Large-Scale CpG Methylation Analysis Identifies Novel Candidate Genes and Reveals Methylation Hotspots in Acute Lymphoblastic Leukemia

Kristen H. Taylor,[1] Keila E. Pena-Hernandez,[1,2] J. Wade Davis,[2,3] Gerald L. Arthur,[2] Deiter J. Duff,[1] Huidong Shi,[1] Farah B. Rahmatpanah,[1] Ozy Sjahputera,[1] and Charles W. Caldwell[1]

[1]Department of Pathology and Anatomical Sciences, Ellis Fischel Cancer Center, [2]Department of Health Management and Informatics, and [3]Department of Statistics, University of Missouri-Columbia School of Medicine, Columbia, Missouri

Abstract

This study examined DNA methylation associated with acute lymphoblastic leukemia (ALL) and showed that selected molecular targets can be pharmacologically modulated to reverse gene silencing. A CpG island (CGI) microarray containing more than 3,400 unique clones that span all human chromosomes was used for large-scale discovery experiments and led to 262 unique CGI loci being statistically identified as methylated in ALL lymphoblasts. The methylation status of 10 clones encompassing 11 genes (*DCC, DLC-1, DDX51, KCNK2, LRP1B, NKX6-1, NOPE, PCDHGA12, RPIB9, ABCB1*, and *SLC2A14*) identified as differentially methylated between ALL patients and controls was independently verified. Consequently, the methylation status of *DDX51* was found to differentiate patients with B- and T-ALL subtypes (P = 0.011, Fisher's exact test). Next, the relationship between methylation and expression of these genes was examined in ALL cell lines (NALM-6 and Jurkat) before and after treatments with 5-aza-2-deoxycytidine and trichostatin A. More than a 10-fold increase in mRNA expression was observed for two previously identified tumor suppressor genes (*DLC-1* and *DCC*) and also for *RPIB9* and *PCDHGA12*. Although the mechanisms that lead to the CGI methylation of these genes are unknown, bisulfite sequencing of the promoter of *RPIB9* suggests that expression is inhibited by methylation within SP1 and AP2 transcription factor binding motifs. Finally, specific chromosomal methylation hotspots were found to be associated with ALL. This study sets the stage for acquiring a better biological understanding of ALL and for the identification of epigenetic biomarkers useful for differential diagnosis, therapeutic monitoring, and the detection of leukemic relapse. [Cancer Res 2007;67(6):2617–25]

Introduction

Acute lymphoblastic leukemia (ALL) arises when B- or T-cell progenitors fail to differentiate into mature cells and is characterized by the rapid proliferation of immature lymphoblasts (1, 2). Several factors may play a role in leukemogenesis, including chromosomal translocations and genetic or epigenetic modifications that modify the function of a gene or set of genes. Many nonrandom chromosomal translocations are known to modify proliferation, differentiation, apoptosis, and gene transcription in ALL (3).

The aberrant methylation of gene promoter-related CpG islands (CGI) is an epigenetic modification of DNA that can inappropriately silence or down-regulate gene expression and could have deleterious effects if the targeted genes function as tumor suppressors. However, the mechanisms by which DNA methylation occurs in ALL, when these events occur, and which genes are involved are not wholly understood. It has been established, however, that hematopoietic progenitor cells undergo several molecular changes on interaction with stromal cells, including the up-regulation of *DNA methyltransferase 1* (4), which is involved in both the establishment and the maintenance of DNA methylation (5). Additionally, previous studies, which investigated aberrant methylation in ALL, have associated hypermethylated promoters with prognosis (6), cytogenetic alterations (7), subtype (8), and relapse (9). Although the elucidation of the aberrant methylation profiles involved in ALL has been limited by the small number of CGIs analyzed to date, these studies suggest that it is common for multiple promoter CGIs to be aberrantly methylated.

Toward a more global view of the extent to which gene promoter and/or first exon DNA methylation is present in ALL, methylation microarray profiles were developed using a CGI microarray consisting of ~9K clones (10) representing more than 3,400 unique sequences spanning all human chromosomes. The patterns of aberrant methylation and mRNA expression levels were independently confirmed in 11 putative candidate genes. Additionally, the chromosomal locations of 262 differentially methylated loci were examined to identify the distribution of genomic regions that were methylated in ALL patients. Although this microarray does not include all of the annotated CGI fragments within the human genome, it has proven to be a valuable discovery tool for elucidating epigenetic mechanisms underlying other cancer types (11, 12) and has been used in this study to generate specific ALL epigenetic profiles and to provide new insights into ALL cell biology.

Materials and Methods

Tissue specimens. Bone marrow samples of patients diagnosed with ALL at the Ellis Fischel Cancer Center (Columbia, MO) were obtained in compliance with the local Institutional Review Board. Leukemic blasts, due to their high proportion, were concentrated to >90% purity using Ficoll density gradient centrifugation. DNA was isolated using the QIAamp DNA Mini kit (Qiagen, Valencia, CA) according to the manufacturer's specifications from 20 specimens: 6 from patients with precursor T-cell ALL, 10 from patients with precursor B-cell ALL, and 4 from healthy donors used as controls.

