

Reduced Genomic Cytosine Methylation and Defective Cellular Differentiation in Embryonic Stem Cells Lacking CpG Binding Protein

Diana L. Carlone,[†] Jeong-Heon Lee,[†] Suzanne R. L. Young,[†] Erika Dobrota,[†]
Jill Sergesketter Butler,[†] Joseph Ruiz, and David G. Skalnik*

Herman B Wells Center for Pediatric Research, Section of Pediatric Hematology/Oncology, Departments of Pediatrics and Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, Indiana 46202

Received 11 February 2005/Returned for modification 11 March 2005/Accepted 21 March 2005

Cytosine methylation at CpG dinucleotides is a critical epigenetic modification of mammalian genomes. CpG binding protein (CGBP) exhibits a unique DNA-binding specificity for unmethylated CpG motifs and is essential for early murine development. Embryonic stem cell lines deficient for CGBP were generated to further examine CGBP function. *CGBP*^{-/-} cells are viable but show an increased rate of apoptosis and are unable to achieve *in vitro* differentiation following removal of leukemia inhibitory factor from the growth media. Instead, *CGBP*^{-/-} embryonic stem cells remain undifferentiated as revealed by persistent expression of the pluripotent markers Oct4 and alkaline phosphatase. *CGBP*^{-/-} cells exhibit a 60 to 80% decrease in global cytosine methylation, including hypo-methylation of repetitive elements, single-copy genes, and imprinted genes. Total DNA methyltransferase activity is reduced by 30 to 60% in *CGBP*^{-/-} cells, and expression of the maintenance DNA methyltransferase 1 protein is similarly reduced. However, *de novo* DNA methyltransferase activity is normal. Nearly all aspects of the pleiotropic *CGBP*^{-/-} phenotype are rescued by introduction of a CGBP expression vector. Hence, CGBP is essential for normal epigenetic modification of the genome by cytosine methylation and for cellular differentiation, consistent with the requirement for CGBP during early mammalian development.

The CpG dinucleotide is an important regulatory component of mammalian genomes. The cytosine of this dinucleotide serves as the target for methylation by DNA methyltransferase (Dnmt) enzymes, which functions as a critical epigenetic modification of DNA. Methylated DNA is correlated with heterochromatin and transcriptionally inactive genes, while actively expressed genes are generally hypomethylated (58). Cytosine methylation may also represent a defense mechanism to silence parasitic repetitive DNA elements present in mammalian genomes (72, 78). In addition, cytosine methylation is involved in the processes of genomic imprinting, in which paternal and maternal alleles of a gene exhibit distinct patterns of cytosine methylation and expression (65), and X chromosome inactivation, in which one X chromosome in each cell of a female becomes irreversibly inactivated during early development (52). The CpG dinucleotide is underrepresented in mammalian genomes (5 to 10% of the expected frequency), presumably due to the propensity of 5-methylcytosine to undergo spontaneous deamination to form thymine (8). Approximately 50% of human and mouse genes reside near unmethylated CpG islands, which contain the statistically expected frequency of CpG dinucleotides.

Global cytosine methylation patterns inherited from gametes are erased during early embryogenesis (morula), fol-

lowed by a wave of *de novo* DNA methylation in the blastocyst upon implantation (44). Dnmt3a and Dnmt3b are *de novo* methyltransferases that preferentially recognize unmethylated CpG motifs (49), while Dnmt1 is a maintenance methyltransferase that recognizes hemimethylated DNA (5), the immediate product of DNA replication. Appropriate cytosine methylation in mammals is essential for normal development. Individual ablation of the *Dnmt1* or *Dnmt3b* gene leads to a lethal disruption of murine embryonic development (38, 49). Mice lacking Dnmt3a develop to birth but become runted and die within 4 weeks of age (49). Furthermore, mutations that are predicted to partially inhibit Dnmt3b function are associated with the ICF (immunodeficiency, centromere instability, and facial anomalies) syndrome in humans (77). Overexpression of Dnmt1 in mice leads to global hypermethylation, loss of genomic imprinting, and embryonic lethality (7).

A number of DNA-binding factors interact with methylated CpG motifs (27). These include MeCP2, methyl binding domain protein 1 (MBD1) and MBD2, which are involved in repression of gene expression, and MBD4, which functions in DNA repair. Each of these factors contains a conserved methyl-CpG binding domain, but otherwise they exhibit little sequence similarity. Mutations in the methyl-CpG binding protein MeCP2 lead to Rett syndrome, a progressive neurodegenerative disorder (2).

Recent reports reveal intricate interrelationships linking cytosine methylation and histone modifications, thus providing a unifying framework for the control of chromatin structure and gene regulation (9). For example, MBD2 and MBD3 are components of the histone deacetylase (HDAC) complexes

* Corresponding author. Mailing address: Cancer Research Building, Room W327, 1044 West Walnut St., Indianapolis, IN 46202. Phone: (317) 274-8977. Fax: (317) 274-8679. E-mail: dskalnik@iupui.edu.

[†] These authors contributed equally to this work.

MeCP1 and Mi-2, respectively (46, 81), and Dnmt proteins also associate with HDAC complexes (20, 21, 55). Furthermore, the chromatin remodeling protein DDM1 in *Arabidopsis* and the related factor LSH in mammals are required for normal cytosine methylation (15, 29, 30). Disruption of the Suv39h histone methyltransferase gene in murine embryonic stem (ES) cells leads to altered localization of Dnmt3b and decreased cytosine methylation at pericentric satellite repeats (35). Hence, DNA methylation and histone modifications appear to be highly integrated and mutually reinforcing mechanisms that serve to maintain heterochromatin structure and repress gene expression.

CpG binding protein (CGBP) exhibits a unique DNA-binding specificity for unmethylated CpG motifs and acts as a transcriptional activator (71). Originally identified in humans, homologues of CGBP have been detected in *Drosophila*, *Caenorhabditis elegans*, and both *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* (41, 71). CGBP contains a cysteine-rich CXXC DNA-binding domain (34, 71) which is present in several other proteins, including Dnmt1 (6); human trithorax (HRX) (also known as ALL-1 or MLL), a histone methyltransferase encoded by a gene frequently involved in chromosomal translocations in leukemia (17, 25, 39, 53, 66, 80); MBD1 (14, 27); leukemia-associated protein LCX (50); and MLL-2, which is often amplified in solid tumors (19). CGBP additionally contains two PHD domains, which are characteristic of chromatin-associated proteins and/or regulators of gene expression (1, 71) and often mediate protein-protein interactions (22, 24, 48). CGBP is a component of the nuclear matrix and localizes to nuclear speckles associated with euchromatin (33). Targeted disruption of the *CGBP* gene results in peri-implantation embryonic lethality in mice (11), a developmental stage associated with global remodeling of chromatin structure and cytosine methylation patterns (32, 37, 54).

The molecular mechanisms involved in targeting methylation to specific CpG motifs during development, as well as maintaining hypomethylation of CpG islands, are not well understood. The binding specificity of CGBP for unmethylated CpG motifs suggests a possible role in these events. The early death of embryos lacking CGBP establishes the importance of this protein for mammalian development. However, the severity of this phenotype makes further analysis of this mutant difficult. In the study reported here, murine ES cells lacking CGBP were isolated to permit a more detailed analysis of the *CGBP*^{-/-} phenotype and provide further insight into CGBP function. The results presented implicate CGBP as a critical regulator of DNA methylation and cellular differentiation.

MATERIALS AND METHODS

Generation and analysis of ES cell lines. Six- to eight-week-old *CGBP*^{+/-} females were superovulated (7.5 to 10 IU pregnant mare serum gonadotropin followed 48 h later with 7.5 to 10 IU of human chorionic gonadotropin [Sigma-Aldrich Co., St. Louis, MO]) and mated with heterozygous males. Blastocysts (3.5 days postcoitum [dpc]) were collected and cultured for 4 days on mitomycin-treated STO feeder layers in ES culture media containing leukemia inhibitory factor (LIF). The inner cell mass was manually isolated, trypsinized, and replated onto mitomycin-treated STO feeder layers. ES cell clusters were expanded on mitomycin-treated STO feeder layers and maintained as ES cell lines. Approximately 20% of the blastocysts gave rise to ES cell lines. For subsequent studies, ES cells were cultured on gelatin-coated dishes in ES cell media containing LIF.

The genotypes of ES cell lines were determined by Southern blot analysis. Twenty micrograms of isolated genomic DNA was digested with NcoI, subjected to electrophoresis and transferred to nylon membrane. The blots were then probed with a 500-bp KpnI-EcoRI fragment corresponding to the 3' region of the *CGBP* gene locus, as previously described (10, 11). To rescue the *CGBP*^{-/-} ES cell phenotype, murine *CGBP* cDNA (10) was subcloned into pcDNA 3.1/Zeo (Invitrogen). *CGBP*^{-/-} ES cells were transfected with linearized plasmid by electroporation (11), and zeomycin-resistant clones were recovered. Cell doubling times were determined by seeding 1×10^4 cells in six-well dishes (in triplicate) and collecting and counting cells at various times. To induce in vitro differentiation, ES cells were trypsinized and single cell suspensions (1×10^6 cells) were transferred to bacterial petri dishes containing 10 ml of ES cell media lacking LIF (73) and cultured for up to 10 days. ES cells differentiated in culture for 10 days were harvested, dispersed, and reseeded, and alkaline phosphatase activity was histochemically detected using an alkaline phosphatase leukocyte detection kit (Sigma).

Analysis of cytosine methylation. Analysis of global 5-methylcytosine in the context of the sequence CCGG was analyzed by thin-layer chromatography as described previously (38). Briefly, genomic DNA was digested with the restriction enzyme MspI or the methyl-sensitive isoschizomer HpaII, labeled with T4 polynucleotide kinase and [γ -³²P]ATP, and digested with nuclease P1. Five-methylcytosine monophosphate and cytosine monophosphate were separated by thin-layer chromatography, visualized by autoradiography, and quantitated by densitometry. Alternatively, global cytosine methylation was assessed utilizing a methyl acceptance assay as described (4). Briefly, 500 ng of genomic DNA was incubated with 2 μ Ci of ³H-methyl-S-adenosyl L-methionine (Perkin-Elmer; 15 Ci/mmol), and 3 units of SssI methylase (New England Biolabs) in 120 mM NaCl, 10 mM Tris-HCl (pH 7.9), 10 mM EDTA, and 1 mM dithiothreitol (DTT) for 1 h at 30°C. In vitro methylated DNA was isolated by filtration through Whatman DE-81 ion-exchange filter, and incorporated radioactivity was measured by scintillation counting.

