LETTERS

Polycomb complexes repress developmental regulators in murine embryonic stem cells

Laurie A. Boyer[1]*, Kathrin Plath[1]*†, Julia Zeitlinger[1], Tobias Brambrink[1], Lea A. Medeiros[1,2], Tong Ihn Lee[1], Stuart S. Levine[1], Marius Wernig[1], Adriana Tajonar[2], Mridula K. Ray[2], George W. Bell[1], Arie P. Otte[3], Miguel Vidal[4], David K. Gifford[5], Richard A. Young[1,2] & Rudolf Jaenisch[1,2]

The mechanisms by which embryonic stem (ES) cells self-renew while maintaining the ability to differentiate into virtually all adult cell types are not well understood. Polycomb group (PcG) proteins are transcriptional repressors that help to maintain cellular identity during metazoan development by epigenetic modification of chromatin structure[1]. PcG proteins have essential roles in early embryonic development[2–6] and have been implicated in ES cell pluripotency[2], but few of their target genes are known in mammals. Here we show that PcG proteins directly repress a large cohort of developmental regulators in murine ES cells, the expression of which would otherwise promote differentiation. Using genome-wide location analysis in murine ES cells, we found that the Polycomb repressive complexes PRC1 and PRC2 co-occupied 512 genes, many of which encode transcription factors with important roles in development. All of the co-occupied genes contained modified nucleosomes (trimethylated Lys 27 on histone H3). Consistent with a causal role in gene silencing in ES cells, PcG target genes were de-repressed in cells deficient for the PRC2 component Eed, and were preferentially activated on induction of differentiation. Our results indicate that dynamic repression of developmental pathways by Polycomb complexes may be required for maintaining ES cell pluripotency and plasticity during embryonic development.

Biochemical and genetic evidence indicates that PcG proteins function in two distinct complexes, PRC1 and PRC2, the core components of which are conserved from fruitfly to human and are essential for PcG activity both *in vitro* and *in vivo*[1]. To gain insights into the role of PcG proteins in ES cells, we identified the genes occupied by PcG proteins in murine ES cells by performing genome-wide location analysis using antibodies against core components of PRC1 (Phc1 and Rnf2) and PRC2 (Suz12 and Eed) (Fig. 1 and Supplementary Fig. S1).

The genomic DNA associated with PcG proteins and total (control) DNA were combined and hybridized to microarrays that contained 60-mer oligonucleotide probes covering the region from −8 kb to +2 kb relative to the transcription start sites for 15,742 annotated mouse genes (see Supplementary Information). Genomic sites occupied by PcG proteins (Supplementary Tables S1–S4) were identified as peaks of chromatin immunoprecipitation (ChIP)-enriched DNA using a previously validated algorithm[7] (Fig. 1a and Supplementary Information). Notably, the vast majority (~90%, $P < 10^{-159}$) of bound sequences were detected within 1 kb of a transcription start site (Supplementary Fig. S2). We focused further analysis on this set of bound regions, as the binding of a regulator in close proximity to a transcription start site is probably associated with the regulation of that gene. Our analysis revealed that PRC1 and PRC2 components occupied an overlapping set of target genes, and identified with high confidence 512 genes bound by all four PcG proteins in ES cells (Fig. 1b, Supplementary Tables S5 and S6, and Supplementary Figs S3–S5). These data show that PRC1 and PRC2 co-occupy the promoter regions of a large set of genes in ES cells.

Because trimethylation of Lys 27 on histone H3 (H3K27me3) is thought to be a marker for repressive chromatin and is associated with PRC2 activity[8–12], we investigated whether PcG occupancy correlated with this histone modification across the genome. We found that H3K27me3 was enriched at all 512 genes occupied by components of both PRC1 and PRC2 complexes (Supplementary Tables S5 and S7). Analysis of the distribution of H3K27me3 at this set of genes revealed that, similar to PRC1 and PRC2 components, H3K27me3 was also associated with probes close to the transcription start site of PcG target genes (Fig. 1c). These data show that PRC1 and PRC2 occupancy is associated with modified nucleosomes, consistent with a repressed transcriptional state.

