highly mutagenic base 8-hydroxyguanine (8-oxo-dG) [9]. The presence of 8-oxo-dG in DNA causes C:G → A:T and T:A → G:C transversions in DNA, since, unless repaired, 8-oxo-dG allows the misincorporation of cytosine and adenine nucleotides opposite the lesion during DNA replication [10].

Recently, several studies suggest the involvement of the deregulation of DNA repair pathways in hypoxia-induced mutagenesis [8,11,12]. Mihaylova et al. reported decreased expression of the DNA mismatch repair gene *Mlh1* under hypoxic conditions [11]. Koshiji et al. demonstrated that HIF-1 induces genetic instability by transcriptionally down-regulating the expression of MutSα which recognizes base mismatches [12]. Hypoxia may cause mutagenesis, at least in part, by hindering repair of ROS-induced DNA damage through down-regulation of DNA mismatch repair enzymes.

To date, several DNA polymerases that are clearly involved in translesion synthesis including Pol η (RAD30), Pol ι (RAD30B), Pol θ, Pol κ, and Rev1 which belong to the Y superfamily have been reported [13]. One of their most distinct features is a high error propensity during DNA synthesis. Among these Y family polymerases, Pol η, Pol ι, and Pol κ can efficiently bypass 8-oxo-dG [14–16]. Pol μ which belongs to the X superfamily has also been reported to possess efficient lesion bypass activities in response to several types of DNA damage including 8-oxo-dG [17].

We hypothesized that translesion DNA polymerases that can bypass 8-oxo-dG might be involved in the generation of mutations after hypoxia/reoxygenation. We report here that hypoxia enhances the expression of *pol ι* gene through HIF-1 interaction with the consensus HRE site in the intron 1 of the gene.

Materials and methods

Cells and cell culture. Human cervical carcinoma HeLa cells, hepatocarcinoma HepG2 cells, mammary carcinoma MCF-7 and MDA-MB-231 cells, lung adenocarcinoma A549 cells, fibrosarcoma HT1080 cells, colon carcinoma LS174T cells, and glioma U87MG cells were cultured at 37 °C in a humidified atmosphere with 21% O_2/5% CO_2 (normoxia) or 1% O_2/5% CO_2 (hypoxia).

Detection of ROS generation. ROS generation was detected with 2′,7′-dichlorofluorescin diacetate (DCFH-DA) (Molecular Probe, Inc., Eugene, OR) as described previously [18].

Immunostaining of 8-oxo-dG. HeLa cells on glass coverslips were cultured in 21% O_2, or 1% O_2 for 24 h, or 1% O_2 for 24 h followed by 21% O_2 for 30 min. Staining for 8-oxo-dG with anti-8-oxo-dG antibody 1F7 (Trevigen, Inc., Gaithersburg, MD) was performed according to the manufacturer's instructions with some modifications. Briefly, the cells fixed with 70% ethanol at −20 °C were treated with RNase (100 μg/ml) in 10 mM Tris–HCl, pH 7.5, and 1 mM EDTA, and 0.4 M NaCl for 1 h at 37 °C. DNA was denatured with 4 N HCl for 7 min at room temperature. After neutralization, the cells were incubated with 10% fetal bovine serum, and then incubated with anti-8-oxo-dG antibody at 4 °C overnight followed by TRITC-conjugated goat anti-mouse IgG. The nuclei were stained with DAPI (1 μg/ml).

SDS–PAGE and Western blotting. Total cell lysates were prepared by directly solubilizing cells in SDS sample buffer. Nuclear extracts were prepared by using Nuclear Extract Kit (Active Motif, Carlsbad, CA). Proteins were resolved by SDS–PAGE and transferred to nitrocellulose membrane. The membrane was incubated with mouse anti-HIF-1α (Novus Biologicals, Littleton, NO), goat anti-pol ι antibody (Santa Cruz Biotechnology, Inc., Santa Cruz, CA) or rabbit anti-E2F-1 antibody (Santa Cruz Biotechnology) followed by incubation with appropriate horseradish peroxidase-conjugated secondary antibody. Proteins were detected using ECL Western blotting detection reagents (Amersham Biosciences, Buckinghamshire, UK).

Semi-quantitative RT-PCR. One microgram of total RNA, which was extracted with guanidinium thiocyanate, was reverse transcribed into cDNA, and the resulting cDNA was used for amplification of target cDNAs using *rTaq* DNA polymerase (TOYOBO, Osaka). The sense and antisense oligonucleotide primers used for PCR were: 5′-GCTGTG CTGGAGTGGCTTCT-3′ and 5′-GCCAGAGCGTGAAGTAGTTG-3′ for *pol ι*, 5′-GCCATGCCAGGATTTATTGCTA-3′ and 5′-CTCCTTTG TTGGTGTTTCCT-3′ for *pol κ*, 5′-ACCCAGGCAACTACCCAAAAC-3′ and 5′-GGGCTCAGTTCCTGTACTTTG-3′ for *pol η*, 5′-AGCCTG-TACCTGTGGAGTGC-3′ and 5′-CCAGGCGGGTAGGGGACTCA-3′ for *pol μ*, 5′-ATGCCTCAACCGTGGACAAT-3′ and 5′-CTTGCT CTCGATGTGCTGC-3′ for *Mlh1*, 5′-GCAGAATCATCACGAA GTGG-3′ and 5′-GCATGGTGATGTTGGACTCC-3′ for *VEGF*, and 5′-TGACGGGGTGACCCACACTGTGCCCATCTA-3′ and 5′-CTAGAA GCATTTGCGGTGGACGATGGAGGG-3′ for *β-actin*, respectively. The PCR conditions were: 95 °C for 2 min, and then 30 cycles with 95 °C for 10 s, 59 °C for 10 s, 72 °C for 1 min for *pol ι*, *pol η*, and *VEGF*, 30 cycles with 95 °C for 5 s, 61 °C for 10 s, 72 °C for 2 min for *pol κ*, 30 cycles with 95 °C for 5 s, 65 °C for 10 s, 72 °C for 1.5 min for *pol μ*, 30 cycles with 95 °C for 5 s, 59 °C for 10 s, 72 °C for 1 min for *Mlh1* or 25 cycles with 95 °C for 5 s, 59 °C for 10 s, 72 °C for 1 min for *β-actin*, and 72 °C for 5 min.

Construction of plasmids. The plasmid pcDNA3.1/HIF-1α^{DN} expressing dominant-negative HIF-1α was prepared essentially as described previously [19]. The plasmid expressing constitutively active HIF-1α, pcDNA3.1/HIF-1α$^{P402A/P564A}$, was constructed by introducing mutations into pcDNA3.1/HIF-1α that change both proline-402 and proline-564 of HIF-1α to alanine by using QuikChange® Site Directed Mutagenesis Kit (Stratagene, La Jolla, CA). A luciferase reporter plasmid harboring the regulatory region (from −1346 to +418) of human *pol ι* gene, where +1 represents the transcription start site, was constructed as follows. First, a DNA fragment was amplified by PCR using genomic DNA, the 5′-primer carrying *Kpn*I site at the 5′ end, 5′-GGTACCCCTCCCTTCTGTCTG TGGA-3′, and the 3′ primer carrying *Sac*I site at the 5′ end, 5′-GAG CTCCTCGGCGTCTTCCTCGTCG-3′, and *Ex Taq*® DNA polymerase (TaKaRa Bio, Shiga). The resulting PCR product was subcloned into pGEM-T Easy vector (Promega Corp., Madison, WI), generating a plasmid pGEM/pol ι. After digesting with *Kpn*I and *Sac*I, the insert was ligated into the *Kpn*I/*Sac*I-cut pGL2-basic (Promega), generating a pGL2/pol ι reporter plasmid. Introduction of a mutation into the core sequence of putative HREs of *pol ι* (named HRE1 (from −400 to −396, 5′-GCGTG-3′), HRE2 (from −176 to −172, 5′-ACGTG-3′), HRE3 (from −149 to −145, 5′-GCGTG-3′), and HRE4 (from +330 to +334, 5′-ACGTG-3′)) that changes the sequence 5′-(A/G)CGTG-3′ to 5′-(A/G)AAAG-3′ was done by using QuikChange® Site Directed Mutagenesis Kit and pGEM/pol ι as a template. The identity of all of the cloned fragments was verified by nucleotide sequence analysis.

Luciferase reporter assays. Transient transfection of the luciferase reporter constructs into HepG2 cells and luciferase reporter assays were carried out as described previously [18]. One day after transfection, the cells were exposed to hypoxia (1% O_2) for 18 h, and luciferase activities in cell extracts were measured.

Electrophoretic mobility shift assay (EMSA). The nuclear proteins for EMSA were prepared from HeLa cells cultured in 21% O_2 or 1% O_2 for 8 h as described previously [20]. The HRE4-specific double-stranded oligonucleotide probe (wtHRE4) or its mutant form (mutHRE4) was prepared by annealing the sense 5′-ACTACAAATACGTGTCGAGGGT-3′ and the antisense 5′-ACCCTCGACACGTATTTGTAGT-3′ oligonucleotides (from +321 to +341) (containing putative HRE (marked in bold type)) or the sense 5′-ACTACAAATAAAAGTCGAGGGT-3′ and the

antisense 5′-ACCCTCGACTTTTATTTGTAGT-3′ oligonucleotides (containing mutated HRE (underlined)), respectively. Ten micrograms of nuclear proteins, ^{32}P-labeled double-stranded probe, 0.4 μg of calf thymus DNA, and binding buffer (10 mM Tris–HCl, pH 7.5, 50 mM NaCl, 1 mM $MgCl_2$, 1 mM EDTA, 5% glycerol, and 5 mM DTT) were mixed in a total volume of 20 μl. In competition assays, 40-fold molar excess amount of unlabeled competitors was included in the reaction mixture. The mixture was incubated at room temperature for 30 min and then loaded on a 4% polyacrylamide gel in TBE buffer (89 mM Tris base, 89 mM boric acid, and 5 mM EDTA). Supershift assay was performed using 1 μg of mouse monoclonal anti-HIF-1α antibody (clone H1alpha67, Novus Biologicals) or rabbit polyclonal anti-HIF-2α antibody (Novus Biologicals).

Chromatin immunoprecipitation (ChIP) assay. HeLa cells cultured in 21% O_2 or 1% O_2 for 8 h were fixed with 1% formaldehyde for 10 min at room temperature. Preparation of chromatin solution was performed essentially as described previously [21]. The chromatin solution was incubated with 5 μg of mouse monoclonal anti-HIF-1α antibody (clone H1alpha67, Novus Biologicals) at 4 °C for 15 h. Normal mouse IgG served as a control. Immunoprecipitation, washing, and elution of immune complexes were carried out with Protein A agarose beads (Upstate, Lake Placid, NY) according to the manufacturer's protocols. After reversing cross-links, the DNA was recovered by phenol:chloroform extraction and precipitated by ethanol. The association of HIF-1α with HRE4 was examined by hot-start PCR using *GoTaq* DNA polymerase (Promega). The sense and the antisense primers used were 5′-GCTGCCTCCCTCTGCCTT-3′ (from +236 to +253) and 5′-GGTTCTGAGCCATCCCTTC-3′ (from +506 to +524), respectively.

Results

To examine whether exposure of cells to hypoxia followed by reoxygenation causes ROS overproduction, HeLa cells were cultured under hypoxic conditions for 24 h and then reoxygenated for 30 min. ROS production was detected with DCFH-DA. The results showed that some of the cells produced ROS under both normoxic and hypoxic conditions (see Supplementary data, S1a). However, at 30 min after reoxygenation, many cells were found to produce a large amount of ROS (S1a). This ROS production was transient and attenuated to the normoxic level within 2 h after reoxygenation (data not shown). Then we examined the formation of 8-oxo-dG by immunostaining the cells with anti-8-oxo-dG antibody. As expected, a larger amount of 8-oxo-dG was formed in the reoxygenated cells than in normoxic and hypoxic cells (S1b).

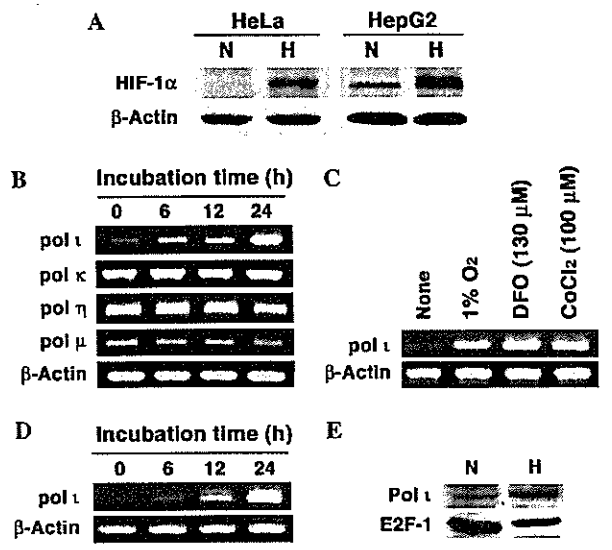

Fig. 1. Effects of hypoxia on the expressions of various error-prone DNA polymerase genes. (A) HIF-1α expression. HeLa and HepG2 cells were exposed to 1% O_2 for 6 h. Total cell lysates were subjected to immunoblot analysis. (B) Expression of mRNAs for various translesion DNA polymerases in HeLa cells exposed to 1% O_2 for the indicated times. Total RNA was subjected to RT-PCR. (C) Expression of *pol ι* mRNA in HeLa cells exposed to 1% O_2 or treated with desferrioxamine (DFO) or $CoCl_2$ for 24 h. (D) Expression of *pol ι* mRNA in HepG2 cells exposed to 1% O_2 for the indicated times. (E) Pol ι protein expression in HepG2 cells exposed to 1% O_2 for 24 h. Nuclear extracts were subjected to immunoblot analysis. E2F-1 was used as a control.