To analyze cytosine methylation at specific loci, genomic DNA was prepared from various ES cell lines, subjected to restriction enzyme digestion and electrophoresis, and transferred to nylon membrane for Southern blot analysis. Blots were probed and washed as previously described (10, 71). Minor satellite and intracisternal A-particle (IAP) probes were provided by En Li (Novartis Institutes for Biomedical Research, Cambridge, MA). The *Rac2* probe was generated from a 300-bp NcoI/KpnI fragment of the proximal murine promoter (51). Additional probes were generated by PCR amplification. A 330-bp fragment was used to analyze the MluI site of region 2 of the *Igf2r* gene (62). The *H19* imprinted region was analyzed by using a 1.5-kb probe corresponding to 1,282 to 2,808 bp (GenBank accession no. U19619) (67). The *Pgk-2* probe corresponded to region III within the 3'-untranslated region (3).

Total DNA methyltransferase activity in ES cells was measured as described by Li et al. (38). Exponentially growing ES cells were lysed and sonicated on ice in 1 to 2 ml of lysis buffer (20 mM Tris-HCl [pH 7.4], 0.4 M NaCl, 25% glycerol, 5 mM EDTA, 0.1% Nonidet P-40, 1 mM dithiothreitol [DTT]) containing a proteinase inhibitor cocktail (Sigma). An equal volume of 50% DEAE-Sepharose 1 slurry equilibrated with lysis buffer was added and incubated for 10 min with shaking at 4°C. The mixture was subjected to centrifugation, the supernatant was collected, and protein concentration was determined by the Bradford method. Thirty micrograms of protein was incubated in DNA methyltransferase assay buffer (20 mM Tris-HCl [pH 7.4], 5 mM EDTA, 25% glycerol, 5 μ Ci of [³H]methyl-S-adenosyl L-methionine, 4 μ g poly(dI-dC), 1 mM DTT, and 200 μ g/ml bovine serum albumin) in a 200- μ l reaction volume at 37°C for 2 h and extracted twice with phenol-chloroform. The aqueous phase was adjusted to 0.1 M NaOH and incubated at 50°C for 2 h. The solution was neutralized with HCl, and radioactivity that incorporated into DNA was measured by scintillation counting after trichloroacetic acid precipitation. Control reactions lacked poly(dI-dC).

Alternatively, DNA methyltransferase activity was assessed using hemimethylated or unmethylated 33-bp double-stranded oligonucleotide substrates: 5'-GATCGCCGATGCGCGAATCGCGATCGATGCGAT-3' (methylated cytosine are underlined) (61). Nuclear extracts were prepared as described previously (16) with DNA methyltransferase lysis buffer and quantitated for protein concentration by the Bradford method. Twenty-five micrograms of nuclear extract was incubated at 37°C for 2 h in DNA methyltransferase assay buffer (20 mM Tris-HCl [pH 7.4], 5 mM EDTA, 25% glycerol) containing 5 μ Ci of [³H]methyl-S-adenosyl L-methionine, 8 μ g oligonucleotide, 1 mM DTT, and 200 μ g/ml bovine serum albumin in a total reaction volume of 200 μ l. Reactions were extracted twice with phenol-chloroform, and DNA was precipitated with ethanol. Reaction products were analyzed by 9% polyacrylamide gel electrophoresis, and autoradiography was performed following fluorography. Band intensities were

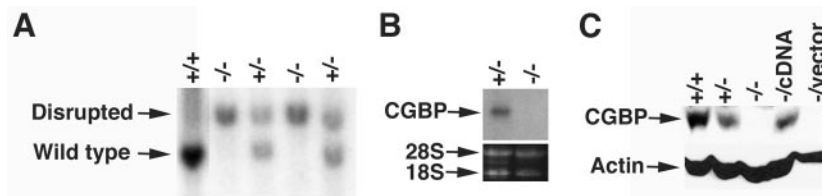


FIG. 1. Generation of ES cell lines lacking CGBP. (A) Southern blot analysis was performed on genomic DNA isolated from ES cell clones derived from blastocysts resulting from heterozygous crosses. DNA was digested with *Nco*I and hybridized with a 500-bp *Kpn*I/*Eco*RI probe downstream of the *CGBP* gene (10, 11). The arrows indicate the wild-type and disrupted alleles, and the deduced genotype at the *CGBP* locus is indicated above each lane. (B) Northern blot analysis was performed using total RNA isolated from either *CGBP*^{+/-} or *CGBP*^{-/-} ES cells and a 347-bp *Bam*HI fragment probe derived from the murine *CGBP* cDNA (10). The arrow indicates the murine *CGBP* transcript of approximately 2.6 kb. As a control for loading, 28S and 18S rRNA bands were visualized by ethidium bromide staining. (C) Western blot analysis was performed on protein extracts isolated from ES cells carrying the indicated *CGBP* alleles using antiserum raised against CGBP (71). The blot was also incubated with an antiactin antibody as a loading control.

determined by densitometry. As a positive control, *Sss*I methylase was substituted for nuclear extract.

De novo DNA methyltransferase activity was also measured following retroviral transduction of ES cells as described previously (36). Briefly, 1×10^6 *CGBP*^{+/+} or *CGBP*^{-/-} ES cells were seeded in six-well dishes and then transduced with Moloney murine leukemia virus retrovirus (generously provided by En Li, Novartis) the following day. Cells were incubated with virus in the presence of polybrene (3.2 μ g/ml) for 7 h. Cells were then grown in normal ES cell media and harvested at various times following transduction for isolation of genomic DNA. The cytosine methylation status of the provirus was assessed by digestion of isolated genomic DNA with *Kpn*I and *Hpa*II and Southern blot analysis using a retroviral probe (36).

Northern and Western analyses. Total RNA was isolated from exponentially growing wild-type (CCE916), *CGBP*^{+/-}, and *CGBP*^{-/-} ES cells using Tri-reagent solution per the manufacturer's recommended protocol (Life Technologies, Carlsbad, CA). Twenty micrograms of total RNA was fractionated by formaldehyde agarose gel electrophoresis and transferred to a nylon membrane as described previously (71). The blot was hybridized with a 347-bp *Bam*HI fragment probe derived from the murine *CGBP* cDNA (10), washed as previously described (71), and exposed to X-ray film.

Whole-cell lysates were prepared in 8 M urea or Laemmli sample buffers. Samples were subjected to electrophoresis and Western blot analysis as previously described (33). Membranes were incubated with antibody directed against actin (Sigma), CGBP (71), Dnmt1, Dnmt3a (Santa Cruz Biotechnology), or Dnmt3b (Imgenex, Inc.), followed by horseradish peroxidase-labeled secondary antibody. Signal was detected with an ECL kit (Amersham) and autoradiography and was quantitated by densitometry.

Reverse transcription-PCR analysis. For detection of developmental and lineage-specific mRNAs expressed during in vitro ES cell differentiation, total RNA was isolated from undifferentiated ($t = 0$) ES cells and differentiated embryoid bodies (5 and 10 days following the removal of LIF) using Tri-reagent. Total RNA (1 μ g) was reverse transcribed using avian myeloblastosis virus reverse transcriptase and random hexamers (Roche, Inc., Indianapolis, IN) at 42°C for 60 min. Single-stranded cDNA (0.1 μ g) was amplified in a 25- μ l reaction mixture that included 0.2 mM of each deoxynucleoside triphosphate, 50 pmol of sense and antisense primers, and 1 U of *Taq* DNA polymerase (Roche) in buffer supplied by the manufacturer. Samples were heat denatured at 94°C for 2 min, followed by 25 to 30 cycles at 94°C for 30 s, 60°C for 30 s, 72°C for 30 s, and finally 10 min at 72°C. PCR was performed for *HPRT* to monitor the integrity of the cDNA produced by reverse transcription. Ten microliters of amplified DNA was subjected to electrophoresis on a 1.5% agarose gel in 0.5 \times Tris-borate-EDTA. Primer pairs used were the following: *Brachyury*, 5'-ATCAAGGAAGGCTTTA GCAAATGGG-3' and 5'-GAACCTCGGATTCACATCGTGAGA-3' (76); *GATA-4*, 5'-CACTATGGGCACAGCAGCTCC-3' and 5'-TTGGAGCTGGCC TGCGATGTC-3' (76); *c-fms*, 5'-CTGAGTCAGAAGCCCTTCGACAAAG-3' and 5'-CTTTGCCAGACCAAAGGCTGTAGC-3' (40); *gp-IIB*, 5'-AGGC AGAGAAGACTCCGGTA-3' and 5'-TACCGAATATCCCCGGTAAC (70); *MHC- β* , 5'-TGCAAAGGCTCCAGGCTGTAGGGC-3' and 5'-GCCAACACCA ACCTGTCCAAGTTC-3' (42); *HPRT*, 5'-CACAGGACTAGAACACCTGC-3' and 5'-GCTGGTAAAAGGACCTCT-3' (31); *Oct4*, 5'-GGCGTTCCTTTGGA AAGGTGTC-3' and 5'-CTCGAACACATCCTTCTCT-3' (generously provided by Rebecca Chan, Indiana University).

Statistical analysis. Statistical significance was assessed by one-tailed *t* tests, with a *P* value of <0.05 interpreted as statistical significance.

Flow cytometric analysis. For cell cycle analysis, exponentially growing asynchronous cells were harvested, washed twice with phosphate-buffered saline (PBS), and incubated in PBS containing 0.3% NP-40, 0.5-mg/ml RNase A, and 50- μ g/ml propidium iodide for 30 min on ice. Following staining, samples were analyzed using a FACScan flow cytometer (Becton Dickinson, San Jose, CA) and ModFit LT software (Verity Software, Topsham, ME). Apoptotic cells were detected using an Annexin V-FLUOS and propidium iodide staining kit (Roche). One million exponentially growing asynchronous cells were harvested, washed once with PBS, and incubated with 100 μ l of 1% Annexin V-FLUOS and 0.5 μ g/ml propidium iodide in 10 mM HEPES (pH 7.5), 140 mM NaCl, and 5 mM $CaCl_2$ at room temperature for 15 min. Binding buffer (400 μ l) was then added prior to flow cytometric analysis. Apoptotic cells were defined as the fraction of Annexin V-positive and propidium iodide-negative cells.

