl4c* may have a role in vesicular transport[17], but its role in lymphoid specification is unknown. Several other genes with DMRs indicative of a role in lymphoid development, such as *Smad7*, *Gcnt2* and *Cited2*, were also identified (Supplementary Fig. 5). *Smad7*, which negatively regulates TGF-beta signalling, is selectively upregulated and hypomethylated at the earliest stages of thymocyte development, indicating a role in promoting lymphopoiesis (Supplementary Fig. 5a). However, it causes myeloid lineage skewing when overexpressed in human cord blood progenitors[18]. *Gcnt2* transcripts were downregulated in thymocyte progenitors, and the locus became hypermethylated progressively in DN1–3 progenitors (Supplementary Fig. 5b), consistent with a role for *Gcnt2* in enabling the myeloid potential that is lost during final lymphoid lineage commitment at the DN3 stage[19,20].

Novel potential regulators of myelopoiesis were also identified. The *Jdp2* locus was hypomethylated and its transcript was upregulated in CMP and GMP relative to thymocyte progenitors (Fig. 3b). *Jdp2* is thought to repress transcription by recruiting histone deacetylases and regulating nucleosome assembly[21]. *Dach1* was also hypomethylated and expressed from MPP^{FL-} through GMP, but was silenced in CLP and DN1–3 thymocyte progenitors (Supplementary Fig. 5d), indicating it may contribute to myelopoiesis. *Dach1* has been implicated in transcriptional repression through association with histone deacetylases and its *Drosophila* homologue is known to play a role in gonadal, limb and ocular development[22]. Thus, *Jdp2* and *Dach1* may feedback on the epigenome to control expression of tissue-specific genes, but their role in haematopoiesis remains uncharacterized.

Our analyses also revealed a set of genes that were progressively hypermethylated and transcriptionally silenced as differentiation progressed towards both myeloid and lymphoid fates, indicating a role in maintenance of a multipotent state. *Meis1*, *2900052L18Rik*, *Hlf*, *Hoxa9* and *Prdm16* are all such candidates (Fig. 3c and Supplementary Fig. 6). *Meis1* is known to be required for haematopoiesis and megakaryocyte lineage development[23] and may function cooperatively with *Hoxa9* to regulate haematopoiesis[24]. Furthermore, both *Hlf and Prdm16* have been implicated in haematopoiesis[25,26].

Lastly, epigenetic chromatin modifiers, including *Hdac7* and *Dnmt3b*, were also differentially methylated during haematopoietic differentiation, indicating feed-forward mechanisms that could expand and lock in epigenetic programming necessary for cell fate commitment (Fig. 3d and Supplementary Fig. 7). *Hdac7*, which encodes a histone deacetylase and represses transcription, was demethylated and upregulated in DN1–3 thymocytes (Fig. 3d). Since *Hdac7* is highly expressed in DN3 cells, which can no longer be reprogrammed towards a myeloid fate by ectopic IL-2R signalling[19], it may actively repress genes responsible for maintaining myeloid lineage potential[19]. In contrast, *Dnmt3b*, a

