Genome-Wide Profiling of CpG Methylation Identifies Novel Targets of Aberrant Hypermethylation in Myeloid Leukemia

Claudia Gebhard, Lucia Schwarzfischer, Thu-Hang Pham, Elmar Schilling, Maja Klug, Reinhard Andreesen, and Michael Rehli

Department of Hematology and Oncology, University Hospital, Regensburg, Germany

Abstract

The methylation of CpG islands is associated with transcriptional repression and, in cancer, leads to the abnormal silencing of tumor suppressor genes. Because aberrant hypermethylation may be used as a marker for disease, a sensitive method for the global detection of DNA methylation events is of particular importance. We describe a novel and robust technique, called methyl-CpG immunoprecipitation, which allows the unbiased genome-wide profiling of CpG methylation in limited DNA samples. The approach is based on a recombinant, antibody-like protein that efficiently binds native CpG-methylated DNA. In combination with CpG island microarrays, the technique was used to identify >100 genes with aberrantly methylated CpG islands in three myeloid leukemia cell lines. Interestingly, within all hypermethylation targets, genes involved in transcriptional regulation were significantly overrepresented. More than half of the identified genes were absent in microarray expression studies in either leukemia or normal monocytes, indicating that hypermethylation in cancer may be largely independent of the transcriptional status of the affected gene. Most individually tested genes were also hypermethylated in primary blast cells from acute myeloid leukemia patients, suggesting that our approach can identify novel potential disease markers. The technique may prove useful for genome-wide comparative methylation analysis not only in malignancies. (Cancer Res 2006; 66(12): 6118-28)

Introduction

The transcriptional state of a gene is largely determined by its local epigenetic code. Pathologic changes in chromatin structure or DNA methylation that lead to abnormal expression or repression of genes are now being recognized as important contributing factors in developmental diseases, cancer, and aging (1–5).

In cancer, the epigenetic silencing of tumor suppressor genes seems to be a common, nonrandom, and tumor type-specific event that often coincides with the aberrant methylation of CpG dinucleotides in so-called CpG islands. These regions of high CpG content are mostly unmethylated in normal cells and frequently contain gene promoters and transcription start sites (2, 6–8).

Hypermethylation of CpG islands seems to be a tumor type-specific event (9, 10), and current efforts are concentrating on finding ways to exploit the diagnostic and therapeutic implications of the abnormalities (11, 12). A comprehensive knowledge of the methylation profile of a given tumor may provide important information for risk assessment, diagnosis, monitoring, and treatment (6, 13).

The investigation of aberrant CpG island methylation has primarily taken a candidate gene approach. Assessment of the clinical potential of hypermethylation profiles and the identification of relevant marker genes, however, require means and methods to detect hypermethylation on a genome-wide level. Here, we apply a novel and relatively simple technique that allows the generation of unbiased, genome-wide profiles of normal or aberrant CpG island methylation. The procedure is independent of methylation-sensitive restriction endonucleases or bisulfite treatment, which are of limited use for genome-wide profiling strategies (14). Our approach is based on a recombinant methyl-CpG-binding, antibody-like protein that was engineered to efficiently bind CpG-methylated DNA fragments in a so-called methyl-CpG immunoprecipitation (MCIp) assay. In combination with appropriate DNA microarrays, MCIp-enriched DNA fragments can be used to detect hypermethylation events on a genome-wide scale. The efficacy of this system was tested by the profiling of three leukemia cell lines using a 12K CpG island microarray, which led to the identification of many novel genes that are hypermethylated in human myeloid leukemia. The novel profiling technique provides a powerful tool to identify changes in DNA methylation as well as novel marker genes of potential diagnostic or prognostic value.

Materials and Methods

Reagents. All chemical reagents used were purchased form Sigma-Aldrich (Taufkirchen, Germany) unless otherwise noted. Oligonucleotides were synthesized and high-pressure liquid chromatography purified by Metabion (Planegg-Martinsried, Germany). DNA sequencing was done by Entelechon (Regensburg, Germany).

Cells. Human monocytes were isolated as described previously (15). The human myeloid leukemia cell lines THP-1, KG-1, and U937 were grown in RPMI 1640 supplemented with 10% FCS (Life Technologies, Eggenstein, Germany). For demethylation experiments, U937 cells were grown in the presence of decitabine (Sigma-Aldrich), and the medium was replaced every 24 hours. *Drosophila* Schneider 2 (S2) cells (American Type Culture Collection, Rockville, MD) were cultured in Insect-Xpress medium (BioWhittaker, Heidelberg, Germany) containing 10% FCS (PAA, Cölbe, Germany) in an incubator at 21°C. Fresh peripheral blood samples from 12 patients with newly diagnosed and untreated *de novo* acute myeloid leukemia (AML) were used for the study. Patients were treated according to the protocol AMLCG-2000 of the German AML Cooperative Group. Written informed consent was obtained from each patient.

Plasmid construction. A cDNA corresponding to the methyl-CpG-binding domain (MBD) of human MBD2 (Genbank accession no. NM_003927; AA 144-230) was PCR amplified from reverse-transcribed human macrophage total RNA using primers MBD2-Nhe_S (5′-AGATGCTAGCACGGAGAGCGGGAAGAGG-3′) and MBD2-Not_AS (5′-ATCACGCGGCCGCCAGAGGATCGTTTCGCAGTCTC-3′) and Herculase DNA polymerase

Note: Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org/).

Requests for reprints: Michael Rehli, Department of Hematology and Oncology, University Hospital, 93042 Regensburg, Germany. Phone: 49-941-944-5587; Fax: 49-941-944-5593; E-mail: michael.rehli@klinik.uni-r.de.

doi:10.1158/0008-5472.CAN-06-0376

(Stratagene, Heidelberg, Germany). The PCR product was directly cloned into *Not*I/*Nhe*I sites of signal-pIg plus vector (Ingenious, R&D Systems, Abingdon, United Kingdom) and sequence verified. An *Apa*I/*Nhe*I fragment of pIg/MBD-Fc, containing the MBD of human MBD2 fused to the Fc tail of human IgG1, was subcloned into *Apa*I/*Spe*I sites of pMTBiP/V5-His B (Invitrogen, Karlsruhe, Germany), resulting in pMTBip/MBD-Fc.

DNA preparation. Genomic DNA from various cell types was prepared using the Blood and Cell Culture DNA Midi kit (Qiagen, Hilden, Germany). At least 1 μg genomic DNA was digested using *Mse*I or a combination of *Mse*I and *Csp*6I. Completion of the digest was controlled using agarose gel electrophoresis, and digested DNA was quantified using PicoGreen dsDNA quantitation reagent (Molecular Probes, Karlsruhe, Germany).

Protein expression. *Drosophila* S2 cells (4×10^6) per 60-mm cell culture dish were stably transfected with a mixture of 1.5 μg pMTBip/MBD-Fc and 0.3 μg pCoHygro (Invitrogen) using Effectene transfection reagent (Qiagen) according to the manufacturer's protocol. For large-scale production, 1×10^8 to 5×10^8 cells were cultured in 100 to 200 mL Insect-Xpress without FCS and hygromycin in 2,000 mL roller bottles for 2 days before the addition of 0.5 mmol/L $CuSO_4$. Culture medium was harvested every 4 to 7 days, and cell culture supernatants were purified using a protein A-Sepharose column.

Reverse Southwestern blot. DNA fragments were separated by agarose gel electrophoresis and directly blotted (without prior denaturation) onto a nylon membrane using a capillary transfer system equivalent to traditional Southern blotting procedures. DNA was UV cross-linked after transfer and blocked overnight with 5% fat-free powdered milk in TBST [20 mmol/L Tris-HCl (pH 7.4), 150 mmol/L NaCl, 0.1% Tween 20]. Blots were washed thrice in TBST for 10 minutes at room temperature. Blots were then incubated with MBD-Fc (1:20,000 dilution) in TBST with 5% milk powder for 1 hour at room temperature, washed thrice (TBST, 10 minutes, room temperature), and incubated with horseradish peroxidase (HRP)–conjugated anti-human Fc (1:10,000) in TBST with 5% milk powder for 1 hour at room temperature. After three more washes (TBST, 10 minutes, room temperature), bands were detected using enhanced chemiluminescence (ECL; Amersham, Little Chalfont, United Kingdom).

Methyl-CpG immunoprecipitation. Purified MBD-Fc protein (usually 7.5 μg) was added to 50 μL protein A-Sepharose 4 Fast Flow beads (Amersham) in 1 mL TBS and rotated overnight on a rotator at 4°C.

On the next day, MBD-Fc beads were transferred into SpinX columns (Sigma-Aldrich) and spin washed twice with buffer A [20 mmol/L Tris-HCl (pH 8), 2 mmol/L $MgCl_2$, 0.5 mmol/L EDTA, 300 mmol/L NaCl, 0.1% NP40]. Digested DNA (usually 300 ng) was added to the washed MBD-Fc beads in 350 μL buffer A and rotated in sealed SpinX columns for 3 hours on a rotator at 4°C. Beads were centrifuged to recover unbound DNA (300 mmol/L fraction) and subsequently washed with increasing NaCl concentrations (400-600 mmol/L). Flow through of each wash step was either discarded or collected in separate tubes for further analyses. Highly CpG-methylated DNA was eluted with 350 μL buffer E [20 mmol/L Tris-HCl (pH 8), 2 mmol/L $MgCl_2$, 0.5 mmol/L EDTA, 1,000 mmol/L NaCl, 0.1% NP40] into a separate tube. Eluted DNA was desalted using QIAquick PCR purification kit (Qiagen). In parallel, 300 ng digested DNA was resuspended in 350 μL buffer D and desalted using QIAquick PCR purification kit. DNA preparations were quantified using PicoGreen dsDNA quantitation reagent.