To gain insights into the biological role of PcG proteins in ES cell pluripotency, we determined which gene ontology (GO) terms were over-represented in the set of genes associated with both Polycomb complexes and H3K27me3 (Fig. 1d and Supplementary Table S8). This analysis revealed an extremely significant enrichment for genes connected to transcription and development hierarchies, including organogenesis, morphogenesis, pattern specification, neurogenesis, cell differentiation, embryonic development and cell-fate commitment, among others (Supplementary Fig. S6). Further analysis showed that the target genes within the development and transcription functional groups overlap significantly ($P < 10^{-96}$; Supplementary Table S8), indicating that most of the PcG target genes are transcription factors with important roles in a variety of developmental processes.

In addition to *Hox* genes, which are the classic PcG target genes, PcG proteins bound to over 100 genes encoding homeodomain-containing transcriptional regulators in murine ES cells, including members of the *Dlx*, *Irx*, *Lhx*, *Pou*, *Pax* and *Six* gene families (Supplementary Table S9 and Supplementary Fig. S7). Homeodomain-containing transcription factors are evolutionarily conserved regulators that specify cell fate during embryonic development through transcriptional control of other developmental regulators[13]. Polycomb complexes also occupied promoters of members of the

[1]Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, Massachusetts 02142, USA. [2]Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. [3]Swammerdam Institute for Life Sciences, University of Amsterdam, 1098 SM Amsterdam, The Netherlands. [4]Developmental and Cell Biology Centro de Investigaciones Biológicas (CSIC), 28040 Madrid, Spain. [5]Computer Science and Artificial Intelligence Laboratories, 32 Vassar Street, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. †Present address: Biological Chemistry Department, David Geffen School of Medicine at UCLA, 36-133 CHS, PO Box 951737, Los Angeles, California 90095-1737, USA.
*These authors contributed equally to this work.

Fox, *Sox*, *Gata* and *Tbx* transcription factor families, which also have essential roles in development and disease[14–17]. These data demonstrate that in ES cells, Polycomb complexes target transcription factors that have key roles in a variety of developmental processes.

To establish a direct functional link between PcG protein binding and repression of target genes, we used real-time polymerase chain reaction (PCR) to compare the expression levels of PcG target genes in wild-type ES cells and ES cells lacking the PRC2 component Eed (Fig. 2). We examined *Eed*-deficient cells ($Eed^{-/-}$) because lack of Eed leads to disruption of PRC2 and H3K27 methylation[18] (Supplementary Fig. S8). Transcript levels for the PcG target genes *Gata3*, *Gata4* and *Gata6* increased significantly in *Eed* mutant cells compared with wild-type ES cells, whereas transcript levels for *Gata1*, which was unbound in the ChIP assay, were unchanged (Fig. 2a). Immunostaining showed that Gata4 protein was not detected in wild-type ES cells, but it was abundant in *Eed* mutant cells, consistent with our PCR results (Fig. 2b). Loss of Eed resulted in at least a twofold increase in transcript abundance for 87% (71/82) of the PcG target genes tested, including members of the *Hox*, *Pax*, *Lhx*, *Sox*, *Fox* and *Fgf* gene families, whereas the expression of control genes remained largely unaffected (Fig. 2c and Supplementary Table S10). These data indicate that genes bound by PcG proteins in ES cells are direct targets of PRC2-mediated repression.