RESULTS

ES cells lacking CGBP are viable. The absence of murine CGBP results in embryonic lethality at the peri-implantation stage of development (11), indicating a critical role for this DNA-binding protein in early mammalian embryogenesis. Due to the difficulty of studying preimplantation embryos, *CGBP*^{-/-} ES cell lines were generated to permit more-detailed studies on the function of CGBP. ES cell lines were derived from blastocysts (3.5 dpc) resulting from matings between *CGBP*^{+/-} mice. Several *CGBP*^{-/-} (null) and *CGBP*^{+/-} (heterozygous) ES cell lines were obtained (Fig. 1A). Loss of CGBP expression in *CGBP*^{-/-} cells was confirmed by Northern (Fig. 1B) and Western (Fig. 1C) blot analyses. Western blot analysis also revealed that *CGBP*^{+/-} ES cells express approximately 50% of the wild-type level of CGBP. A rescued cell line, the *CGBP*^{-/-cDNA} cell line, was produced by stable transfection of *CGBP*^{-/-} ES cells with a *CGBP* cDNA expression vector. Western blot analysis confirmed the expression of CGBP in the *CGBP*^{-/-cDNA} ES cell line (approximately 50% of wild type) but not in *CGBP*^{-/-} ES cells transfected with the parental vector (*CGBP*^{-/-vector}) (Fig. 1C).

CGBP-deficient ES cells exhibit an extended doubling time due to increased apoptosis. Initial observations indicated that *CGBP*^{-/-} ES cells grow slower than wild-type ES cells. *CGBP*^{-/-} ES cells exhibit a doubling time of approximately 14.5 h, compared to 10.7 h for *CGBP*^{+/+} ES cells (Fig. 2A). Similar growth characteristics were observed for a second independent *CGBP*^{-/-} ES cell line (data not shown). *CGBP*^{+/-} cells exhibit a growth rate indistinguishable from that of wild-

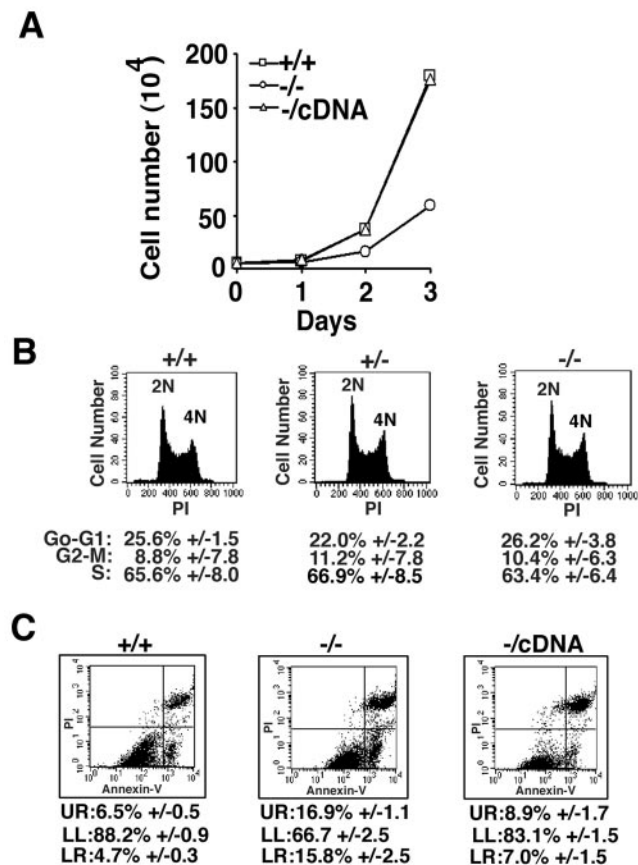


FIG. 2. *CGBP*^{-/-} ES cells exhibit an extended doubling time due to an elevated rate of apoptosis. (A) Growth rates were measured for ES cells carrying the indicated *CGBP* alleles. The curves for the *CGBP*^{+/+} and rescued *CGBP*^{-/cDNA} cells are overlapping. (B) Exponentially growing ES cells of the indicated *CGBP* alleles were analyzed for cell cycle distribution by propidium iodide (PI) staining and flow cytometry. Numerical values represent the summary of data from three experiments. 2N and 4N represent diploid (G₀-G₁) and tetraploid (G₂-M) genome content, respectively. (C) Exponentially growing ES cells of the indicated *CGBP* alleles were analyzed for apoptosis using Annexin V and propidium iodide staining and flow cytometry. UR, upper right panel corresponding to dead cells; LL, lower left panel corresponding to healthy cells; LR, lower right panel corresponding to apoptotic cells. Numerical values represent the summary of data from three experiments.

type cells (data not shown). In addition, normal doubling time was restored in rescued *CGBP*^{-/cDNA} ES cells (Fig. 2A), indicating that this phenotype is a consequence of *CGBP* deficiency. Despite an extended doubling time, *CGBP*^{-/-} ES cells exhibit a normal cell cycle distribution (Fig. 2B). Instead, staining for Annexin V reveals that approximately 33% of exponentially growing *CGBP*^{-/-} ES cells are dead or undergoing apoptosis, in contrast to approximately 11% of *CGBP*^{+/+} ES cells (Fig. 2C). Hence, the extended doubling time of *CGBP*^{-/-} ES cells results from an increased rate of apoptosis. Importantly, apoptosis rates return to near-normal levels in rescued *CGBP*^{-/cDNA} ES cells, since 83% of *CGBP*^{-/cDNA} ES cells are viable, compared to 67% of *CGBP*^{-/-} ES cells and 88% of wild-type ES cells.

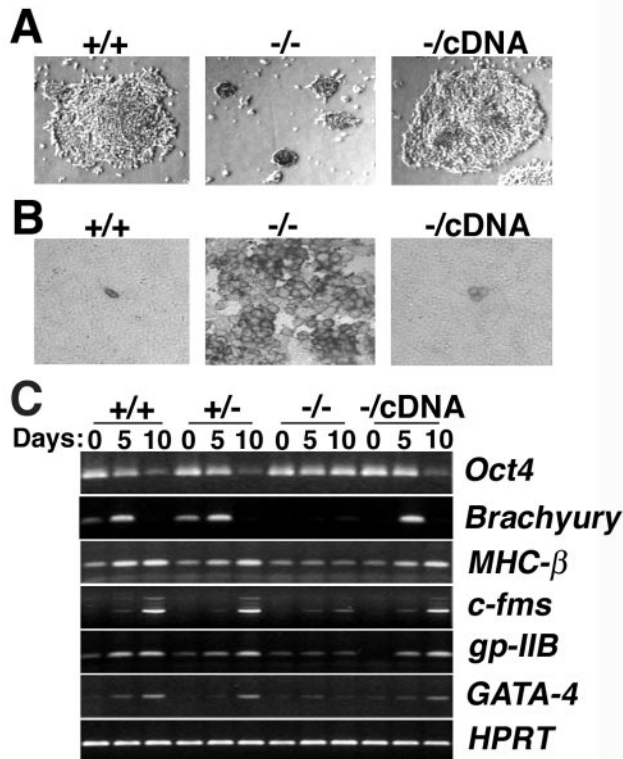


FIG. 3. Failure of in vitro cellular differentiation in the absence of *CGBP*. (A) Colony morphology following induction of differentiation. Embryoid body outgrowths carrying the indicated *CGBP* alleles were cultured for 5 days in the absence of LIF. Magnification, $\times 20$. (B) ES cells grown in the absence of LIF for 10 days were harvested, disaggregated, reseeded, and stained for alkaline phosphatase activity. One hundred cells of each genotype were scored for staining. Magnification, $\times 64$. (C) ES cells carrying the indicated *CGBP* alleles were cultured for various times (0, 5, or 10 days) in the absence of LIF to induce differentiation. Total RNA was isolated, and reverse transcription-PCR was performed to examine the expression of lineage and development-specific gene markers. *Oct4*, marker of undifferentiated ES cells; *Brachyury*, mesoderm; *MHC-β*, muscle; *c-fms*, myeloid; *gp-IIB*, megakaryocyte; *GATA-4*, endoderm; *HPRT*, control for RNA quantity and integrity.

***CGBP* is required for in vitro ES cell differentiation.** Because murine embryos lacking *CGBP* fail to gastrulate (11), in vitro differentiation assays were performed to determine the developmental potential of *CGBP*^{-/-} ES cells. Both *CGBP*^{-/-} and *CGBP*^{+/+} ES cells formed embryoid bodies at 2 days following removal of LIF from the growth medium (data not shown). However, after further culture under differentiation conditions, *CGBP*^{+/+} colonies grew in size and produced a prominent outgrowth, while *CGBP*^{-/-} colonies remained small and failed to produce outgrowths (Fig. 3A). Gross morphological evidence of in vitro differentiation was restored in *CGBP*^{-/cDNA} cells.

Molecular markers of differentiation were also examined to assess the ability of *CGBP*^{-/-} cells to achieve in vitro differentiation. Alkaline phosphatase is a marker of pluripotency in undifferentiated ES cells and is down-regulated upon cellular differentiation (74). As expected, very few *CGBP*^{+/+} ES cells (2 to 3%) express alkaline phosphatase activity 10 days after

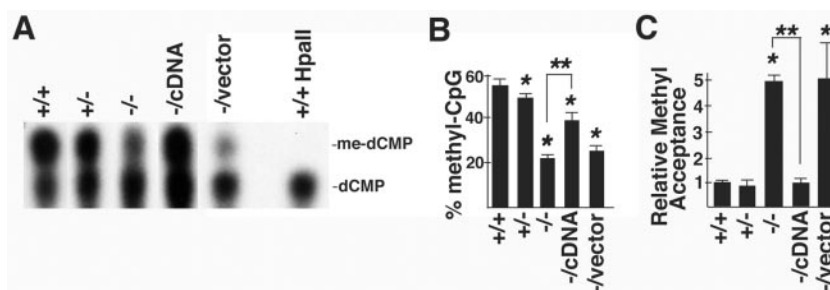


FIG. 4. CGBP is essential for proper DNA methylation in murine ES cells. (A) Global cytosine methylation levels in ES cells carrying the indicated *CGBP* alleles were determined in the context of CCGG by thin-layer chromatography following digestion of genomic DNA with the restriction enzymes *MspI* or *HpaII*. A representative experiment is shown. (B) Summary of thin-layer chromatography data from three experiments. Error bars represent standard error, and asterisks denote a statistically significant ($P < 0.05$) difference compared to the wild type. Double asterisks denote a statistically significant ($P < 0.05$) difference between *CGBP*^{-/-} cells and *CGBP*^{-/cDNA} cells. (C) Global cytosine methylation levels in ES cells carrying the indicated *CGBP* alleles were determined by methyl acceptance assay. Error bars represent standard error, and asterisks denote a statistically significant ($P < 0.05$) difference compared to the wild type. Double asterisks denote a statistically significant ($P < 0.05$) difference between *CGBP*^{-/-} cells and *CGBP*^{-/cDNA} cells. The experiment was performed three times.

induction of differentiation. In contrast, alkaline phosphatase activity was detected in >96% of similarly treated *CGBP*^{-/-} cells, indicating a failure to differentiate (Fig. 3B). However, alkaline phosphatase expression was extinguished in >96% of rescued *CGBP*^{-/cDNA} cells, indicating that restoration of CGBP expression permits in vitro differentiation.