Real-time genomic PCR. Enrichment of a specific *Mse*I fragment in the MCIp eluate or in MCIp amplicons was detected and quantified relative to the genomic input by real-time LightCycler PCR using the QuantiTect kit (Qiagen) according to the manufacturer's instructions. Primer sequences are given in the Supplementary Table S2. Cycling parameters were as follows: denaturation, 95°C, 15 minutes and amplification, 95°C, 15 seconds, 57°C, 20 seconds, and 72°C, 25 seconds for 42 cycles. The product size was initially controlled by agarose gel electrophoresis, and melting curves were analyzed to control for specificity of the PCRs.

Real-time reverse transcription-PCR. Total RNA was prepared using the RNeasy Midi kit. Contaminating genomic DNA was removed from the samples using TURBO DNA-free kit (Ambion, Huntingdon, United Kingdom) before 1 μg total RNA was reverse transcribed using Moloney murine leukemia virus reverse transcriptase, RNase H minus, point mutant (Promega, Mannheim, Germany). Real-time PCR was done using the LightCycler (Roche, Mannheim, Germany) with the QuantiTect kit according to the manufacturer's instructions. Primers used are given in Supplementary Table S4. Cycling parameters were as above.

CpG island microarray analysis. To generate fluorescently labeled DNA for CpG island microarray hybridization, *Mse*I-compatible unidirectional ligation-mediated PCR (LMPCR) linker (LMPCR_S-L: 5′-GCGGTGACC-CGGGAGATCTCTTAAG-3′ and LMPCR_AS-L: 5′-TACTTAAGAGATC-3′, 20 μmol/L) was ligated to the MCIp-eluted DNA and in a separate reaction to an equal amount of input DNA (0.5 μL linker/ng DNA) in 60 μL reactions using 1,200 units T4 ligase (New England Biolabs, Frankfurt am Main, Germany) at 16°C overnight. Linker-ligated DNA was desalted using QIAquick PCR purification kit. Amplification of linker-ligated DNA preparations was done using LMPCR primer (5′-GTGACCCGGGA-GATCTCTTAAG-3′) and Taq polymerase (Roche) in the presence of 1.3 mol/L betaine. Amplicons were desalted using QIAquick PCR purification kit and quantified (PicoGreen dsDNA quantitation reagent). Labeling and hybridization of MCIp amplicons were done by the Kompetenzzentrum für fluoreszente Bioanatytik (KFB; Regensburg, Germany) according to the protocol provided by the CpG island microarray manufacturer (University Health Network Microarray Centre, Toronto, Ontario, Canada) with modifications. Briefly, normal and tumor MCIp amplicons (4 μg) were directly labeled with Cy5-dUTP and Cy3-dUTP, respectively, using the BioPrime Array Comparative Genomic Hybridization Genomic Labeling System (Invitrogen). Each fluorescently labeled and purified DNA amplicon (4 μg) in 300 μL digoxygenin Easy Hyp solution (Roche) supplemented with 25 μg Cot-1 DNA (Invitrogen) and 30 μg yeast tRNA was hybridized to human CpG 12K arrays (HCGI12K, University Health Network Microarray Centre) in Gene Frames measuring 60 × 21 mm (ABgene, Hamburg, Germany) at 37°C for overnight. Slides were washed thrice in 1× SSC, 0.1% SDS at 50°C for 10 minutes. After two more rinses with 0.1× SSC, slides were dried and scanned using the Affymetrix 428 scanner (Wooburn Green, United Kingdom). Images were analyzed using the ImaGene 5.6 and Gene Sight Lite software (BioDiscovery, Inc., El Segundo, CA). Locally weighted scatterplot smoothing normalization was used to normalize Cy3 and Cy5 signals. Clones that produced reproducible differential signals on the CpG island microarray were sequenced by the University Health Network Microarray Centre.

Affymetrix microarray analysis. RNA from KG-1, U937, and THP-1 cells as well as from freshly isolated human blood monocytes of a health donor was analyzed using Affymetrix HG-U133_Plus_2 arrays. Hybridization, cRNA labeling, and data handling were done by the KFB.

Sodium bisulfite sequencing. Modification of DNA with sodium bisulfite was done as described previously (16) with modifications (17). Bisulfite-treated DNA was amplified in individual PCRs using the primers given in Supplementary Table S3. PCR products (representing the sense strand) were cloned using the TOPO cloning kit (Invitrogen), and several individual clones were sequenced.

Databases for sequence analysis. Mapping of CpG island clones was done using the University of California at Santa Cruz genome browser (http://genome.ucsc.edu/). Expression profiles for myeloid cell types other than monocytes were obtained from Genomics Institute of the Novartis Research Foundation SymAtlas (http://symatlas.gnf.org/SymAtlas/).

Gene Expression Omnibus accession numbers. Microarray data have been deposited with the Gene Expression Omnibus database (accession numbers for Affymetrix gene arrays: GSM73641, GSM73642, GSM73644, and GSM73645; accession numbers for CpG island microarrays: GSM91623-30).

Results

Design, generation, and properties of a recombinant antibody-like MBD-Fc. To enable an antibody-like detection of double-stranded CpG-methylated DNA, we constructed a vector encoding a fusion protein comprising the MBD of human MBD2, a flexible linker polypeptide, and the Fc portion of human IgG1. The MBD-Fc polypeptide was expressed under the control of a metal-inducible promoter in *Drosophila* S2 cells and collected from the

supernatant via protein A affinity chromatography. The purified protein had the expected molecular weight of 40 kDa. To test whether MBD-Fc was able to detect CpG-methylated DNA, we blotted *in vitro*–methylated PCR fragments with different CpG density onto a nylon membrane using a capillary transfer system equivalent to traditional Southern blotting, however, without denaturing the DNA before blotting. As shown in Fig. 1*A*, using standard immunoblot conditions and MBD-Fc as equivalent to the primary antibody, methylated DNA could be detected on nylon membranes in a linear fashion.

We next tested MBD-Fc binding of *in vitro* generated and differentially methylated DNA fragments in an immunoprecipitation-like approach. PCR fragments of human promoters with varying CpG density were generated using PCR and *in vitro* CpG methylated using *Sss*I or left unmethylated as described in Supplementary Data. The mixture of methylated and unmethylated DNA fragments was bound to MBD-Fc protein A-Sepharose beads and eluted using increasing concentrations of NaCl. Fractions were collected, spin purified, and subjected to agarose gel electrophoresis. As shown in Supplementary Fig. S1, the affinity of a methylated fragment increased with the density of methylated CpG dinucleotide, with unmethylated DNA eluting at relatively low salt concentrations and highly methylated DNA eluting at high salt concentrations. Variation of the amount of input DNA within the binding capacity of the MBD-Fc polypeptide did not significantly change the elution profile. However, the salt-dependent affinity of methylated DNA was contingent on the density of the MBD-Fc fusion protein on the protein A-Sepharose beads. To clarify whether this approach allows the detection of differential methylation of a single gene locus, a CpG island promoter fragment was cloned into a CpG-free vector. The DNA was partially *in vitro* methylated, fractionated by MCIp, and subjected to bisulfite sequencing. As shown in Supplementary Fig. S2, partially *in vitro*–methylated DNA fragments were separable according to their methylation degree in a fractionating approach using increasing concentrations of NaCl.

Combination of MCIp and real-time PCR to detect the methylation status of specific CpG island promoters. We assumed that MCIp may be used to discriminate methylated and unmethylated DNA fragments from genomic DNA. To explore this type of application on single-gene level, we used the above procedure to precipitate *Mse*I-restricted genomic DNA of *in vitro* *Sss*I methylated and untreated normal DNA from monocytes of a healthy donor. *Mse*I was chosen for DNA fragmentation because it is known to preferentially cut in regions of low CpG content while leaving many CpG islands uncut (18).