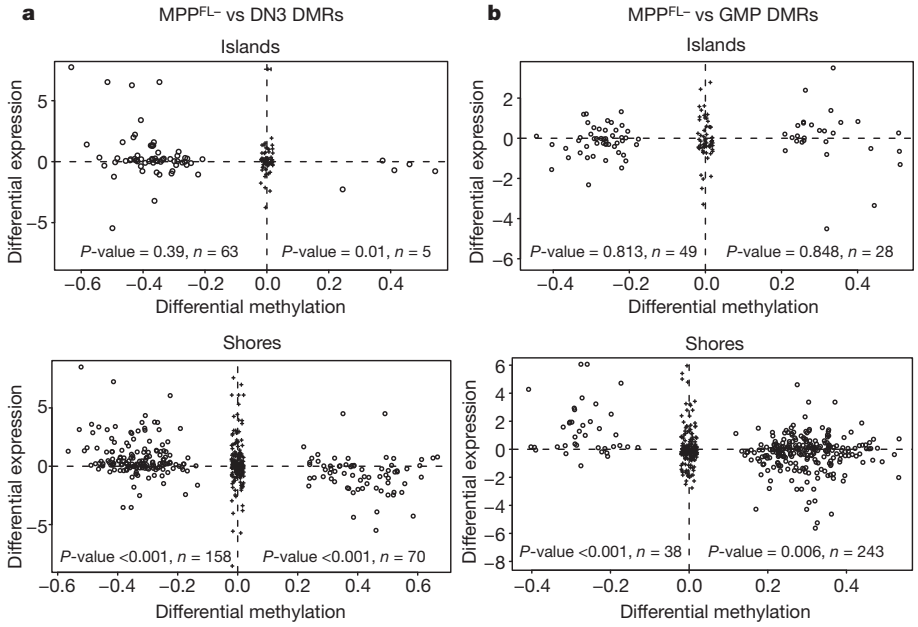

Figure 2 | Gene expression correlates strongly with DMRs at shores. DMRs within 2 kb of gene TSSs (black circles) were divided into two groups: Island (inside, cover, or overlap more than 50% of a CpG island), and Shores (up to 2 kb away from a CpG island). After robust multi-array average preprocessing, the $\log_2$ ratios of the gene expression differences (from left to right) were plotted against Δp (left group minus right group). Black pluses represent random DMR-gene pairs more than 2 kb apart. Wilcoxon rank-sum tests were performed to test the null hypothesis. **a**, MPP^{FL-} versus DN3 DMRs. **b**, MPP^{FL-} versus GMP DMRs.