The expression patterns of additional developmental and lineage-specific gene markers were assessed in undifferentiated ES cells and 5 or 10 days following the removal of LIF (Fig. 3C). Consistent with the alkaline phosphatase results, *CGBP*^{-/-} ES cells fail to down-regulate *Oct4*, a marker of pluripotent cells that is down-regulated upon cellular differentiation (47). In contrast, *CGBP*^{+/+} and *CGBP*^{+/-} cells down-regulate *Oct4* levels by day 10 of differentiation, as do rescued *CGBP*^{-/cDNA} cells. Following induction of in vitro differentiation, *CGBP*^{-/-} ES cells also fail to normally induce the expression of the cardiac lineage marker *MHC-β* (42), the hematopoietic markers *gp-IIB* (megakaryocyte) (70) and *c-fms* (myeloid) (40), and markers of earlier stages of differentiation, such as *GATA-4* (visceral/parietal endoderm) (31) and *Brachyury* (mesoderm) (73) (Fig. 3C). *CGBP*^{+/+} and *CGBP*^{+/-} ES cells induce the expression of each of these markers following removal of LIF. Rescued *CGBP*^{-/cDNA} ES cells induce expression of all differentiation markers that were examined (Fig. 3C). These results indicate that CGBP plays a critical role in the initiation and execution of in vitro ES cell differentiation.

Loss of CGBP results in reduced levels of 5-methylcytosine in the genome. Studies were performed to gain insight into the molecular basis for the developmental defects exhibited by *CGBP*^{-/-} ES cells. CGBP has the unique property of specifically binding to DNA sequences containing unmethylated CpG motifs (71). Hence, we reasoned that CGBP might function to modulate the methylation state of the genome. Global cytosine methylation levels were quantitated by thin-layer chromatography following digestion with *MspI* to determine if loss of CGBP affects DNA methylation in the context of the sequence CCGG. As previously demonstrated (38), *CGBP*^{+/+} ES cells exhibit a high degree of cytosine methylation (approximately 55% of CpG dinucleotides) (Fig. 4A and B). Loss of a single *CGBP* allele led to a slight but statistically significant reduction ($P < 0.05$) in the level of global cytosine methylation. How-

ever, a dramatic, approximately 60% reduction in global cytosine methylation was observed in *CGBP*^{-/-} ES cells. Absence of signal for 5-methylcytosine following digestion with the methyl-sensitive enzyme *HpaII* (Fig. 4A) demonstrates the specificity of the assay. A similar decrease in global cytosine methylation was observed in a second, independently derived *CGBP*^{-/-} ES cell line (data not shown). Cytosine methylation levels were significantly increased in rescued *CGBP*^{-/cDNA} cells (compared to *CGBP*^{-/-} cells) but not in *CGBP*^{-/vector} ES cells (Fig. 4A and B).

Global cytosine methylation was also assessed by determining the ability of genomic DNA to accept methyl groups (4). These experiments confirm the involvement of CGBP in regulating global cytosine methylation (Fig. 4C). DNA derived from *CGBP*^{-/-} cells accepts fivefold more methyl groups than do DNA samples isolated from *CGBP*^{+/+} or *CGBP*^{+/-} ES cells, indicating an 80% decline in global cytosine methylation in cells lacking CGBP. The capacity to accept methyl groups was reduced to wild-type levels in *CGBP*^{-/cDNA} ES cells but not *CGBP*^{-/vector} cells. These results indicate that CGBP is required for normal global cytosine methylation.

Additional studies were conducted to examine cytosine methylation at specific genomic loci. Southern blot analysis reveals that minor satellite repetitive elements are highly methylated in *CGBP*^{+/+} and *CGBP*^{+/-} ES cells, as indicated by the absence of a low-molecular-weight ladder in *HpaII*-digested genomic DNA (Fig. 5A). In the absence of CGBP, these sequences become hypomethylated, as revealed by increased *HpaII* digestion and appearance of a low-molecular-weight ladder. A similar disruption in cytosine methylation is observed at IAP retroviral sequences (Fig. 5B). Four distinct *HpaII*-digested fragments were detected in DNA isolated from *CGBP*^{-/-} ES cells but not from *CGBP*^{+/+} ES cells. However, *CGBP*^{-/-} ES cells did not exhibit a complete loss of cytosine methylation at these repetitive elements, since the levels of low-molecular-weight DNA fragments produced by *HpaII* digestion were not comparable in intensity with those generated by digestion with the methyl-insensitive isoschizomer *MspI*. These results are consistent with the partial decrease of global cytosine methylation observed in *CGBP*^{-/-} ES cells (Fig. 4). Methylation of repetitive sequences was restored in rescued

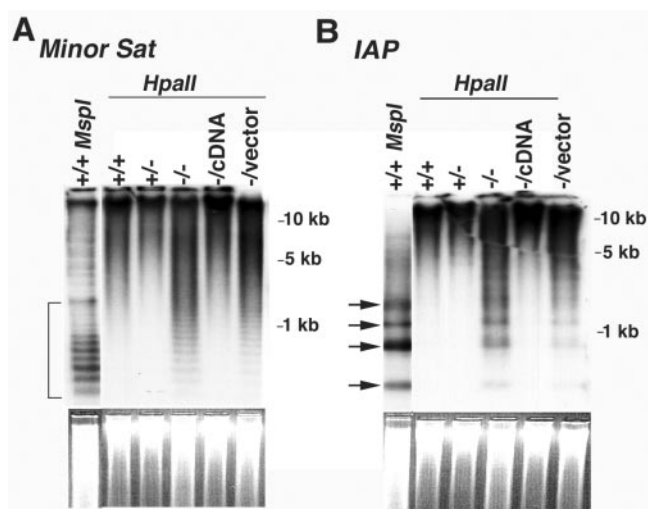


FIG. 5. *CGBP*^{-/-} ES cells exhibit reduced cytosine methylation at repetitive genomic elements. (A) Genomic DNA was isolated from ES cells carrying the indicated *CGBP* alleles, digested with *Msp*I or *Hpa*II, and Southern blot analysis was performed using a probe for the minor satellite repetitive element. The bracket indicates the region of low-molecular-weight bands that reflect cytosine hypomethylation. The ethidium bromide-stained gel is shown below to illustrate relative DNA loading. (B) Same as in panel A, except the probe corresponds to the IAP retrovirus repetitive element. Arrows indicate bands that reflect hypomethylation. The ethidium bromide-stained gel is shown below to illustrate relative DNA loading.

CGBP^{-cDNA} ES cells, further demonstrating a functional role for *CGBP* in the regulation of DNA methylation.

The pattern of cytosine methylation at single-copy genes was also assessed. A probe derived from the hematopoietic cell-restricted *Rac2* promoter (51) was used to analyze the methylation status of a 2-kb *Nco*I genomic fragment that contains several *Hpa*II restriction sites. Analysis of genomic DNA isolated from *CGBP*^{+/+} or *CGBP*^{+/-} ES cells reveals a dominant 2-kb band following digestion with *Nco*I and *Hpa*II, indicating heavy cytosine methylation throughout this region (Fig. 6A). The intensity of this band is dramatically reduced in *CGBP*^{-/-} ES cells, indicating reduced cytosine methylation, although partial methylation persists because a band representing a fully demethylated region (300 bp) is not apparent. Similar results were found for the *Pgk-2* gene. A probe derived from the 3'-flanking region detects three dominant high-molecular-weight bands when genomic DNA isolated from *CGBP*^{+/+} or *CGBP*^{+/-} ES cells is digested with *Bam*HI and the methyl-sensitive enzyme *Hha*I (Fig. 6B). However, the intensity of these bands is reduced when DNA isolated from *CGBP*^{-/-} ES cells is similarly analyzed, and four low-molecular-weight bands become more abundant. Introduction of the *CGBP* cDNA into *CGBP*^{-/-} ES cells (*CGBP*^{-cDNA}) restores cytosine methylation at both of these gene loci.

The methylation status of imprinted genes in *CGBP*^{-/-} ES cells was next examined (Fig. 7). The upstream region of the paternally imprinted *H19* gene, which contains the differentially methylated domain that controls expression of the *Igf2r* and *H19* genes (64), was examined by Southern blot analysis. A large DNA fragment corresponding to the paternally methyl-

ated *H19* allele and a smaller fragment representing the hypomethylated maternal allele were observed in *CGBP*^{+/+} and *CGBP*^{+/-} ES cells, while a loss of methylation at the paternal allele was observed in *CGBP*^{-/-} ES cells (Fig. 7A). Methylation at the *H19* locus is increased following restoration of *CGBP* expression (*CGBP*^{-cDNA}), suggesting that epigenetic marks persist elsewhere within this locus in *CGBP*^{-/-} ES cells. Analysis of region 2 of the maternally methylated *Igf2r* gene (60) revealed methylated and unmethylated DNA fragments corresponding to the maternal and paternal alleles, respectively, in *CGBP*^{+/+} and *CGBP*^{+/-} ES cells. In *CGBP*^{-/-} ES cells, however, only the unmethylated fragment was detected, indicating a loss of maternal imprinting (Fig. 7B). In contrast to the *H19* locus, the *Igf2r* locus was not remethylated in *CGBP*^{-cDNA} ES cells.