The salt concentration-dependent enrichment of four different CpG island promoters and a promoter with low CpG density was determined in *Sss*I-methylated and untreated DNA relative to the input DNA using LightCycler real-time PCR. As a positive control for DNA methylation, we used the *SNRPN* gene promoter that is subject to maternal imprinting with one of its two copies being methylated also in normal cells (19). In normal DNA, the two differentially methylated allele fragments of *SNRPN* were enriched in two separate fractions (Fig. 1*B*). Only one enriched fraction was observed with *Sss*I-methylated DNA. For *CDKN2B* gene (also known as $p15^{INK4b}$), which is frequently methylated in leukemia cells (Fig. 1*C*; refs. 20–22), the fragment was detected mainly in a low salt fraction from normal DNA and in the high salt fraction from *Sss*I-methylated DNA. Similar results were obtained for the human *estrogen receptor 1* (*ESR1*) and human *TLR2* genes (data not shown). The profiles of methylated and unmethylated DNA at the *CHI3L1* locus were significantly different from those of the above-tested CpG island promoters (Fig. 1*D*). Most of the untreated *CHI3L1* fragment was recovered at lower NaCl concentrations, and a shift was observed toward higher NaCl concentrations when the DNA was *Sss*I methylated. The average difference between *Sss*I-treated and untreated monocyte DNA at the *CHI3L1* locus is ~5 to 6 (out of 12) methylated CpG residues (data not shown), suggesting that the fractionated approach is able to discriminate relatively small differences in CpG methylation. Analysis of the above elution

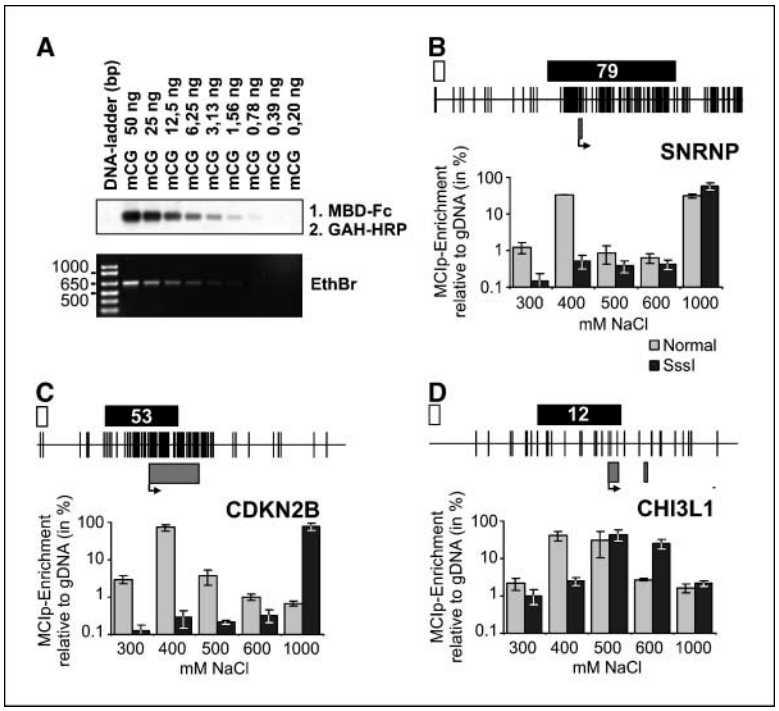

Figure 1. Detection of CpG-methylated DNA using the recombinant MBD-Fc protein. *A,* PCR fragment of a CpG island promoter (*ICSBP* gene) was methylated using *Sss*I, subjected to agarose gel electrophoresis [ethidium bromide (*EthBr*) staining is control], and directly blotted onto a nylon membrane. Membranes were stained using MBD-Fc, HRP-conjugated anti-human Fc, and ECL reagent (Amersham). *B* to *D,* fractionated MCIp was used in combination with real-time LightCycler PCR to detect the methylation status of the indicated genes from untreated (*gray columns*) and *Sss*I-methylated and *Mse*I-restricted genomic DNA fragments (*black columns*). Gene fragments recovered from MCIp eluates [NaCl concentrations (in mmol/L)] and an equivalent amount of genomic input DNA (*gDNA*) were amplified by LightCycler PCR. A standard curve (3 log scales) was used to obtain the concentration of a specific gene fragment in the MCIp eluate relative to its concentration in the genomic input DNA. Values of individual fractions represent the percentage of recovery and are calculated relative to the amount of PCR product generated from the respective genomic input DNA (100%). *Columns,* mean ($n = 4$) using at least two different preparations of MBD-Fc (10 log scales); *bars,* SD. A 3-kb region of the corresponding CpG island is represented above each figure. *Vertical lines,* CpG dinucleotide. *Gray boxes,* positions of exons. *Arrows,* transcription start sites. *White boxes,* 100 bp; *black boxes,* positions of the *Mse*I fragments that are detected. *Numbers,* number of CpG dinucleotides within the *Mse*I fragment.

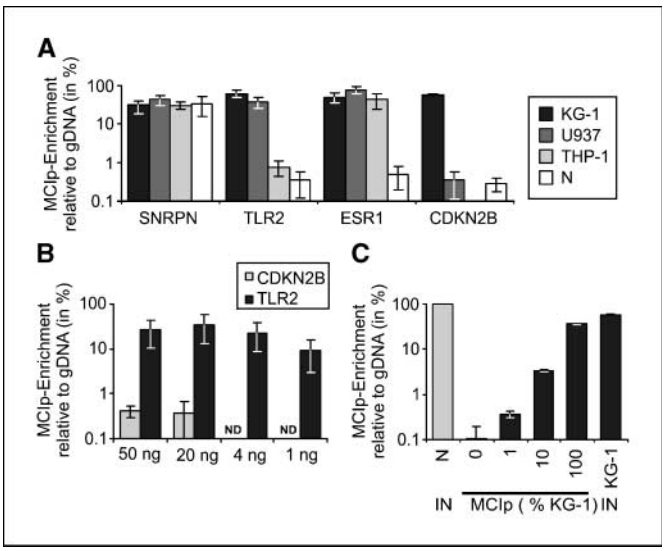

Figure 2. MCIp detection of CpG island methylation in specific CpG island promoters using real-time PCR. *A, SNRPN, TLR2, ESR1*, and *CDKN2B* gene fragments in the high salt (1,000 mmol/L after 600 mmol/L) MCIp fraction of three human myeloid leukemia cell lines (KG-1, U937, and THP-1) as well as normal human blood monocytes (*N*) were analyzed by real-time PCR as described in Fig. 1. *B,* decreasing amounts of *Mse*I-treated U937 DNA were subjected to MCIp. *CDKN2B* and *TLR2* gene fragments were quantified as above. *C, Mse*I-treated DNA of normal human blood monocytes and KG-1 cells was mixed at the indicated ratios, the mixture was subjected to MCIp, and the *TLR2* gene fragment was quantified in the 1,000 mmol/L fraction using LightCycler PCR as described above.

profiles suggests that (*a*) a 200- to 300-fold enrichment of stronger over less methylated genomic fragments can be obtained in either low or high salt fractions, (*b*) fragments with low CpG density are largely excluded from the high salt fraction, and (*c*) the fractionated MCIp approach may allow the resolution of relatively small differences in CpG methylation density.

To test whether MCIp can detect aberrant hypermethylation in tumor samples, DNA from three leukemia cell lines, KG-1 (AML), U937 (histiocytic malignancy, monocytic), and THP-1 (acute monocytic leukemia), as well as from monocytes of a healthy donor were analyzed for *SNRNP, CDKN2B, ESR1*, and *TLR2* promoter enrichment in the high salt fraction (Fig. 2*A*). The *TLR2* gene promoter was enriched in KG-1 and U937 cells but not in THP-1 or normal cells. The methylation pattern of *TLR2* was confirmed by bisulfite sequencing (data not shown; ref. 17). Results for *CDKN2B* (methylated in KG-1 and unmethylated in U937) and *ESR1* (methylated in KG-1) were in line with previously published studies (20, 22, 23). None of the three *Mse*I fragments were significantly enriched in the DNA from normal cells. In concordance with its imprinting-related methylation status, the *SNRPN* gene promoter was significantly enriched in all leukemia cell lines as well as in normal cells. These experiments established that the high salt MCIp fraction specifically enriches CpG-rich genomic DNA fragments with a high degree of CpG methylation.

To test the sensitivity of our approach, decreasing amounts of U937 DNA were analyzed using the MCIp approach. The enrichment of *TLR2* (strong methylation) and *CDKN2B* gene fragments (no methylation) was determined by LightCycler real-time PCR. As shown in Fig. 2*B*, a significant enrichment of the TLR2 fragment was achieved using as little as 1 ng genomic DNA fragments (equivalent to ~150 tumor cells) for the MCIp procedure.

Samples derived from tumors may contain significant numbers of normal cells, which are expected to be unmethylated at most CpG islands. To test the linearity of methyl-CpG detection with respect to cell purity, MCIp was done using mixtures of DNA from normal blood cells and from the leukemia cell line KG-1 showing high levels of CpG island methylation at several promoters. As shown in Fig. 2*C*, the *TLR2* promoter fragment was only detected in samples containing KG-1 DNA and the signal gradually increased with the proportion of methylated DNA in the sample. Similar results were obtained for the *ESR1* locus (data not shown).

Combination of MCIp and CpG island microarray analysis to generate unbiased genome-wide promoter methylation profiles. To achieve a genome-wide analysis of CpG island DNA methylation, we combined MCIp with CpG island microarray hybridization. A small amount of *Mse*I-restricted DNA (300 ng) of three leukemia cell lines (KG-1, U937, and THP-1) and of normal human peripheral blood monocytes was subjected to MCIp. The isolated DNA was amplified using LMPCR.

The resulting amplicons were directly labeled with Cy5 (normal DNA) and Cy3 (leukemia cell DNA), and each leukemia sample was cohybridized with the normal control sample to CpG island microarrays (University Health Network Microarray Centre; Supplementary Fig. S3). This array contains 12,192 CpG island clones from a *Mse*I CpG DNA library that was originally prepared by MeCP2 column purification of nonmethylated CpG island fragments (18). Representative scatter plots of microarray hybridizations are presented in Supplementary Fig. S3C. In hybridizations using the amplicons from tumor cell lines, signals corresponding to both hypomethylated and hypermethylated fragments were observed. Because we were interested in tumor-specific epigenetic silencing, we focused on the analysis of hypermethylated fragments. To identify CpG island fragments that were affected by hypermethylation, results of three independent MCIp experiments (using two different MBD-Fc preparations and three independent DNA preparations) were analyzed in conjunction. Hybridization signals that were consistently different (>2-fold enriched in the leukemia sample) in at least one cell line were selected for further analysis.