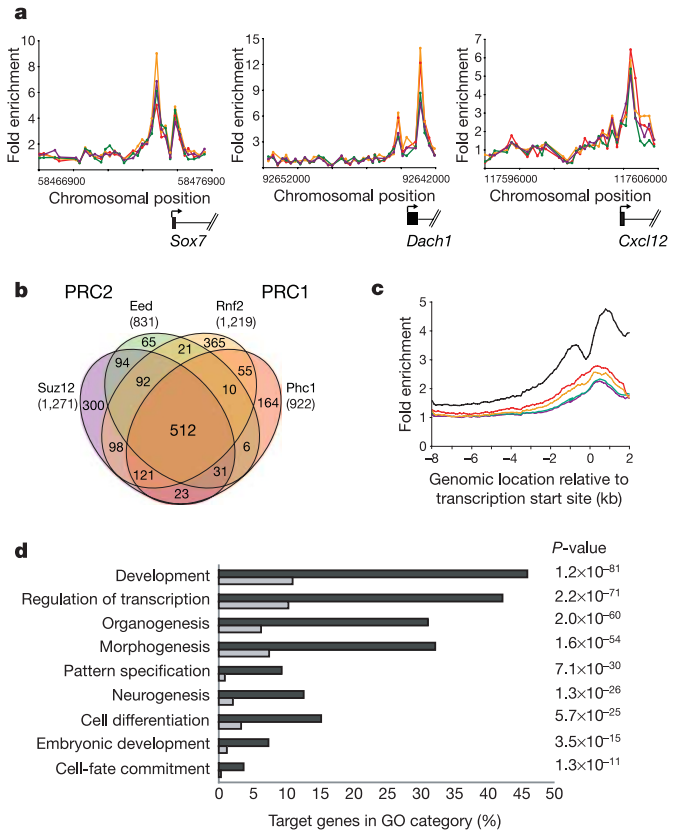

Figure 1 | PRC1 and PRC2 colocalize at genes encoding developmental regulators. **a**, Unprocessed enrichment ratios for all probes within a genomic region (ChIP-enriched versus total genomic DNA) for Suz12 (green), Eed (purple), Rnf2 (orange) and Phc1 (red). Example genes are drawn to scale below plots, and the start and direction of transcription are noted by arrows. **b**, Venn diagram showing the overlap among genes bound by PcG proteins within 1 kb of a transcription start site. Numbers in parentheses represent the total number of genes bound by the respective PcG protein. **c**, Average unprocessed enrichment ratios for each oligonucleotide probe within the −8 kb to +2 kb genomic region for all co-bound RefSeq genes relative to the transcription start site. Colours represent PcG proteins as in **a**, and the black line shows H3K27me3 enrichment. **d**, Gene Ontology analysis of PcG target genes. Black bars represent the observed percentage of PcG target genes in a particular GO category. Grey bars represent the percentage expected on the basis of all GO-annotated genes on the oligonucleotide array. The significance (*P*-value) of this enrichment is based on a hypergeometric distribution.

To investigate whether PRC1 promoter occupancy is dependent on PRC2 at target genes in ES cells, we performed ChIP experiments and site-specific PCR for a subset of genes in wild-type and *Eed* mutant cells (Fig. 2d). Previous work has shown that loss of PRC2 components results in disruption of PRC1 binding at *Hox* genes in somatic cells[8,19]. In *Eed* mutant cells, we observed a significant reduction in binding at PcG target genes for two PRC1 components, Rnf2 and Cbx2, as well as the PRC2 subunit Suz12, even though the levels of these proteins had not dramatically changed in the absence of Eed (Supplementary Fig. S8). Our results suggest that an intact PRC2 complex and/or its associated histone methyl mark (H3K27me3) are required for the association of PRC1 at target genes in ES cells.

The association of PcG components with repressive chromatin structure and developmental regulators suggests that genes targeted by Polycomb complexes are globally repressed in ES cells but must be activated during differentiation. To test this, we analysed the expression of PcG target genes during ES cell differentiation (Fig. 3), and found that most PcG target gene transcripts were lower in abundance in ES cells relative to differentiated cells (Fig. 3a and Supplementary Table S11). Specifically, 93% of PcG target transcripts were upregulated during ES cell differentiation, compared with only 59% of all transcripts represented on the array (Fig. 3b). Further analysis using the Gene Set Enrichment Analysis (GSEA) tool (see Supplementary Information), which tests for non-random distribution of a subset of genes within a ranked expression data set, confirmed our hypothesis that PcG target genes are preferentially upregulated during ES cell differentiation (Fig. 3c). This suggests that PcG proteins have specialized roles in silencing genes in ES cells, the activation of which correlates with differentiation and loss of pluripotency.