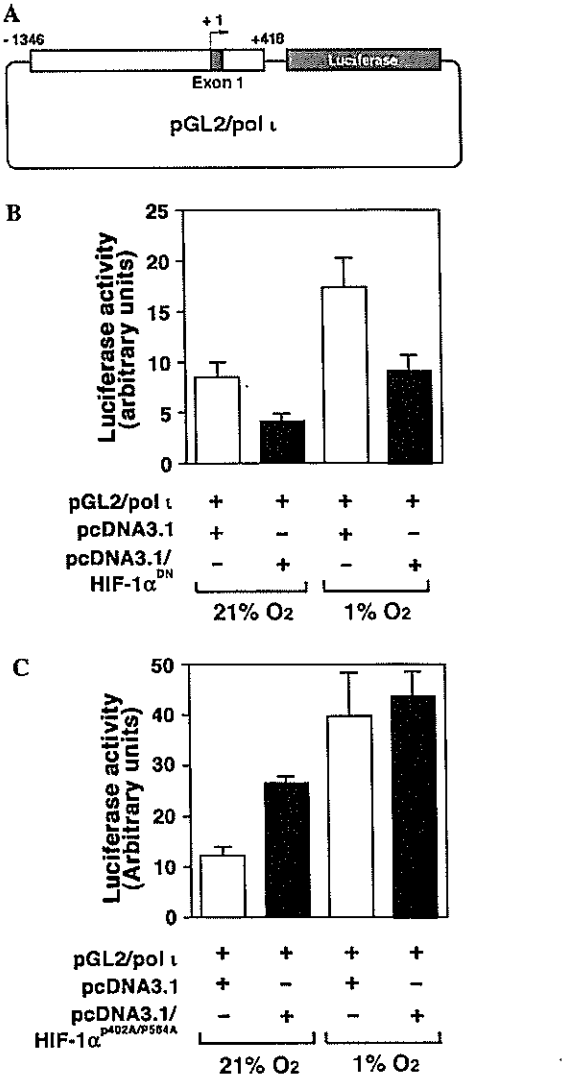

Fig. 2. Responsiveness of the luciferase reporter plasmid harboring *pol ι* regulatory region to hypoxia. (A) Construct of pGL2/pol ι luciferase reporter plasmid. (B) Effects of hypoxia and dominant-negative HIF-1α on luciferase activity. The reporter plasmid pGL2/pol ι was co-transfected with pcDNA3.1 or pcDNA3.1/HIF-1α^{DN} into HepG2 cells. The cells were exposed to 1% O_2 for 18 h. (C) Effects of hypoxia and constitutively active HIF-1α on luciferase activity. The reporter plasmid pGL2/pol ι was co-transfected with pcDNA3.1 or pcDNA3.1/HIF-1α$^{P402A/P564A}$ into HepG2 cells. Bars, SD.

Exposure of HeLa and HepG2 cells to hypoxia for 6 h resulted in the accumulation of HIF-1α (Fig. 1A). We then examined the expressions of translesion DNA polymerases that can bypass 8-oxo-dG in hypoxic cells by semi-quantitative RT-PCR. The results clearly showed that the expression of *pol ι* mRNA, but not of *pol κ*, *pol η*, or *pol μ* mRNA, was significantly increased by hypoxic stress in HeLa and HepG2 cells, depending on the incubation periods (6–24 h) (Fig. 1B and D). The hypoxia mimetics, desferrioxamine and $CoCl_2$, also increased the expression of *pol ι* mRNA (Fig. 1C). Accordingly, the amount of Pol ι protein was elevated in the nuclear extracts of HepG2 cells (Fig. 1E). Furthermore, hypoxia increased the level of *pol ι* mRNA in other cell lines such as A549, HT1080, LS174T, MCF7, MDA-MB-231, and U87MG (S2a). The induction of *pol ι* mRNA by hypoxia was observed in parallel with that of *VEGF* mRNA that is a well-known hypoxia-inducible gene. On the other hand, the decrease in the expression of *Mlh1* mRNA was detectable in both HeLa and HepG2 cells after a 48-h incubation under hypoxic conditions (S2b).

To examine whether HIF mediates the expression of *pol ι* mRNA in hypoxia, we made a pGL2/pol ι luciferase reporter construct harboring *pol ι* gene regulatory region (from −1346 to +418) (Fig. 2A). Transfection of the construct into HepG2 cells followed by exposure to hypoxia resulted in an approximately 2- to 3-fold increase in the luciferase activity (Fig. 2B and C). Co-transfection of pcDNA3.1/HIF-1α^{DN} abolished the increase (Fig. 2B). On the other hand, co-transfection with pcDNA3.1/HIF-1α$^{P402A/P564A}$ increased luciferase activity even under normoxic conditions (Fig. 2C), suggesting the involvement of HIF in the expression of *pol ι* mRNA in hypoxia.

Searching HRE consensus sequence (5′-(A/G)CGTG-3′) within this region showed that there are four putative HREs, three of them locating upstream of the transcription start site, named HRE1, HRE2, and HRE3, and the one locating in the intron 1, named HRE4. To determine which is functional, we generated a series of luciferase reporter plasmids in which three or all of the putative HREs were destroyed by introducing mutations (Fig. 3). Co-transfection of each of them with pcDNA3.1/HIF-1α$^{P402A/P564A}$ into HepG2 cells revealed that the reporter plasmid with intact HRE4 was most responsive while that with intact

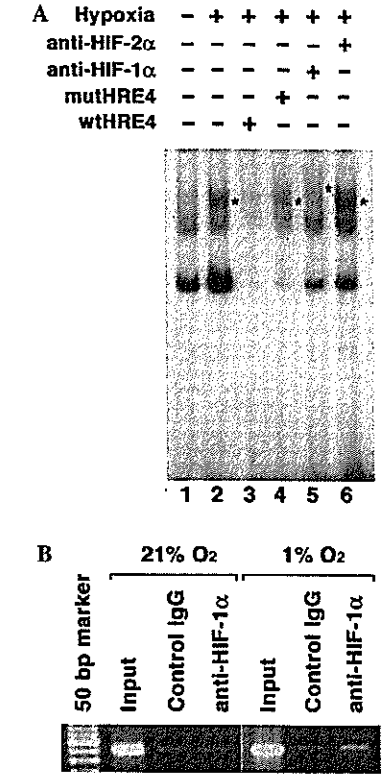

Fig. 4. HIF-1 binds to the HRE located in the intron 1 of the *pol ι* gene. (A) EMSA. Hela cells were cultured under 21% O_2 or 1% O_2 for 8 h. Nuclear extracts were subjected to EMSA using ^{32}P-labeled wtHRE4 as a probe. Asterisks indicate binding activity. (B) ChIP assay. HeLa cells were cultured under 21% O_2 or 1% O_2 for 8 h. ChIP assay was performed with mouse monoclonal anti-HIF-1α antibodies or mouse IgG as a control.

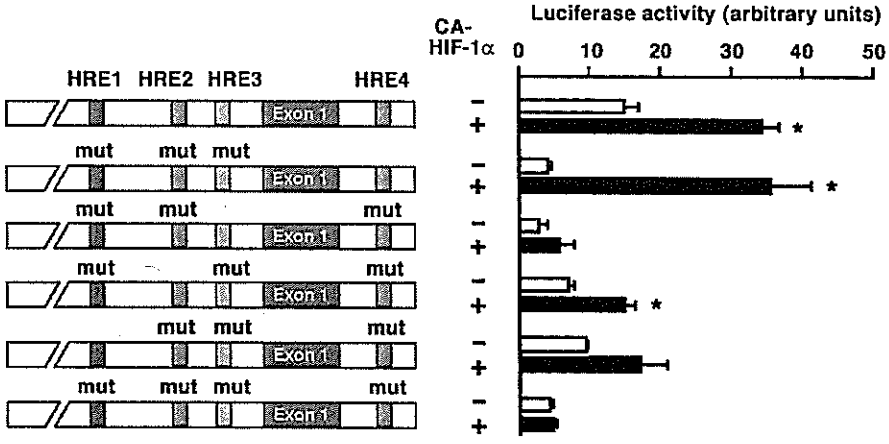

Fig. 3. Analysis of functional HRE in the *pol ι* regulatory region by luciferase reporter assays. Each of pGL2/pol ι plasmids with or without mutated (mut) HREs as indicated was co-transfected with pcDNA3.1 or pcDNA3.1/HIF-1α$^{P402A/P564A}$ (CA-HIF-1α) into HepG2 cells. The cells were exposed to 1% O_2 for 18 h. Bars, SD. *$p < 0.01$.

HRE1, HRE2, or HRE3 and that without all of the putative HREs showed weak and no responsiveness, respectively. To obtain evidence that HIF-1 binds to HRE4, we carried out EMSA using wtHRE4 as a DNA probe. The results showed that hypoxia enhanced binding activity (Fig. 4A, lane 2) that was competed with excess wtHRE4 (lane 3), but not with its mutant form mutHRE4 (lane 4). Addition of antibodies directed against HIF-1α, but not HIF-2α, induced a supershift of the binding activity (lanes 5 and 6). Furthermore, ChIP assay demonstrated that HIF-1α bound to the region containing HRE4 (Fig. 4B). From these results, we concluded that HIF-1 indeed interacts with the consensus HRE site (from +330 to +334) in the intron 1 of the *pol ι* gene.

Discussion

In this study, we confirmed that ROS were transiently overproduced in HeLa cells during reoxygenation and indeed caused 8-oxo-dG formation in the cells. We then focused on the expressions of translesion DNA polymerases that can bypass 8-oxo-dG, and found that hypoxia enhanced the expression of *pol ι* mRNA in various tumor cell lines.

Human Pol ι has low processivity and lacks an intrinsic 3′–5′ exonuclease activity, and has the lowest fidelity among so far reported eukaryotic polymerases [22]. Purified Pol ι has been observed to be able to efficiently bypass oxidized guanine and cytosine residues, as well as a variety of uracil lesions [23]. Not only can Pol ι mediate translesion replication in damaged DNA, but it also can misincorporate bases in a template-dependent manner in undamaged DNA [24,25]. Although these data are based on the *in vitro* studies, Yang et al. recently reported that Pol ι is overexpressed in human breast carcinoma cells and, importantly, that the expression level of Pol ι correlates with a significant decrease in DNA replication fidelity [26]. Therefore, up-regulation of Pol ι under hypoxic conditions might contribute to hypoxia/reoxygenation-induced mutagenesis.

Three lines of evidence suggested the involvement of HIF-1 in the mechanisms by which hypoxia induces *pol ι* mRNA expression. First, desferrioxamine and $CoCl_2$ also increased the expression of *pol ι* mRNA. Second, the reporter assays showed that dominant-negative HIF-1α suppressed the hypoxia-enhanced luciferase activity. Third, constitutively active HIF-1α enhanced the luciferase activity under normoxic conditions. Sequence analysis revealed the presence of four putative HREs which could explain the described effect of hypoxia on the induction of *pol ι* mRNA. Introduction of mutations in these HREs revealed that the reporter construct with intact HRE present in the intron 1 showed the maximal response to the co-transfected constitutively active HIF-1α construct, pointing out the major contribution of this HRE. Binding of HIF-1 to the HRE sequence was corroborated by EMSA. A shifted band was detected in hypoxic nuclear extracts, and a supershifted band was also detected after the incubation of the probe with anti-HIF-1α antibody. Since anti-HIF-2α antibody did not induce a supershifted band, HIF-2 is unlikely to bind the HRE. HIF-1 binding to the same sequence was further strengthened by the results of ChIP assay. Altogether, these results demonstrate that *pol ι* is a hypoxia-inducible gene through HIF-1 interaction with the consensus HRE site located at +330 in the intron 1 of the gene.

Recent studies highlight the deregulation of DNA mismatch repair enzymes in hypoxia as a mechanism of hypoxia-induced mutagenesis [11,12]. Loss of DNA mismatch repair renders cells hypersensitive to the mutagenic effect of oxidative stress [27]. In addition to this mechanism, our results may provide another mechanism underlying hypoxia/reoxygenation-induced mutagenesis. Through these mechanisms, hypoxia could lead to the genetic instability in tumor tissues.

Acknowledgments

This work was supported by Grants-in-Aid from the Ministry of Health, Labour, and Welfare for Third Term Comprehensive Control Research for Cancer and from the Ministry of Education, Culture, Sports, Science and Technology for Scientific Research, Japan.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bbrc.2006.10.048.

References

[1] A.L. Harris, Hypoxia—a key regulatory factor in tumour growth, Nat. Rev. Cancer 2 (2002) 38–47.
[2] G.L. Semenza, HIF-1 and tumor progression: pathophysiology and therapeutics, Trends Mol. Med. 8 (2002) S62–S67.
[3] L.E. Huang, J. Gu, M. Schau, H.F. Bunn, Regulation of hypoxia-inducible factor 1α is mediated by an O_2-dependent degradation domain via the ubiquitin-proteasome pathway, Proc. Natl. Acad. Sci. USA 95 (1998) 7987–7992.
[4] B.H. Jiang, G.L. Semenza, C. Bauer, H.H. Marti, Hypoxia-inducible factor 1 levels vary exponentially over a physiologically relevant range of O_2 tension, Am. J. Physiol. 271 (1996) C1172–C1180.
[5] G.L. Wang, G.L. Semenza, Desferrioxamine induces erythropoietin gene expression and hypoxia-inducible factor 1 DNA-binding activity: implications for models of hypoxia signal transduction, Blood 82 (1993) 3610–3615.
[6] B.L. Ebert, H.F. Bunn, Regulation of the erythropoietin gene, Blood 94 (1999) 1864–1877.
[7] G.L. Semenza, B.H. Jiang, S.W. Leung, R. Passantino, J.P. Concordet, P. Maire, A. Giallongo, Hypoxia response elements in the aldolase A, enolase 1, and lactate dehydrogenase A gene promoters contain essential binding sites for hypoxia-inducible factor 1, J. Biol. Chem. 271 (1996) 32529–32537.
[8] J. Yuan, L. Narayanan, S. Rockwell, P.M. Glazer, Diminished DNA repair and elevated mutagenesis in mammalian cells exposed to hypoxia and low pH, Cancer Res. 60 (2000) 4372–4376.
[9] A.P. Breen, J.A. Murphy, Reactions of oxyl radicals with DNA, Free Radic. Biol. Med. 18 (1995) 1033–1077.
[10] S. Shibutani, M. Takeshita, A.P. Grollman, Insertion of specific bases during DNA synthesis past the oxidation-damaged base 8-oxodG, Nature 349 (1991) 431–434.

[11] V.T. Mihaylova, R.S. Bindra, J. Yuan, D. Campisi, L. Narayanan, R. Jensen, F. Giordano, R.S. Johnson, S. Rockwell, P.M. Glazer, Decreased expression of the DNA mismatch repair gene *Mlh1* under hypoxic stress in mammalian cells, Mol. Cell Biol. 23 (2003) 3265–3273.

[12] M. Koshiji, K.K. To, S. Hammer, K. Kumamoto, A.L. Harris, P. Modrich, L.E. Huang, HIF-1α induces genetic instability by transcriptionally downregulating MutSα expression, Mol. Cell 17 (2005) 793–803.

[13] M.F. Goodman, Error-prone repair DNA polymerases in prokaryotes and eukaryotes, Annu. Rev. Biochem. 71 (2002) 17–50.

[14] Y. Zhang, F. Yuan, X. Wu, O. Rechkoblit, J.S. Taylor, N.E. Geacintov, Z. Wang, Error-prone lesion bypass by human DNA polymerase η, Nucleic Acids Res. 28 (2000) 4717–4724.