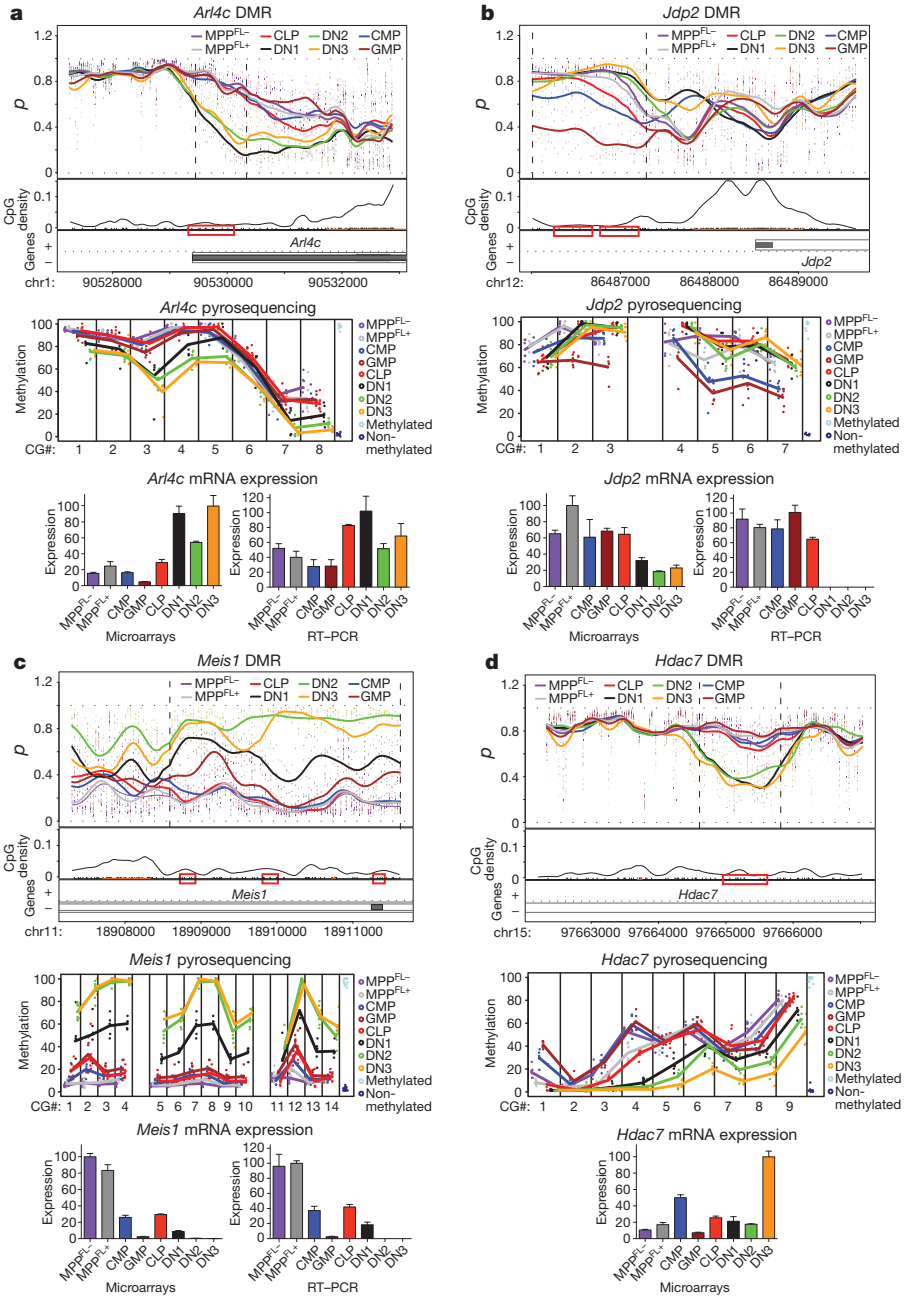

Figure 3 | CHARM identified genes with previously unknown functions in lymphoid/myeloid lineage commitment and pluripotency maintenance. **a** and **b**, Examples of DMRs with methylation changes in lymphoid/myeloid progenitors. **a**, the DMR in *Arl4c*. **b**, the DMR in *Jdp2*. **c**, the DMR in *Meis1*. **d**, the DMR in *Hdac7*. The CHARM plots, pyrosequencing, Affymetrix GeneChip, and RT–PCR data are organized and displayed as in Fig. 1b.

methyltransferase responsible for *de novo* CpG methylation, is hypermethylated and downregulated progressively in CMPs and GMPs (Supplementary Fig. 7). *Dnmt3a* and *Dnmt3b* were shown to be essential for HSC self-renewal, but their roles in lineage commitment remain inconclusive[27]. Downregulation of *Dnmt3b* in myeloid-committed cells could prevent new DNA methylation, helping to maintain the observed hypomethylated state associated with myelopoiesis. In addition, the upregulation of *Dnmt3b* in DN1 independent of DNA methylation changes might explain the marked acquisition of DNA methylation from CLP to DN1 (Table 1).

In summary, these data provide a comprehensive map of the methylome during myeloid and lymphoid commitment from haematopoietic progenitors. To facilitate the general accessibility of the methylome for these haematopoietic progenitors, we also provide here a novel web platform with which the methylation status of any genomic locus of interest can be easily queried to output methylation plots. In addition to identifying candidate genes for further investigation, the data indicate several important themes for the epigenetics of lineage-specific differentiation. First, myelopoiesis and lymphopoiesis achieve markedly different methylation endpoints in differentiation, with lymphopoiesis depending much more heavily on the acquisition of DNA methylation marks, and myelopoiesis depending much more on their loss. Besides providing a mechanism for the proposed *DNMT1*-dependence of lymphopoiesis, these results may also explain the therapeutic specificity of DNA demethylating drug treatment of myelodysplasia, in which malignant cells arrested in early development may be induced to differentiate by DNA demethylation[28]. In addition, the results show a remarkable dynamic plasticity in methylation during lineage development. The changes are evocative of Waddington's illustrations of hills and valleys in the epigenetic landscape of development. We have recently proposed that development depends on dynamic stochastic variation in the epigenetic landscape in a given genetic environment[29], and the maturation of undifferentiated progenitors to progressively more differentiated states could restrict that variation. Support for this idea is provided in an accompanying manuscript in this journal examining the epigenetic memory in iPS cells derived from fibroblasts and blood[30]. In that paper, lymphocyte-derived iPS cells cluster with CLP

but not the myeloid lineage using DNA methylation differences we identified, suggesting the existence of lymphocyte memory in these iPS cells consistent with the DNA methylation profiles described in this paper.

METHODS SUMMARY

Flow cytometry. Bone marrow cells and thymocytes were stained with monoclonal antibodies, then analysed and sorted using a FACSAria. Antibody details are provided in Methods.

CHARM DNA methylation analysis. Genomic DNA was isolated from samples, fractionated, digested, purified, labelled and subjected to CHARM array analysis as described previously[7]. Details are provided in Methods.