In total, THP-1 cells showed 277 differential hybridization signals, U937 cells showed 454, and KG-1 cells showed 330. One hundred ninety-one of 535 spots in total were unambiguously annotated and located in the close proximity (approximately ±3,000 bp) to predicted transcriptional start sites and were chosen for further analysis. Several sequences were represented more than once on the CpG island microarray. The final, nonredundant list of differentially represented CpG island DNA fragments contained 131 entries that were in the close proximity to 134 genes (Table 1). At least nine of the listed genes [*LMX1A* (24), *TFAP2A* (25), *CR2* (26), *DCC* (27), *MYOD1* (28), *DLEC1* (29), *AKAP12* (30), *SSIAH2* (*LOC283514*), and *FOXF1* (31)] have been identified previously as targets of hypermethylation in cancer.

The hypermethylated genes listed in Table 1 are involved in many biological functions. Most strikingly, half of the genes with an assigned molecular function (46 of 89) are involved in DNA binding and transcriptional regulation.

Experimental validation of microarray results. A representative number of gene fragments that were identified using combined MCIp chip analyses were selected for further validation. LightCycler real-time PCR was used to measure the MCIp enrichment of 29 candidate genes (Fig. 3; Supplementary Figs. S4 and S5). Out of these, 26 gene fragments were enriched in a manner comparable with the results obtained by microarray analysis. To

Table 1. Hypermethylated gene fragments in myeloid leukemia cell lines

Gene			CpG−methylation			Positon	mRNA expression				
Name	Symbol	Location	KG1	U937	THP1		KG1	U937	THP1	N	Probe Set ID
hypothetical gene	LOC400027	12q12	1.82	1.16	1.19	down	NC/P	NC/P	NC/P	A	226413_at
branched chain aminotransferase 2, mitochondrial	BCAT2	19q13	1.24	1.18	1.29	proximal	1	1.8	NC/P	P	203576_at
chondrolectin	CHODL	21q11.2	1.52	1.49	1.44	down	NC/A	NC/A	NC/A	A	219867_at
collagen, type XIV, alpha 1 (undulin)	***COL14A1***	8q23	2.02	2.97	2.03	down	NC/A	NC/A	NC/A	A	1562189_at
cytochrome P450, family 27, subfamily B, polypeptide 1	***CYP27B1***	12q13.1−q13.3	1.55	1.87	2.15	down	NC/A	NC/P	0.8	A	205676_at
v-erb-a erythroblastic leukemia viral oncogene homolog 4	ERBB4	2q33.3−q34	1.63	1.34	1.36	down	NC/A	NC/A	NC/A	A	241581_at
family with sequence similarity 5, member B	FAM5B	1q24.1−q25.3	2.55	1.32	1.92	proximal	NC/A	NC/A	NC/A	A	214822_at
fibroblast growth factor 12	FGF12	3q28	1.20	1.76	1.52	proximal/down	NC/A	NC/A	NC/A	A	240067_at
hypothetical gene	FLJ13192	15q14	1.79	1.49	1.37	down	NC/A	NC/A	NC/A	A	233382_at
hypothetical gene	FLJ20366	8q23.2	1.90	1.39	1.45	up/down	NC/A	NC/A	NC/A	A	218692_at
hypothetical gene	FLJ36633					up	ND	ND	ND	ND	NA
hypothetical protein	FLJ20972	1p34.2	1.46	1.29	1.82	down	NC/M	NC/P	NC/P	P	230897_at
hypothetical protein	FLJ35074	6p24	2.07	1.09	1.59	down	NC/A	1.6	NC/A	A	1560503_a_at
transcription factor AP-2 alpha	TFAP2A					up	NC/A	2.4	NC/A	A	204653_at
laeverin	FLJ90650	5q23.1	2.14	3.00	2.77	up/down	NC/A	NC/A	NC/A	A	235382_at
forkhead box F1	*FOXF1*	16q24	1.40	1.84	2.72	down	2.7	NC/A	NC/A	A	205935_at
glycoprotein M6A	GPM6A	4q34	1.30	2.05	1.19	down	NC/A	NC/A	NC/A	A	209469_at
GS homeobox 2	GSH2	4q11−q12	1.34	2.29	1.26	down	NC/A	NC/A	NC/A	A	230338_x_at
hypocretin (orexin) receptor 2	HCRTR2	6p11−q11	1.08	1.46	1.62	down	NC/A	NC/A	NC/A	A	207393_at
Hey-like transcriptional repressor	HELT	4q35.1	1.37	3.51	1.52	up	ND	ND	ND	ND	NA
homeo box C10	HOXC10	12q13.3	1.13	1.62	1.01	up	NC/A	0.8	NC/P	A	214562_at
iroquois homeobox protein 1	IRX1	5p15.3	2.20	2.32	2.07	down	NC/A	NC/A	NC/A	A	230472_at
hypothetical protein	KIAA1024	15q25.1	1.04	2.06	1.73	down	NC/A	NC/A	NC/A	A	215081_at
LIM homeobox transcription factor 1, alpha	LMX1A	1q22−q23	2.42	2.08	2.29	down	NC/A	NC/A	NC/A	A	1553541_at
similar to seven in absentia 2 (SSIAH2)	***LOC283514***	13q14.13	2.43	2.73	1.54	proximal	NC/A	NC/A	NC/A	A	1560676_at
hypothetical protein	MGC12982	1p33	2.17	2.36	1.92	up	NC/P	NC/A	NC/A	A	207653_at
forkhead box D2	FOXD2					down	NC/P	NC/A	NC/A	A	224457_at
hypothetical protein	MGC42090	7p21.1	2.62	4.03	1.97	proximal	NC/A	NC/A	NC/A	A	1552293_at
hypothetical protein	MGC4767	12q24.31	1.68	2.58	2.23	proximal	1	1.2	NC/P	P	223114_at
myogenic differentiation 1	*MYOD1*	11p15.4	2.49	2.57	1.14	down	NC/A	NC/A	NC/A	A	206657_s_at
NK2 transcription factor related, locus 3 (Drosophila)	NKX2-3	10q24.2	2.96	4.06	1.30	down	NC/A	NC/A	3.3	A	1553808_a_at
one cut domain, family member 1	ONECUT1	15q21.1−q21.2	1.58	2.64	2.49	up	NC/A	NC/A	NC/A	P	210745_at
protocadherin gamma subfamily B, 1	PCDHGB1	5q31	1.12	1.69	3.19	proximal	ND	ND	ND	ND	NA
phospholipase A2, group VII	***PLA2G7***	6p21.2−p12	1.39	1.36	2.27	proximal	−8.9	−5.1	−5.7	P	206214_at
phospholipase D family, member 5	PLD5	1q43	1.43	1.69	1.92	down	NC/A	NC/P	NC/A	A	1563933_a_at
scinderin	SCIN	7p21.3	1.52	1.42	1.15	proximal/up	NC/A	NC/A	NC/A	A	239365_at
SLIT and NTRK-like family, member 3	***SLITRK3***	3q26.1	2.56	1.42	2.56	down	NC/A	NC/A	NC/A	A	206732_at
Sp5 transcription factor	SP5	2q31.1	1.30	1.89	1.00	down	NC/A	NC/A	NC/A	A	235845_at
transcription factor AP-2 gamma	TFAP2C	20q13.2	1.60	1.02	1.07	up	NC/P	NC/A	NC/A	A	205286_at
transmembrane protein 39A	TMEM39A	3q13.33	1.61	1.48	1.64	prom	NC/P	NC/P	0.6	P	222690_s_at
zinc finger protein 483	ZNF483	9q31.3	1.98	1.71	2.43	down	NC/A	NC/P	NC/P	A	1570534_a_at
zinc finger protein 565	ZNF565	19q13.12	2.08	1.74	1.80	down	NC/P	NC/P	0.9	P	228305_at
hypothetical gene	AF086288	9p24	1.75	1.34	−0.52	proximal	NC/A	NC/A	NC/A	A	237421_at
hypothetical gene	AY358245	15q24	1.02	1.33	0.01	proximal	ND	ND	ND	ND	NA
hypothetical gene	BC026095	11q12.1	1.40	1.26	−0.11	down	NC/A	NC/A	NC/A	A	1570068_at
bone morphogenetic protein 4	BMP4	14q22−q23	2.07	1.10	0.82	down	NC/A	NC/A	NC/A	A	211518_s_at

(*Continued on the following page*)

Table 1. Hypermethylated gene fragments in myeloid leukemia cell lines (Cont'd)