These findings suggest that PRC2 binding and H3K27 methylation should be decreased at target genes that are activated during lineage commitment. To test this directly, we analysed expression levels and H3K27me3 status for a subset of genes during directed differentiation of ES cells to neural precursor (NP) cells using real-time PCR upon reverse transcription and ChIP with site-specific real-time PCR, respectively, as we noted that many PcG target genes have known roles in neural development (Table 1 and Supplementary Fig. S9). We found that H3K27me3 was depleted at least 100-fold at neural-specific genes (for example, *Olig1*, *Olig2*, *Olig3* and *Nes*), and that expression of these genes was significantly increased in NP cells relative to ES cells (Table 1, group A). The loss of H3K27me3 was concomitant with RNA polymerase II occupancy and an increase in H3K4me3, the histone methylation mark associated with active transcription (data not shown). In contrast, PcG target genes that were not expressed in either ES or NP cells retained high levels of H3K27me3 in their promoter regions and were not associated with RNA polymerase II (Table 1, group B and data not shown). We also found that the pluripotency genes *Oct4* (also known as *Pou5f1*) and *Nanog* were associated with low levels of H3K27me3 in ES cells, consistent with their high expression (Table 1, group C). Although these genes become silenced on ES cell differentiation, H3K27me3 levels were similar in ES and NP cells, consistent with previous studies suggesting that *Oct4* is repressed by other epigenetic silencing mechanisms[20–22]. Together, these data demonstrate that genes repressed by PcG proteins in ES cells maintain the potential to become activated on lineage commitment, revealing a dynamic role for PcG complexes and their chromatin modifications during differentiation.

This dynamic role for PcG-mediated gene repression seems to be

different from other epigenetic silencing mechanisms. For example, DNA methylation is thought to be a stable silencing mechanism that is required for irreversibly locking-in the repressed transcriptional state. It has recently been shown that PRC2 recruits DNA methylation machinery to silence target genes in somatic cells[23]. Our finding that PcG proteins repress genes that are poised for activation on differentiation suggests that PcG-mediated repression is not linked to DNA methylation in ES cells. Consistent with this idea, DNA methylation does not have an essential role in maintaining ES cells in an undifferentiated state, but is essential for the survival of somatic cells[24,25]. In contrast, ES cells cannot be derived from blastocysts deficient for the PRC2 component Ezh2 (ref. 2), and ES cells lacking Eed have a strong propensity to differentiate (Supplementary Fig. S8), indicating that PcG-mediated gene repression in ES cells and during early development may be important for developmental plasticity.

PcG proteins are recruited to target sites through interaction with site-specific DNA-binding proteins in *Drosophila*[1]. It is currently unknown how PcG proteins are specifically recruited to target genes in mammals. Many of the developmental PcG target genes identified in this study in murine ES cells have recently been shown to be bound by three key pluripotency transcription factors—*OCT4* (*POU5F1*), *SOX2* and *NANOG*—in human ES cells[7] (see Supplementary Information). Moreover, it has recently been shown that PcG proteins target a similar set of developmental regulators in human ES cells[26]. This raises the possibility that PcG proteins, at least at a subset of genes, act as transcriptional repressors by collaborating with a specific set of transcription factors. Improved understanding of the mechanisms by which these transcription factors and PcG proteins contribute to maintenance of the pluripotent state and repression of developmental genes remain important issues for future investigation.