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

Requests for reprints: Charles W. Caldwell, Department of Pathology and Anatomical Sciences, Ellis Fischel Cancer Center, University of Missouri School of Medicine, One Hospital Drive, DC055.07, Columbia, MO 65212. Phone: 573-884-2642; Fax: 573-884-4612; E-mail: caldwellc@health.missouri.edu.

doi:10.1158/0008-5472.CAN-06-3993

Cell lines. Precursor B-cell ALL cell lines representing various immunophenotypic stages of precursor B-cell development, NALM-6 ($CD10^+$, $CD19^+$, and $CD20^-$), MN-60 ($CD10^+$, $CD19^+$, and $CD20^+$), and SD-1 ($CD10^-$, $CD19^+$, and $CD20^+$) and the precursor T-cell ALL cell line, Jurkat, were purchased from the Deutsche Sammlung von Mikroorganismen und Zellkulturen (Braunschweig, Germany) and grown in appropriate medium and resupplemented as necessary with fresh medium. Cells were harvested and DNA was extracted using the QIAamp DNA Mini kit.

Preparation of the CGI microarray. The microarray panel containing 8,640 CGI clones was prepared as described (13). Amplified PCR products were spotted in the presence of 20% DMSO, on UltraGap slides (Corning Life Science, Acton, MA). The slides were postprocessed immediately before the hybridization using Pronto Universal Microarray Reagents (Corning Life Science). All CGI clones present on the microarray were sequenced recently by the Microarray Centre of University Health Network (Toronto, Ontario, Canada; ref. 14).

Amplicon development and differential methylation hybridization. Amplicons were generated and differential methylation hybridization (DMH) was done as described previously (10–12). Briefly, 2 μg of genomic DNA were digested with *Mse*I followed by ligation of PCR linkers and digestion with methylation sensitive endonucleases (*Hpa*II and *Bst*UI). PCR was then done to amplify only methylated fragments or fragments containing no internal *Hpa*II or *Bst*UI sites. The amplicons from the ALL and sex-matched normal control sample, which comprised pooled DNA from multiple donors (Promega, Madison, WI), were labeled with Cy5 or Cy3 fluorescence dye, respectively, and cohybridized to a panel of 8,640 short CGI tags arrayed on a glass slide. The slides were scanned with a GenePix 4200A scanner and the signal intensities of the hybridized probes were analyzed using GenePix 5.1 (Molecular Devices Corp., Sunnyvale, CA).

CGI microarray analysis. The 9K chip developed by Huang et al. (10) includes 8,640 *Mse*I fragments. Our method for generating amplicons relies on the presence of *Bst*UI or *Hpa*II recognition sequences; therefore, all of the *Mse*I fragments lacking a recognition sequence were removed from the analysis resulting in a reduction of the number of CGIs analyzed. Because DNA methylation can vary among individuals depending on age or gender and even in different tissue types within the same individual, we reduced the gender-based variability and the variability due to tissue type by comparing each patient sample (16 total) and each normal bone marrow sample (4 total) with peripheral blood DNA from a pool of sex-matched individuals. To determine which clones were differentially methylated between ALL and normal samples, we globally normalized each microarray and then used the nonparametric Kruskal-Wallis test to do an across-array analysis for each probe.

Clone sequences. Sequences of the differentially methylated CGI clones were extracted from a Web site.[4] Next, BLAST searches were done to determine if these clone sequences were associated with the promoter region of known genes and if these regions contained CGIs. Finally, primers were developed for real-time reverse-transcription (RT-PCR) and for PCR of bisulfite-treated DNA using Primer3 (15) and MethPrimer (16), respectively.

Methylation-specific PCR and combined bisulfite restriction analysis. Two micrograms of DNA were treated with sodium bisulfite according to the manufacturer's recommendations (Ez DNA methylation kit, Zymo Research, Orange, CA). Bisulfite-treated DNA was then used as a template for PCR with primers designed using MethPrimer that were specific to the CGI regions of each tested gene. Purified amplicons produced using combined bisulfite restriction analysis (COBRA) primers (Table 1) were restricted with *Bst*U1, *Taq*aI, or *Hpy*CH4IV according to the manufacturer's recommendations (New England Biolabs, Ipswich, MA). The methylation-specific PCR (MSP) primers (as reported previously; ref. 11) were used in PCR to differentiate methylated and unmethylated sequences in low-density *lipoprotein receptor-related protein 1B* (*LRP1B*). Electrophoresis was done using a 3% agarose gel stained with SYBR Green or a 1.5% agarose gel stained with ethidium bromide to visualize COBRA and MSP products, respectively.

[4] http://data.microarrays.ca/cpg/

Quantitative real-time MSP. Quantitative real-time MSP (qMSP) was done with *deleted in liver cancer 1* (*DLC-1*) primers as described previously (12). Briefly, 100 ng of bisulfite-treated DNA and ABgene QPCR Mix (ABgene, Inc., Rochester, NY) were used for PCR amplification as recommended by the manufacturer. The reaction was carried out in 40 to 45 cycles using a SmartCycler (Cepheid, Kingwood, TX).