Bisulphite pyrosequencing. Genomic DNA was isolated from cells, treated with bisulphite and amplified by PCR. DNA methylation was measured by quantitative pyrosequencing. Details are provided in Methods.

Affymetrix microarray expression analysis. Genome-wide gene expression analysis was performed using the Affymetrix GeneChip Mouse Genome 430 2.0 Array. Details are provided in Methods.

OP9:OP9DL1 stromal co-cultures. Double sorted progenitors (50) were grown in wells containing confluent 1:1 OP9:OP9DL1 stromal cells in the presence of cytokines. At day 6, the cells in each well were stained and analysed by flow cytometry. Details are provided in Methods.

Quantitative PCR. Cells were sorted into TRIzol, RNA was isolated and cDNA was synthesized. Real-time PCR was performed using SYBR Green reagents. Details are provided in Methods.

DNA methylation query website. DNA methylation in any region from the CHARM array can be plotted at http://charm.jhmi.edu/hsc. Details are provided in Methods.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 8 February; accepted 26 July 2010.
Published online 15 August 2010.

1. Bröske, A. M. *et al.* DNA methylation protects hematopoietic stem cell multipotency from myeloerythroid restriction. *Nature Genet.* **41,** 1207–1215 (2009).
2. Chao, M. P., Seita, J. & Weissman, I. L. Establishment of a normal hematopoietic and leukemia stem cell hierarchy. *Cold Spring Harb. Symp. Quant. Biol.* **73,** 439–449 (2008).
3. Trowbridge, J. J., Snow, J. W., Kim, J. & Orkin, S. H. DNA methyltransferase 1 is essential for and uniquely regulates hematopoietic stem and progenitor cells. *Cell Stem Cell* **5,** 442–449 (2009).
4. Gardiner-Garden, M. & Frommer, M. CpG islands in vertebrate genomes. *J. Mol. Biol.* **196,** 261–282 (1987).
5. Irizarry, R. A. *et al.* Comprehensive high-throughput arrays for relative methylation (CHARM). *Genome Res.* **18,** 780–790 (2008).
6. Doi, A. *et al.* Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. *Nature Genet.* **41,** 1350–1353 (2009).
7. Irizarry, R. A. *et al.* The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nature Genet.* **41,** 178–186 (2009).
8. Molina, T. J. *et al.* Profound block in thymocyte development in mice lacking $p56^{lck}$. *Nature* **357,** 161–164 (1992).
9. Klebanoff, S. J. Myeloperoxidase: friend and foe. *J. Leukoc. Biol.* **77,** 598–625 (2005).
10. Cacalano, G. *et al.* Neutrophil and B cell expansion in mice that lack the murine IL-8 receptor homolog. *Science* **265,** 682–684 (1994).
11. Gupta, S. K., Gupta, M., Hoffman, B. & Liebermann, D. A. Hematopoietic cells from gadd45a-deficient and gadd45b-deficient mice exhibit impaired stress responses to acute stimulation with cytokines, myeloablation and inflammation. *Oncogene* **25,** 5537–5546 (2006).
12. Barreto, G. *et al.* Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation. *Nature* **445,** 671–675 (2007).
13. Rai, K. *et al.* DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and Gadd45. *Cell* **135,** 1201–1212 (2008).
14. Engel, N. *et al.* Conserved DNA methylation in $Gadd45a^{-/-}$ mice. *Epigenetics* **4,** 98–99 (2009).
15. Bell, J. J. & Bhandoola, A. The earliest thymic progenitors for T cells possess myeloid lineage potential. *Nature* **452,** 764–767 (2008).
16. Wada, H. *et al.* Adult T-cell progenitors retain myeloid potential. *Nature* **452,** 768–772 (2008).
17. Wei, S. M., Xie, C. G., Abe, Y. & Cai, J. T. ADP-ribosylation factor like 7 (ARL7) interacts with α-tubulin and modulates intracellular vesicular transport. *Biochem. Biophys. Res. Commun.* **384,** 352–356 (2009).
18. Chadwick, K., Shojaei, F., Gallacher, L. & Bhatia, M. Smad7 alters cell fate decisions of human hematopoietic repopulating cells. *Blood* **105,** 1905–1915 (2005).
19. King, A. G., Kondo, M., Scherer, D. C. & Weissman, I. L. Lineage infidelity in myeloid cells with TCR gene rearrangement: a latent developmental potential of proT cells revealed by ectopic cytokine receptor signaling. *Proc. Natl Acad. Sci. USA* **99,** 4508–4513 (2002).