***CGBP*^{-/-} ES cells exhibit decreased DNA methyltransferase activity.** Additional experiments were performed to directly assess the DNA methylation machinery in *CGBP*^{-/-} ES cells. Protein extracts isolated from these cells exhibited a 30% reduction in total DNA methyltransferase activity towards a poly(dI-dC) substrate (Fig. 8A). Activity was restored in *CGBP*^{-cDNA} ES cells. Similar experiments were performed using hemimethylated or unmethylated oligonucleotide substrates to distinguish between maintenance and de novo DNA methyltransferase activity (Fig. 8A). Extracts derived from *CGBP*^{-/-} cells exhibit a 60% decline in DNA methyltransferase activity towards the hemimethylated oligonucleotide substrate but normal activity towards the unmethylated substrate. These results indicate a deficiency in maintenance cytosine methylation in the absence of *CGBP*. As an independent method to determine whether *CGBP* is necessary for de novo cytosine methylation, *CGBP*^{-/-} ES cells were assessed for their ability to methylate a newly introduced retroviral transgene. The retroviral provirus begins to acquire cytosine methylation within 48 h of transduction in both *CGBP*^{+/+} and *CGBP*^{-/-} ES cells (Fig. 8B), indicating normal de novo DNA methyltransferase activity. However, in contrast to *CGBP*^{+/+} cells, significant levels of unmethylated provirus persist in *CGBP*^{-/-} ES cells. The intensity of the unmethylated band is 70% of that of the fully methylated band in *CGBP*^{-/-} ES cells 4 days following transduction, compared to 30% for *CGBP*^{+/+} ES cells, consistent with a defect in maintenance DNA methyltransferase activity in the absence of *CGBP*. A similar pattern of proviral cytosine methylation was previously found in ES cells lacking *Dnmt1* (36).

Western blot analysis was performed to determine the expression level of *Dnmt* enzymes (Fig. 9). These studies reveal a 50% reduction of the *Dnmt1* protein in *CGBP*^{-/-} ES cells. *Dnmt1* levels recover in *CGBP*^{-cDNA} ES cells but not in *CGBP*^{-vector} cells. Significantly reduced levels of *Dnmt3a* are also observed in *CGBP*^{-/-} cells, but this does not correlate with the absence of *CGBP*, since similarly reduced levels of *Dnmt3a* are observed in *CGBP*^{+/-} and *CGBP*^{-cDNA} ES cells. However, heterozygous and rescued ES cells contain only approximately 50% of wild-type *CGBP* levels (Fig. 1C), so it remains possible that this partial deficiency leads to decreased *Dnmt3a* expression. The significance of this finding is unclear, since *CGBP*^{+/-} and *CGBP*^{-cDNA} ES cells contain normal genomic cytosine methylation and *CGBP*^{-/-} cells exhibit normal de novo methyltransferase activity. However, *Dnmt3a* has

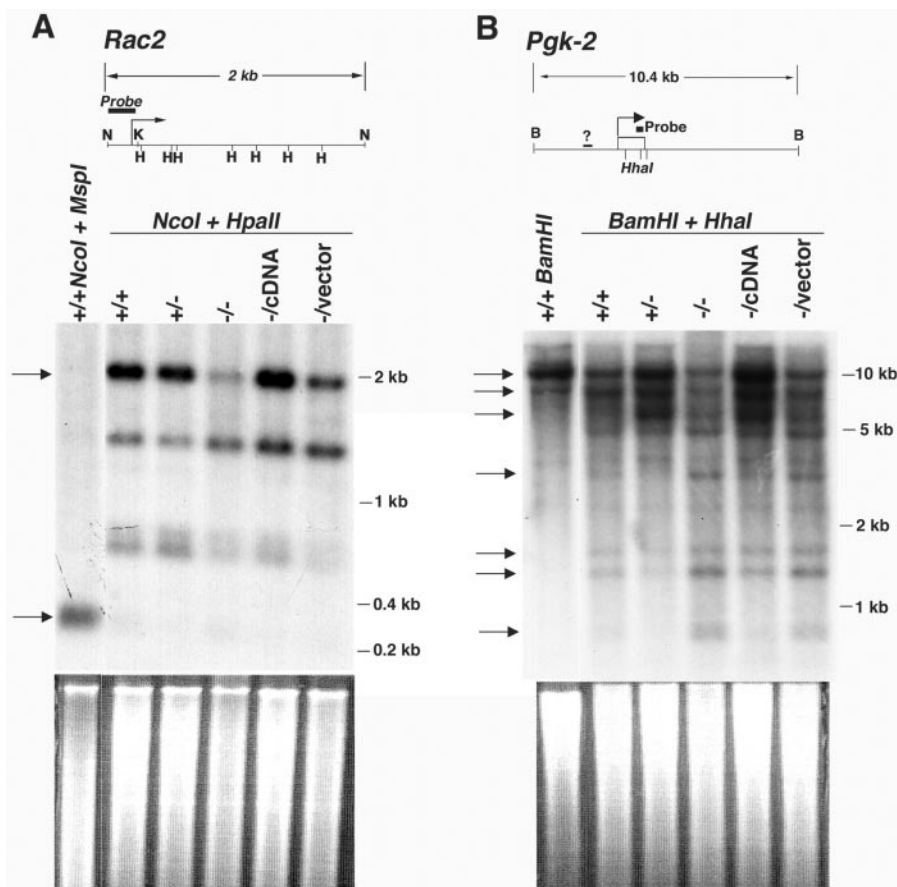


FIG. 6. *CGBP*^{-/-} ES cells exhibit reduced cytosine methylation at single-copy genes. (A) Genomic DNA was isolated from ES cells carrying the indicated *CGBP* alleles and digested with *Nco*I and *Hpa*II, and Southern blot analysis was performed using a probe for the *Rac2* gene (shown schematically). Arrows indicate a 2-kb band produced by *Hpa*II and *Nco*I digestion of the fully methylated promoter and a 300-bp band produced by *Msp*I digestion or by *Hpa*II digestion if the restriction sites are unmethylated. The bent arrow indicates site of transcription initiation. N, *Nco*I; H, *Hpa*II. The ethidium bromide-stained gel is shown below to illustrate relative DNA loading. (B) Same as in panel A, except the genomic DNA was digested with *Hha*I and *Bam*HI, and the probe corresponds to the *Pgk-2* gene (shown schematically). The top three arrows indicate high-molecular-weight bands that correspond to heavy cytosine methylation which are dominant in *CGBP*^{+/+}, *CGBP*^{+/-}, and *CGBP*^{-/cDNA} cells. The bottom four arrows indicate low-molecular-weight bands that reflect hypomethylation and increase in intensity in *CGBP*^{-/-} and *CGBP*^{-/vector} cells. The bent arrow indicates site of transcription initiation. B, *Bam*HI. *Hha*I sites are denoted by hatch marks, and the question mark indicates a region of undetermined nucleotide sequence. The ethidium bromide-stained gel is shown below to illustrate relative DNA loading.

been found to be required for maintenance of cytosine methylation patterns in ES cells (13). No significant decrease in *Dnmt3b* was detected in *CGBP*^{-/-} ES cells. In summary, these data indicate that *CGBP*^{-/-} ES cells exhibit reduced levels of genomic cytosine methylation and reduced maintenance DNA methyltransferase activity.

DISCUSSION

CGBP is a transcriptional activator that exhibits a unique DNA-binding specificity for sequences containing unmethylated CpG motifs (34, 71). Disruption of the *CGBP* gene in mice leads to peri-implantation death (11). Although this finding establishes the importance of CGBP for normal mammalian development, the early time of death makes further analysis of this mutant difficult. In the study reported here, murine ES cells lacking CGBP were generated to permit a more-

detailed analysis of the *CGBP*^{-/-} phenotype and provide further insight into normal CGBP function.

The ability to isolate *CGBP*^{-/-} ES cell lines demonstrates that this protein is not essential for ES cell viability. However, consistent with the inability of *CGBP*^{-/-} embryos to gastrulate, *CGBP*^{-/-} ES cells are unable to achieve in vitro differentiation. Instead, they remain undifferentiated following removal of LIF from the growth medium, as indicated by persistent expression of Oct4 and alkaline phosphatase, markers of the undifferentiated state. These in vitro data suggest that *CGBP*^{-/-} mice exhibit a peri-implantation death due to the requirement of CGBP for the execution of specific differentiation programs, rather than exhaustion of maternally inherited stores of CGBP protein. Methylation of the Oct4 promoter is essential for its down-regulation during ES cell differentiation (23, 26). It remains to be determined whether persistent Oct4 expression, possibly as a consequence of de-

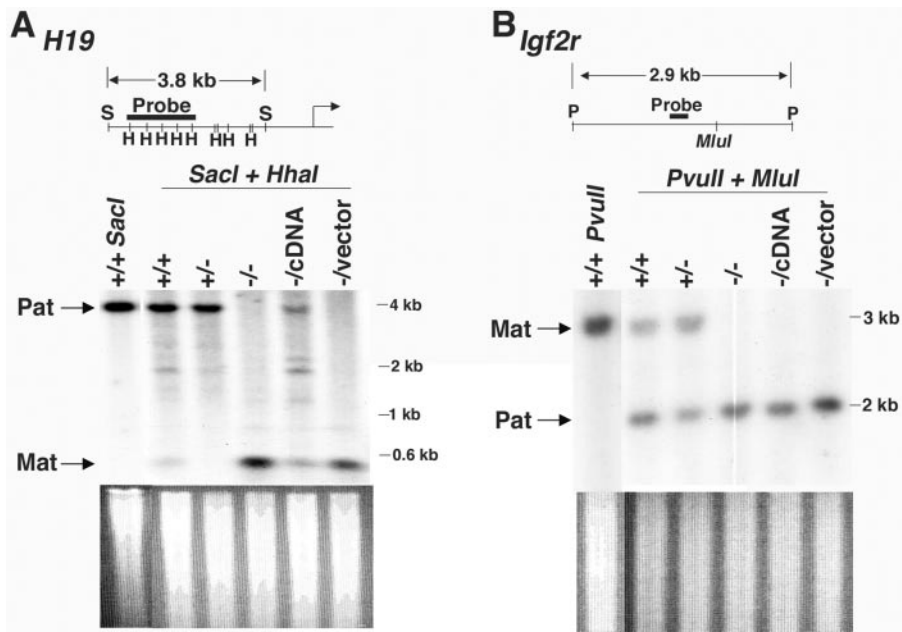


FIG. 7. *CGBP*^{-/-} ES cells exhibit reduced cytosine methylation at imprinted genes. (A) Genomic DNA was isolated from ES cells carrying the indicated *CGBP* alleles and digested with *HhaI* and *SacI*, and Southern blot analysis was performed using a probe for the *H19* gene (shown schematically). Arrows indicate position of paternally derived (imprinted) and maternally derived (nonimprinted) alleles. The bent arrow indicates site of transcription initiation. S, *SacI*; H, *HhaI*. The ethidium bromide-stained gel is shown below to illustrate relative DNA loading. (B) Same as in panel A, except the DNA was digested with *MluI* and *PvuII* and the probe corresponds to the *Igf2r* locus (shown schematically). Arrows indicate position of maternally derived (imprinted) allele and paternally derived (nonimprinted) allele. P, *PvuII*. The ethidium bromide-stained gel is shown below to illustrate relative DNA loading.

fective promoter methylation, is causally related to the inability of *CGBP*^{-/-} ES cells to achieve in vitro differentiation. Alternatively, persistent Oct4 expression could be a secondary consequence of the failure of these cells to effectively initiate the differentiation program, possibly due to a global derangement of epigenetic modifications.