Cell line treatment. The Jurkat (precursor T cell) and NALM-6 (precursor B cell) ALL cell lines were grown in flasks with RPMI 1640 supplemented with 10% fetal bovine serum, L-glutamine, and gentamicin. Treatment was conducted during the log phase of growth with 5-aza-2-deoxycytidine (5-aza) and/or trichostatin A (TSA), whereas the control cells were not treated. Jurkat cells were seeded at 8×10^6/mL and NALM-6 cells were seeded at 5×10^6/mL. In culture, TSA was added at a 1 μmol/L final concentration and incubated for 6 h, whereas 5-aza was added at a 1 μmol/L concentration and incubated for 54 and 78 h in Jurkat and NALM-6 (based on different dividing times), respectively, with a medium change every 24 h. The cell culture that received both TSA and 5-aza treatment was first incubated with 5-aza as described previously, which was followed by an additional 6-h incubation with TSA. RNA and DNA from the cultured cells were extracted for use in RT-PCR and COBRA, respectively, using the previously mentioned kits.

Quantitative real-time RT-PCR. Total RNA (2 μg) from treated and untreated cell lines was pretreated with DNase I to remove potential DNA contaminants and was then reverse transcribed in the presence of SuperScript II reverse transcriptase (Invitrogen, Carlsbad, CA). The generated cDNA was used for PCR amplification with appropriate reagents in the reaction mix with SYBR Green and fluorescein (ABgene) as recommended by the manufacturer. *Glyceraldehyde-3-phosphate dehydrogenase* (*GAPDH*) and *hypoxanthine phosphoribosyltransferase 1* (*HPRT1*; ref. 17) were used as the housekeeping genes in the Taqman and SYBR Green real-time RT-PCR assays, respectively. The *DLC-1* and *GAPDH* Taqman probe and primer sets for real-time PCR were purchased from Applied Biosystems Assay-on-Demand services. The reaction was carried out using an iCycler real-time PCR instrument (Bio-Rad Laboratories, Hercules, CA). The cycling conditions included an initial 15 min hot start at 95°C followed by 45 cycles at 95°C for 15 s and 60°C for 1 min. Primers were developed for SYBR Green assays using Primer3 (Table 1). The reactions were carried out using an iCycler. The cycling conditions included an initial 15 min hot start at 95°C followed by 50 cycles at 95°C for 15 s, 58°C for 30 s, and 72°C for 30 s. All samples were analyzed in triplicate and fold changes were determined using the $2^{-\Delta\Delta CT}$ method (18).

Bisulfite genomic sequencing analysis. Genomic DNA was treated with sodium bisulfite as described previously. Primer sequences and PCR conditions were as for the previously described COBRA assay. Amplified PCR products for *Rap2-binding protein 9* (*RPIB9*) fragment 1, *RPIB9* fragment 2, and *protocadherin-γ subfamily A member 12* (*PCDHGA12*) were subcloned using the TOPO-TA cloning system (Invitrogen). Plasmid DNA from 6 to 10 insert-positive clones was isolated using the QIAquick Plasmid Mini Prep kit (Qiagen) and sequenced using an ABI 3730 DNA Analyzer (Applied Biosystems, Foster City, CA).

Statistical methods. The geometric mean was used to summarize average fold changes in all RT-PCR analyses. All hypothesis tests were two sided and conducted at the 5% significance level unless otherwise stated. Statistical analysis was done using SAS 9.1 (SAS Institute, Cary NC) and R (R Foundation for Statistical Computing, Vienna, Austria).

Spearman's nonparametric correlation was used to evaluate the agreement between bisulfite sequencing and COBRA for a given gene. Only loci that could be detected with both technologies were considered. The proportion of methylated clones (based on bisulfite sequencing) for these loci was computed and correlated with an ordinal measure of COBRA methylation defined with three levels: none, partial, and complete. 'None' refers to those cases in which no cuts were observed, 'partial' refers to those cases in which partial banding was observed, and 'complete' refers to those cases in which all possible bands were observed (i.e., the banding pattern matched that of the *Sss*I and bisulfite-treated control).

Despite the fact that more powerful parametric methods exist to detect localized clusters on a chromosome and due to the heterogeneous coverage