20. Kondo, M. *et al.* Cell-fate conversion of lymphoid-committed progenitors by instructive actions of cytokines. *Nature* **407,** 383–386 (2000).
21. Jin, C. *et al.* Regulation of histone acetylation and nucleosome assembly by transcription factor JDP2. *Nature Struct. Mol. Biol.* **13,** 331–338 (2006).
22. Popov, V. M. *et al.* The *Dachshund* gene in development and hormone-responsive tumorigenesis. *Trends Endocrinol. Metab.* **21,** 41–49 (2010).
23. Pillay, L. M., Forrester, A. M., Erickson, T., Berman, J. N. & Waskiewicz, A. J. The Hox cofactors Meis1 and Pbx act upstream of *gata1* to regulate primitive hematopoiesis. *Dev. Biol.* **340,** 306–317 (2010).
24. Hu, Y. L., Fong, S., Ferrell, C., Largman, C. & Shen, W. F. HOXA9 modulates its oncogenic partner *Meis1* to influence normal hematopoiesis. *Mol. Cell. Biol.* **29,** 5181–5192 (2009).
25. Crable, S. C. & Anderson, K. P. A. PAR domain transcription factor is involved in the expression from a hematopoietic-specific promoter for the human *LMO2* gene. *Blood* **101,** 4757–4764 (2003).
26. Du, Y., Jenkins, N. A. & Copeland, N. G. Insertional mutagenesis identifies genes that promote the immortalization of primary bone marrow progenitor cells. *Blood* **106,** 3932–3939 (2005).
27. Tadokoro, Y., Ema, H., Okano, M., Li, E. & Nakauchi, H. De novo DNA methyltransferase is essential for self-renewal, but not for differentiation, in hematopoietic stem cells. *J. Exp. Med.* **204,** 715–722 (2007).
28. Claus, R., Almstedt, M. & Lubbert, M. Epigenetic treatment of hematopoietic malignancies: *in vivo* targets of demethylating agents. *Semin. Oncol.* **32,** 511–520 (2005).
29. Feinberg, A. P. & Irizarry, R. A. Evolution in Health and Medicine Sackler Colloquium: stochastic epigenetic variation as a driving force of development, evolutionary adaptation, and disease. *Proc. Natl Acad. Sci. USA* **107,** 1757–1764 (2005).
30. Kim, K. *et al.* Epigenetic memory in induced pluripotent stem cells. *Nature* advance online publication, doi:10.1038/nature09342 (19 July 2010).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank L. Jerabek for laboratory management, C. Richter and N. Teja for antibody production, A. Mosley, J. Dollaga and D. Escoto for animal care, E. Zuo and the Stanford PAN facility for microarray processing, and E. Briem and A. N. Allen for CHARM array processing. This investigation was supported by National Institutes of Health grants R37CA053458 and P50HG003233 (to A.P.F), R01AI047457 and R01AI047458 (to I.L.W.), and a grant from the Thomas and Stacey Siebel Foundation (to I.L.W). L.I.R.E. was supported by Special Fellow Career Development award from the Leukemia and Lymphoma Society; J.S. was supported by a fellowship from the California Institute for Regenerative Medicine (T1-00001); D.J.R. was supported by National Institutes of Health grant R00AGO29760; M.A.I. was supported by National Institutes of Health grant CA09151 and a fellowship from the California Institute for Regenerative Medicine (T1-00001); T.S. was supported by a fellowship from the National Institutes of Health (F32AI058521).

Author Contributions H.J. performed genome-scale and gene-specific DNA methylation analysis; L.I.R.E. and J.S. performed cell-sorting, generated microarray datasets, and performed gene expression analysis; P.L. assisted L.I.R.E. and H.L. assisted H.J.; A.D. and H.J. performed statistical analysis with P.M., M.J.A. and R.A.I; D.J.R., M.A.I., T.S., H.K. and L.H. generated microarray datasets; A.P.F. and I.L.W. designed the experiment, and A.P.F., H.J., L.I.R.E. and J.S. wrote the paper with the assistance of K.K. and G.Q.D.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/nature. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to A.P.F. (afeinberg@jhu.edu).