[15] Y. Zhang, F. Yuan, X. Wu, J.S. Taylor, Z. Wang Z, Response of human DNA polymerase ι to DNA lesions, Nucleic Acids Res. 29 (2001) 928–935.

[16] L. Haracska, L. Prakash, S. Prakash, Role of human DNA polymerase κ as an extender in translesion synthesis, Proc. Natl. Acad. Sci. USA 99 (2002) 16000–16005.

[17] Y. Zhang, X. Wu, D. Guo, O. Rechkoblit, J.S. Taylor, N.E. Geacintov, Z. Wang, Lesion bypass activities of human DNA polymerase μ, J. Biol. Chem. 277 (2002) 44582–44587.

[18] N. Enomoto, N. Koshikawa, M. Gassmann, J. Hayashi, K. Takenaga, Hypoxic induction of hypoxia-inducible factor-1α and oxygen-regulated gene expression in mitochondrial DNA-depleted HeLa cells, Biochem. Biophys. Res. Commun. 297 (2002) 346–352.

[19] J. Chen, S. Zhao, K. Nakada, Y. Kuge, N. Tamaki, F. Okada, J. Wang, M. Shindo, F. Higashino, K. Takeda, M. Asaka, H. Katoh, T. Sugiyama, M. Hosokawa, M. Kobayashi, Dominant-negative hypoxia-inducible factor-1 α reduces tumorigenicity of pancreatic cancer cells through the suppression of glucose metabolism, Am. J. Pathol. 162 (2003) 1283–1291.

[20] T. Fink, A. Kazlauskas, L. Poellinger, P. Ebbesen, V. Zachar, Identification of a tightly regulated hypoxia-response element in the promoter of human plasminogen activator inhibitor-1, Blood 99 (2002) 2077–2083.

[21] Y. Ge, T.L. Jensen, L.H. Matherly, J.W. Taub, Transcriptional regulation of the cystathionine-β-synthase gene in Down syndrome and non-Down syndrome megakaryocytic leukemia cell lines, Blood 101 (2003) 1551–1557.

[22] J.P. McDonald, A. Tissier, E.G. Frank, S. Iwai, F. Hanaoka, R. Woodgate, DNA polymerase ι and related rad30-like enzymes, Philos. Trans R. Soc. Lond. B Biol. Sci. 356 (2001) 53–60.

[23] A. Vaisman, R. Woodgate, Unique misinsertion specificity of polι may decrease the mutagenic potential of deaminated cytosines, EMBO J. 20 (2001) 6520–6529.

[24] T.A. Kunkel, Y.I. Pavlov, K. Bebenek, Functions of human DNA polymerases η, κ and ι suggested by their properties, including fidelity with undamaged DNA templates, DNA Repair (Amst) 2 (2003) 135–149.

[25] A. Tissier, J.P. McDonald, E.G. Frank, R. Woodgate, Polι, a remarkably error-prone human DNA polymerase, Genes Dev. 14 (2000) 1642–1650.

[26] J. Yang, Z. Chen, Y. Liu, R.J. Hickey, L.H. Malkas, Altered DNA polymerase ι expression in breast cancer cells leads to a reduction in DNA replication fidelity and a higher rate of mutagenesis, Cancer Res. 64 (2004) 5597–5607.

[27] X. Lin, K. Ramamurthi, M. Mishima, A. Kondo, S.B. Howell, p53 interacts with the DNA mismatch repair system to modulate the cytotoxicity and mutagenicity of hydrogen peroxide, Mol. Pharmacol. 58 (2000) 1222–1229.

Cell

Control of Developmental Regulators by Polycomb in Human Embryonic Stem Cells

Tong Ihn Lee,[1,8] Richard G. Jenner,[1,8] Laurie A. Boyer,[1,8] Matthew G. Guenther,[1,8] Stuart S. Levine,[1,8] Roshan M. Kumar,[1] Brett Chevalier,[1] Sarah E. Johnstone,[1,2] Megan F. Cole,[1,2] Kyo-ichi Isono,[3] Haruhiko Koseki,[3] Takuya Fuchikami,[4] Kuniya Abe,[4] Heather L. Murray,[1] Jacob P. Zucker,[6] Bingbing Yuan,[1] George W. Bell,[1] Elizabeth Herbolsheimer,[1] Nancy M. Hannett,[1] Kaiming Sun,[1] Duncan T. Odom,[1] Arie P. Otte,[5] Thomas L. Volkert,[1] David P. Bartel,[1,2] Douglas A. Melton,[6] David K. Gifford,[1,7] Rudolf Jaenisch,[1,2] and Richard A. Young[1,2,*]

[1]Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, MA 02142, USA

[2]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[3]Developmental Genetics Group, RIKEN Center for Allergy and Immunology, 1-7-22, Suehiro, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

[4]Technology and Development Team for Mammalian Cellular Dynamics, BioResource Center, RIKEN Tsukuba Institute, 3-1-1, Koyadai, Tsukuba, Ibaraki 230-0045, Japan

[5]Swammerdam Institute for Life Sciences, University of Amsterdam, 1098 SM Amsterdam, The Netherlands

[6]Howard Hughes Medical Institute, Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138, USA

[7]MIT CSAIL, 32 Vassar Street, Cambridge, MA 02139, USA

[8]These authors contributed equally to this work.

*Contact: young@wi.mit.edu

DOI 10.1016/j.cell.2006.02.043

SUMMARY

Polycomb group proteins are essential for early development in metazoans, but their contributions to human development are not well understood. We have mapped the Polycomb Repressive Complex 2 (PRC2) subunit SUZ12 across the entire nonrepeat portion of the genome in human embryonic stem (ES) cells. We found that SUZ12 is distributed across large portions of over two hundred genes encoding key developmental regulators. These genes are occupied by nucleosomes trimethylated at histone H3K27, are transcriptionally repressed, and contain some of the most highly conserved noncoding elements in the genome. We found that PRC2 target genes are preferentially activated during ES cell differentiation and that the ES cell regulators OCT4, SOX2, and NANOG cooccupy a significant subset of these genes. These results indicate that PRC2 occupies a special set of developmental genes in ES cells that must be repressed to maintain pluripotency and that are poised for activation during ES cell differentiation.

INTRODUCTION

Embryonic stem (ES) cells are a unique self-renewing cell type that can give rise to the ectodermal, endodermal, and mesodermal germ layers during embryogenesis. Human ES cells, which can be propagated in culture in an undifferentiated state but selectively induced to differentiate into many specialized cell types, are thought to hold great promise for regenerative medicine (Thomson et al., 1998; Reubinoff et al., 2000; Mayhall et al., 2004; Pera and Trounson, 2004). The gene expression program of ES cells must allow these cells to maintain a pluripotent state but also allow for differentiation into more specialized states when signaled to do so. Learning how this is accomplished may be key to realizing the therapeutic potential of ES cells and further understanding early development.

Among regulators of development, the Polycomb group proteins (PcG) are of special interest. These regulators were first described in *Drosophila*, where they repress the homeotic genes controlling segment identity in the developing embryo (Lewis, 1978; Denell and Frederick, 1983; Simon et al., 1992; Orlando and Paro, 1995; Pirrotta, 1998; Kennison, 2004). The initial repression of these genes is carried out by DNA binding transcriptional repressors, and PcG proteins modify chromatin to maintain these genes in a repressed state (Duncan, 1986; Bender et al., 1987; Strutt et al., 1997; Horard et al., 2000; Hodgson et al., 2001; Mulholland et al., 2003).

The PcG proteins form multiple Polycomb Repressive Complexes (PRCs), the components of which are conserved from *Drosophila* to humans (Franke et al., 1992; Shao et al., 1999; Birve et al., 2001; Tie et al., 2001; Cao et al., 2002; Czermin et al., 2002; Kuzmichev et al., 2002; Levine et al., 2002). The PRCs are brought to the site of initial repression and act through epigenetic modification of chromatin structure to promote gene silencing (Pirrotta,

1998; Levine et al., 2004; Lund and van Lohuizen, 2004; Ringrose and Paro, 2004). PRC2 catalyzes histone H3 lysine-27 (H3K27) methylation, and this enzymatic activity is required for PRC2-mediated gene silencing (Cao et al., 2002; Czermin et al., 2002; Kuzmichev et al., 2002; Muller et al., 2002; Kirmizis et al., 2004). H3K27 methylation is thought to provide a binding surface for PRC1, which facilitates oligomerization, condensation of chromatin structure, and inhibition of chromatin remodeling activity in order to maintain silencing (Shao et al., 1999; Francis et al., 2001; Cao et al., 2002; Czermin et al., 2002).

Components of PRC2 are essential for the earliest stages of vertebrate development (Faust et al., 1998; O'Carroll et al., 2001; Pasini et al., 2004). PRC2 and its related complexes, PRC3 and PRC4, contain the core components EZH2, SUZ12, and EED (Kuzmichev et al., 2004; Kuzmichev et al., 2005). EZH2 is a H3K27 methyltransferase, and SUZ12 (Suppressor of zeste 12) is required for this activity (Cao and Zhang, 2004; Pasini et al., 2004). ES cell lines cannot be established from Ezh2-deficient blastocysts (O'Carroll et al., 2001), suggesting that PRC2 is involved in regulating pluripotency and self-renewal. Although the PRCs are known to repress individual *HOX* genes (van der Lugt et al., 1996; Akasaka et al., 2001; Wang et al., 2002; Cao and Zhang, 2004), it is not clear how these important PcG regulators contribute to early development in vertebrates.

Because the nature of PRC2 target genes in ES cells might reveal why PRC2 is essential for early embryonic development, pluripotency, and self-renewal, we have mapped the sites occupied by the SUZ12 subunit throughout the genome in human ES cells. This genome-wide map reveals that PRC2 is associated with a remarkable cadre of genes encoding key regulators of developmental processes that are repressed in ES cells. The genes occupied by PRC2 contain nucleosomes that are trimethylated at histone H3 lysine-27 (H3K27me3), a modification catalyzed by PRC2 and associated with the repressed chromatin state. Both PRC2 and nucleosomes with histone H3K27me3 occupy surprisingly large genomic domains around these developmental regulators and are frequently associated with highly conserved noncoding sequence elements previously identified by comparative genomic methods. The transcription factors OCT4, SOX2, and NANOG, which are also key regulators of ES cell pluripotency and self-renewal, occupy a significant subset of these genes. Thus, the model of epigenetic regulation of homeotic genes extends to a large set of developmental regulators whose repression in ES cells appears to be key to pluripotency. We suggest that PRC2 functions in ES cells to repress developmental genes that are preferentially activated during differentiation.

RESULTS AND DISCUSSION

Mapping Genome Occupancy in ES Cells

We mapped the location of both RNA polymerase II and the SUZ12 subunit of PRC2 genome-wide in human ES cells (Figure 1). The initiating form of RNA polymerase II was mapped to test the accuracy of the method and provide a reference for comparison with sites occupied by PRC2. The SUZ12 subunit of PRC2 is critical for the function of the complex and was selected for these genome-wide experiments. Human ES cells (H9, NIH code WA09) were analyzed by immunohistochemistry for characteristic stem cell markers, tested for their ability to generate cell types from all three germ layers upon differentiation into embryoid bodies, and shown to form teratomas in immunocompromised mice (Supplemental Data; Figures S1–S3).

DNA sequences bound by RNA polymerase II were identified in replicate chromatin-immunoprecipitation (ChIP) experiments using DNA microarrays that contain over 4.6 million unique 60-mer oligonucleotide probes spanning the entire nonrepeat portion of the human genome (Figure 1 and Supplemental Data). To obtain a probabilistic assessment of binding events, an algorithm was implemented that incorporates information from multiple probes representing contiguous regions of the genome, and threshold criteria were established to identify a dataset with minimal false positives and false negatives. RNA polymerase II was associated with the promoters of 7,106 of the approximately 22,500 annotated human genes, indicating that one-third of protein-coding genes are prepared to be transcribed in ES cells. Three lines of evidence suggest this dataset is of high quality. Most of the RNA polymerase II sites (87%) occurred at promoters of known or predicted genes. Transcripts were detected for 88% of the genes bound by RNA polymerase II in previous expression experiments in ES cells. Finally, independent analysis using gene-specific PCR (Supplemental Data) indicated that the frequency of false positives was approximately 4% and the frequency of false negatives was approximately 30% in this dataset. A detailed analysis of the RNA polymerase II dataset, including binding to miRNA genes, can be found in Supplemental Data (Tables S1–S6 and Figures S4 and S5).

The sites occupied by SUZ12 were then mapped throughout the entire nonrepeat genome in H9 ES cells using the same approach described for RNA polymerase II (Figure 1C). SUZ12 was associated with the promoters of 1,893 of the approximately 22,500 annotated human genes, indicating that ~8% of protein-coding genes are occupied by SUZ12 in ES cells (Supplemental Data; Tables S7 and S8). Independent site-specific analysis indicated that the frequency of false positives was approximately 3% and the frequency of false negatives was approximately 27% in this dataset.