Gene			CpG_methylation			Positon	mRNA expression				
Name	Symbol	Location	KG1	U937	THP1		KG1	U937	THP1	N	Probe Set ID
chromosome 16 open reading frame 45	C16orf45	16p13.11	1.00	1.17	−0.23	proximal	NC/A	NC/A	NC/A	A	239971_at
chromosome 1 open reading frame 126	C1orf126	1p36.21	1.84	1.44	−0.40	proximal	ND	ND	ND	ND	NA
chromosome 20 open reading frame 39	C20orf39	20p11.21	1.08	1.70	0.08	down	NC/P	NC/A	NC/A	A	231619_at
calponin 1, basic, smooth muscle	CNN1	19p13.2-p13.1	1.31	1.27	−0.37	proximal	NC/A	NC/A	NC/A	A	203951_at
hypothetical gene	CR611340	6p22.1	2.15	1.70	0.38	proximal	ND	ND	ND	ND	NA
chemokine (C-X-C motif) ligand 5	CXCL5	4q12-q13	1.22	1.35	0.84	proximal	NC/A	NC/A	NC/A	A	207852_at
cytochrome P450, family 1, subfamily B, polypeptide 1	***CYP1B1***	2p21	1.25	2.83	0.90	down	1.7	−6.6	−2.1	P	202435_s_at
fatty acid desaturase 3	FADS3	11q12-q13.1	1.01	1.15	−0.13	up	NC/A	NC/A	NC/A	A	204257_at
hypothetical protein	FLJ42262	8q12.3	2.67	2.22	0.96	up	NC/A	NC/A	−3.4	A	242193_at
homeo box D10	***HOXD10***	2q31.1	1.91	1.70	0.72	up	−1.5	−1.2	NC/A	A	229490_at
hypothetical protein	KIAA1465	15q24.1	1.34	1.49	0.30	up/down	NC/A	NC/A	NC/A	A	232208_at
Kruppel-like factor 5	***KLF5***	13q22.1	1.82	2.09	0.29	down	−3.7	NC/A	NC/A	A	209212_s_at
ladybird homeobox homolog 1 (Drosophila)	LBX1	10q24	1.03	1.64	0.97	up	NC/A	NC/A	NC/A	A	208380_at
LIM homeobox 9	LHX9	1q31-q32	2.65	1.60	0.59	up/down	NC/A	NC/A	NC/A	A	1565407_at
hypothetical protein	LOC282992	10q24.32	1.52	2.34	0.60	down	NC/A	NC/A	NC/A	A	244209_at
myeloid leukemia factor 1	***MLF1***	3q25.1	1.68	1.79	0.23	proximal	−2.9	−1.3	2	A	204784_s_at
5′-nucleotidase, cytosolic IA	NT5C1A	1p34.3-p33	1.53	1.57	0.80	up	NC/A	NC/A	NC/A	A	224529_s_at
phosphodiesterase 4B, cAMP-specific	***PDE4B***	1p31	1.64	1.01	−0.28	proximal	−4.7	−3.6	−4.8	P	211302_s_at
properdin P factor, complement	***PFC***	Xp11.3-p11.23	1.31	1.79	0.44	down	−5.9	−7.8	−8.2	P	206380_s_at
retina and anterior neural fold homeobox	***RAX***	18q21.32	1.54	1.27	−0.26	down	NC/A	NC/A	NC/A	A	208242_at
RGM domain family, member A	***RGMA***	15q26.1	1.33	1.92	0.46	up	−0.9	−0.8	NC/A	A	223468_s_at
Rap2-binding protein 9	***RPIB9***	7q21.12	1.31	1.68	0.20	up	6.7	NC/A	NC/A	A	215321_at
SHC (Src homology 2 domain containing) family, member 4	SHC4	15q21.1-q21.2	1.45	1.48	−0.11	up	NC/A	NC/P	NC/A	A	230538_at
SET binding protein 1	***SETBP1***	18q21.1	1.47	1.75	−0.07	down	−0.9	−0.7	−2.8	P	205933_at
SRY (sex determining region Y)-box 9	SOX9	17q24.3-q25.1	1.49	1.00	0.71	up	NC/A	NC/A	NC/A	A	202935_s_at
transcription factor 2, hepatic	TCF2	17cen-q21.3	1.99	2.92	−0.35	up	NC/A	NC/A	NC/A	A	205313_at
ELAV (embryonic lethal, abnormal vision, Drosophila)-like 2	ELAVL2	9p21	1.15	0.61	1.55	up/down	NC/A	NC/A	NC/A	A	1560905_at
forkhead box A1	FOXA1	14q12-q13	1.03	0.92	1.47	up	NC/A	NC/A	NC/A	A	204667_at
potassium channel, subfamily T, member 2	KCNT2	1q31.3	2.52	0.63	1.45	up	NC/A	NC/A	NC/A	A	244455_at
multiple PDZ domain protein	MPDZ	9p24-p22	2.14	0.93	1.79	down	NC/A	NC/A	NC/A	A	213306_at
paired box gene 9	***PAX9***	14q12-q13	1.41	0.59	1.48	up	NC/M	NC/A	NC/A	A	207059_at
serum deprivation response	SDPR	2q32-q33	1.19	0.52	1.61	down	1.9	−5.4	−1.9	P	222717_at
complement component (3d/Epstein Barr virus) receptor 2	*CR2*	1q32	1.45	0.55	−0.17	down	NC/A	NC/A	NC/A	A	244097_at
hypothetical protein	FLJ40542	22q11.21	1.36	0.43	0.60	down	−1.5	−1	−2.8	P	1556072_at
glial cell derived neurotrophic factor	GDNF	5p13.1-p12	1.35	0.37	0.04	up/down	NC/A	NC/A	NC/A	A	221359_at
Kruppel-like factor 11	***KLF11***	2p25	1.79	0.86	−0.18	up	−3.8	−3.1	−1.4	P	218486_at
LIM homeobox 4	LHX4	1q25.2	2.26	0.89	0.51	down	NC/A	NC/P	NC/A	A	1553157_at
zinc finger protein 215	ZNF215	11p15.4	1.82	0.78	0.43	proximal	NC/A	NC/A	NC/A	A	1555510_at
A kinase (PRKA) anchor protein (gravin) 12	*AKAP12*	6q24-q25	0.90	1.50	1.66	up	−3.3	−6.1	−5.4	P	210517_s_at
hypothetical gene	LOC389372	6p22.1	0.46	2.99	2.21	proximal	NC/A	NC/A	NC/A	A	1568826_at
hypothetical protein	FLJ10159	6q21	0.78	2.25	1.46	up	NC/A	NC/A	NC/A	A	1563906_at
formin binding protein 1	***FNBP1***	9q34	0.47	1.18	1.09	up	NC/P	−2.7	−3	P	230389_at
homeo box A9	HOXA9	7p15-p14	−0.02	2.14	1.09	up	5.8	6.3	4.6	A	209905_at
zinc finger protein 312-like	LOC389549	7q31.32	ND	1.78	1.09	up	ND	ND	ND	ND	NA
v-maf musculoaponeurotic fibrosarcoma oncogene B	***MAFB***	20q11.2-q13.1	0.40	1.38	1.12	up	−10.9	−10.6	−6.1	P	218559_s_at

(*Continued on the following page*)

Table 1. Hypermethylated gene fragments in myeloid leukemia cell lines (Cont'd)