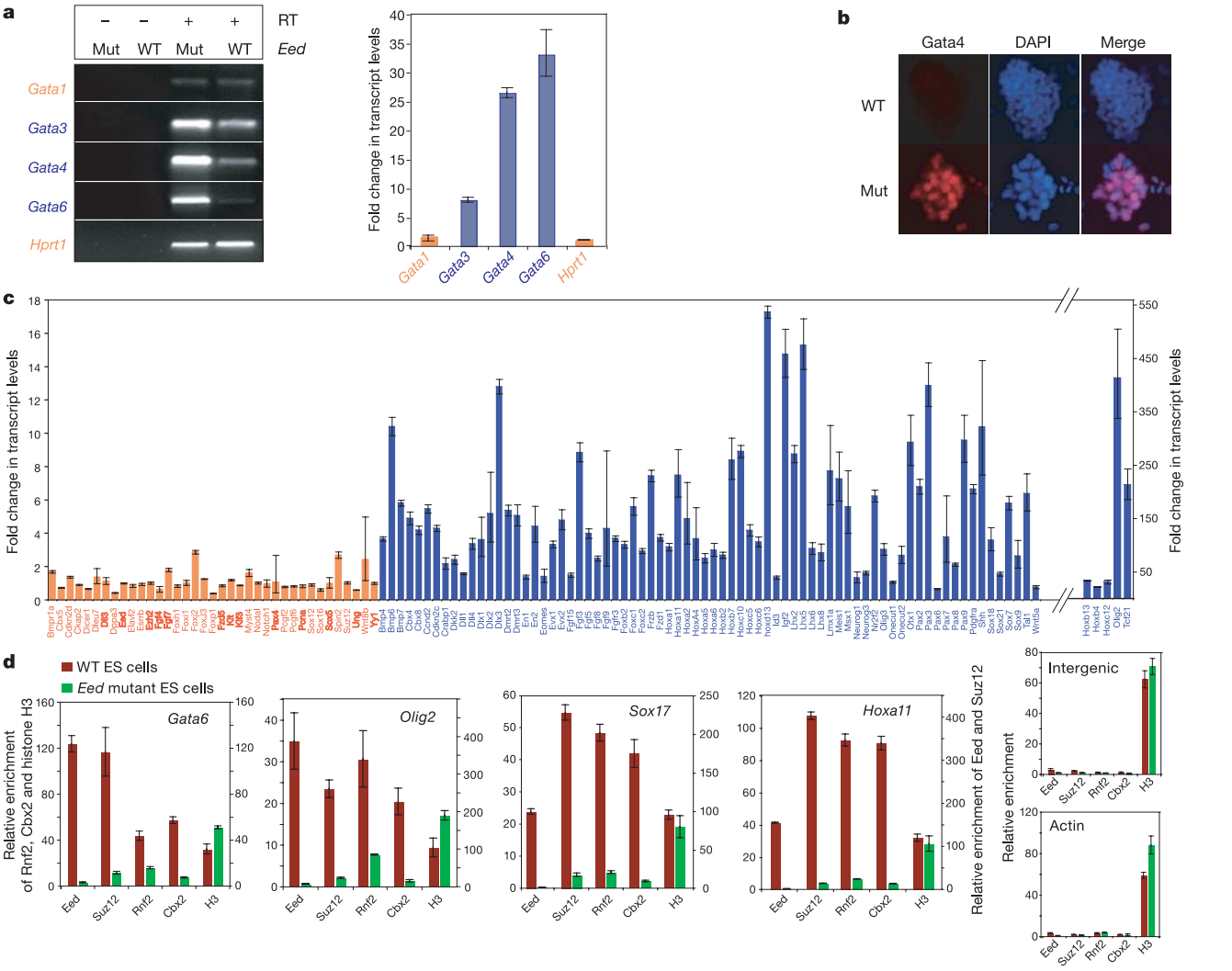

Figure 2 | De-repression of PcG target genes and loss of PRC1 binding in the absence of the PRC2 component Eed. **a**, Real-time PCR analysis for *Gata* and *Hprt1* gene transcripts in wild-type (WT) and *Eed* mutant (Mut) ES cells (left panel). In control reactions, reverse transcriptase (RT) was omitted (−). Orange and blue depict unbound and bound genes, respectively, as determined by genome-scale location analysis. Transcript levels were quantified by real-time PCR, normalized to a *Gapdh* control and depicted as a fold change between *Eed* mutant and wild-type ES cells. Error bars are based on the standard deviation derived from triplicate PCR reactions. **b**, Immunostaining for Gata4 in wild-type and *Eed* mutant ES cells. Nuclei are stained with DAPI. **c**, Quantification of transcript levels in *Eed* mutant ES cells relative to wild-type ES cells, as described in **a**. Genes with the highest relative expression changes are shown at the right of the graph. **d**, Association of PcG components with promoter regions of representative target genes, an intergenic control region, and the highly transcribed actin (*Actb*) gene was determined by ChIP and site-specific real-time PCR in wild-type (brown) and *Eed* mutant (green) ES cells. Histone H3 enrichment levels indicate that an equal amount of input material was used. Error bars represent standard deviations determined from three independent experiments.