Comparison of the genes occupied by SUZ12 with those occupied by RNA polymerase II revealed that the two sets were largely exclusive (Figure 1D; Supplemental Data; Table S8). There were, however, genes where SUZ12 and RNA polymerase II cooccupied promoters. At these genes, PRC complexes may fail to block assembly of the preinitiation complex (Dellino et al., 2004), consistent with the observation that Polycomb group proteins

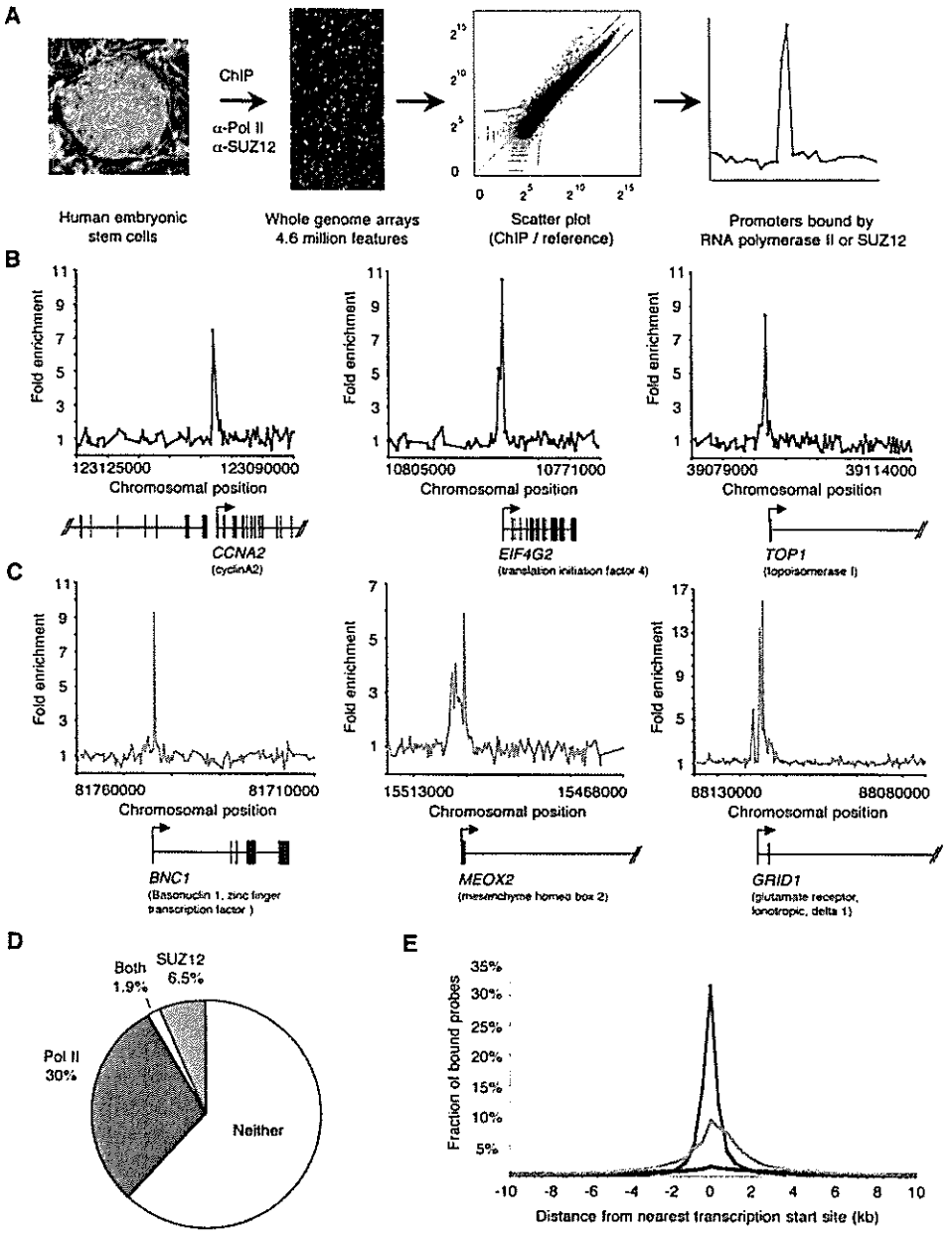

Figure 1. Genome-Wide ChIP-Chip in Human Embryonic Stem Cells

(A) DNA segments bound by the initiation form of RNA polymerase II or SUZ12 were isolated using chromatin-immunoprecipitation (ChIP) and identified with DNA microarrays containing over 4.6 million unique 60-mer oligonucleotide probes spanning the entire nonrepeat portion of the human genome. ES cell growth and quality control, the antibodies, ChIP protocol, DNA microarray probe design, and data analysis methods are described in detail in Supplemental Data.

(B) Examples of RNA polymerase II ChIP signals from genome-wide ChIP-Chip. The plots show unprocessed enrichment ratios (blue) for all probes within a genomic region (ChIP versus whole genomic DNA). Chromosomal positions are from NCBI build 35 of the human genome. Genes are shown to scale below plots (exons are represented by vertical bars). The start and direction of transcription are noted by arrows.

(C) Examples of SUZ12 ChIP signals from genome-wide ChIP-Chip. The plots show unprocessed enrichment ratios (green) for all probes within a genomic region (ChIP versus whole genomic DNA). Chromosomal positions, genes, and notations are as described in (B).

(D) Chart showing percentage of all annotated genes bound by RNA polymerase II (blue), SUZ12 (green), both (yellow), or neither (gray).

(E) Distribution of the distance between bound probes and the closest transcription start sites from RefSeq, Ensembl, MGC, UCSC Known Genes and H-Inv databases for SUZ12 (green line), and RNA polymerase II (blue line). The number of bound probes is given as the percentage of total probes and is calculated for 400 bp intervals from the start site. The null-distribution of the distance between all probes and the closest transcription are shown as a black line.

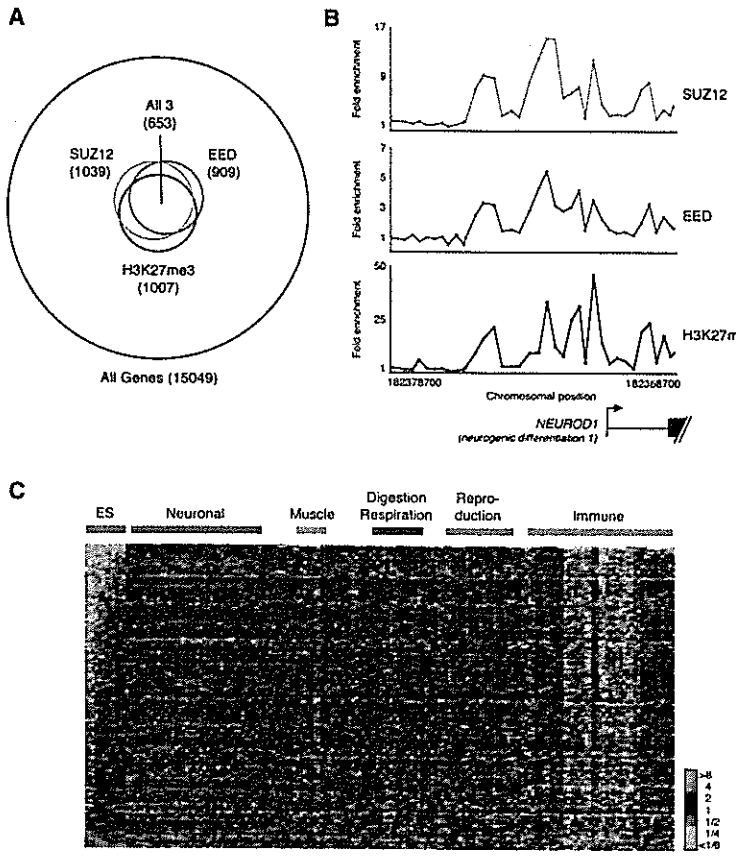

Figure 2. SUZ12 Is Associated with EED, histone H3K27me3 Modification, and Transcriptional Repression in ES Cells

(A) Venn diagram showing the overlap of genes bound by SUZ12 at high-confidence, genes bound by EED at high-confidence, and genes trimethylated at H3K27 at high-confidence. The data are from promoter microarrays that contain probes tiling −8 kb and +2 kb around transcription start. 72% of the genes bound by SUZ12 at high-confidence are also bound by EED at high-confidence; others are bound by EED at lower confidence (Figure S6).

(B) SUZ12 (top), EED (middle), and H3K27me3 (bottom) occupancy at *NEUROD1*. The plots show unprocessed enrichment ratios for all probes within this genomic region (SUZ12 ChIP versus whole genomic DNA, EED ChIP versus whole genomic DNA, and H3K27me3 ChIP versus total H3 ChIP). Chromosomal positions are from NCBI build 35 of the human genome. *NEUROD1* is shown to scale below plots (exons are represented by vertical bars). The start and direction of transcription are noted by arrows.

(C) Relative expression levels of 604 genes occupied by PRC2 and trimethylated at H3K27 in ES cells. Comparisons were made across four ES cell lines and 79 differentiated cell types. Each row corresponds to a single gene that is bound by SUZ12, associated with EED and H3K27me3, and for which Affymetrix expression data are available. Each column corresponds to a single expression microarray. ES cells are in the following order: H1, H9, HSF6, HSF1. For each gene, expression is shown relative to the average expression level of that gene across all samples, with shades of red indicating higher than average expression and green lower than average expression according to the scale on the right. Cell types are grouped by tissue or organ function, and genes are ranked according the significance of their relative level of gene expression in ES cells.

can associate with components of the general transcription apparatus (Breiling et al., 2001; Saurin et al., 2001).

The vast majority of SUZ12 bound sites were found at gene promoters (Figure 1E). Ninety-five percent of the SUZ12 bound regions were found within 1 kb of known or predicted transcription start sites (Supplemental Data and Table S7). This suggests that SUZ12 functions in human ES cells primarily at promoters rather than at distal regulatory elements. It is interesting that 40% of all SUZ12 bound regions are within 1 Kb of CpG islands (Table S7), given the recent discovery of a mechanistic link between PcG proteins and DNA methyltransferoses (Vire et al., 2006).

Global Transcriptional Repression by PRC2

PRC2 is composed of three core subunits, SUZ12, EED, and EZH2, and has been shown to mediate histone H3K27 methylation at specific genes in vivo. To confirm that SUZ12 is associated with active PRC2 at target genes, we used chromatin immunoprecipitation with antibodies against EED and the histone H3K27me3 mark and analyzed the results with promoter microarrays. We found that EED and the histone H3K27me3 mark cooccurred with SUZ12 at most genes using a high-confidence binding threshold (Figure 2). The false negative rates of thresholded data can lead to an underestimate of the similarity between different datasets. Plotting raw enrichment ratios for genes associated with SUZ12, EED, or H3K27me3 demonstrates that SUZ12 binding represents PRC2 binding at almost all target genes (Figure S6).

Genetic and biochemical studies at selected genes indicate that PRC2-mediated H3K27 methylation represses gene expression, but it has not been established if it acts as a repressor genome-wide. If genes occupied by SUZ12 are repressed by PRC2, then transcripts from these genes should generally be present at lower levels in ES cells than in differentiated cell types. To test this prediction, we compared the expression levels of PRC2-occupied genes in four different ES cell lines with the expression level of these genes in 79 differentiated human cell and tissue types (Sato et al., 2003; Abeyta et al., 2004; Su et al., 2004). We found that PRC2 occupied genes were generally underexpressed in ES cells relative to other cell types (Figure 2C). A small fraction of the genes occupied by PRC2 were relatively overexpressed in ES cells (Figure 2C); these tended to show less extensive SUZ12 occupancy and were more likely to be cooccupied by

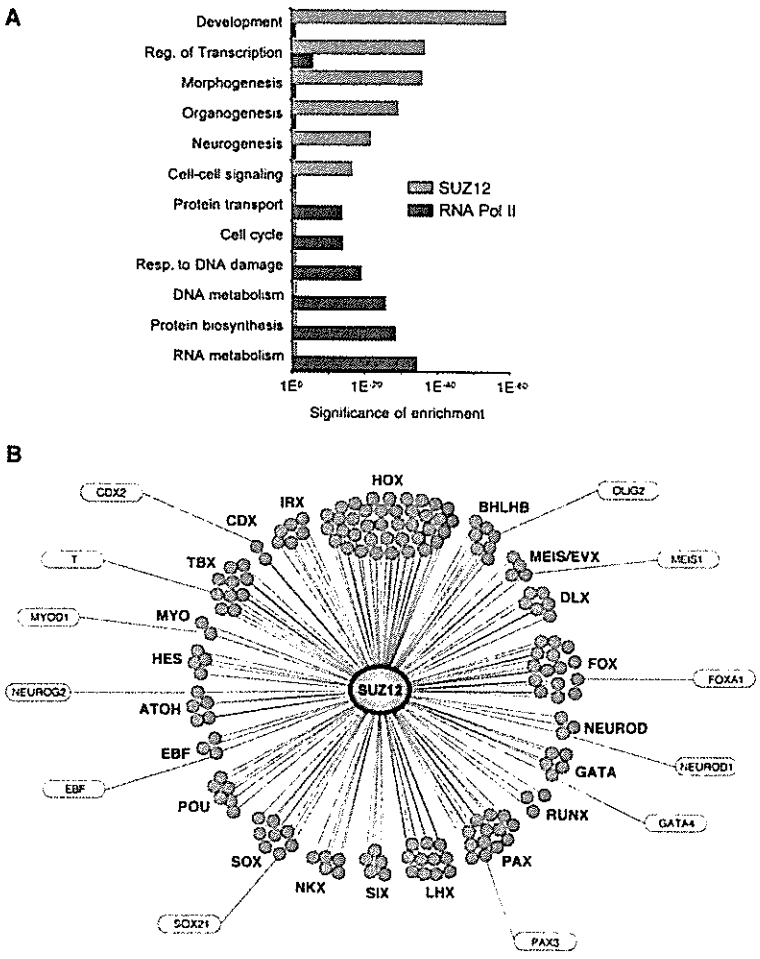

Figure 3. Cellular Functions of Genes Occupied by SUZ12

(A) Genes bound by SUZ12 or RNA polymerase II were compared to biological process gene ontology categories; highly represented categories are shown. Ontology terms are shown on the *y* axis; p-values for the significance of enrichment are graphed along the *x* axis (SUZ12 in green, RNA polymerase II in blue).

(B) Selected examples of developmental transcription factor families bound by SUZ12. SUZ12 is represented by the green oval; individual transcription factors are represented by circles and grouped by family as indicated. Examples of transcription factors with defined roles in development are labeled. Transcription factor families include homeobox protein (HOX), basic helix-loop-helix domain containing, class B (BHLHB), HOX cofactors (MEIS/EVX), distal-less homeobox (DLX), Forkhead box (FOX), NEUROD, GATA binding protein (GATA), runt related transcription factor (RUNX), paired box and paired-like (PAX), LIM homeobox (LHX), sine oculis homeobox homolog (SIX), NK transcription factor related (NKX), SRY box (SOX), POU domain containing, classes 3 and 4 (POU), early B-cell factor (EBF), atonal homolog (ATOH), hairy and enhancer of split protein (HES), myogenic basic domain (MYO), T-box (TBX), caudal type homeobox (CDX), and iroquois homeobox protein (IRX).

RNA polymerase II (Supplemental Data). These results are consistent with the model that PRC2-mediated histone H3K27 methylation promotes gene silencing at the majority of its target genes throughout the genome in ES cells.

Key Developmental Regulators Are Targets of PRC2

Examination of the targets of SUZ12 revealed that they were remarkably enriched for genes that control development and transcription (Figure 3) and that SUZ12 tended to occupy large domains at these genes (Figure 4). Although only 8% of all annotated genes were occupied by SUZ12, ~50% of those encoding transcription factors associated with developmental processes were occupied by SUZ12. By comparison, RNA polymerase II preferentially occupied genes involved in a broad spectrum of cell proliferation functions such as nucleic acid metabolism, protein synthesis, and cell cycle (Figure 3A and examples in Figure 1B; Supplemental Data; Table S10).