Gene			CpG−methylation			Positon	mRNA expression				
Name	Symbol	Location	KG1	U937	THP1		KG1	U937	THP1	N	Probe Set ID
hypothetical protein	MGC33530	7p11.2	0.99	1.06	1.84	proximal	NC/A	NC/A	NC/A	A	1554530_at
netrin 4	NTN4	12q22-q23	0.91	2.82	2.43	up	NC/A	NC/A	NC/A	A	234202_at
orthopedia homolog (Drosophila)	OTP	5q13.3	0.89	1.45	1.40	down	NC/A	NC/A	NC/A	A	237906_at
protocadherin 19	PCDH19	Xq13.3	0.45	1.45	1.51	down	NC/A	NC/A	NC/A	A	227282_at
protein tyrosine phosphatase, receptor type, K	PTPRK	6q22.2-23.1	0.80	1.52	1.75	down	NC/A	NC/A	NC/A	A	233770_at
toll-IL 1 receptor (TIR) domain containing adaptor protein	TIRAP	11q24.2	0.83	1.12	1.46	up	NC/P	NC/P	NC/P	P	1552360_a_at
zinc finger protein 37 homolog (mouse)	ZFP37	9q32	0.91	1.71	1.24	proximal	NC/A	NC/A	NC/A	A	207068_at
zinc finger protein 229	ZNF229	19q13.31	0.23	2.64	1.37	proximal	NC/P	NC/A	NC/A	A	1562789_at
zinc finger protein 312	ZNF312	3p14.2	0.22	2.02	1.05	up	NC/A	NC/A	NC/A	A	221086_s_at
zinc finger protein 629	ZNF629	16p11.2	0.98	1.27	1.50	down	−1.7	NC/P	1.2	P	213196_at
deleted in colorectal carcinoma	*DCC*	18q21.3	0.71	1.30	0.34	proximal	NC/A	NC/A	NC/A	A	206939_at
deleted in lung and esophageal cancer 1	*DLEC1*	3p22-p21.3	0.91	1.32	−0.12	proximal	NC/A	NC/A	NC/A	A	207896_s_at
distal-less homeo box 3	DLX3	17q21	0.21	1.22	−0.03	down	NC/A	NC/A	NC/A	A	231778_at
dual oxidase 2	DUOX2	15q15.3	0.82	1.97	−0.25	down	NC/A	NC/A	NC/A	A	219727_at
endothelial PAS domain protein 1	EPAS1	2p21-p16	0.21	2.57	−0.46	down	−3.9	1.8	1.6	P	200878_at
EPH receptor A10	EPHA10	1p34.3	0.65	1.15	0.54	up	NC/A	NC/A	NC/A	A	243717_at
ES cell expressed Ras	ERAS	Xp11.23	0.99	1.00	0.62	up	ND	ND	ND	ND	NA
FERM, RhoGEF and pleckstrin domain protein 1	***FARP1***	13q32.2	0.44	2.57	0.44	proximal/up	NC/A	NC/A	NC/A	A	227996_at
hypothetical protein	FLJ42461	17p13.2	0.45	2.18	0.40	proximal	NC/A	NC/A	NC/A	A	229730_at
gamma-glutamyl hydrolase	GGH	8q12.3	0.28	3.10	−0.10	proximal	4.4	−2.9	2.6	P	203560_at
glycoprotein V (platelet)	GP5	3q29	0.08	1.36	−0.18	down	NC/P	NC/P	NC/A	A	207926_at
hyperpolarization activated cyclic nucleotide-gated K+4	HCN4	15q24-q25	0.68	1.28	0.53	up	NC/A	NC/A	NC/A	A	206946_at
histone 1, H4l	HIST1H4L	6p22-p21.3	−0.63	3.21	ND	proximal	NC/A	NC/A	NC/A	A	214562_at
v-jun sarcoma virus 17 oncogene homolog (avian)	***JUN***	1p32-p31	0.11	2.07	-0.08	up	−4.5	−5.1	−4.3	P	201466_s_at
potassium channel beta 3 chain	KCNAB1	3q26.1	0.66	1.48	0.98	down/down	NC/A	5.4	NC/A	A	210471_s_at
protocadherin 8	PCDH8	13q14.3-q21.1	0.79	1.66	−0.09	up	NC/P	NC/P	NC/P	A	206935_at
RAB38, member RAS oncogene family	RAB38	11q14	0.53	1.37	0.55	down	NC/P	NC/P	NC/P	A	234666_at
RAB3C, member RAS oncogene family	***RAB3C***	5q13	0.00	1.57	0.19	proximal	NC/A	NC/A	1.7	A	242328_at
ribonuclease P/MRP 30kDa subunit	***RPP30***	10q23.31	−0.22	1.68	−0.14	up	0.4	−1	NC/P	P	203436_at
secretogranin III	SCG3	15q21	0.96	2.30	0.12	down	NC/A	NC/A	NC/A	A	219196_at
sonic hedgehog homolog (Drosophila)	SHH	7q36	0.99	1.29	0.50	down	NC/A	NC/A	NC/M	A	207586_at
synaptojanin 2	SYNJ2	6q25.3	0.01	1.15	0.07	up	2.3	2.2	1.3	P	212828_at
zinc finger protein 36, C3H type-like 1	***ZFP36L1***	14q22-q24	−0.02	2.07	-0.03	down	−1.2	−5.5	−3.8	P	211965_at
zinc finger protein 222	ZNF222	19q13.2	0.06	1.56	0.00	proximal	2.2	NC/A	3.4	A	206175_x_at
zinc finger protein 516	***ZNF516***	18q23	−0.51	1.86	−0.12	down	NC/P	−1.6	NC/P	P	203604_at
zinc finger protein 582	ZNF582	19q13.43	0.10	1.27	0.47	proximal	NC/P	NC/P	NC/P	A	1553221_at
zinc finger protein 610	ZNF610	19q13.41	0.96	1.72	0.47	up/down	1.9	NC/A	NC/P	A	235953_at
hypothetical gene	AK055761	12q24.2	0.43	0.75	2.18	down	ND	ND	ND	ND	NA
doublesex and mab-3 related transcription factor 2	***DMRT2***	9p24.3	0.03	0.12	2.08	up	NC/A	NC/A	2.9	A	223704_s_at
NK6 transcription factor related, locus 1 (Drosophila)	NKX6-1	4q21.2-q22	−0.09	0.86	2.02	down	NC/A	NC/A	NC/A	A	221366_at

NOTE: Hybridization results of CpG island microarrays are presented as mean $\log_2$ ratios between normal and tumor cell lines of three independent microarray experiments ($\log_2$ ratios above 1 are boxed in black). Results of expression array analysis are presented as mean $\log_2$ ratios between normal and tumor cell lines if a significant change was detected (negative $\log_2$ ratios indicate lower expression in tumor cell lines and are boxed in black). Genes that were independently analyzed by MCIp and real-time PCR (Fig. 4; Supplementary Fig. S2) are indicated in bold lettering. Confirmed targets of aberrant hypermethylation are indicated in bold italics. Genes that have been shown to be hypermethylated in other types of tumors are in italics. Abbreviations: P, present; A, absent; NC, no change; ND, not detected; NA, not available.

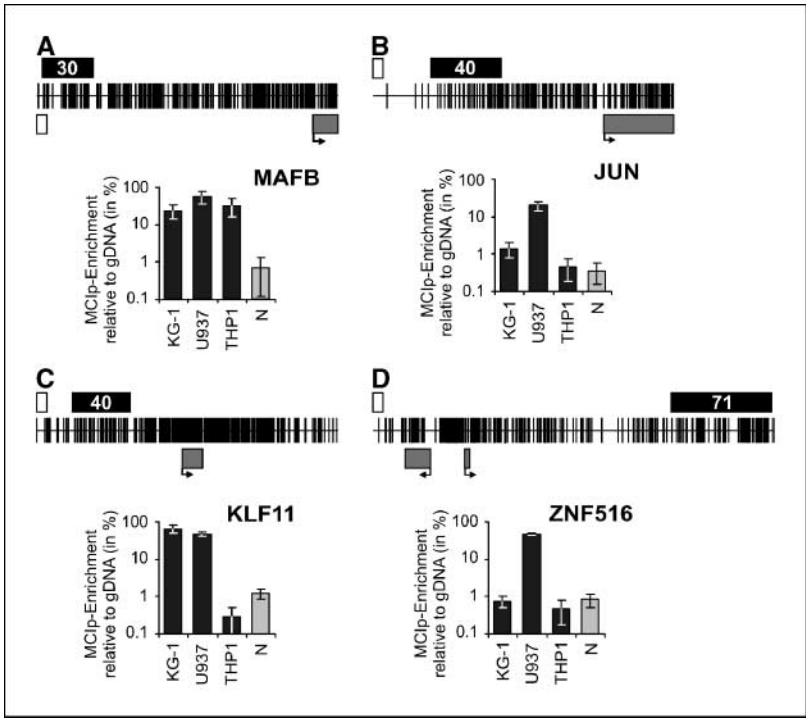

Figure 3. Validation of CpG island microarray results by single-gene MCIp. *A* to *D,* schematic representation of the MCIp enrichment detected by single-gene real-time LightCycler PCR for *MAFB, JUN, KLF11*, and *ZNF516 Mse*I gene fragments in the three human myeloid leukemia cell lines (KG-1, U937, and THP-1) as well as in the normal human blood monocytes. Results are relative to the amount of PCR product generated from the genomic input DNA (100%) of each cell type. *Columns,* mean (n = 4) using at least two different preparations of MBD-Fc; *bars,* SD. Corresponding CpG island on top of each figure as described in Fig. 1.

validate MCIp-detected methylation differences using an independent approach, the methylation status of six CpG island fragments [*JUN, RAB3C, MAFB, KLF11, ZNF516*, and *SSIAH2* (*LOC283514*)] was additionally determined using bisulfite sequencing. As shown in Supplementary Fig. S6, the degree of methylation as determined by bisulfite sequencing correlated well with the results obtained by MCIp. In several cases, the *Mse*I fragment represented on the CpG island microarray did not include the proximal promoter. Because transcription factors may have a particular role in leukemogenesis, DNA fragments that included transcriptional start sites of the four transcription factor genes (*MAFB, KLF11, JUN*, and *ZNF516*) were additionally analyzed by MCIp. Whereas *JUN* promoter fragments were not significantly methylated in any of the samples (data not shown), *MAFB, KLF11*, and *ZNF516* promoter fragments also showed significant methylation (Supplementary Fig. S5).

Global comparison of CpG island methylation and mRNA expression. Complementary mRNA expression data of the above cell lines were generated by microarray analysis using the human HG-133 Plus 2 Array (Affymetrix). A side by side comparison of CpG methylation and mRNA expression data is presented in Table 1. Interestingly, more than half of the genes (69 of 125) were undetectable ('absent') by microarray analysis in all samples. In cases where significant signals were detected, the mRNA expression of genes that were found to be hypermethylated in tumor cell lines was often low relative to normal monocytes (e.g., *PLA2G7, FNBP1, MAFB*, or *ZNF516*). In some cases, the degree of methylation showed no correlation with gene expression detected by microarray analysis (e.g., *HOXA10, EPAS1*, or *SYNJ2*), suggesting that CpG methylation in those cases may target regions not relevant for transcription.

Quantitative real-time PCR analysis of transcripts for five of the previously tested genes with hypermethylated CpG islands (*JUN, MAFB, KLF11, SSIAH2*, and *ZNF516*) confirmed their down-regulation in leukemia cell lines relative to human blood monocytes (data not shown) and showed a significant derepression by treatment with decitabine (5-aza-2′-deoxycytidine) in U937 cells (Figure 4). The effect of demethylation was most striking for *MAFB* and *SSIAH2* (data not shown) that were induced up to 100-fold in treated cells.