Given the binding specificity of *CGBP* for DNA sequences containing unmethylated CpG motifs (34, 71), genomic cytosine methylation patterns in *CGBP*^{-/-} ES cells were analyzed. *CGBP*^{-/-} ES cells exhibit a 60 to 80% decrease in global cytosine methylation, including reduced cytosine methylation of repetitive elements, single-copy genes, and imprinted genes. This deficiency is correlated with reduced maintenance DNA methyltransferase activity, because *CGBP*^{-/-} cell extracts exhibit a 60% reduction in DNA methyltransferase activity towards a hemimethylated DNA substrate and a 50% reduction in the level of Dnmt1 protein. In contrast, de novo DNA methyltransferase activity in *CGBP*^{-/-} ES cells is normal. To our knowledge, *CGBP*^{-/-} ES cells represent the first example of reduced maintenance DNA methyltransferase activity without direct abrogation of *Dnmt* gene function.

Importantly, nearly all of the epigenetic perturbations detected in *CGBP*^{-/-} ES cells are corrected upon introduction of the *CGBP* cDNA into these cells, illustrating the plasticity of the epigenome. The exception is the maternally imprinted *Igf2r* gene, which remains hypomethylated following restoration of *CGBP* expression, DNA methyltransferase activity, and global cytosine methylation. These results are consistent with previous reports of variable degrees of corrected genomic imprinting upon restoration of Dnmt expression in DNA methyltrans-

ferase-deficient cells (7, 13, 68, 69). The ability of wild-type *CGBP* to rescue defects in *CGBP*^{-/-} ES cells offers an attractive system with which to probe structure/function relationships of this novel factor.

The deficiencies observed for cytosine methylation in *CGBP*^{-/-} cells, although dramatic, cannot account for the severity of the observed phenotype. For example, ES cells lacking Dnmt1 exhibit a 90% reduction in DNA methyltransferase activity and cytosine methylation yet exhibit normal growth prior to in vitro differentiation (36). In contrast, undifferentiated *CGBP*^{-/-} ES cells exhibit a 35% increase in doubling time as a consequence of increased apoptosis. Others have reported the existence of an epigenetic surveillance mechanism which induces apoptosis or cell cycle arrest in response to aberrations in cytosine methylation patterns or reduced levels of Dnmt1, respectively (28, 43, 59). In addition, mouse embryos lacking Dnmt1 die later in gestation (8.5 to 9.5 dpc) (36) than *CGBP*^{-/-} embryos (4.5 to 6.5 dpc) (11). Finally, decreased DNA methyltransferase activity and Dnmt1 protein levels detected in *CGBP*^{-/-} ES cells are unlikely to fully explain the observed deficiency in genomic cytosine methylation, since *Dnmt1*^{+/-} ES cells expressing reduced Dnmt1 protein retain normal levels of cytosine methylation (12, 38).

The existence of *CGBP* homologues in lower eukaryotes that lack CpG methylation, such as yeast and *C. elegans*, provides circumstantial evidence for a function of *CGBP* that is independent of cytosine methylation. Interestingly, sequence alignment reveals that *CGBP* homologues in organisms that lack CpG methylation lack the CXXC DNA-binding domain (data not shown). The yeast *CGBP* homologue, Spp1, is a

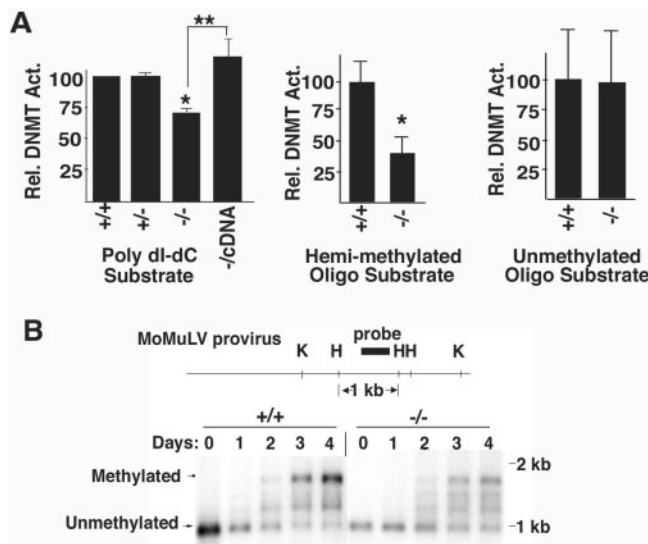


FIG. 8. *CGBP*^{-/-} ES cells exhibit reduced maintenance DNA methyltransferase activity. (A) Protein extracts were prepared from ES cells carrying the indicated *CGBP* alleles and assayed for in vitro DNA methyltransferase activity using poly(dI-dC), hemimethylated oligonucleotide, or unmethylated oligonucleotide substrates. Error bars represent standard error, and asterisks denote a statistically significant ($P < 0.05$) difference compared to the wild type. Double asterisks denote a statistically significant ($P < 0.05$) difference between *CGBP*^{-/-} cells and *CGBP*^{-/cDNA} cells. Each experimental value represents a summary from three experiments, with the exception of the analysis of *CGBP*^{+/-} and *CGBP*^{-/-} extracts using the poly(dI-dC) substrate, which was performed seven times. (B) *CGBP*^{+/-} or *CGBP*^{-/-} ES cells were transduced with the murine Moloney leukemia virus retrovirus. Genomic DNA was isolated at the indicated times following transduction and the provirus analyzed for cytosine methylation by Southern blot analysis. The structure of the provirus is illustrated schematically above the blot. H, HpaII; K, KpnI. Arrows indicate position of unmethylated and fully methylated provirus fragments.

component of the megadalton Set1 histone methyltransferase complex (41). Spp1 is dispensable for Set1 histone methyltransferase activity in *Saccharomyces cerevisiae* but is necessary for histone methyltransferase activity in *Schizosaccharomyces pombe* (56). In addition, human CGBP colocalizes to an identical set of nuclear speckles with the human trithorax (HRX, ALL-1, MLL1) histone methyltransferase (33). The composition of the megadalton HRX histone methyltransferase complex has been reported, but CGBP was not detected (45, 79). Similarly, CGBP was not detected as a component of the mammalian Set1/Ash2 histone methyltransferase complex (75). However, these studies reported the composition of soluble complexes. CGBP is localized nearly exclusively in the nuclear matrix and hence might not have been recovered by the extraction methods utilized. Whether CGBP interacts with histone methyltransferase complexes at the nuclear matrix remains to be determined. However, given the presence of the yeast CGBP homologue in a histone-modifying complex, it is tempting to speculate that mammalian CGBP may also play a role in the control of histone modification and chromatin structure.

Indeed, there are several examples of perturbed patterns of cytosine methylation as a consequence of altered chromatin structure. For example, cytosine methylation in *Neurospora* is dependent on methylation of histone H3 (63), and inhibition of

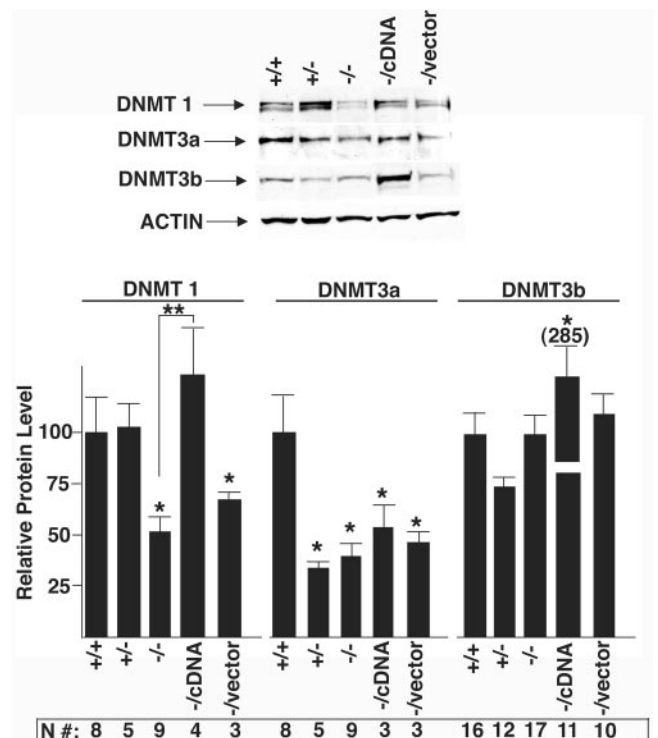


FIG. 9. *CGBP*^{-/-} ES cells express reduced levels of Dnmt1 protein. Protein extracts isolated from ES cells carrying the indicated *CGBP* alleles were subjected to Western blot analysis using antiserum directed against Dnmt1, Dnmt3a, or Dnmt3b. An antiserum directed against actin was utilized as an internal control for protein loading. A representative experiment is presented, and the summary of data from numerous experiments is presented below. The number of replicates for each experimental value is indicated (N#). Error bars represent standard errors, and asterisks denote a statistically significant ($P < 0.05$) difference compared to the wild type. Double asterisks denote a statistically significant ($P < 0.05$) difference between *CGBP*^{-/-} cells and *CGBP*^{-/cDNA} cells.

HDAC activity by trichostatin A results in a loss of cytosine methylation (57). Furthermore, the chromatin remodeling protein DDM1 in *Arabidopsis* and the related factor LSH in mammals are required for normal cytosine methylation (15, 29, 30), and disruption of the *Suv39h* histone methyltransferase gene in murine ES cells leads to altered localization of Dnmt3b and decreased cytosine methylation at pericentric satellite repeats (35). Conversely, deficiency in Dnmt1 leads to increased histone acetylation and decreased histone H3-Lys9 methylation at pericentromeric sequences (18). It will therefore be interesting to determine if *CGBP*^{-/-} ES cells exhibit aberrations in histone modifications.