It was striking that SUZ12 occupied many families of genes that control development and transcription (Figures 3B and S7 and Table S11). These included 39 of 40 of the homeotic genes found in the *HOX* clusters and the majority of homeodomain genes. SUZ12 bound homeodomain genes included almost all members of the *DLX*, *IRX*, *LHX*, and *PAX* gene families, which regulate early developmental steps in neurogenesis, hematopoiesis, axial patterning, tissue patterning, organogenesis, and cell-fate specification. SUZ12 also occupied promoters for large subsets of the *FOX, SOX, and TBX* gene families. The forkhead family of *FOX* genes is involved in axial patterning and tissue development from all three germ layers (Lehmann et al., 2003). Mutations in members of the *SOX* gene family alter cell-fate specification and differentiation and are linked to several developmental diseases (Schepers et al., 2002). The *TBX* family of genes regulates a wide variety of developmental processes such as gastrulation, early pattern formation, organogenesis, and limb formation (Showell et al., 2004). Thus, the genes preferentially bound by SUZ12 have functions that, when expressed, promote differentiation. This is likely to explain, at least in part, why PRC2 is essential for early development and ES cell pluripotency.

A remarkable feature of PRC2 binding at most genes encoding developmental regulators was the extensive span over which the regulator occupied the locus (Figures 4, S8, and S9). For the majority (72%) of bound sites

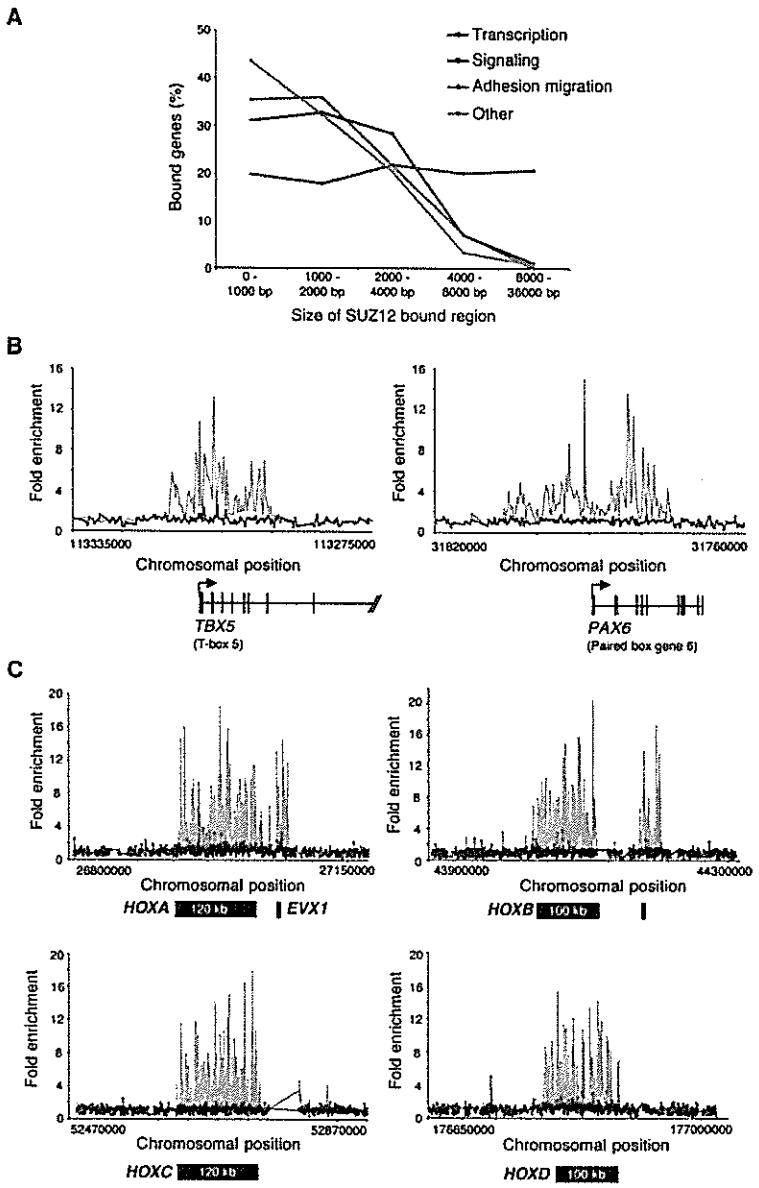

Figure 4. SUZ12 Occupies Large Portions of Genes Encoding Transcription Factors with Roles in Development

(A) The fraction of SUZ12 target genes associated with different sizes of binding domains. Genes are grouped into four categories according to their function: Signaling, Adhesion/migration, Transcription, and Other.

(B) Examples of SUZ12 (green) and RNA polymerase II (blue) binding at the genes encoding developmental regulators TBX5 and PAX6. The plots show unprocessed enrichment ratios for all probes within a genomic region (ChIP versus whole genomic DNA). Genes are shown to scale below plots (exons are represented by vertical bars). The start and direction of transcription are noted by arrows.

(C) Binding profiles of SUZ12 (green) and RNA polymerase II (blue) across ~500 kb regions encompassing *HOX* clusters *A–D*. Unprocessed enrichment ratios for all probes within a genomic region are shown (ChIP versus whole genomic DNA). Approximate *HOX* cluster region sizes are indicated within black bars.

across the genome, SUZ12 occupied a small region of the promoter similar in size to regions bound by RNA polymerase II (Figure 1). For the remaining bound regions, SUZ12 occupancy encompassed large domains spanning 2–35 kb and extending from the promoter into the gene. A large portion of genes encoding developmental regulators (72%) exhibited these extended regions of SUZ12 binding. In some cases, binding encompassed multiple contiguous genes. For instance, SUZ12 binding extended ~100 kb across the entire *HOXA, HOXB, HOXC,* and *HOXD* clusters but did not bind to adjacent genomic sequences, yielding a highly defined spatial pattern (Figure 4B). In contrast, clusters of unrelated genes, such as the interleukin 1-β cluster, were not similarly bound by SUZ12. Thus, genes encoding developmental regulators showed an unusual tendency to be occupied by PRC2 over much or all of their transcribed regions.

PRC2 and Highly Conserved Elements

Previous studies have noted that many highly conserved noncoding elements of vertebrate genomes are associated with genes encoding developmental regulators (Bejerano et al., 2004; Siepel et al., 2005; Woolfe et al., 2005). Given SUZ12's strong association with this class of genes, we investigated the possibility that SUZ12 bound regions are associated with these highly conserved elements. Inspection of individual genes suggested that SUZ12 occupancy was associated with regions of sequence conservation (Figure 5A). Eight percent of the approximately 1,400 highly conserved noncoding DNA elements

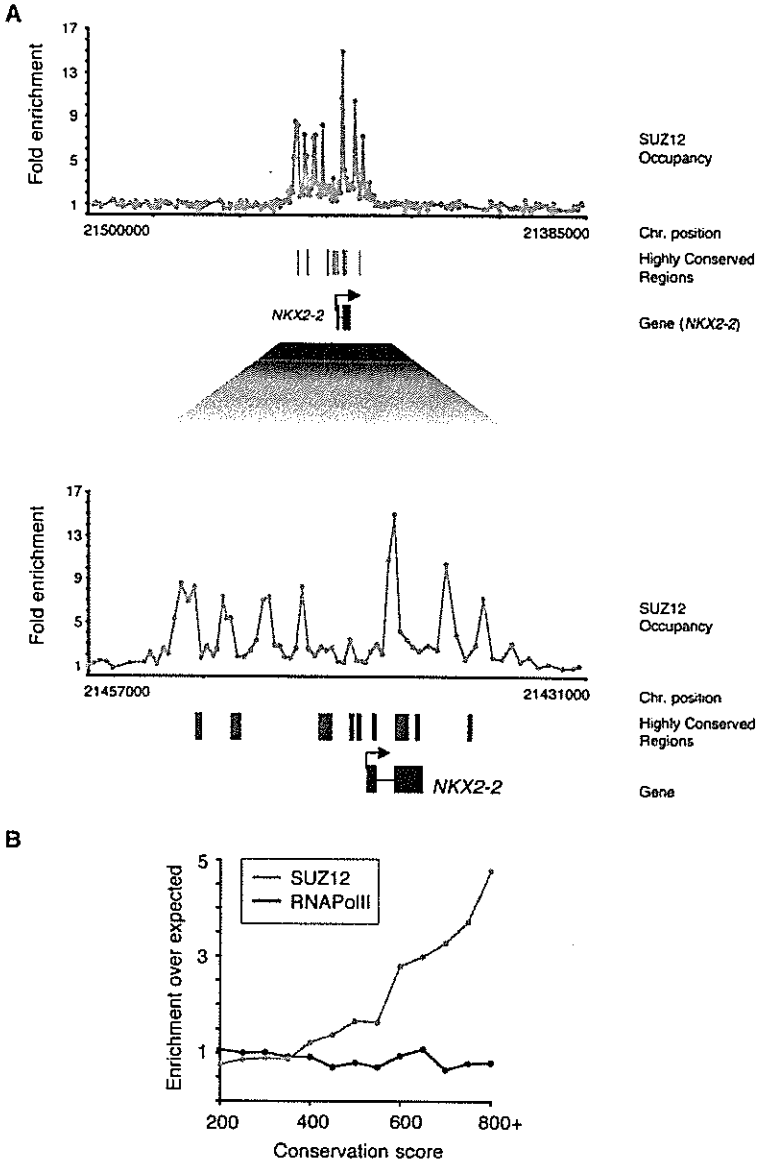

Figure 5. SUZ12 Binding Is Associated with Highly Conserved Regions

(A) SUZ12 occupancy (green) and conserved elements are shown at *NKX2-2* and adjacent genomic regions. The plots show unprocessed enrichment ratios for all probes within this genomic region (SUZ12 ChIP versus whole genomic DNA). Conserved elements (red) with LoD scores > 160 derived from the PhastCons program (Siepel et al., 2005) are shown to scale above the plot. Genes are shown to scale below plots (exons are represented by vertical bars). A higher resolution view is also shown below.

(B) Enrichment of conserved noncoding elements within SUZ12 (green) and RNA polymerase II (blue) bound regions. The maximum nonexonic PhastCons conservation score was determined for each bound region. For comparison, the same parameter was determined using a randomized set of genomic regions with the same size distribution. The graph displays the ratio of the number of bound regions with that score versus the number of randomized genomic regions with that score.

described by Woolfe and colleagues (Woolfe et al., 2005) were found to be associated with the SUZ12 bound developmental regulators (p-value 10^{-14}). Using entries from the PhastCons database of conserved elements (Siepel et al., 2005), we found that SUZ12 occupancy of highly conserved elements was highly significant (using highly conserved elements with a LoD conservation score of 100 or better, the p-value for significances was less than 10^{-85}). Since PRC2 has not been shown to directly bind DNA sequences, we expect that specific DNA binding proteins occupy the highly conserved DNA sequences and may associate with PRC2, which spreads and occupies adjacent chromatin. Thus, the peaks of SUZ12 occupancy might not be expected to precisely colocate with the highly conserved elements, even if these elements are associated with PRC2 recruitment.

Remarkably, the degree of the association between SUZ12 binding and conserved sequences increases when considering sequences with an increasing degree of conservation (Figure 5B). By comparison, RNA polymerase II showed no such enrichment. These results suggest that the subset of highly conserved noncoding elements at genes encoding developmental regulators may be associated with PcG-mediated silencing of these regulators.

Signaling Genes Are among PRC2 Targets

The targets of SUZ12 were also enriched for genes that encode components of signaling pathways (Figure 3A and Table S12). There is evidence that transforming growth factor-β (TGFβ), bone morphogenic protein (BMP), wingless-type MMTV integration site (Wnt), and fibroblast growth factor (FGF) signaling pathways, which

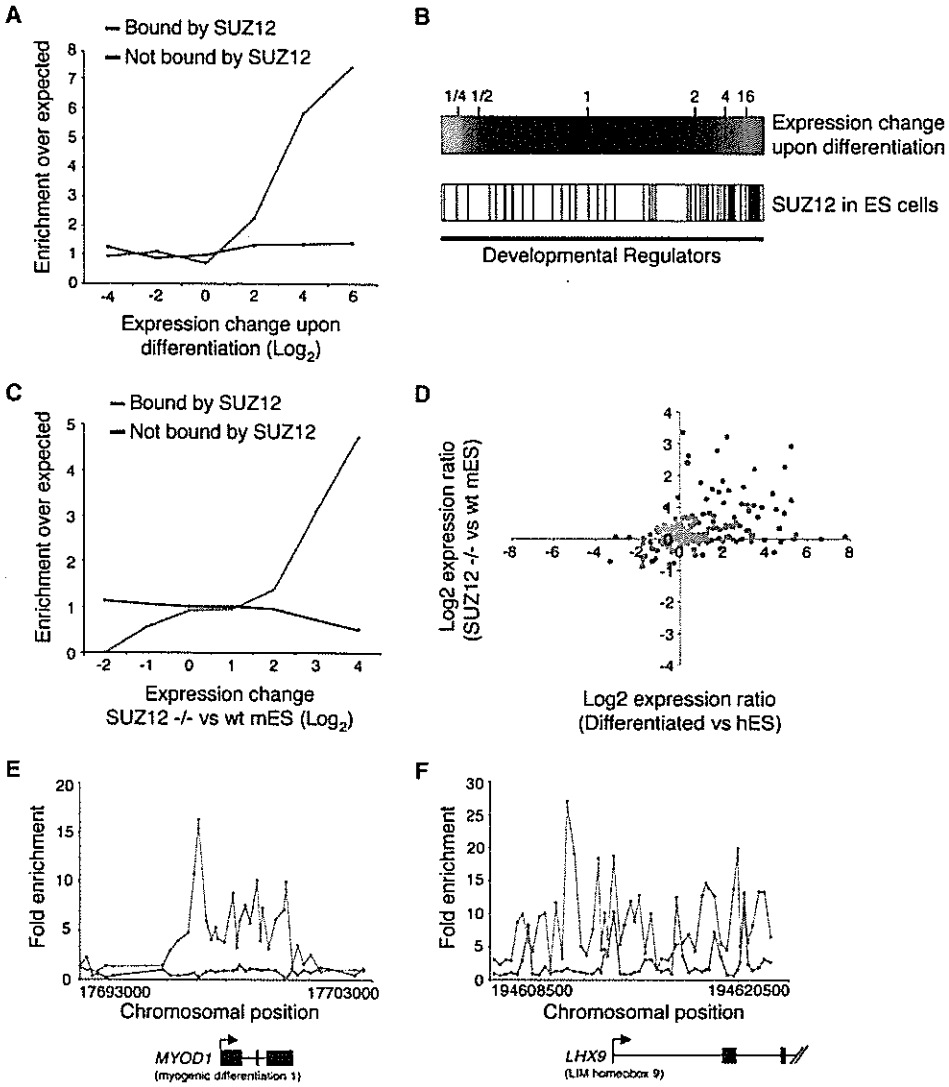

Figure 6. Preferential Activation of PRC2 Target Genes during ES Cell Differentiation

(A) Fold enrichment in the number of genes induced or repressed during ES cell differentiation. The change in gene expression is given as the log(2) transformed ratio of the signals in differentiated H1 cells versus pluripotent H1 cells and is binned into six groups. The upper limit of each bin is indicated on the *x* axis. The two lines show genes transcriptionally inactive in ES cells (absence of RNA polymerase II) and bound by SUZ12 (green) and genes transcriptionally inactive in ES cells and repressed by other means (blue). In both cases, fold enrichment is calculated against the total population of genes and normalized for the number of genes present in each group.