METHODS

Flow cytometry. Bone marrow cells and thymocytes were stained with monoclonal antibodies, then analysed and sorted on a FACSAria (Beckton Dickinson). The following monoclonal antibodies were purified and conjugated using hybridomas maintained in the Weissman laboratory: anti-CD8 (53.6.7) conjugated to Alexa Fluor 488, anti-CD4 (GK1.5) conjugated to Alexa Fluor 647, anti-CD44 (IM781) conjugated to Alexa Fluor 680, anti-CD25 (PC.61) conjugated to Pacific Orange, anti-FcγRII/III (2.4G2) conjugated to Pacific Orange, anti-Sca-1 (E13-161-7) conjugated to Pacific Blue, and anti-Ly6D (49H4.3, courtesy of Herzenberg laboratory) conjugated to Pacific Orange. The following antibodies were purchased from eBioscience: anti-CD34 (RAM34) conjugated to FITC (fluorescein isothiocyanate); anti-C135/Flk2 (A2F10) conjugated to phycoerythrin (PE); anti-CD127/Il7ra (A7R34) conjugated to PE-Cy5; anti-CD4 (GK1.5), -CD8 (53-6.7), -B220 (RA3-6B2), -Ter119 (TER119), -Mac-1 (M1/70), and anti-Gr-1 (RB6-8C5) conjugated to PE-Cy7; anti-CD27 (LG.7F9) conjugated to allophycocyanin (APC), anti-c-Kit (2B8) conjugated to APC-Alexa Fluor 750.

CHARM DNA methylation analysis. Genomic DNA from each sample was purified using the MasterPure DNA purification kit (Epicentre) as recommended by the manufacturer. Genomic DNA (1.5–2 μg) was fractionated, digested with McrBC, gel-purified, labelled and hybridized to a CHARM microarray as described. CHARM microarrays (CHARM 1.0) are prepared as described previously using custom-designed Nimblegen HD2 microarrays[5,7]. For the new CHARM arrays used in this study (CHARM 1.1), ~11% of probes with lowest CpG density on CHARM 1.0 were substituted with probes in promoters that were not previously covered. For each probe, the average methylation values across the same cell type were computed and converted to the percentage of methylation (p). p was used to find regions of differential methylation (Δp) for each pairwise cell type comparison. The absolute area of each region was calculated by multiplying the number of probes by mean Δp and DMRs were ranked based on this absolute area. In the CHARM plots, the upper panel shows the extent of methylation across a region of the genome. The top half of the panel is a plot of the percentage of CpG methylation versus genomic location, where the curve represents averaged smoothed p values from each cell population indicated (four replicates of MPP^{FL-} and MPP^{FL+}, and three replicates of the remaining cell populations; 40,000–100,000 cells of each population were sorted for each replicate). Two vertical dotted lines mark the range of the DMR identified. The lower half of the panel illustrates the location of CpG dinucleotides (black tick marks), CpG density (curve), location of CpG islands (orange line) and the gene annotation. + or – on the left side of the bottom panel indicates the orientation of genes and grey boxes represent exons with location indicated. Detailed information of DMRs identified in this study is listed in Supplementary Table 1. CHARM microarray data are deposited at the Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo) under accession number GSE23110.

Bisulphite pyrosequencing. Genomic DNA from each sample (200 ng; 6 samples for each progenitor) was treated with bisulphite using an EZ DNA methylation-Gold Kit (ZYMO research) according to the manufacturer's specifications. The bisulphite-treated genomic DNA was amplified by PCR using unbiased nested primers and DNA methylation was measured by quantitative pyrosequencing using a PSQ HS96 (Biotage). The DNA methylation percentage at each CpG site was determined using the Q-CpG methylation software (Biotage). SssI-treated mouse genomic DNA was used as 100% methylation control and mouse genomic DNA amplified by GenomePlex Complete Whole Genome Amplification (WGA) Kit (Sigma) was used as the non-methylated DNA control. Primer sequences used for the bisulphite pyrosequencing reactions are shown in Supplementary Table 3, as well as the chromosomal coordinates in the University of California at Santa Cruz March 2006 mouse genome assembly for each CpG site measured. The annealing temperature used for all PCR reactions was between 50 and 55 °C.