Table 1. Primers used for COBRA and real-time SYBR Green analyses

	Sense primer (5′–3′)	Antisense primer (5′–3′)	Annealing temperature (°C)	Product size*
COBRA †				
DCC	GGATATTTTAGAAAAGTGAGAG	CAAATCATCAATAAACCACATCCAAA	55	300
DDX51	TTTTTTATTTGTTTTATTTAAGGTGTT	TCTACTAAACTTACCCCTATCCTCC	56	250
KCNK2	TTTAGTAAAGGGGTTTTGTTTTGAG	AACCCTAACTTCTTCCAATCTACAC	56	230
NKX6-1	TTTTGTATATTTGGAGGGATAGGTAT	CCTTTTATTCATCAAAAATTTACCC	54	210
NOPE	TTTTTTGTTTTATTTATTTTAGTTTTAGTT	AAAACCCATCTCCACAAATATCAT	56	210
PCDHGA12	AATGTTTAGATTTAATGTATATTTGATGGT	CTCCAAAAACCTAAAACTAAAACCC	56	180
RP1B9	ATTGGAATTGATATAAAGTTTAGGGTT	ACCCCCTTAAACAAATATAAAAAAC	56	400
SLC2A14	GGTTTTAAGGTTAGTTTTTTAGAGT	AAACAATTAATAAATCCCAAC	54	270
Real time				
ABCB1	TGTATGCTCAGAGTTTGCAGGT	TTCCAAAGATGTGTGCTTTCC	58	60
DCC	CCGAAAGTCCCTTACACACC	CATGGGTCTTAGGAAGAGTGG	58	60
DDX51	CACACTGCTCCTGAAAGTGC	TTCAGTTAGCATTCGGAGGAA	58	50
HPRT1 ‡	TGACACTGGCAAAACAATGCA	GGTCCTTTTCACCAGCAAGCT	58	90
KCNK2	TAACAACTATTGGATTTGGTGACTAC	GCCCTACAAGGATCCAGAAC	58	100
LRP1B	CATGATCACAACGATGGAGGT	CTTGAAAGCACTGGGTCCTC	58	90
NKX6-1	CTTCTGGCCCGGAGTGAT	TCTTCCCGTCTTTGTCCAAC	58	100
NOPE	ACAGGGCTGAAGTGCACAG	CTTGGTTGAGCCCAGGAGA	58	90
PCDHGA12	TGCTGTCAGGTGATTCGGTA	AGAAACGCCAGTCCGTGTT	58	80
RPIB9	GGCCAGTCACAAGAAGGAGA	GAGATCCACAGAGGCCAAGT	58	100
SLC2A14	TCCACGCTCATGACTGTTTC	CAGGCCACAAAGACCAAGAT	58	90

*Product sizes are approximate.
† All COBRA amplifications were digested with *Bst*UI except for *DDX51* (*Taq*aI) and *KCNK2* (*Hpy*CH4IV).
‡ *HPRT1* primer sequences were previously published (17).

of loci on the CGI microarray, a nonparametric randomization test was devised and used based on its simplicity and general applicability. Let n_k denote the number of loci on the CGI microarray, which are located on chromosome k and let $S_k \leq n_k$ denote the number of differentially methylated loci on chromosome k. The locations (bp) of the S_k methylated loci on chromosome k are given by $x_1,\ldots,x_{Sk}$ and the distances between consecutively methylated loci are given by $d_i = x_{i+1} - x_i$ where $i = 1,\ldots,S_{k-1}$. For each chromosome, the median of these distances was computed to serve as the test statistic, t^*. Chromosomes, which have clustered loci, should have smaller consecutive distances and therefore smaller medians. To determine the null distribution for chromosome k, a random sample of S_k loci was drawn from the n_k loci with this set of S_k loci representing the methylated loci under the null hypothesis of loci methylated at random with respect to chromosomal location. For each sample, the median of the consecutive distances was computed as described previously. This process was repeated 1 million times and the distribution of the 1 million medians formed the distribution of the test statistic under the null hypothesis for each chromosome. The empirical P value was taken as the proportion of the 1 million medians less than the observed test statistic, t^*, for each chromosome. Note that our method requires no assumption about the distribution of the n_k loci on the CGI microarray. Furthermore, because the null distribution for each chromosome is generated from 1 million random samples of size S_k selected from among the n_k loci, the spatial density of the loci are implicitly incorporated.

Results

Global methylation profiles in ALL patients. DNA was extracted from bone marrow aspirate obtained from patients at the time of diagnosis and from four healthy donors. The amplicons from primary ALL cells and from the four healthy donors were compared with amplicons generated from a pool of sex-matched normal peripheral blood lymphocyte DNA by cohybridization to a CGI microarray. After global normalization of each microarray, the nonparametric Mann-Whitney test was used in an interarray analysis to identify those CGI clones that were methylated in ALL patient samples but not in the pooled normal control samples. To avoid the inclusion of false-positive loci, only those clones with at least a 1.5-fold difference in the normalized Cy5/Cy3 ratio between patient and pooled normal DNA and present in at least 25% of the ALL samples and in none of the normal bone marrow samples were deemed a methylated locus. Of the 262 differentially methylated loci meeting these criteria, 148 were found within 2,500 bases of a transcriptional start site and 131 were also within 500 bases of an annotated CGI (Supplementary Table S1). This set of genes includes *RPIB9*, which was also reported recently to be methylated in AML patients (19). *RPIB9* overlaps in genomic sequence but is transcribed in the opposite direction to *ATP-binding cassette, subfamily B, member 1* (*ABCB1*), which has been shown previously to be aberrantly methylated in ALL patients (20). The set also contains genes that are aberrantly methylated in other malignancies, including deleted in *colorectal carcinoma* (*DCC*), *DLC-1*, and *LRP1B* (21–23).