Table 1 | Analysis of transcript levels and H3K27 methylation in ES cells and ES-cell-derived NP cells

Group*	Gene	Ratio of transcript levels between NP and ES cells†	Relative H3K27me3 enrichment‡				
			ES cells	U/E	NP cells	U/E	Ratio in NP/ES cells
A	*Olig1*§	232 (221–245)	186 ± 7	E	2.1 ± 0.2	U	0.01 ± 0.001
	Olig2§	98 (92–105)	518 ± 6	E	1.5 ± 0.4	U	0.003 ± 0.0008
	Olig3§	69 (66–72)	320 ± 26	E	1.4 ± 0.05	U	0.004 ± 0.0004
	Egfr§	109 (105–114)	107 ± 5	E	1.1 ± 0.2	U	0.01 ± 0.0015
	Nes§	248 (237–261)	190 ± 9	E	1.6 ± 1.0	U	0.009 ± 0.005
	Sox9§	719 (690–749)	1021 ± 24	E	1.0 ± 0.2	U	0.001 ± 0.0002
B	*Gata6*§	0.21 (0.16–0.26)	797 ± 84	E	127 ± 10	E	0.16 ± 0.02
	Hoxa11§	2.02 (1.91–2.14)	442 ± 5	E	58 ± 3	E	0.13 ± 0.01
	Lmx1a§	1.53 (1.21–1.92)	327 ± 79	E	26 ± 0.4	E	0.08 ± 0.02
	Sox17§	1.91 (1.37–2.68)	663 ± 19	E	66 ± 35	E	0.1 ± 0.05
C	*Oct4*	0.0008 (0.0007–0.0009)	2.1 ± 0.2	U	1.8 ± 0.2	U	0.88 ± 0.12
	Nanog	0.0016 (0.0015–0.016)	7.0 ± 0.5	U	2.6 ± 0.1	U	0.37 ± 0.03
D	*Sox2*	1.61 (1.57–1.65)	7.8 ± 0.6	U	1.8 ± 0.3	U	0.22 ± 0.04
Other	Intergenic, chr. 8	NA	0.7 ± 0.2	U	2.9 ± 0.8	U	4.25 ± 1.51
	Intergenic, chr. 6	NA	10.6 ± 1.9	U	0.8 ± 0.5	U	0.07 ± 0.05

Chr., chromosome; E, enriched; U, not enriched; NA, not applicable.
*Group A genes are repressed in ES cells and strongly expressed in NP cells. Group B genes are repressed in ES and NP cells. Group C genes are expressed at high levels in ES cells and repressed in NP cells. Group D genes are expressed at high levels in ES and NP cells.
†Determined by real-time PCR with reverse transcription. The error range (in parentheses) is determined based on the standard deviation derived from triplicate PCR reactions.
‡Determined by ChIP and site-specific real-time PCR. The U/E status of H3K27me3 enrichment at the regions analysed by site-specific PCR is based on the 'not enriched' threshold level determined by intergenic regions. Error ranges indicate s.d. of triplicate PCR reactions.
§Genes identified in our genome-wide location analysis as PcG and H3K27me3 targets in undifferentiated ES cells.