Aberrant hypermethylation in AML. Tumor cell lines represent *in vitro* models of primary tumors, which may have acquired additional alterations on both genetic and epigenetic levels. To test whether genes that were found to be hypermethylated in the leukemia cell lines are also affected in primary tumors, we analyzed DNA of blast cells obtained from 12 AML patients for hypermethylation at promoter fragments of *CDKN2B, MAFB* (promoter

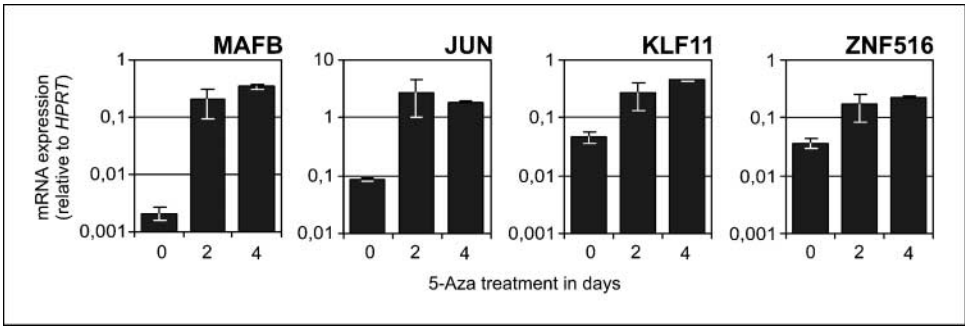

Figure 4. Derepression of hypermethylated target genes by decitabine. *MAFB, JUN, KLF11*, and *ZNF516* mRNA expression levels were quantified in U937 cells after decitabine treatment (5 μmol/L) using real-time PCR. The relative units were calculated from a standard curve with four different concentrations of log dilutions to the PCR cycle number at which the measured fluorescence intensity reaches a fixed value. Expression levels are relative to *HPRT* expression. *Columns,* mean of two experiments; *bars,* SE.

and upstream fragment), *KLF11, ZNF516, KLF5, PFC*, and *SSIAH2*. With the exception of *ZNF516* (data not shown), a significant number of patients (ranging from two to nine) were markedly hypermethylated at each of the above loci (Fig. 5; Supplementary Fig. S6). These results suggest that the CpG island fragments identified in tumor cell lines can also be subject to hypermethylation in primary tumor cells and may represent novel biological markers for leukemia. Notably, the youngest patient (P20) was hypermethylated at all loci tested, whereas the oldest patient (P07) was only significantly methylated at the *PFC* CpG island, indicating that methylation of the above-tested CpG fragments does not correlate with aging.

Discussion

We describe a novel reagent and applications that allow the rapid and sensitive screening of DNA methylation. The central technique, called MCIp, consists of the binding of methylated DNA fragments to the bivalent, antibody-like fusion protein MBD-Fc (a methyl-binding domain fused to a Fc tail) in an immunoprecipitation-like approach. Enriched methylated DNA fragments can be efficiently detected on single-gene level or throughout the genome. The power of these novel techniques was shown by the identification of large number of genes that are affected by aberrant hypermethylation in myeloid leukemias.

Comparison with existing protocols. At present, many techniques are known, which are used for the detection of the CpG methylation of single known candidate gene loci (14). Commonly used assays rely on two basic approaches to distinguish methylated and unmethylated DNA: digestion with methylation-sensitive restriction enzymes or bisulfite treatment of DNA. Methods allowing the analysis of the CpG methylation throughout the genome are less well established. In most techniques [e.g., restriction landmark genomic scanning (RLGS; ref. 32) or methylated CpG island amplification (33, 34)], methylation-sensitive restriction enzymes are used as a component of the method. Here, a major disadvantage is that the analyses only inform on the methylation status of the cytosine residues, which have been recognized by the methylation-sensitive restriction enzymes used. In addition, the selection of the restriction enzymes automatically limits the number of detectable sequences—a global analysis of CpG methylation is therefore not achieved.

The use of naturally occurring MBD proteins to separate methylated and unmethylated DNA fragments are known for more than a decade. Already in 1994, the laboratory of A. Bird developed a method for enriching methylated DNA fragments by means of affinity chromatography using recombinant MeCP2 (18). The technique has been used, improved, and combined with further techniques by other laboratories (35, 36). A disadvantage of MeCP2 affinity chromatography is the large amount of genomic DNA required (50-100 μg) and the relatively time-consuming procedure. In addition, a recent report by Klose et al. (37) showed that MeCP2 requires an A/T run adjacent to the methylated CpG dinucleotide for efficient DNA binding, suggesting that MeCP2 affinity chromatography will be biased toward certain CpG motifs. No binding requirements or preferences of MBD2 were detected in this and in previous studies.

We believe that the high methyl-CpG affinity of MBD2 (38), combined with the bivalent, antibody-like structure of the recombinant MBD-Fc polypeptide, largely increases its binding capacity, enabling the efficient retention of DNA fragments in

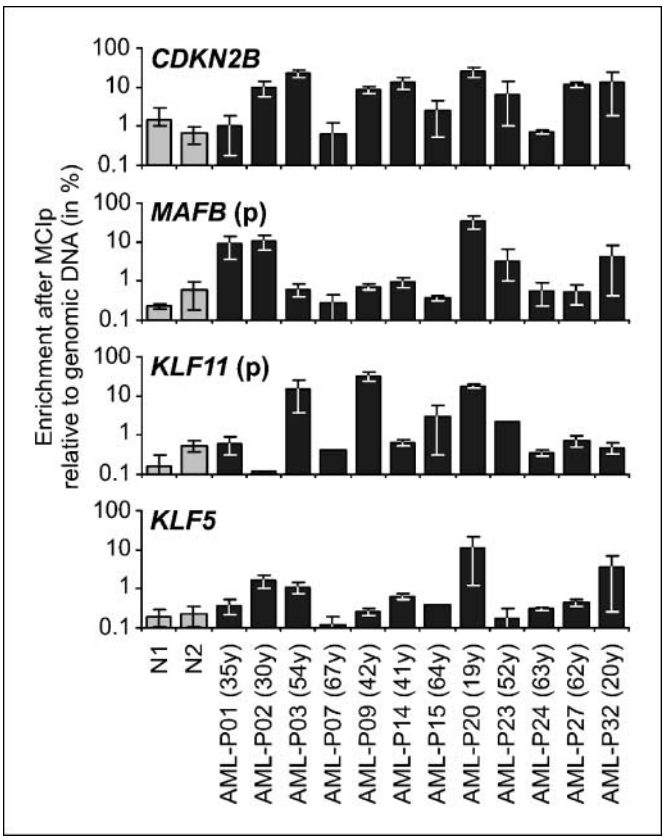

Figure 5. Methylation profiles of AML patients. Schematic representation of the MCIp enrichment detected by single-gene real-time LightCycler PCR for *CDKN2B, MAFB* (promoter), *KLF11* (promoter), and *KLF5* in *MseI*/*Csp*6I-restricted DNA in the 12 AML samples (AML-P no.; age in years) as well as two normal human blood monocytes samples (*N1* and *N2*). Results are relative to the amount of PCR product generated from the genomic input DNA (100%) of each cell type. *Columns,* mean of at least three LightCycler amplifications using at least two different preparations of MBD-Fc; *bars,* SE.

dependence on their methylation degree. The properties of the recombinant MBD-Fc polypeptide allow its application in small-scale assays requiring only little amounts of DNA. This may actually permit the profiling of DNA methylation of candidate genes from very limited cell numbers, including biopsy samples or cells collected by laser-mediated microdissection. The MCIp approach presented in this report requires only little amounts (>300 ng) of genomic DNA for a complete genome-wide methylation profile when combined with an unspecific LMPCR amplification step (the latter step may be omitted if sufficient starting material is available; data not shown). An unmethylated DNA fragment may be 200- to 500-fold depleted, and up to 80% of a highly methylated fragment may be recovered in the high salt MCIp fraction, showing the high affinity of our recombinant polypeptide. In addition to enzymatic restriction, the DNA may also be fragmented by sonication, resulting in a similar enrichment of methylated fragments in the high salt fraction (data not shown).

A recent paper by Weber et al. (31) describes a related approach using a 5-methylcytosine (5mC) antibody and a denaturing step before the immunoprecipitation of DNA fragments. Their analysis revealed only a small set of promoters being methylated differentially in a normal and a transformed cell line, suggesting that aberrant methylation of CpG island promoters in malignancy might be less frequent than previously hypothesized. In contrast to

their observations, we detected a much higher percentage of differentially methylated genes, much in line with previous estimates, using the same CpG island microarray platform. This may reflect an inherent property of the cell lines used and, however, may also point to a lesser sensitivity of the 5mC antibody approach compared with our fractionated MCIp approach.

Hypermethylated genes in leukemia cell lines. Our profiling of three leukemia cell lines identified a large number of gene fragments that are likely to be methylated in neoplastic cells. To our knowledge, this study provides one of the largest published collections of potentially hypermethylated genes in cancer cells.

Interestingly, most genes that were identified as hypermethylated in leukemia cell lines showed extremely low or undetectable mRNA expression levels in our microarray experiments. A comparison with published expression profiles for human bone marrow, CD33-positive bone marrow cells, as well as mature myeloid cells (http://symatlas.gnf.org/SymAtlas/; data not shown) indicates that a large proportion of these genes may not be significantly transcribed in myeloid cell types. A hypothetical (thus far unknown) targeting mechanism may therefore induce CpG methylation of genes independent of their transcriptional status during cellular differentiation. Although such genes may not have a significant suppressor role in tumor development and/or progression, they may still serve as valuable biomarkers, provided that the targeting mechanism is specific for the disease. Methylation profiling of larger sample groups, using the described (or similar) approaches, may help to clarify whether aberrant methylation of CpG islands in malignancies is random or specific.