Validation of microarray results. Results of the CGI microarray experiments were validated either by COBRA, MSP, or qMSP by verifying the presence of methylation in 10 clones encompassing 11 candidate genes in these patient samples and also in the 4 ALL cell lines. The genes chosen for validation were not chosen based on fold differences but were chosen to cover the spectrum of genes meeting the set criteria and included genes that were methylated in 31.3% to 75% of the patients according to DMH. MSP and qMSP assays were developed previously in our laboratory to study

promoter methylation in *LRPIB* and *DLC-1*, respectively. All other validations were conducted using COBRA. Cell line gel images were analyzed and samples were scored as methylated (Fig. 1*B, gray blocks*), unmethylated (Fig. 1*B, white blocks*), or partially methylated (Fig. 1*B, hatched blocks*). The results for the cell line studies varied, with methylation detected for all of the candidate genes in NALM-6 (3), 80% in Jurkat (1), 80% in MN-60 (2), but only 20% in SD-1 (4; Fig. 1*A* and *B*). The percentage of patients for which candidate gene CGI methylation was detected ranged from 43.8% for *DEAD box peptide 51* (*DDX51*) to 100% for *NK6 transcription factor related, locus 1* (*NKX6-1*; Fig. 1*B*). In all instances except one (*DDX51*), validation revealed a higher number of methylated patients than indicated by the DMH results at the 1.5-fold cutoff.

Examination of the effects of gene promoter demethylation *in vitro* by RT-PCR. To determine whether the promoter methylation detected in the validated gene set was responsible for the down-regulation of these genes, we examined the *in vitro* effects of treatment with a demethylating agent (5-aza) and a histone deacetylase inhibitor (TSA) both individually and in combination in a B-ALL cell line (NALM-6) and a T-ALL cell line (Jurkat) using RT-PCR. There was little to no mRNA detected for 9 of the 11 genes [except *DDX51* and *solute carrier family 2, member 14* (*SLC2A14*)] in the untreated (control) NALM-6 cell line (Fig. 2*A*) and in 7 of the 11 genes (except *DDX51, DLC-1, RPIB9*, and *SLC2A14*) in the untreated Jurkat cell line (Fig. 2*B*). Overall, there was an increase in mRNA expression post-treatment in both cell lines. However, the post-treatment increase in mRNA expression was >10-fold for *ABCB1, DCC, DLC-1, PCDHGA12*, and *RPIB9* in the Jurkat cell line and for *DCC, DLC-1, PCDHGA12*, and *RPIB9* in the NALM-6 cell line. All examined mRNAs were expressed in normal bone marrow with the exception of *NKX6-1* (data not shown). In general, we observed an inverse relationship between promoter methylation and mRNA expression in the interrogated genes.

Bisulfite sequencing of *PCDHGA12* and *RPIB9/ABCB1*. To gain insight into mechanisms responsible for the silencing of two genes and to further investigate the methylation status of these genes, we did bisulfite sequencing of the first exon of *PCDHGA12* and of the predicted promoter and the first exon of *RPIB9* (this region is also located within the first intron of *ABCB1*) in ALL cell lines and patient samples (Fig. 3*A*). These regions were chosen because (*a*) they harbor the clone sequences interrogated on the 9K CGI microarray used in the DMH experiments; (*b*) they include transcription factor binding motifs located within the putative promoter of *RPIB9*; and (*c*) these genes had substantial increases in mRNA expression after treatment with a demethylating agent. For each gene, we collected sequencing data for two cell lines (NALM-6 and Jurkat) and three patient samples (*RPIB9*: B10, T2, and B5; *PCDHGA12*: B3, T3, and B4) determined to be positive for methylation according to COBRA and one cell line (SD-1) and one patient sample (*RPIB9*: T5; *PCDHGA12*: B1) determined to be negative for methylation according to COBRA. Consistent with the