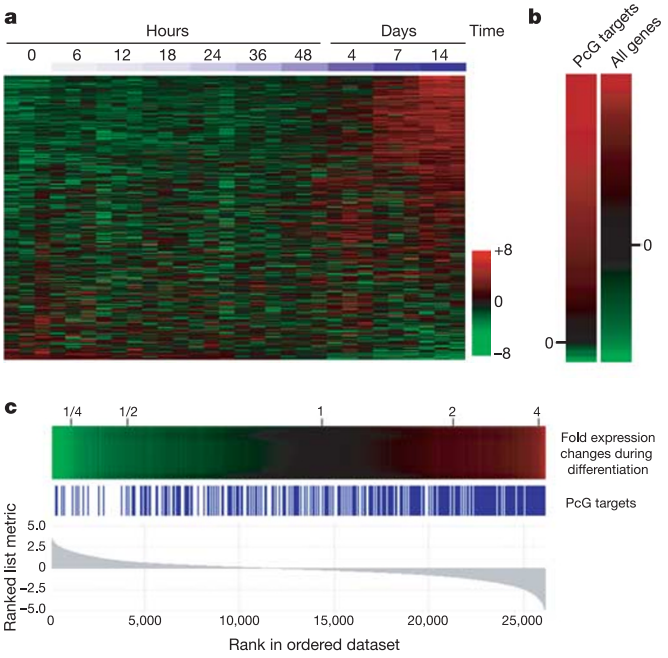

Figure 3 | PcG target genes are preferentially upregulated during ES cell differentiation. a, Triplicate expression profiling data sets were analysed at ten time points during the differentiation of ES cells into embryoid bodies. Log_2-transformed probe signal levels are shown relative to the average signal level of each probe set across all samples for PcG target genes. Genes were ordered by the ratio of average signal value in undifferentiated ES cells (0 h) over the average signal value of differentiated ES cells (14 days). Green and red represent lower-than-average and higher-than-average signal levels, respectively. **b**, The ratio of the average signal value for undifferentiated ES cells and ES cells differentiated for 14 days was determined for all probes representing PcG target genes analysed in **a**, and for all probes on the array. Shown is the distribution of log_2-transformed ratios for transcripts that are up- or down-regulated by more than 1.5-fold in both experimental groups. **c**, Triplicate sets of expression profiles were compared between undifferentiated ES cells and cells after 14 days of differentiation. All probe sets in the data set were assigned a score based on a signal-to-noise ratio algorithm (grey bars) and rank-ordered by this score. PcG target genes in the ordered data set are shown as blue lines, and the change in probe signal level is indicated for each probe using the colorimetric scale as in **a**.

METHODS

A detailed description of all materials and methods used can be found in Supplementary Information.

Growth of murine ES cells and derivation of NP cells. E14 (ola/129) ES cells were plated without irradiated murine embryonic fibroblasts (MEFs) and grown under typical ES cell conditions on gelatinized tissue-culture plates. The *Eed* mutant ES cell line (17Rn5-3354SB) obtained from T. Magnuson was grown on irradiated MEFs under standard conditions. V6.5 (C57BL/6-129) ES cells were differentiated along the neural lineage using standard protocols (see Supplementary Information).

Antibodies and chromatin immunoprecipitation assays. Detailed descriptions of antibodies, antibody specificity and ChIP methods used in this study are provided in Supplementary Information. Purified immuno-enriched and input genomic DNA was used for real-time sequence-specific PCR reactions.

Array design and data extraction. The design of the oligonucleotide-based promoter array set and data extraction methods are described in Supplementary Information. The microarrays used in this study were manufactured by Agilent Technologies (http://www.agilent.com).

Gene Ontology classification of bound genes. Gene Ontology analysis was performed using BiNGO (http://www.psb.ugent.be/cbd/papers/BiNGO/).

Comparing binding and expression data. For expression analysis of PcG target genes during ES cell differentiation, expression data from a time course study comparing V6.5 ES cells with several time points on induction of embryoid body differentiation were retrieved from the NCBI Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/) and are accessible through GEO Series accession number GSE3231.

RNA isolation, real-time PCR and analysis of transcript levels. To determine transcript levels in ES and NP cells, RNA was isolated, reverse-transcribed and subjected to real-time PCR using the SYBR Green PCR master mix and the 7000 ABI Detection System (ABI). Detailed information and all oligonucleotide sequences are provided in Supplementary Information.

Received 30 January; accepted 20 March 2006.
Published online 19 April 2006.

1. Ringrose, L. & Paro, R. Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. *Annu. Rev. Genet.* **38,** 413–443 (2004).
2. O'Carroll, D. *et al.* The *Polycomb*-group gene *Ezh2* is required for early mouse development. *Mol. Cell. Biol.* **21,** 4330–4336 (2001).
3. Pasini, D., Bracken, A. P., Jensen, M. R., Lazzerini Denchi, E. & Helin, K. Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. *EMBO J.* **23,** 4061–4071 (2004).
4. Shumacher, A., Faust, C. & Magnuson, T. Positional cloning of a global regulator of anterior–posterior patterning in mice. *Nature* **383,** 250–253 (1996).
5. Voncken, J. W. *et al. Rnf2* (*Ring1b*) deficiency causes gastrulation arrest and cell cycle inhibition. *Proc. Natl Acad. Sci. USA* **100,** 2468–2473 (2003).