(B) Expression changes of genes encoding developmental regulators during ES cell differentiation. Expression ratio (differentiated/pluripotent) is represented by color, with shades of red indicating upregulation and shades of green downregulation according to the scale shown above. Genes are ordered according to change in gene expression, with genes exhibiting higher expression in pluripotent ES cells to the left and genes exhibiting higher expression in differentiated cells to the right. Genes bound by SUZ12 in undifferentiated ES cells are indicated by blue lines in the lower panel.

(C) Fold enrichment in the number of genes induced or repressed in SUZ12-deficient mouse cells. The change in gene expression is given as the log(2) transformed ratio of the signals in Suz12-deficient cells versus wild-type ES cells. The two lines show genes transcriptionally inactive in human ES cells (absence of RNA polymerase II) and bound by SUZ12 (green) and genes transcriptionally inactive in human ES cells and repressed by other means (blue). In both cases, fold enrichment is calculated against the total population of genes.

(D) Gene expression ratios (log base 2) of Suz12 target genes in differentiated human H1 ES cells relative to pluripotent H1 ES cells (*x* axis) and in Suz12-deficient mouse cells relative to wild-type mouse ES cells (*y* axis). Upper right quadrant: genes upregulated during human ES cell differentiation and in Suz12-deficient mouse cells; lower right: genes upregulated during ES cell differentiation and downregulated in Suz12-deficient cells; lower left: genes downregulated during ES cell differentiation and in Suz12-deficient cells; upper left: genes downregulated during ES cell differentiation and upregulated in Suz12-deficient cells.

(E) SUZ12 binding profiles across the gene encoding muscle regulator MYOD1 in H9 human ES cells (green) and primary human skeletal myotubes (gray). The plots show unprocessed enrichment ratios for all probes within a genomic region (ChIP versus whole genomic DNA). Genes are shown to scale below plots (exons are represented by vertical bars). The start and direction of transcription are noted by arrows.

are required for gastrulation and lineage differentiation in the embryo, are also essential for self-renewal and differentiation of ES cells in culture (Loebel et al., 2003; Molofsky et al., 2004). SUZ12 generally occupied the promoters of multiple components of these pathways, but it occupied larger domains within a group of signaling genes that contained highly conserved elements. This group contained members of the Wnt family (*WNT1*, *WNT2*, *WNT6*) as well as components of the TGFβ superfamily (*BMP2*, *GDF6*). Recent studies have shown that Wnt signaling plays a role in pluripotency and self-renewal in both mouse and human ES cells (Sato et al., 2004), and our results suggest that it is important to maintain specific family members in a repressed state in ES cells.

Activation of PRC2 Target Genes during Differentiation

PRC2 is associated with an important set of developmental regulators that must be silent in ES cells but activated during differentiation. This observation suggests that PRC2 ultimately functions to repress occupied genes in ES cells and that these genes may be especially poised for transcriptional activation during ES cell differentiation. We reasoned that if this model is correct, genes bound by SUZ12 should be preferentially activated upon ES cell differentiation or in cells that lack SUZ12. Furthermore, in differentiated cells, SUZ12 might continue to be observed at silent genes but must be removed from genes whose expression is essential for that cell type.

We first examined gene expression in ES cells stimulated to undergo differentiation (Sato et al., 2003). We found that genes occupied by SUZ12 were more likely to be activated during ES cell differentiation than genes that were not occupied by SUZ12 (Figure 6A; Supplemental Data; Table S13), indicating that SUZ12-occupied genes show preferential activation during differentiation under these conditions. Thirty-six percent of genes bound by SUZ12 showed greater than 2-fold increases in expression during ES cell differentiation, whereas only 16% of genes not bound by SUZ12 showed such an increase. This effect was particularly striking at the set of developmental regulators (Figure 6B). SUZ12 occupied most (83%) of the developmental regulators that were induced more than 10-fold during ES cell differentiation.

We next examined the expression of SUZ12 target genes in Suz12-deficient cell lines derived from homozygous mutant blastocysts (Supplemental Data). We reasoned that genes bound by SUZ12 in human ES cells have orthologs in mice that should be upregulated in Suz12-deficient mouse cells, although we expected the overlap in these sets of genes to be imperfect because of potential differences between human and mouse ES cells, the possible repression of PRC2 target genes by additional mechanisms, and pleiotropic effects of the Suz12 knockout on genes downstream of Suz12-target genes. Differences in gene expression between Suz12 homozygous mutant cells and wild-type ES cells were measured using gene expression microarrays and the human SUZ12 binding data mapped to orthologous mouse genes using HomoloGene (www.ncbi.nlm.nih.gov/HomoloGene). We found that a significant portion of mouse genes whose counterparts were bound by SUZ12 in human ES cells were upregulated in Suz12-deficient mouse cells (70 of 346 genes, $p = 6 \times 10^{-4}$); these genes are listed in Table S14. Orthologs of genes occupied by SUZ12 in human ES cells were more likely to be activated and less likely to be repressed in Suz12-deficient mouse cells than orthologs of genes not occupied by SUZ12 (Figure 6C). Furthermore, we found that orthologs of Suz12 target genes that were induced upon human ES cell differentiation were generally also induced upon loss of Suz12 in mouse cells (Figure 6D). Genes that were activated during ES cell differentiation and in Suz12-deficient cells included those encoding transcriptional regulators (*GATA2*, *GATA3*, *GATA6*, *HAND1*, *MEIS2*, and *SOX17*) signaling proteins (*WNT5A*, *DKK1*, *DKK2*, *EFNA1*, *EFNB1*, *EPHA4*, and *EPHB3*) and the cell-cycle inhibitor *CDKN1A*. These data indicate that Suz12 is necessary to fully repress the genes that are occupied by PRC2 in wild-type ES cells and have since been confirmed with binding data and knockout studies of a second PRC subunit in mouse (Boyer et al., 2006).

If PRC2 functions to repress genes in ES cells that are activated during differentiation, then in differentiated tissues SUZ12 occupancy should be diminished at genes encoding developmental regulators that have a role in specifying the identity of that tissue, similar to results seen with Ezh2 at specific genes in mouse (Caretti et al., 2004). To test this, we designed an array focused on the promoters of developmental regulators and used ChIP-Chip to investigate SUZ12 occupancy at these promoters in primary differentiated muscle cells. The results demonstrated that genes encoding key regulators of muscle differentiation, including *MYOD1*, displayed greatly diminished SUZ12 occupancy when compared to ES cells (Figure 6E). MYOD1 is a master regulator for muscle differentiation (Tapscott, 2005), and the gene encoding this transcription factor displayed no significant SUZ12 occupancy when compared to the levels of SUZ12 occupancy observed in ES cells. Genes encoding other transcriptional regulators that play a central role in muscle development, such as *PAX3* and *PAX7* (Brand-Saberi, 2005), showed reduced levels of SUZ12 occupancy in muscle cells relative to ES cells (Supplemental Data and Figure S11). In contrast, other developmental regulators important for differentiation of nonmuscle tissues remained occupied by SUZ12 in differentiated muscle cells (Figure 6F and Table S15). These data

(F) Suz12 binding profiles across the gene encoding LHX9 in H9 human ES cells (green) and primary human skeletal myotubes (gray). The plots show unprocessed enrichment ratios for all probes within a genomic region (ChIP versus whole genomic DNA). Genes are shown to scale below plots (exons are represented by vertical bars). The start and direction of transcription are noted by arrows.

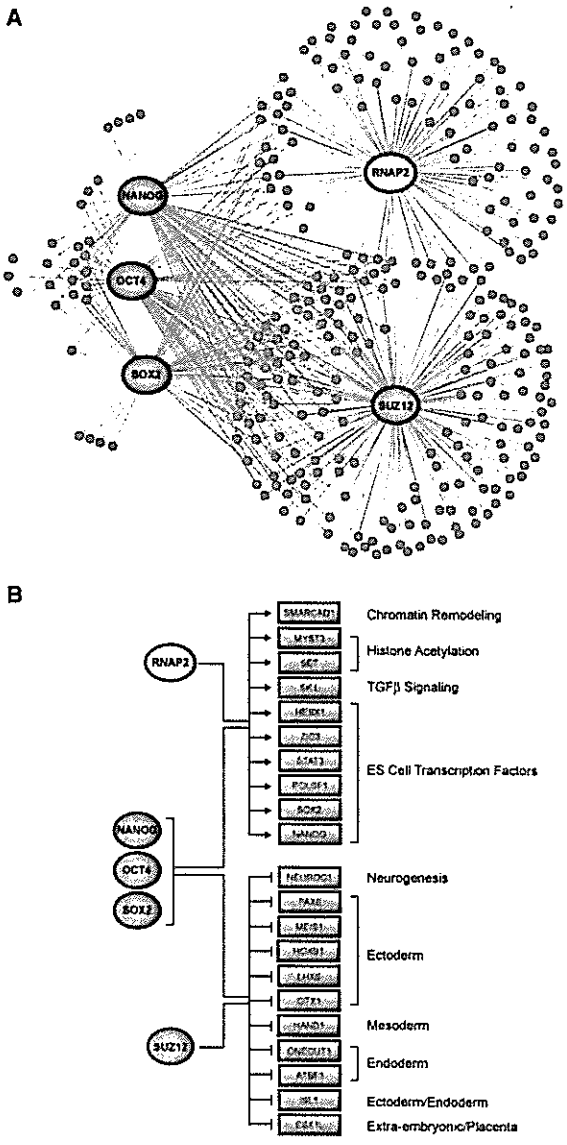

Figure 7. SUZ12 Is Localized to Genes also Bound by ES Cell Transcriptional Regulators

(A) Transcriptional regulatory network model of developmental regulators governed by OCT4, SOX2, NANOG, RNA polymerase II, and SUZ12 in human ES cells. The ES cell transcription factors each bound to approximately one-third of the PRC2-occupied, developmental transcription factor genes. Developmental regulators were selected based on gene ontology. Regulators are represented by dark blue circles; RNA polymerase II is represented by a light blue circle; SUZ12 is represented by a green circle; gene promoters for developmental regulators are represented by small red circles.

(B) SUZ12 occupies a set of repressed developmental regulators also bound by OCT4, SOX2, and NANOG in human ES cells. Genes annotated as bound by OCT4, SOX2, and NANOG previously and identified as active or repressed based on expression data (Boyer et al., 2005) were tested to see if they were bound by SUZ12 or RNA polymerase II. Ten of eleven previously identified active genes were found to be bound by RNA polymerase II at known promoters, while eleven of twelve previously identified repressed genes were bound by SUZ12. Regulators are represented by dark blue circles, RNA polymerase II by a light blue circle, and SUZ12 by a green circle. Gene promoters are represented by red rectangles.

support a model where PRC2 binding in ES cells represses key developmental regulators that are later expressed during differentiation.

Targets of PRC2 Are Shared with Key ES Cell Regulators

The transcription factors OCT4, SOX2, and NANOG have essential roles in early development and are required for the propagation of undifferentiated ES cells in culture (Nichols et al., 1998; Avilion et al., 2003; Chambers et al., 2003; Mitsui et al., 2003). We recently reported that these transcription factors occupied promoters for many important developmental regulators in human ES cells (Boyer et al., 2005). This led us to compare the set of genes encoding developmental regulators and occupied by OCT4, SOX2, and NANOG with those occupied by PRC2 (Figure 7 and Supplemental Data). We found that each of the three DNA binding transcription factors occupied approximately one-third of the PRC2-occupied genes that encode developmental transcription factors (Figure 7A; Supplemental Data; Table S11). Remarkably, we found that the subset of genes encoding developmental regulators that were occupied by OCT4, SOX2, and NANOG and repressed in the regulatory circuitry highlighted in Boyer et al. were almost all occupied by PRC2 (Figure 7B). These included genes for transcription factors known to be important for differentiation into extraembryonic, endodermal, mesodermal, and ectodermal lineages (e.g., *ESX1L, ONECUT1, HAND1, HOXB1*). As expected, active genes encoding ES cell transcription factors (e.g., *ZIC3, STAT3, OCT4, NANOG*) were occupied by OCT4, SOX2, NANOG, and RNA polymerase II but not by PRC2 (Figure 7B).

The observation that OCT4, SOX2, and NANOG are bound to a significant subset of developmental genes occupied by PRC2 supports a link between repression of developmental regulators and stem cell pluripotency. Like PRC2, OCT4 and NANOG have been shown to be important for early development and ES cell identity. It is possible, therefore, that inappropriate regulation of developmental regulators that are common targets of OCT4, NANOG, and PRC2 contributes to the inability to establish ES cell lines in OCT4, NANOG, and EZH2 mutants (Nichols et al., 1998; O'Carroll et al., 2001; Chambers et al., 2003; Mitsui et al., 2003).

Concluding Remarks

We have mapped the sites occupied by SUZ12 throughout the genome to gain insights into how PRC2 contributes to pluripotency in human embryonic stem cells. ES cells proliferate in an undifferentiated state yet remain poised to respond to development cues. Genes encoding the transcriptional regulators that promote differentiation must therefore be repressed in ES cells but activated

upon receiving signals to differentiate. We found that PRC2 occupies large domains at genes encoding a key set of repressed developmental regulators that are preferentially activated upon cellular differentiation, thus implicating this complex directly in the maintenance of the pluripotent state.

Transcription factors and chromatin regulators contribute to the transcriptional regulatory circuitry responsible for pluripotency and self-renewal in human ES cells. Understanding this circuitry is fundamental to understanding human development and realizing the therapeutic potential of these cells. In this context, we find it exciting that the outlines of the core transcriptional regulatory circuitry of human ES cells are emerging. The transcription factors OCT4, SOX2, and NANOG are associated with actively transcribed genes that contribute to growth and self-renewal (Boyer et al., 2005). These factors also occupy genes encoding key developmental regulators that are transcriptionally repressed, due at least in part to their association with PRC2 and nucleosomes modified at histone H3K27me3. Further study of transcription factors and chromatin regulators genome-wide will allow investigators to produce a more comprehensive map of transcriptional regulatory circuitry in ES cells and to test models that emerge from the circuitry. This information may provide insights into approaches by which pluripotent cells can be stimulated to differentiate into different cell types.

EXPERIMENTAL PROCEDURES

Cells and Cell Culture

Human H9 ES cells (WiCell, Madison, WI) were cultured as described (Boyer et al., 2005). Primary human skeletal muscle cells were obtained from Cell Applications (San Diego, CA) and expanded and differentiated into myotubes according to the supplier's protocols. Suz12 −/− mouse cell lines were derived from blastocysts from crosses between heterozygous Suz12 mutant animals, as described in Supplemental Data.