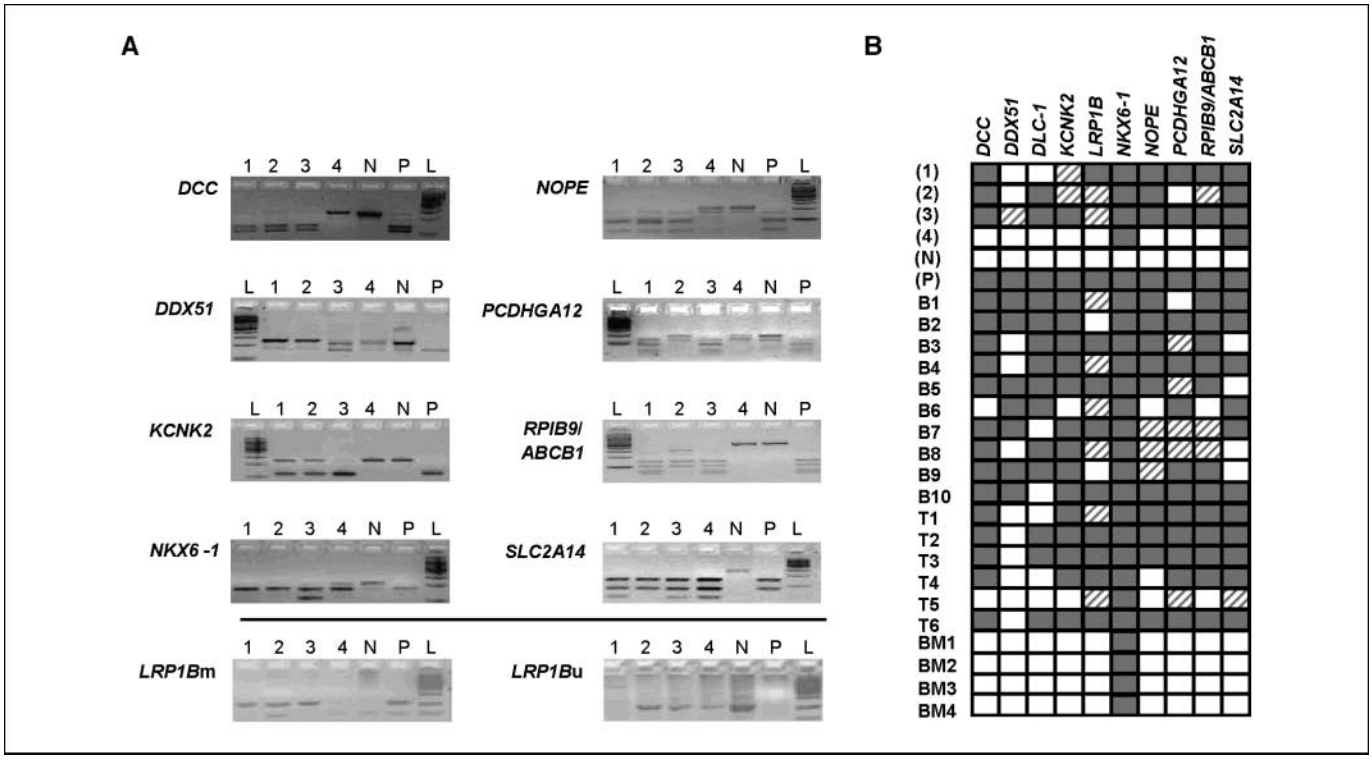

Figure 1. *A,* gel images and summary of results of methylation present in 10 CGI clones associated with 11 genes (*DLC-1* was validated using qMSP; therefore, no gel images are available) in four ALL cell lines. *1,* Jurkat; *2,* MN-60; *3,* NALM-6; *4,* SD-1; *N,* bisulfite-treated normal PBL DNA; *P, Sss*I and bisulfite-treated normal PBL DNA; *L,* ladder. The gel images above the solid line are the results of COBRA analysis and the gel images below the solid line are the results of MSP. *LRP1B*m: assay for methylated allele; *LRP1B*u: assay for unmethylated allele. *B,* validation summary of methylation present in 10 CGI clones associated with 11 genes identified by the DMH analysis in 4 ALL cell lines (*1*, Jurkat; *2*, MN-60; *3*, NALM-6; *4*, SD-1), a negative control (*N*, normal peripheral blood DNA), a positive control (*P, Sss*I-treated normal peripheral blood DNA), 10 B-ALL patients (*B1–B10*), 6 T-ALL patients (*T1–T6*), and 4 normal controls (*BM1–BM4*). *DLC-1* was validated by real-time qMSP assay, *LRP1B* was validated by qMSP, and the remaining genes were validated by COBRA. Analysis of gel images (see *A*) resulted in scoring a sample as methylated (*gray blocks*), partially methylated (*hatched blocks*), or not methylated (*white blocks*). *Column,* gene; *row,* sample.