6. Isono, K. *et al.* Mammalian polyhomeotic homologues Phc2 and Phc1 act in synergy to mediate polycomb repression of *Hox* genes. *Mol. Cell. Biol.* **25,** 6694–6706 (2005).
7. Boyer, L. A. *et al.* Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122,** 947–956 (2005).
8. Cao, R. *et al.* Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* **298,** 1039–1043 (2002).
9. Czermin, B. *et al. Drosophila* enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. *Cell* **111,** 185–196 (2002).
10. Kuzmichev, A., Nishioka, K., Erdjument-Bromage, H., Tempst, P. & Reinberg, D. Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. *Genes Dev.* **16,** 2893–2905 (2002).
11. Muller, J. *et al.* Histone methyltransferase activity of a *Drosophila* Polycomb group repressor complex. *Cell* **111,** 197–208 (2002).
12. Kirmizis, A. *et al.* Silencing of human polycomb target genes is associated with methylation of histone H3 Lys 27. *Genes Dev.* **18,** 1592–1605 (2004).
13. Hombria, J. C. & Lovegrove, B. Beyond homeosis—HOX function in morphogenesis and organogenesis. *Differentiation* **71,** 461–476 (2003).
14. Schepers, G. E., Teasdale, R. D. & Koopman, P. Twenty pairs of sox: extent, homology, and nomenclature of the mouse and human sox transcription factor gene families. *Dev. Cell* **3,** 167–170 (2002).
15. Lehmann, O. J., Sowden, J. C., Carlsson, P., Jordan, T. & Bhattacharya, S. S. Fox's in development and disease. *Trends Genet.* **19,** 339–344 (2003).
16. Showell, C., Binder, O. & Conlon, F. L. T-box genes in early embryogenesis. *Dev. Dyn.* **229,** 201–218 (2004).
17. Burch, J. B. Regulation of GATA gene expression during vertebrate development. *Semin. Cell Dev. Biol.* **16,** 71–81 (2005).
18. Montgomery, N. D. *et al.* The murine polycomb group protein Eed is required for global histone H3 lysine-27 methylation. *Curr. Biol.* **15,** 942–947 (2005).
19. Cao, R., Tsukada, Y. & Zhang, Y. Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. *Mol. Cell* **20,** 845–854 (2005).
20. Gidekel, S. & Bergman, Y. A unique developmental pattern of *Oct-3/4* DNA methylation is controlled by a *cis*-demodification element. *J. Biol. Chem.* **277,** 34521–34530 (2002).
21. Hattori, N. *et al.* Epigenetic control of mouse *Oct-4* gene expression in embryonic stem cells and trophoblast stem cells. *J. Biol. Chem.* **279,** 17063–17069 (2004).
22. Feldman, N. *et al.* G9a-mediated irreversible epigenetic inactivation of *Oct-3/4* during early embryogenesis. *Nature Cell Biol.* **8,** 188–194 (2006).
23. Vire, E. *et al.* The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* **439,** 871–874 (2006).
24. Meissner, A. *et al.* Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Res.* **33,** 5868–5877 (2005).
25. Li, E., Bestor, T. H. & Jaenisch, R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69,** 915–926 (1992).
26. Lee, T. *et al.* Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell* **125,** 1–13 (2006).

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

Acknowledgements We thank Biology and Research Computing (BaRC), especially B. Yuan, T. Dicesare and T. Volkert, at the Whitehead Institute's Center for Microarray Technology, as well as E. Herbolsheimer, for computational and technical support. We also thank members of the Gifford, Young and Jaenisch laboratories for discussions and critical review of the manuscript. We are grateful to T. Magnuson for the *Eed* mutant (17Rn5-3354SB) ES cell lines. L.A.B. was supported in part by an NIH NRSA fellowship. K.P. is a special fellow of the Leukemia and Lymphoma Society. M.W. is supported by a Human Frontiers Science Program Organization Fellowship. M.V. is supported by a grant from the Spanish Ministry of Science and Education. This work was supported in part by grants from the NIHGRI and NIGMS to R.A.Y. and from the NIH and NCI to R.J.

Author Information All microarray data from this study are available from ArrayExpress at the EBI (http://www.ebi.ac.uk/arrayexpress) under accession code E-WMIT-12. Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare competing financial interests: details accompany the paper at www.nature.com/nature. Correspondence and requests for materials should be addressed to R.J. (jaenisch@wi.mit.edu).