Acute leukemia is characterized by a block of differentiation of early progenitors, which leads to the accumulation of immature cells in bone marrow and blood. The frequent mutation or down-regulation of a relatively small number of transcription factors in AML patients suggests that the inactivation of transcriptional regulators may be critically involved into the malignant transformation process. Our methylation profiling of leukemia cell lines preferentially identified genes that are involved in transcriptional regulation. Half of the listed genes with an assigned molecular function (46 of 89) are involved in DNA binding and transcriptional regulation, which represents a significant overrepresentation. Aberrant hypermethylation of these transcription factor genes may lead to their epigenetic down-regulation and likely contributes to the observed differentiation arrest in leukemia cells. This observation is in line with a previous study from Rush et al. (39) that investigated the methylation status of a large set of CpG islands in AML patients using RLGS and also found that a large proportion of the known methylated promoters (4 of 11) corresponded to genes involved in transcriptional regulation.

The list of hypermethylation targets contains several transcription factor genes, including *MAFB, JUN*, and *KLF11*, which are highly expressed in normal myeloid cells. A good tumor suppressor candidate, for example, is represented by the bZip transcription factor MAFB, which is expressed specifically in the myeloid lineage of the hemopoietic system. Its expression is up-regulated successively during myeloid differentiation from multipotent progenitors to macrophages, suggesting an essential role of *MAFB* in early myeloid and monocytic differentiation (40).

In summary, the data provided validate the experimental and possibly diagnostic potential of the MCIp technique. The recombinant MBD-Fc protein and its application in DNA methylation analysis may represent an important step toward genome-wide CpG methylation profiling not only in cancer diagnostics. The identification of those hypermethylation targets that are relevant in cancer development and/or progression will be a major future challenge.

Acknowledgments

Received 1/31/2006; revised 3/21/2006; accepted 4/19/2006.

Grant support: Deutsche Forschungsgemeinschaft grant Re1310/2 and Wilhelm Sander-Stiftung (M. Rehli).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

We thank Drs. Krishna Mondal and Stefan Krause (Department of Hematology, University Hospital, Regensburg, Germany) for reading the article, Dr. Ilja Hagen (KFB) for his support in microarray analysis, Neil Winegarden (University Health Network Microarray Centre) for the sequencing of CpG island clones, Dr. Andreas Mackensen (Department of Hematology, University Hospital) for providing patient material, Dr. Sven Heinz (University of California at San Diego, San Diego, CA), Andreas Waha (Department of Neuropathology, University of Bonn, Bonn Germany) for advices, and laboratory members for helpful discussions.

References

1. Costello JF, Plass C. Methylation matters. J Med Genet 2001;38:285–303.
2. Esteller M, Fraga MF, Paz MF, et al. Cancer epigenetics and methylation. Science 2002;297:1807–8.
3. Ng HH, Bird A. DNA methylation and chromatin modification. Curr Opin Genet Dev 1999;9:158–63.
4. Razin A. CpG methylation, chromatin structure, and gene silencing—a three-way connection. EMBO J 1998; 17:4905–8.
5. Robertson KD. DNA methylation and human disease. Nat Rev Genet 2005;6:597–610.
6. Herman JG, Baylin SB. Gene silencing in cancer in association with promoter hypermethylation. N Engl J Med 2003;349:2042–54.
7. Momparler RL. Cancer epigenetics. Oncogene 2003;22: 6479–83.
8. Plass C. Cancer epigenomics. Hum Mol Genet 2002;11: 2479–88.
9. Costello JF, Fruhwald MC, Smiraglia DJ, et al. Aberrant CpG-island methylation has non-random and tumour-type-specific patterns. Nat Genet 2000;24:132–8.
10. Esteller M, Corn PG, Baylin SB, Herman JG. A gene hypermethylation profile of human cancer. Cancer Res 2001;61:3225–9.
11. Esteller M. Dormant hypermethylated tumour suppressor genes: questions and answers. J Pathol 2005;205: 172–80.
12. Kalebic T. Epigenetic changes: potential therapeutic targets. Ann N Y Acad Sci 2003;983:278–85.
13. Esteller M. Relevance of DNA methylation in the management of cancer. Lancet Oncol 2003;4:351–8.
14. Dahl C, Guldberg P. DNA methylation analysis techniques. Biogerontology 2003;4:233–50.
15. Krause SW, Rehli M, Kreutz M, Schwarzfischer L, Paulauskis JD, Andreesen R. Differential screening identifies genetic markers of monocyte to macrophage maturation. J Leukoc Biol 1996;60:540–5.
16. Frommer M, McDonald LE, Millar DS, et al. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. Proc Natl Acad Sci U S A 1992;89:1827–31.
17. Haehnel V, Schwarzfischer L, Fenton MJ, Rehli M. Transcriptional regulation of the human toll-like receptor 2 gene in monocytes and macrophages. J Immunol 2002;168:5629–37.
18. Cross SH, Charlton JA, Nan X, Bird AP. Purification of CpG islands using a methylated DNA binding column. Nat Genet 1994;6:236–44.
19. Zeschnigk M, Schmitz B, Dittrich B, Buiting K, Horsthemke B, Doerfler W. Imprinted segments in the human genome: different DNA methylation patterns in the Prader-Willi/Angelman syndrome region as determined by the genomic sequencing method. Hum Mol Genet 1997;6:387–95.
20. Chim CS, Wong AS, Kwong YL. Epigenetic inactivation of INK4/CDK/RB cell cycle pathway in acute leukemias. Ann Hematol 2003;82:738–42.
21. Dodge JE, List AF, Futscher BW. Selective variegated methylation of the p15 CpG island in acute myeloid leukemia. Int J Cancer 1998;78:561–7.
22. Dodge JE, Munson C, List AF. KG-1 and KG-1a model the p15 CpG island methylation observed in acute myeloid leukemia patients. Leuk Res 2001;25:917–25.
23. Issa JP, Zehnbauer BA, Civin CI, et al. The estrogen receptor CpG island is methylated in most hematopoietic neoplasms. Cancer Res 1996;56:973–7.
24. Paz MF, Wei S, Cigudosa JC, et al. Genetic unmasking of epigenetically silenced tumor suppressor genes in colon cancer cells deficient in DNA methyltransferases. Hum Mol Genet 2003;12:2209–19.

25. Douglas DB, Akiyama Y, Carraway H, et al. Hypermethylation of a small CpGuanine-rich region correlates with loss of activator protein-2α expression during progression of breast cancer. Cancer Res 2004;64:1611–20.
26. Schwab J, Illges H. Regulation of CD21 expression by DNA methylation and histone deacetylation. Int Immunol 2001;13:705–10.
27. Sato K, Tamura G, Tsuchiya T, et al. Frequent loss of expression without sequence mutations of the DCC gene in primary gastric cancer. Br J Cancer 2001;85:199–203.
28. Jones PA, Wolkowicz MJ, Rideout WM III, et al. *De novo* methylation of the MyoD1 CpG island during the establishment of immortal cell lines. Proc Natl Acad Sci U S A 1990;87:6117–21.
29. Yuan BZ, Durkin ME, Popescu NC. Promoter hypermethylation of DLC-1, a candidate tumor suppressor gene, in several common human cancers. Cancer Genet Cytogenet 2003;140:113–7.
30. Choi MC, Jong HS, Kim TY, et al. AKAP12/Gravin is inactivated by epigenetic mechanism in human gastric carcinoma and shows growth suppressor activity. Oncogene 2004;23:7095–103.
31. Weber M, Davies JJ, Wittig D, et al. Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. Nat Genet 2005;37:853–62.
32. Costello JF, Smiraglia DJ, Plass C. Restriction landmark genome scanning. Methods 2002;27:144–9.
33. Smith RJ, Dean W, Konfortova G, Kelsey G. Identification of novel imprinted genes in a genome-wide screen for maternal methylation. Genome Res 2003;13:558–69.
34. Yan PS, Chen CM, Shi H, et al. Dissecting complex epigenetic alterations in breast cancer using CpG island microarrays. Cancer Res 2001;61:8375–80.
35. Brock GJ, Huang TH, Chen CM, Johnson KJ. A novel technique for the identification of CpG islands exhibiting altered methylation patterns (ICEAMP). Nucleic Acids Res 2001;29:E123.
36. Shiraishi M, Chuu YH, Sekiya T. Isolation of DNA fragments associated with methylated CpG islands in human adenocarcinomas of the lung using a methylated DNA binding column and denaturing gradient gel electrophoresis. Proc Natl Acad Sci U S A 1999;96:2913–8.
37. Klose RJ, Sarraf SA, Schmiedeberg L, McDermott SM, Stancheva I, Bird AP. DNA binding selectivity of MeCP2 due to a requirement for A/T sequences adjacent to methyl-CpG. Mol Cell 2005;19:667–78.
38. Fraga MF, Ballestar E, Montoya G, Taysavang P, Wade PA, Esteller M. The affinity of different MBD proteins for a specific methylated locus depends on their intrinsic bindingproperties. Nucleic Acids Res 2003;31:1765–74.
39. Rush LJ, Dai Z, Smiraglia DJ, et al. Novel methylation targets in *de novo* acute myeloid leukemia with prevalence of chromosome 11 loci. Blood 2001;97:3226–33.
40. Kelly LM, Englmeier U, Lafon I, Sieweke MH, Graf T. MafB is an inducer of monocytic differentiation. EMBO J 2000;19:1987–97.

Genome-Wide Profiling of CpG Methylation Identifies Novel Targets of Aberrant Hypermethylation in Myeloid Leukemia

Claudia Gebhard, Lucia Schwarzfischer, Thu-Hang Pham, et al.

Cancer Res 2006;66:6118-6128.