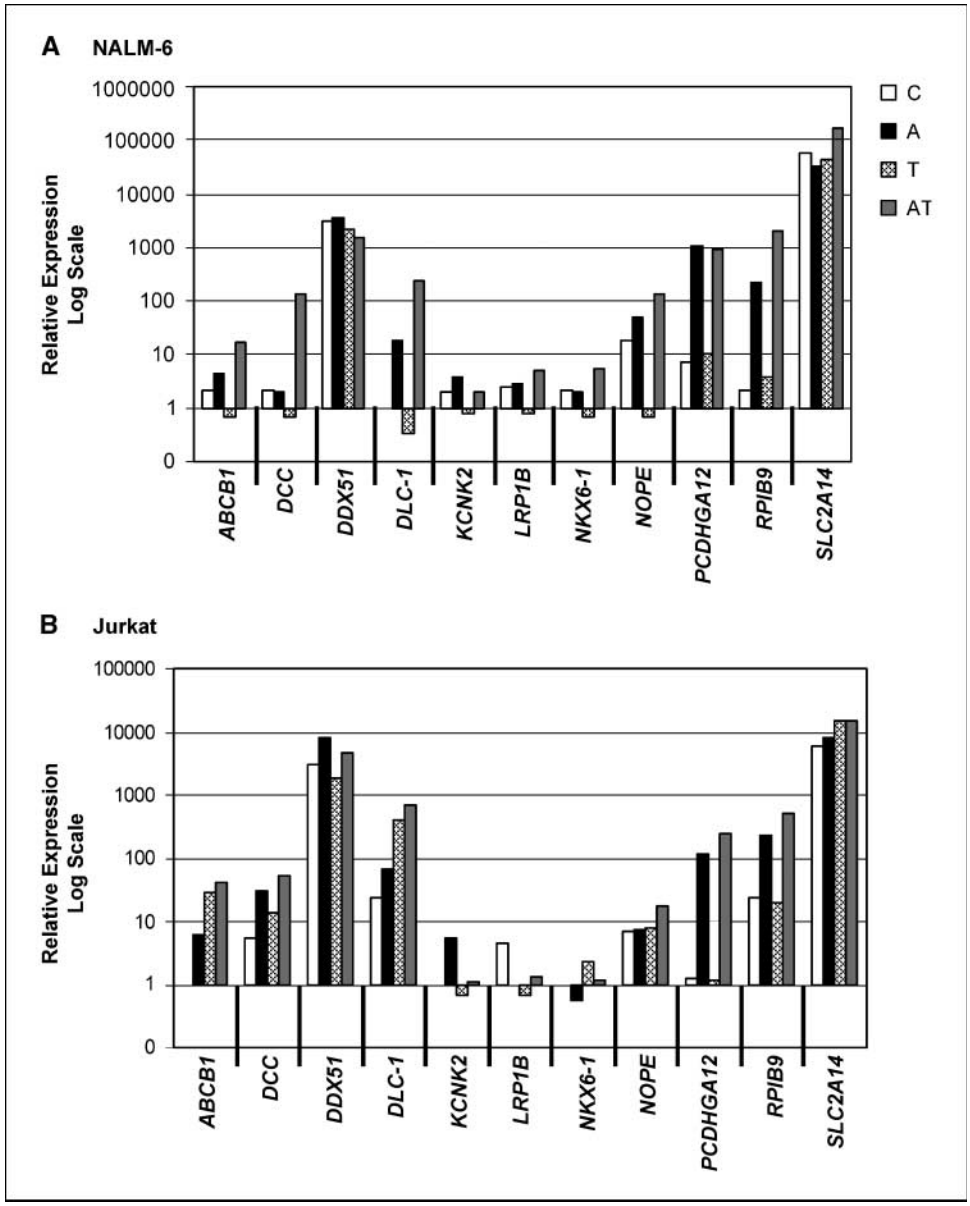

Figure 2. Change in mRNA expression in NALM-6 and Jurkat cell lines following treatment with a demethylating agent and a histone deacetylase inhibitor. *ABCB1, RPIB9, PCDHGA12, DCC, KCNK2, NOPE, NKX6-1, LRP1B, DDX51*, and *SLC2A14* expression levels were measured using SYBR Green assays. *DLC-1* expression levels were measured using Taqman assays. *A,* NALM-6 mRNA relative expression levels in 11 candidate genes. Each gene includes expression in the cell line before treatment (*C*); expression post-treatment with 5-aza (*A*); expression post-treatment with TSA (*T*); and expression post-treatment with a combination of 5-aza and TSA (*AT*). No bar present indicates no amplification. *B,* Jurkat mRNA relative expression levels in 11 candidate genes. Each gene includes expression in the cell line before treatment (*C*); expression post-treatment with 5-aza (*A*); expression post-treatment with TSA (*T*); and expression post-treatment with a combination of 5-aza and TSA (*AT*). No bar present indicates no amplification.

COBRA results, these genes had significant DNA methylation in NALM-6; Jurkat; patients B3, T3, and B4 (*PCDHGA12*); and patients B10, T2, and B5 (*RPIB9*) and little or no methylation in SD-1, normal bone marrow, patient B1 (*PCDHGA12*), and patient T5 (*RPIB9*; Fig. 3*B*). The proportions of methylated clones detected by bisulfite sequencing and the samples that were considered methylated by the COBRA assays were positively correlated for both *PCDHGA12* ($r = 0.68$; $P = 0.022$, one-sided hypothesis) and *RPIB9* ($r = 0.87$; $P = 0.001$, one-sided hypothesis).