GO annotation. We analysed GO annotation using NIA Mouse Gene Index (http://lgsun.grc.nia.nih.gov/geneindex/mm9/upload.html). Genes identified from our analysis were compared to genes on arrays to calculate the enrichment ratio and significantly enriched gene ontology functional categories (FDR <0.05) are included in Supplementary Table 2.

Affymetrix microarray expression analysis. Genome-wide gene expression analysis was performed using Affymetrix GeneChip Mouse Genome 430 2.0 Array. For each sample, 1 μg of high-quality total RNA was amplified, labelled and hybridized onto the microarray according to Affymetrix's specifications, and data were normalized by GC robust multi-array average method and analysed on R/Bioconductor (GEO accession number GSE20244).

OP9:OP9DL1 stromal co-cultures. OP9 and OP9DL1 cells (3,000 of each) were plated in each well of 96 well plates in MEMa + 10% FBS. The next day, 50 double-sorted progenitors were plated per well in the presence of 5 ng ml^{-1} IL-7 and Flt3L, and 10 ng ml^{-1} IL-3, IL-6, M-CSF, GM-CSF and G-CSF (PeproTech). 5-aza-2′-deoxycytide (50 nM, Sigma) or vehicle (50% acetic acid) was added to the wells as indicated. At day 3, half of the media plus cytokines and drugs was replaced. At day 6, progeny from each well were stained and analysed by flow cytometry to identify lymphoid versus myeloid progeny.

Quantitative PCR. Cells were sorted into TRIzol (Invitrogen), and RNA was isolated according to manufacturer's instruction. cDNA was synthesized using the SuperScript III kit (Invitrogen) using random hexamers. Amplifications were performed using SYBR Green PCR core reagents (Applied Biosystems), and transcript levels were quantified using an ABI 7900 Sequence Detection Systems (Applied Biosystems). Mean C_t value of triplicate reaction was normalized against mean C_t value of beta-actin. Amplification efficiency of each primer pair was validated prior application using cDNA libraries of mouse ES cells, whole BM cells, and whole spleen cells. Primer sequences are as follows: *β-actin*, 5′-GTCTGAGGCCTCCCTTTTT-3′ and 5′-GGGAGACCAAAGCCTTCATA-3′; *Lck*, 5′-TGGAGAACATTGACGTGTGTG-3′ and 5′-ATCCCTCATAGGTG ACCAGTG-3′; *Mpo*, 5′-CCACGGAGCTCCTGTTTTAC-3′ and 5′-CAGCTT CCTCTTCAGCAGGT-3′; *Gcnt2*, 5′-TGCTCATCTTTCATCGACGGA-3′ and 5′-AGTGGCTTTGGGTCACATATTC-3′; *Arl4c*, 5′-AGTCTCTGCACATCGT TATGC-3′ and 5′-GGTGTTGAAGCCGATAGTGGG-3′; *Dach1*, 5′-CCTGGG AAACCCGTGTACTC-3′ and 5′-AGATCCACCATTTTGCACTCATT-3′; *Jdp2*, 5′-AGCTGAAATACGCTGACATCC-3′ and 5′-CTCACTCTTCACGGGTTGGG-3′; *Meis1*, 5′-CATGATAGACCAGTCCAACCGA-3′ and 5′-ATTGGCTGTCC ATCAGGGTTA-3′; *Prdm16*, 5′-TGACGGATACAGAGGTGTCAT-3′ and 5′-ACGCTACACGGATGTACTTGA-3′; *Dnmt3b*, 5′-GTTAATGGGAACTTCAGT GACCA-3′ and 5′-CTGCGTGTAATTCAGAAGGCT-3′; *Hdac7*, 5′-TTCCCT ACAGAACTCTTGAGCC-3′ and 5′-GGGGCACTCTCCTTCCTGA-3′.

DNA methylation query website. We have created a website at http://charm.jhmi.edu/hsc that allows plotting of DNA methylation in any region from the CHARM array. Regions of interest are uploaded as a tab- or comma-separated file. Top 50 DMRs plots from the complete sets are listed.