Chromatin Immunoprecipitation and DNA Microarray Analysis

ChIP was combined with DNA microarray analysis as described (Boyer et al., 2005). The antibodies used here were specific for hypophosphorylated RNA polymerase II (8WG16) (Thompson et al., 1989), SUZ12 (Upstate, 07-379), EED (Hamer et al., 2002), H3K27me3 (Abcam, AB6002), and total histone H3 (Abcam, AB1791). The design of the oligo-based arrays, which were manufactured by Agilent Technologies, is described in detail in Supplemental Data. A whole-chip error model was used to calculate confidence values from the enrichment ratio and the signal intensity of each probe (probe p-value) and of each set of three neighboring probes (probe-set p-value). Probe-sets with significant probe-set p-values ($p < 0.001$) and significant individual probe p-values were judged to be bound (see Supplemental Data for additional information). Bound regions were assigned to genes if they were within 1 kb of the transcription start site from one of five genomic databases; RefSeq, MGC, Ensembl, UCSC Known Gene, or H-Inv. All microarray data is available at ArrayExpress under the accession designation E-WMIT-7.

Gene Expression Analysis

Gene expression data were collated from H1 ES cells (Sato et al., 2003), H9, HSF1, and HSF6 ES cells (Abeyta et al., 2004), and 79 differentiated human cell and tissue types (Su et al., 2004) and analyzed as described in detail in Supplemental Data. Replicate gene expression data was obtained for wild-type mouse ES cells and Suz12-deficient cells using Agilent Mouse Development arrays and were analyzed as described in Supplemental Data.

Supplemental Data

Supplemental Data include fifteen figures, fifteen tables, Experimental Procedures, and References and can be found with this article online at http://www.cell.com/cgi/content/full/125/2/301/DC1/.

ACKNOWLEDGMENTS

We thank Elizabeth Jacobsen for technical assistance and Robert Brady for help with array design. L.A.B. and H.L.M. were supported by NRSA postdoctoral fellowships. M.G.G. is an Amgen Fellow of LSRF. R.M.K. was supported by a fellowship from the ACS. D.T.O. was supported by NIH award DK070813. This work was supported by NIH grants HG002668 and GM069400. T.L., T.L.V., D.K.G., and R.A.Y. consult for Agilent Technologies.

Received: October 25, 2005
Revised: January 20, 2006
Accepted: February 23, 2006
Published: April 20, 2006

REFERENCES

Abeyta, M.J., Clark, A.T., Rodriguez, R.T., Bodnar, M.S., Pera, R.A., and Firpo, M.T. (2004). Unique gene expression signatures of independently-derived human embryonic stem cell lines. Hum. Mol. Genet. *13*, 601–608.

Akasaka, T., van Lohuizen, M., van der Lugt, N., Mizutani-Koseki, Y., Kanno, M., Taniguchi, M., Vidal, M., Alkema, M., Berns, A., and Koseki, H. (2001). Mice doubly deficient for the Polycomb Group genes Mel18 and Bmi1 reveal synergy and requirement for maintenance but not initiation of Hox gene expression. Development *128*, 1587–1597.

Avilion, A.A., Nicolis, S.K., Pevny, L.H., Perez, L., Vivian, N., and Lovell-Badge, R. (2003). Multipotent cell lineages in early mouse development depend on SOX2 function. Genes Dev. *17*, 126–140.

Bejerano, G., Pheasant, M., Makunin, I., Stephen, S., Kent, W.J., Mattick, J.S., and Haussler, D. (2004). Ultraconserved elements in the human genome. Science *304*, 1321–1325.

Bender, M., Turner, F.R., and Kaufman, T.C. (1987). A development genetic analysis of the gene regulator of postbithorax in *Drosophila melanogaster*. Dev. Biol. *119*, 418–432.

Birve, A., Sengupta, A.K., Beuchle, D., Larsson, J., Kennison, J.A., Rasmuson-Lestander, A., and Muller, J. (2001). Su(z)12, a novel *Drosophila* Polycomb group gene that is conserved in vertebrates and plants. Development *128*, 3371–3379.

Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., et al. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. Cell *122*, 947–956.

Boyer, L.A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L.A., Lee, T.I., Levine, S.S., Wernig, M., Tajonar, A., Ray, M.K., Otte, A.P., Vidal, M., Gifford, D.K., Young, R.A., and Jaenisch, R. (2006). Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature, in press.

Brand-Saberi, B. (2005). Genetic and epigenetic control of skeletal muscle development. Ann. Anat. *187*, 199–207.

Breiling, A., Turner, B.M., Bianchi, M.E., and Orlando, V. (2001). General transcription factors bind promoters repressed by Polycomb group proteins. Nature *412*, 651–655.

Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R.S., and Zhang, Y. (2002). Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science *298*, 1039–1043.

Cao, R., and Zhang, Y. (2004). SUZ12 is required for both the histone methyltransferase activity and the silencing function of the EED-EZH2 complex. Mol. Cell *15*, 57–67.

Caretti, G., Di Padova, M., Micales, B., Lyons, G.E., and Sartorelli, V. (2004). The Polycomb Ezh2 methyltransferase regulates muscle gene expression and skeletal muscle differentiation. Genes Dev. *18*, 2627–2638.

Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S., and Smith, A. (2003). Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. Cell *113*, 643–655.

Czermin, B., Melfi, R., McCabe, D., Seitz, V., Imhof, A., and Pirrotta, V. (2002). *Drosophila* enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. Cell *111*, 185–196.

Dellino, G.I., Schwartz, Y.B., Farkas, G., McCabe, D., Elgin, S.C., and Pirrotta, V. (2004). Polycomb silencing blocks transcription initiation. Mol. Cell *13*, 887–893.

Denell, R.E., and Frederick, R.D. (1983). Homoeosis in Drosophila: a description of the Polycomb lethal syndrome. Dev. Biol. *97*, 34–47.

Duncan, I. (1986). Control of bithorax complex functions by the segmentation gene fushi tarazu of D. melanogaster. Cell *47*, 297–309.

Faust, C., Lawson, K.A., Schork, N.J., Thiel, B., and Magnuson, T. (1998). The Polycomb-group gene eed is required for normal morphogenetic movements during gastrulation in the mouse embryo. Development *125*, 4495–4506.

Francis, N.J., Saurin, A.J., Shao, Z., and Kingston, R.E. (2001). Reconstitution of a functional core polycomb repressive complex. Mol. Cell *8*, 545–556.

Franke, A., DeCamillis, M., Zink, D., Cheng, N., Brock, H.W., and Paro, R. (1992). Polycomb and polyhomeotic are constituents of a multimeric protein complex in chromatin of Drosophila melanogaster. EMBO J. *11*, 2941–2950.

Hamer, K.M., Sewalt, R.G., den Blaauwen, J.L., Hendrix, T., Satijn, D.P., and Otte, A.P. (2002). A panel of monoclonal antibodies against human polycomb group proteins. Hybrid. Hybridomics *21*, 245–252.

Hodgson, J.W., Argiropoulos, B., and Brock, H.W. (2001). Site-specific recognition of a 70-base-pair element containing d(GA)(n) repeats mediates bithoraxoid polycomb group response element-dependent silencing. Mol. Cell. Biol. *21*, 4528–4543.

Horard, B., Tatout, C., Poux, S., and Pirrotta, V. (2000). Structure of a polycomb response element and in vitro binding of polycomb group complexes containing GAGA factor. Mol. Cell. Biol. *20*, 3187–3197.

Kennison, J.A. (2004). Introduction to Trx-G and Pc-G genes. Methods Enzymol. *377*, 61–70.

Kirmizis, A., Bartley, S.M., Kuzmichev, A., Margueron, R., Reinberg, D., Green, R., and Farnham, P.J. (2004). Silencing of human polycomb target genes is associated with methylation of histone H3 Lys 27. Genes Dev. *18*, 1592–1605.

Kuzmichev, A., Nishioka, K., Erdjument-Bromage, H., Tempst, P., and Reinberg, D. (2002). Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. Genes Dev. *16*, 2893–2905.

Kuzmichev, A., Jenuwein, T., Tempst, P., and Reinberg, D. (2004). Different EZH2-containing complexes target methylation of histone H1 or nucleosomal histone H3. Mol. Cell *14*, 183–193.

Kuzmichev, A., Margueron, R., Vaquero, A., Preissner, T.S., Scher, M., Kirmizis, A., Ouyang, X., Brockdorff, N., Abate-Shen, C., Farnham, P., and Reinberg, D. (2005). Composition and histone substrates of polycomb repressive group complexes change during cellular differentiation. Proc. Natl. Acad. Sci. USA *102*, 1859–1864.

Lehmann, O.J., Sowden, J.C., Carlsson, P., Jordan, T., and Bhattacharya, S.S. (2003). Fox's in development and disease. Trends Genet. *19*, 339–344.

Levine, S.S., Weiss, A., Erdjument-Bromage, H., Shao, Z., Tempst, P., and Kingston, R.E. (2002). The core of the polycomb repressive complex is compositionally and functionally conserved in flies and humans. Mol. Cell. Biol. *22*, 6070–6078.

Levine, S.S., King, I.F., and Kingston, R.E. (2004). Division of labor in polycomb group repression. Trends Biochem. Sci. *29*, 478–485.

Lewis, E.B. (1978). A gene complex controlling segmentation in Drosophila. Nature *276*, 565–570.

Loebel, D.A., Watson, C.M., De Young, R.A., and Tam, P.P. (2003). Lineage choice and differentiation in mouse embryos and embryonic stem cells. Dev. Biol. *264*, 1–14.

Lund, A.H., and van Lohuizen, M. (2004). Polycomb complexes and silencing mechanisms. Curr. Opin. Cell Biol. *16*, 239–246.

Mayhall, E.A., Paffett-Lugassy, N., and Zon, L.I. (2004). The clinical potential of stem cells. Curr. Opin. Cell Biol. *16*, 713–720.

Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., and Yamanaka, S. (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell *113*, 631–642.

Molofsky, A.V., Pardal, R., and Morrison, S.J. (2004). Diverse mechanisms regulate stem cell self-renewal. Curr. Opin. Cell Biol. *16*, 700–707.

Mulholland, N.M., King, I.F., and Kingston, R.E. (2003). Regulation of Polycomb group complexes by the sequence-specific DNA binding proteins Zeste and GAGA. Genes Dev. *17*, 2741–2746.

Muller, J., Hart, C.M., Francis, N.J., Vargas, M.L., Sengupta, A., Wild, B., Miller, E.L., O'Connor, M.B., Kingston, R.E., and Simon, J.A. (2002). Histone methyltransferase activity of a *Drosophila* Polycomb group repressor complex. Cell *111*, 197–208.

Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Scholer, H., and Smith, A. (1998). Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. Cell *95*, 379–391.

O'Carroll, D., Erhardt, S., Pagani, M., Barton, S.C., Surani, M.A., and Jenuwein, T. (2001). The polycomb-group gene Ezh2 is required for early mouse development. Mol. Cell. Biol. *21*, 4330–4336.

Orlando, V., and Paro, R. (1995). Chromatin multiprotein complexes involved in the maintenance of transcription patterns. Curr. Opin. Genet. Dev. *5*, 174–179.

Pasini, D., Bracken, A.P., Jensen, M.R., Denchi, E.L., and Helin, K. (2004). Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. EMBO J. *23*, 4061–4071.

Pera, M.F., and Trounson, A.O. (2004). Human embryonic stem cells: prospects for development. Development *131*, 5515–5525.

Pirrotta, V. (1998). Polycombing the genome: PcG, trxG, and chromatin silencing. Cell *93*, 333–336.

Reubinoff, B.E., Pera, M.F., Fong, C.Y., Trounson, A., and Bongso, A. (2000). Embryonic stem cell lines from human blastocysts: somatic differentiation in vitro. Nat. Biotechnol. *18*, 399–404.

Ringrose, L., and Paro, R. (2004). Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. Annu. Rev. Genet. *38*, 413–443.

Sato, N., Sanjuan, I.M., Heke, M., Uchida, M., Naef, F., and Brivanlou, A.H. (2003). Molecular signature of human embryonic stem cells and its comparison with the mouse. Dev. Biol. *260*, 404–413.

Sato, N., Meijer, L., Skaltsounis, L., Greengard, P., and Brivanlou, A.H. (2004). Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by a pharmacological GSK-3-specific inhibitor. Nat. Med. *10*, 55–63.

Saurin, A.J., Shao, Z., Erdjument-Bromage, H., Tempst, P., and Kingston, R.E. (2001). A Drosophila Polycomb group complex includes Zeste and dTAFII proteins. Nature *412*, 655–660.

Schepers, G.E., Teasdale, R.D., and Koopman, P. (2002). Twenty pairs of sox: extent, homology, and nomenclature of the mouse and human sox transcription factor gene families. Dev. Cell *3*, 167–170.

Shao, Z., Raible, F., Mollaaghababa, R., Guyon, J.R., Wu, C.T., Bender, W., and Kingston, R.E. (1999). Stabilization of chromatin structure by PRC1, a Polycomb complex. Cell *98*, 37–46.

Showell, C., Binder, O., and Conlon, F.L. (2004). T-box genes in early embryogenesis. Dev. Dyn. *229*, 201–218.

Siepel, A., Bejerano, G., Pedersen, J.S., Hinrichs, A.S., Hou, M., Rosenbloom, K., Clawson, H., Spieth, J., Hillier, L.W., Richards, S., et al. (2005). Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Res. *15*, 1034–1050.

Simon, J., Chiang, A., and Bender, W. (1992). Ten different Polycomb group genes are required for spatial control of the abdA and AbdB homeotic products. Development *114*, 493–505.

Strutt, H., Cavalli, G., and Paro, R. (1997). Co-localization of Polycomb protein and GAGA factor on regulatory elements responsible for the maintenance of homeotic gene expression. EMBO J. *16*, 3621–3632.

Su, A.I., Wiltshire, T., Batalov, S., Lapp, H., Ching, K.A., Block, D., Zhang, J., Soden, R., Hayakawa, M., Kreiman, G., et al. (2004). A gene atlas of the mouse and human protein-encoding transcriptomes. Proc. Natl. Acad. Sci. USA *101*, 6062–6067.

Tapscott, S.J. (2005). The circuitry of a master switch: Myod and the regulation of skeletal muscle gene transcription. Development *132*, 2685–2695.

Thompson, N.E., Steinberg, T.H., Aronson, D.B., and Burgess, R.R. (1989). Inhibition of in vivo and in vitro transcription by monoclonal antibodies prepared against wheat germ RNA polymerase II that react with the heptapeptide repeat of eukaryotic RNA polymerase II. J. Biol. Chem. *264*, 11511–11520.

Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. Science *282*, 1145–1147.

Tie, F., Furuyama, T., Prasad-Sinha, J., Jane, E., and Harte, P.J. (2001). The *Drosophila* Polycomb Group proteins ESC and E(Z) are present in a complex containing the histone-binding protein p55 and the histone deacetylase RPD3. Development *128*, 275–286.

van der Lugt, N.M., Alkema, M., Berns, A., and Deschamps, J. (1996). The Polycomb-group homolog Bmi-1 is a regulator of murine Hox gene expression. Mech. Dev. *58*, 153–164.

Vire, E., Brenner, C., Deplus, R., Blanchon, L., Fraga, M., Didelot, C., Morey, L., Van Eynde, A., Bernard, D., Vanderwinden, J.M., Bollen, M., Esteller, M., Di Croce, L., de Launoit, Y., and Fuks, F. (2006). The Polycomb group protein EZH2 directly controls DNA methylation. Nature *439*, 871–874.

Wang, J., Mager, J., Schnedier, E., and Magnuson, T. (2002). The mouse PcG gene eed is required for Hox gene repression and extraembryonic development. Mamm. Genome *13*, 493–503.

Woolfe, A., Goodson, M., Goode, D.K., Snell, P., McEwen, G.K., Vavouri, T., Smith, S.F., North, P., Callaway, H., Kelly, K., et al. (2005). Highly conserved non-coding sequences are associated with vertebrate development. PLoS Biol. *3*, e7.

Development 133, 2371-2381 (2006) doi:10.1242/dev.02405

Distinct roles of Polycomb group gene products in transcriptionally repressed and active domains of *Hoxb8*

Yu-ichi Fujimura[1,*], Kyo-ichi Isono[1,*], Miguel Vidal[2], Mitsuhiro Endoh[1], Hiroshi Kajita[1], Yoko Mizutani-Koseki[1], Yoshihiro Takihara[3], Maarten van Lohuizen[4], Arie Otte[5], Thomas Jenuwein[6], Jacqueline Deschamps[7] and Haruhiko Koseki[1,†]

To address the molecular mechanisms underlying Polycomb group (PcG)-mediated repression of Hox gene expression, we have focused on the binding patterns of PcG gene products to the flanking regions of the *Hoxb8* gene in expressing and non-expressing tissues. In parallel, we followed the distribution of histone marks of transcriptionally active H3 acetylated on lysine 9 (H3-K9) and methylated on lysine 4 (H3-K4), and of transcriptionally inactive chromatin trimethylated on lysine 27 (H3-K27). Chromatin immunoprecipitation revealed that the association of PcG proteins, and H3-K9 acetylation and H3-K27 trimethylation around *Hoxb8* were distinct in tissues expressing and not expressing the gene. We show that developmental changes of these epigenetic marks temporally coincide with the misexpression of Hox genes in PcG mutants. Functional analyses, using mutant alleles impairing the PcG class 2 component Rnf2 or the *Suz12* mutation decreasing H3-K27 trimethylation, revealed that interactions between class 1 and class 2 PcG complexes, mediated by trimethylated H3-K27, play decisive roles in the maintenance of Hox gene repression outside their expression domain. Within the expression domains, class 2 PcG complexes appeared to maintain the transcriptionally active status via profound regulation of H3-K9 acetylation. The present study indicates distinct roles for class 2 PcG complexes in transcriptionally repressed and active domains of *Hoxb8* gene.

KEY WORDS: Polycomb, Hox, Mouse, Chromatin, Immunoprecipitation

INTRODUCTION

In *Drosophila*, the regionally restricted expression of homeotic genes (Hox genes) is maintained by two groups of proteins: the trithorax group (trxG) and Polycomb group (PcG). The trxG gene products are required to maintain the activity of Hox genes in the appropriate segments, whereas the PcG are involved in their repression (Paro, 1995; Pirrotta, 1997). Indeed, some of the PcG genes have been identified as suppressors of trxG mutations (Kennison and Tamkun, 1988). The *trxG* and *PcG* genes encode nuclear factors that, by forming multimeric protein complexes on the chromatin, freeze transcriptional states determined early in embryogenesis. A major function of trxG gene products appears to concern the remodeling of chromatin structure, as several of these genes encode subunits of the SWI/SNF chromatin remodeling complex and associate with catalytic activities that modify core histone tails (Milne et al., 2002; Nakamura et al., 2002; Tamkun et al., 1992). PcG gene products have been shown to form at least two types of multimeric protein complexes (Shao et al., 1999; Czermin et al., 2002; Muller et al., 2002; Cao et al., 2002). The first type of complex, known as class 1, includes Extra sex combs (Esc) and Enhancer of zeste [E(z)], and its association with histone deacetylase and methyltransferase activities suggests a function in the modification of histone tails (Czermin et al., 2002; Muller et al., 2002; Cao et al., 2002). The second type of complex (class 2) contains Polycomb (Pc), Posterior sex combs (Psc), Polyhomeotic (Ph), Dring (Sex comb extra; Sce) and Sex comb on midleg (Scm) in *Drosophila* (Shao et al., 1999). It has been shown that, by interacting with DNA regulatory sequences termed PcG response elements (PREs), this complex can maintain the transcriptional silence of target genes. Recently, it has been reported that class 1 PcG-mediated trimethylation of histone H3 lysine 27 (H3-K27) serves as a signal for the recruitment of class 2 PcG complexes, which in turn may silence target genes flanked by PREs (Cavalli and Paro, 1998; Wang et al., 2004b). Purified class 2 core complexes efficiently block SWI/SNF-dependent remodeling of nucleosomal arrays and subsequent transcription in vitro (King et al., 2002; Levine et al., 2002; Shao et al., 1999).

In mammals, it has been demonstrated that the expression of Hox cluster genes depends on the control of mammalian homologues of trxG and both classes of PcG gene products. Silencing of several Hox gene expressions has been observed in mice lacking *Mll*, a mammalian homologue of *Drosophila trx*, between 8.5 and 9.5 days post coitus (dpc) (Yu et al., 1998). Conversely, a hypomorphic mutation of *Eed*, a homologue of *Drosophila Esc*, revealed its involvement in repressing Hox gene expression (Schumacher et al., 1996). As Mll and class 1 PcG complex are known to be associated with activities that modify histone tails (van der Vlag and Otte, 1999), such modifications could be involved in maintaining spatially restricted expression of Hox cluster genes. Indeed, it has been shown that the Mll protein regulates the acetylation of lysine 4 of histone H3 (H3-K4) at several Hox genes, whereas Ezh2, a homologue of E(z), mediates H3-K27 trimethylation (Milne et al., 2002; Nakamura et al., 2002). The involvement of class 2 PcG in this maintenance has been revealed previously by the effect of mutations

[1]RIKEN Research Center for Allergy and Immunology, 1-7-22 Suehiro, Tsurumi-ku, Yokohama 230-0045, Japan. [2]Centro de Investigaciones Biologicas, Department of Developmental and Cell Biology, Ramiro de Maeztu 9, 28040 Madrid, Spain. [3]Department of Stem Cell Biology, Research Institute for Radiation Biology and Medicine, Hiroshima University, 1-2-3 Kasumi, Minami-ku, Hiroshima 734-8553, Japan. [4]Division of Molecular Genetics, The Netherlands Cancer Institute, 1066CX Amsterdam, The Netherlands. [5]Swammerdam Institute for Life Sciences, University of Amsterdam, Kruislaan 406, 1098 SM Amsterdam, The Netherlands. [6]Research Institute of Molecular Pathology, The Vienna Biocenter, Dr Bohrgasse 7, A-1030 Vienna, Austria. [7]Hubrecht Laboratory, Uppsalalaan 8 3584CT Utrecht, The Netherlands.

*These authors contributed equally to this work
[†]Author for correspondence (e-mail: fujimara@rcai.riken.jp)

Accepted 12 April 2006

in *Bmi1* (*Pcgf4* – Mouse Genome Informatics), *Rnf110* (*Mel18*; *Pcgf2* – Mouse Genome Informatics), *Phc1* (*rae28*) and *Rnf2* (*Ring1B*) on Hox gene expression (Akasaka et al., 1996; van der Lugt et al., 1996; Suzuki et al., 2002; Takihara et al., 1997). Notably, in mice doubly deficient for *Rnf110* and *Bmi1*, homologues of *Drosophila Psc*, *Hoxb6* expression was normally localized in caudal tissues of 8.5 dpc embryos, but was progressively de-repressed cranially thereafter (Akasaka et al., 2001). Thus, class 2 PcG proteins may participate in maintaining transcriptionally silent states of Hox genes outside their expression domains. However, to date, the molecular mechanisms used by mammalian trxG and PcG gene products to regulate Hox genes is not well understood, in part because the relationship between PcG protein binding and histone tail modifications has not yet been widely documented around the Hox loci in developing embryos. We have addressed this issue by documenting the association of PcG proteins, H3-K9 acetylation, H3-K4 methylation (marks of transcriptionally active chromatin) and H3-K27 trimethylation (a mark of transcriptionally inactive chromatin) to the genomic region flanking *Hoxb8*, a Hox gene known to require the class 2 PcG proteins for its posterior restriction.

We show that the association of PcG proteins, H3-K9 acetylation, H3-K4 methylation and H3-K27 trimethylation around *Hoxb8* differs in embryonic tissues expressing and not expressing the gene. By using mutant alleles for *Rnf2*, which encode constituents of class 2 PcG complexes, and *Suz12* causing a strong decrease in H3-K27 methylation, we show that the recruitment of the class 2 PcG complex, mediated by trimethylated H3-K27, plays a decisive role in maintaining the repression of Hox genes outside their expression domain, as it is the case in *Drosophila*. The positive role of class 2 PcG complex proteins in the transcriptionally active domain was shown to involve the regulation of H3-K9 acetylation.

MATERIALS AND METHODS

Mice and cell lines

The generation of Rnf2^{fl} allele and Rnf2 mutant ES cells has been described previously (de Napoles et al., 2004). The generation of *Suz12*-deficient mice and ES cells will be described in a separate publication. CRE-mediated deletion of Rnf2 in MEF cell lines using adenovirus expressing the CRE recombinase (AdCre) was carried out as described (Kanegae et al., 1995). AdCre virus was concentrated to be 2.0×10^{-9} i.f.u./ml and mouse embryonic fibroblast (MEFs) were infected with AdCre virus at MOI5.

Antibodies

Antibodies used are listed in the Table 1.

Chromatin immunoprecipitation (ChIP) analysis

Embryonic tissues of interest were dissected from embryos, at the different developmental stages, under a stereomicroscope. Chromatin fraction from the embryonic tissues was purified by CsCl isopycnic centrifugation as described in Fig. S1 in the supplementary material (Orlando et al., 1997). This process was bypassed in some cases. Correct amounts of NaCl and NP-40 were added to the chromatin fraction or whole cell lysates (WCE) in order to perform optimal immunoprecipitation for each antibody. Precleared protein extracts were incubated with the correct amount of antibodies, at 4°C with rocking, for 2 hours to overnight. Immune complexes were captured after 3 hours incubation with Protein A Sepharose beads.

To isolate genomic DNA from immune complexes, beads were treated with 50 μg/ml of RNaseA at 37°C for 30 minutes followed by overnight incubation with 500 μg/ml proteinase K/0.5% SDS at 37°C. After 3 hours heating at 65°C for reverse crosslinking, supernatants were collected, extracted by phenol-chloroform and concentrated by ethanol precipitation. Genomic DNA was also isolated from the original chromatin fraction or WCE through the same procedure as described above and designated as 'Input' DNA (see Figs 1-3). To measure the DNA yield after immunoprecipitation, the aliquots of immunoprecipitated DNA were electrophoresed for 5 minutes in an agarose gel, next to serially diluted input DNA and band-intensities were compared after ethidium bromide staining (see Fig. S2A in the supplementary material).

Equivalent amounts of immunoprecipitated DNA to that of 'Input' DNA loaded in lane '1' were subjected to PCR reactions. Usually, 10 to 20 ng of genomic DNA was used. Mock-immunoprecipitated DNA (A– and P–) derived from the same volume of the chromatin fraction as used for anti-Rnf2 immunoprecipitation were subjected to the PCR. To carry out semi-quantitative PCR, serially diluted 'Input' DNA and immunoprecipitated DNA were used as templates. The relative quantity of each genomic region in immunoprecipitated genomic DNA was estimated by referring to the serial dilutions of 'Input' DNA isolated from the initial lysates and an enrichment value was determined. Every series of experiments were performed at least three times (see Fig. S3 in the supplementary material). Primers used in this study are listed in Table 2. ChIP analysis by using ES cells was performed as described (Isono et al., 2005a).

Expression analyses for RNA and protein

RNA extraction from ES cells, reverse transcription and PCR reaction was performed as described previously (Isono et al., 2005a). Quantity of total cellular RNA subjected to reverse transcriptase (RT)-PCR for Hox genes was adjusted by the expression of β-actin. *Hoxb8* expression shown in Fig. 3A was quantified by referring to *Gapdh* expression by real-time PCR analyses using Mx3005P multiplex quantitative PCR systems (Stratagene). Preparation of whole-cell extracts from embryos, ES cells and MEFs and western blot analysis were performed as described previously (Isono et al., 2005a).

Table 1. A list of antibodies used in this study

Specificity	Species	Monoclonal or antiserum?	Company or reference
Rnf2	Mouse	Monoclonal	Atsuta et al. (2001)
Ring1	Rabbit	Antiserum	Schoorlemmer et al. (1997)
Phc1	Mouse	Monoclonal	Miyagishima et al. (2003)
Phc1	Rabbit	Antiserum	Suzuki et al. (2002)
Cbx2	Rabbit	Antiserum	Schoorlemmer et al. (1997)
Rnf110	Rabbit	Antiserum	Abcam (Cambridge, UK)
Eed	Mouse	Monoclonal	Hamer et al. (2002)
Ezh2	Mouse	Monoclonal	Hamer et al. (2002)
Suz12	Rabbit	Antiserum	Upstate Biotechnology (Lake Placid, NY)
Acetylated H3-K9	Rabbit	Antiserum	Upstate Biotechnology (Lake Placid, NY)
Dimethylated H3-K9	Rabbit	Antiserum	Peters et al. (2003)
Trimethylated H3-K9	Rabbit	Antiserum	Peters et al. (2003)
Dimethylated H3-K27	Rabbit	Antiserum	Peters et al. (2003)
Trimethylated H3-K27	Rabbit	Antiserum	Peters et al. (2003)
Dimethylated H3-K4	Rabbit	Antiserum	Upstate Biotechnology (Lake Placid, NY)
Trimethylated H3-K4	Rabbit	Antiserum	Upstate Biotechnology (Lake Placid, NY)