Many details regarding the mechanism of CGBP function remain unknown. For example, how does CGBP influence the expression of Dnmt1? What are the target genes to which CGBP binds in vivo? Does CGBP also influence the activity of histone-modifying complexes? And why is the action of this transcriptional activator necessary for normal cytosine methylation, a modification usually associated with repressed gene expression? The data reported here indicate that CGBP facilitates the activity of the DNA methylation machinery and implicate CGBP as an important epigenetic regulator. A better

understanding of CGBP function will provide important insights into epigenetic regulation.

ACKNOWLEDGMENTS

We thank En Li for the minor satellite and IAP probes, Rebecca Chan for providing the Oct4 primers, and Merv Yoder and Loren Field for helpful discussions.

This work is supported by the Riley Children's Foundation, NIH grant HL69974 (D.G.S.), a grant from the 21st Century Fund from the State of Indiana (J.R.), an NRSA from the NIH (D.L.C.), an American Heart Association postdoctoral fellowship (D.L.C.), an American Heart Association and GAANN predoctoral fellowships (S.R.L.Y.), and an Indiana University Cancer Biology Training Fellowship to J.S.B.

REFERENCES

- Aasland, R., T. J. Gibson, and F. Stewart. 1995. The PHD finger: implications for chromatin-mediated transcriptional regulation. *Trends Biochem. Sci.* **20**:56–59.
- Amir, R. E., I. B. Van den Veyver, M. Wan, C. Q. Tran, U. Francke, and H. Y. Zoghbi. 1999. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat. Genet.* **23**:185–188.
- Ariel, M., H. Cedar, and J. McCarrey. 1994. Developmental changes in methylation of spermatogenesis-specific genes include reprogramming in the epididymis. *Nat. Genet.* **7**:59–63.
- Balaghi, M., and C. Wagner. 1993. DNA methylation in folate deficiency: use of CpG methylase. *Biochem. Biophys. Res. Commun.* **193**:1184–1190.
- Bestor, T. H. 1992. Activation of mammalian DNA methyltransferase by cleavage of a Zn binding regulatory domain. *EMBO J.* **11**:2611–2617.
- Bestor, T. H., and G. L. Verdine. 1994. DNA methyltransferases. *Curr. Opin. Cell Biol.* **6**:380–389.
- Biniszkiewicz, D., J. Gribnau, B. Ramsahoye, F. Gaudet, K. Eggan, D. Humphreys, M.-A. Mastrangelo, Z. Jun, J. Walter, and R. Jaenisch. 2002. Dnmt1 overexpression causes genomic hypermethylation, loss of imprinting, and embryonic lethality. *Mol. Cell. Biol.* **22**:2124–2135.
- Bird, A. P. 1987. CpG islands as gene markers in the vertebrate nucleus. *Trends Genet.* **3**:342–347.
- Burgers, W. A., F. Fuks, and T. Kouzarides. 2002. DNA methyltransferases get connected to chromatin. *Trends Genet.* **18**:275–277.
- Carlone, D. L., S. R. L. Hart, P. D. Ladd, and D. G. Skalnik. 2002. Cloning and characterization of the gene encoding the mouse homologue of CpG binding protein. *Gene* **295**:71–77.
- Carlone, D. L., and D. G. Skalnik. 2001. CpG binding protein is crucial for early embryonic development. *Mol. Cell. Biol.* **21**:7601–7606.
- Chan, M. F., R. Van Amerongen, T. Jijjar, E. Cuppen, P. A. Jones, and P. W. Laird. 2001. Reduced rates of gene loss, gene silencing, and gene mutation in Dnmt1-deficient embryonic stem cells. *Mol. Biol. Cell.* **21**:7587–7600.
- Chen, T., Y. Ueda, J. E. Dodge, Z. Wang, and E. Li. 2003. Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. *Mol. Cell. Biol.* **23**:5594–5605.
- Cross, S. H., R. R. Meehan, X. Nan, and A. Bird. 1997. A component of the transcriptional repressor MeCP1 shares a motif with DNA methyltransferase and HRX proteins. *Nat. Genet.* **16**:256–259.
- Dennis, K., T. Fan, T. Geiman, Q. Yan, and K. Muegge. 2001. Lsh, a member of the SNF2 family, is required for genome-wide methylation. *Genes Dev.* **15**:2940–2944.
- Dignam, J. D., R. M. Lebovitz, and R. G. Roeder. 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* **11**:1475–1489.
- Domer, P. H., S. S. Fakhrazadeh, C. S. Chen, J. Jockel, L. Johansen, G. A. Silverman, J. H. Kersey, and S. J. Korsmeyer. 1993. Acute mixed-lineage leukemia t(4;11)(q21;q23) generates an MLL-AF4 fusion product. *Proc. Natl. Acad. Sci. USA.* **90**:7884–7888.
- Espada, J., E. Ballestar, M. F. Fraga, A. Villar-Garea, A. Juarranz, J. C. Stockert, K. D. Robertson, F. Fuks, and M. Esteller. 2004. Hum. DNA methyltransferase 1 is required for maintenance of the histone H3 modification pattern. *J. Biol. Chem.* **279**:37175–37184.
- FitzGerald, K. T., and M. O. Diaz. 1999. MLL2: a new mammalian member of the trx/MLL family of genes. *Genomics* **59**:187–192.
- Fuks, F., W. A. Burgers, A. Brehm, L. Hughes-Davies, and T. Kouzarides. 2000. DNA methyltransferase Dnmt1 associates with histone deacetylase activity. *Nature Genet.* **24**:88–91.
- Fuks, F., W. A. Burgers, N. Godin, M. Kasai, and T. Kouzarides. 2001. Dnmt3a binds deacetylases and is recruited by a sequence-specific repressor to silence transcription. *EMBO J.* **20**:2536–2544.
- Fuks, F., P. J. Hurd, R. Deplus, and T. Kouzarides. 2003. The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. *Nucleic Acids Res.* **31**:2305–2312.
- Gidekel, S., and Y. Bergman. 2002. A unique developmental pattern of Oct-3/4 DNA methylation is controlled by a cis-demodification element. *J. Biol. Chem.* **277**:34521–34530.
- Goarani, O., P. Karuman, D. R. Jones, D. Ivanov, J. Cha, A. A. Lugovskoy, C. L. Baird, H. Zhu, S. J. Field, S. L. Lessnick, J. Villasenor, B. Mehrotra, J. Chen, V. R. Rao, J. S. Brugge, C. G. Ferguson, B. Payrastre, D. G. Myska, L. C. Cantley, G. Wagner, N. Divecha, G. D. Prestwich, and J. Yuan. 2003. The PHD finger of the chromatin-associated protein ING2 functions as a nuclear phosphoinositide receptor. *Cell* **114**:99–111.
- Gu, Y., T. Nakamura, H. Alder, R. Prasad, O. Canaani, G. Cimino, C. M. Croce, and E. Canaani. 1992. The t(4;11) chromosome translocation of human acute leukemias fuses the ALL-1 gene, related to *Drosophila* trithorax, to the AF-4 gene. *Cell* **71**:701–708.
- Hattori, N., K. Nishino, Y. Ko, N. Harroiri, J. Ohgane, S. Tanaka, and K. Shiota. 2004. Epigenetic control of mouse Oct-4 gene expression in ES cells and TS cells. *J. Biol. Chem.* **279**:17063–17069.
- Hendrich, B., and A. Bird. 1998. Identification and characterization of a family of mammalian methyl-CpG binding proteins. *Mol. Cell. Biol.* **18**:6538–6547.
- Jackson-Grusby, L., C. Beard, R. Possemato, M. Tudor, D. Fambrough, G. Csankovszki, J. Dausman, P. Lee, C. Wilson, E. S. Lander, and R. Jaenisch. 2001. Loss of genomic methylation causes p53-dependent apoptosis and epigenetic deregulation. *Nat. Genet.* **27**:31–39.
- Jeddeloh, J. A., J. Bender, and E. J. Richards. 1998. The DNA methylation locus DDM1 is required for maintenance of gene silencing in Arabidopsis. *Genes Dev.* **12**:1714–1725.
- Jeddeloh, J. A., T. L. Stokes, and E. J. Richards. 1999. Maintenance of genomic methylation requires a SWI2/SNF-like protein. *Nat. Genet.* **22**:94–97.
- Keller, G., M. Kennedy, T. Papayannopoulou, and M. Wiles. 1993. Hematopoietic commitment during embryonic stem cell differentiation in culture. *Mol. Cell. Biol.* **13**:473–486.
- Lee, J.-H., S. R. L. Hart, and D. G. Skalnik. 2004. Histone deacetylase activity is required for embryonic stem cell differentiation. *Genesis* **38**:32–38.
- Lee, J.-H., and D. G. Skalnik. 2002. CpG binding protein is a nuclear matrix- and euchromatin-associated protein localized to nuclear speckles containing human trithorax: identification of nuclear matrix targeting signals. *J. Biol. Chem.* **277**:42259–42267.
- Lee, J.-H., K. S. Voo, and D. G. Skalnik. 2001. Identification and characterization of the DNA binding domain of CpG-binding protein. *J. Biol. Chem.* **276**:44669–44676.
- Lehnertz, B., Y. Ueda, A. A. H. A. Derijck, U. Braunschweig, L. Perez-Burgos, S. Kubicek, T. Chen, E. Li, T. Jenwein, and A. Peters. 2003. Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr. Biol.* **13**:1192–1200.
- Lei, H., S. P. Oh, M. Okano, R. Juttermann, K. A. Goss, R. Jaenisch, and E. Li. 1996. De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. *Development* **122**:3195–3205.
- Li, E. 2002. Chromatin modification and epigenetic reprogramming in mammalian development. *Nat. Rev. Genet.* **3**:662–673.
- Li, E., T. H. Bestor, and R. Jaenisch. 1992. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69**:915–926.
- Ma, Q., H. Alder, K. K. Nelson, D. Chatterjee, Y. Gu, T. Nakamura, E. Canaani, C. M. Croce, L. D. Siracusa, and A. M. Buchberg. 1993. Analysis of the murine ALL-1 gene reveals conserved domains with human ALL-1 and identifies a motif shared with DNA methyltransferases. *Proc. Natl. Acad. Sci. USA* **90**:6350–6354.
- McClanahan, T., S. Dalrymple, M. Barkett, and F. Lee. 1993. Hematopoietic growth factor receptor genes as markers of lineage commitment during in vitro development of hematopoietic cells. *Blood* **81**:2903–2915.
- Miller, T., N. J. Krogan, J. Dover, H. Erdjument-Bromage, P. Tempst, M. Johnston, J. F. Greenblatt, and A. Shilatifard. 2001. COMPASS: A complex of proteins associated with a trithorax-related SET domain protein. *Proc. Natl. Acad. Sci. USA* **98**:12902–12907.
- Miller-Hance, W. C., M. LaCorbiere, S. J. Fuller, S. M. Evans, G. Lyons, C. Schmidt, J. Robbins, and K. R. Chien. 1993. In vitro chamber specification during embryonic stem cell cardiogenesis. *J. Biol. Chem.* **268**:25244–25252.
- Milutinovic, S., Q. Zhuang, A. Niveleau, and M. Szyf. 2003. Epigenomic stress response: knockdown of DNA methyltransferase 1 triggers an intra-S-phase arrest of DNA replication and induction of stress response genes. *J. Biol. Chem.* **278**:14985–14995.
- Monk, M., M. Boubelik, and S. Lehnertz. 1987. Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. *Development* **99**:371–382.
- Nakamura, T., T. Mori, S. Tada, W. Krajewski, T. Rozovskaia, R. Wassell, G. Dubois, A. Mazo, C. M. Croce, and E. Canaani. 2002. ALL-1 is a histone methyltransferase that assembles a supercomplex of proteins involved in transcriptional regulation. *Mol. Cell* **10**:1119–1128.
- Ng, H.-H., Y. Zhang, B. Hendrich, C. A. Johnson, B. M. Turner, H. Erdjument-Bromage, P. Tempst, D. Teinberg, and A. Bird. 1999. MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex. *Nat. Genet.* **23**:58–61.

47. Niwa, H., J. Miyazaki, and A. G. Smith. 2000. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat. Genet.* **24**:372–376.
48. O'Connell, S., L. Wang, S. Robert, C. A. Jones, R. Saint, and R. S. Jones. 2001. Polycomblike PHD fingers mediate conserved interaction with enhancer of zeste protein. *J. Biol. Chem.* **276**:43065–43073.
49. Okano, M., D. W. Bell, D. A. Haber, and E. Li. 1999. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**:247–257.
50. Ono, R., T. Taki, T. Taketani, M. Taniwaki, H. Kobayashi, and Y. Hayashi. 2002. LXX, leukemia-associated protein with a CXXC domain, is fused to MLL in acute myeloid leukemia with trilineage dysplasia having t(10;11)(q22;q23). *Cancer Res.* **62**:4075–4080.
51. Ou, X., J. Pollock, M. C. Dinuer, E. Gharehbaghi-Schnell, and D. G. Skalnik. 1999. Identification and functional characterization of the murine Rac2 gene promoter. *DNA Cell Biol.* **18**:253–263.
52. Pfeifer, G. P., R. L. Tanguay, S. D. Steigerwald, and A. D. Riggs. 1990. In vivo footprint and methylation analysis by PCR-aided genomic sequencing: comparison of active and inactive X chromosomal DNA at the CpG island and promoter of human PGK-1. *Genes Dev.* **4**:1277–1287.
53. Prasad, R., T. Yano, C. Sorio, T. Nakamura, R. Rallapalli, Y. Gu, D. Leshkowitz, C. M. Croce, and E. Canaani. 1995. Domains with transcriptional regulatory activity within the ALL1 and AF4 proteins involved in acute leukemia. *Proc. Natl. Acad. Sci. USA* **92**:12160–12164.
54. Reik, W., W. Dean, and J. Walter. 2001. Epigenetic reprogramming in mammalian development. *Science* **293**:1089–1093.
55. Robertson, K. D., S. Ait-Si-Ali, T. Yokochi, P. A. Wade, P. L. Jones, and A. P. Wolffe. 2000. DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters. *Nat. Genet.* **25**:338–342.
56. Roguev, A., D. Schaft, A. Shevchenko, R. Aasland, A. Shevchenko, and A. F. Stewart. 2003. High conservation of the Set1/Rad6 axis of histone 3 lysine 4 methylation in budding and fission yeasts. *J. Biol. Chem.* **278**:8487–8493.
57. Selker, E. U. 1998. Trichostatin A causes selective loss of DNA methylation in *Neurospora*. *Proc. Natl. Acad. Sci. USA* **95**:9430–9435.
58. Singal, R., and G. D. Ginder. 1999. DNA methylation. *Blood* **93**:4059–4070.
59. Stancheva, I., C. Hensey, and R. R. Meehan. 2001. Loss of the maintenance methyltransferase, xDNMT1, induced apoptosis in *Xenopus* embryos. *EMBO J.* **20**:1963–1973.
60. Stoger, R., P. Kubicka, C.-G. Liu, T. Kafri, A. Razin, H. Cedar, and D. P. Barlow. 1993. Maternal-specific methylation of the imprinted mouse *Igf2r* locus identifies the expressed locus as carrying the imprinting signal. *Cell* **73**:61–71.
61. Szyf, M., V. Bozovic, and G. Tanigawa. 1991. Growth regulation of mouse DNA methyltransferase gene expression. *J. Biol. Chem.* **266**:10027–10030.
62. Tada, M., T. Tada, L. Lefebvre, S. C. Barton, and M. A. Surani. 1997. Embryonic germ cells induce epigenetic reprogramming of somatin nucleus in hybrid cells. *EMBO J.* **16**:6510–6520.
63. Tamaru, H., X. Zhang, D. McMillen, P. B. Singh, J. I. Nakayama, S. I. Grewal, C. D. Allis, X. Cheng, and E. U. Selker. 2003. Trimethylated lysine 9 of histone H3 is a mark for DNA methylation in *Neurospora crassa*. *Nat. Genet.* **34**:75–79.
64. Thorvaldsen, J. L., K. L. Duran, and M. S. Bartolomei. 1998. Deletion of the H19 differentially methylated domain results in loss of imprinted expression of H19 and *Igf2*. *Genes Dev.* **12**:3693–3702.
65. Tilghman, S. M. 1999. The sins of the fathers and mothers: genomic imprinting in mammalian development. *Cell* **96**:185–193.
66. Tkachuk, D. C., S. Kohler, and M. L. Cleary. 1992. Involvement of a homolog of *Drosophila* Trithorax by 11q23 chromosomal translocations in acute leukemias. *Cell* **71**:691–700.
67. Tremblay, K. D., K. L. Duan, and M. S. Bartolomei. 1997. A 5' 2-kilobase-pair region of the imprinted mouse H19 gene exhibits exclusive paternal methylation throughout development. *Mol. Cell. Biol.* **17**:4322–4329.
68. Tucker, K. L., C. Beard, J. Dausmann, L. Jackson-Grusby, P. W. Laird, H. Lei, E. Li, and R. Jaenisch. 1996. Germ-line passage is required for establishment of methylation and expression patterns of imprinted but not of nonimprinted genes. *Genes Dev.* **10**:1008–1020.
69. Tucker, K. L., D. Talbot, M. A. Lee, H. Leonhardt, and R. Jaenisch. 1996. Complementation of methylation deficiency in embryonic stem cells by a DNA methyltransferase minigene. *Proc. Natl. Acad. Sci. USA* **93**:12920–12925.
70. Vannucchi, A. M., S. Linari, C. Cellai, M. J. Koury, and F. Paoletti. 1997. Constitutive and inducible expression of megakaryocyte-specific genes in Friend erythroleukaemia cells. *Br. J. Haematol.* **99**:500–508.
71. Voo, K. S., D. L. Carlone, B. M. Jacobsen, A. Flodin, and D. G. Skalnik. 2000. Cloning of a mammalian transcriptional activator that binds unmethylated CpG motifs and shares a CXXC domain with DNA methyltransferase, human trithorax, and methyl-CpG binding domain protein 1. *Mol. Cell. Biol.* **20**:2108–2121.
72. Walsh, C. P., J. R. Chaillet, and T. Bestor. 1998. Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. *Nat. Genet.* **20**:116–117.
73. Wilkinson, D. G., S. Bhatt, and B. G. Herrmann. 1990. Expression pattern of the mouse T gene and its role in mesoderm formation. *Nature* **343**:657–659.
74. Wobus, A., G. Kaomei, J. Shan, M.-C. Wellner, J. Rohwedel, J. Guanju, B. Fleischmann, H. A. Katus, J. Hescheler, and W.-M. Franz. 1997. Retinoic acid accelerates embryonic stem cell-derived cardiac differentiation and enhances development of ventricular cardiomyocytes. *J. Mol. Cell. Cardiol.* **29**:1525–1539.
75. Wysocka, J., M. P. Myers, C. D. Laherty, R. N. Eisenmann, and W. Herr. 2003. Human Sin3 deacetylase and trithorax-related Set1/Ash2 histone H3/K4 methyltransferase are tethered together selectively by the cell-proliferation factor HCF-1. *Genes Dev.* **17**:896–911.
76. Xu, C., G. Liguori, E. D. Adamson, and M. G. Persico. 1998. Specific arrest of cardiogenesis in cultured embryonic stem cells lacking *Cripto-1*. *Dev. Biol.* **196**:237–247.
77. Xu, G.-L., T. H. Bestor, D. Bourc'his, C.-L. Hsieh, N. Tommerup, M. Bugge, M. Hulten, X. Qu, J. J. Russo, and E. Viegas-Pequignot. 1999. Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. *Nature* **402**:187–191.
78. Yoder, J. A., C. P. Walth, and T. H. Bestor. 1997. Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet.* **13**:335–340.
79. Yokoyama, A., Z. Wang, J. Wysocka, M. Sanyal, D. J. Aufiero, I. Kitabayashi, W. Herr, and M. L. Cleary. 2004. Leukemia proto-oncoprotein MLL forms a SET1-like histone methyltransferase complex with menin to regulate Hox gene expression. *Mol. Cell. Biol.* **24**:5639–5649.
80. Zeleznik-Le, N., A. M. Harden, and J. D. Rowley. 1994. 11q23 translocations split the "AT-hook" cruciform DNA-binding region and the transcriptional repression domain from the activation domain of the mixed-lineage leukemia (MLL) gene. *Proc. Natl. Acad. Sci. USA* **91**:10610–10614.
81. Zhang, Y., H.-H. Ng, H. Erdjument-Bromage, P. Tempst, A. Bird, and D. Reinberg. 1999. Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. *Genes Dev.* **13**:1924–1935.