Physical mapping of methylated loci reveals methylation "hotspots." The total number of unique CGI loci represented on the microarray was determined by merging clones with overlapping sequences into individual loci. First, each clone sequence was examined for the presence of *Bst*UI and/or *Hpa*II restriction sites. Only those clones with at least 50 bp of sequence that could be unambiguously aligned to the human genome sequence (chromosomes 1–22; X or Y) and which contained at least one restriction site were used for this process. Merging overlapping clones resulted in 3,435 unique CGI loci spanning all of the autosomes, the X and Y chromosomes (Table 2). These loci include regions of the genome that are found in proximity to known genes and annotated CGIs and also regions that are not. The reduction of unique loci reported in this study compared with the 4,564 originally reported by Heisler et al. (14) can be attributed to the more stringent inclusion criteria used in this study. The 3,435 unique loci, which include the 262 differentially methylated candidate loci detected by DMH, were mapped according to their chromosomal locations using MapChart (24) to examine the nature of their physical distribution within the genome (Fig. 4). This process revealed that the unique loci represented on this microarray are distributed across all chromosomes, albeit with a sparse representation on the Y chromosome, centromeric regions, and nucleolar organizing regions, which is consistent with the observations of others (14, 25, 26). The genomic distribution of the loci that are hypermethylated in ALL is not random across the ~3.5K possible chromosomal sites represented on the

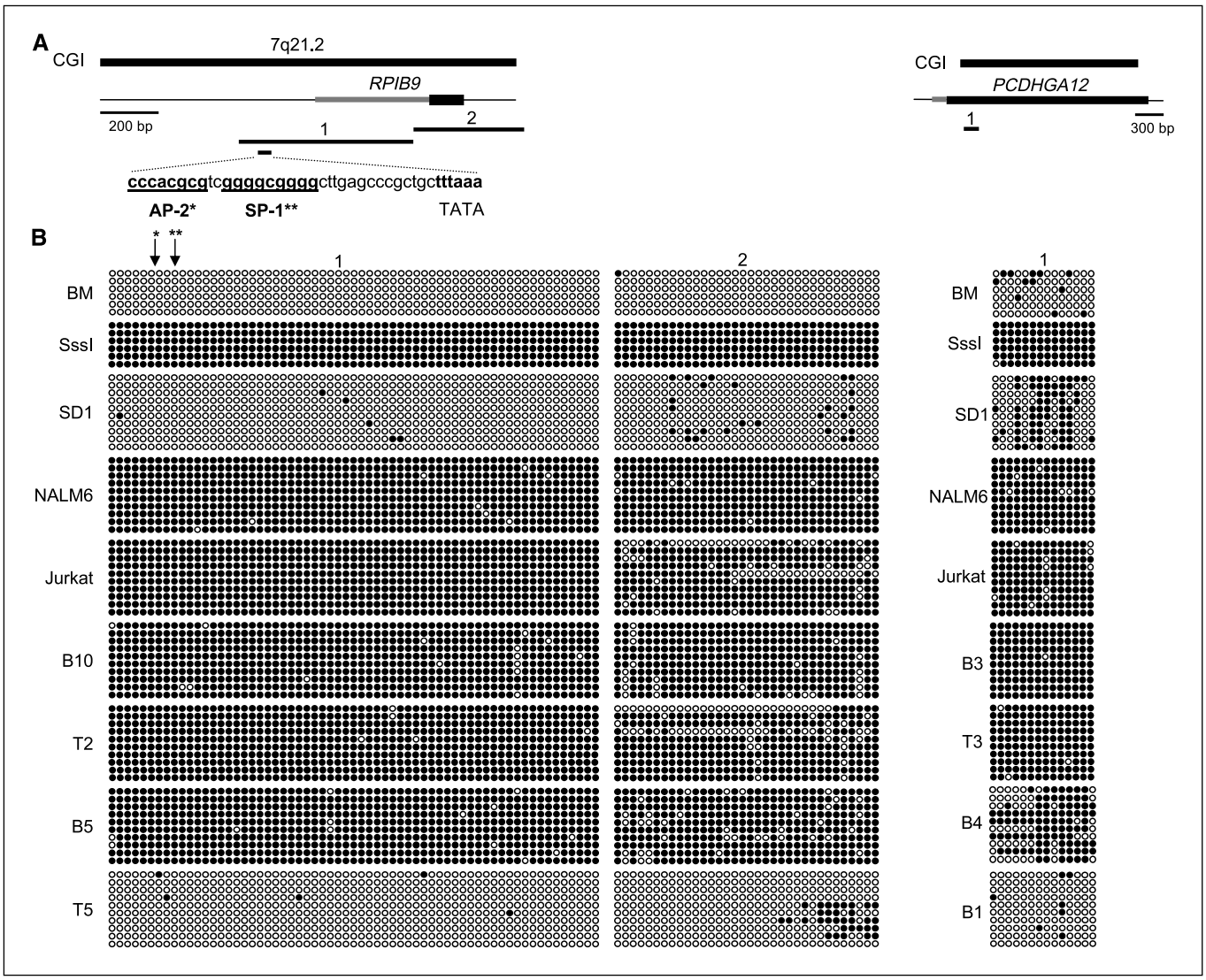

Figure 3. Methylation maps of *RPIB9/ABCB1* and *PCDHGA12. A,* genomic region encompassed by differentially methylated CGI. *Left,* the CGI encompassing the predicted promoter, first exon, and a portion of the first intron of *RPIB9* (this region is also contained within the first intron of *ABCB1*); *right,* the CGI located solely within the first exon of *PCDHGA12*. The CGI is indicated at the top of each image with a heavy black line. Next is a depiction of exon 1 for both genes with the 5′-untranslated region (*thinner gray box*) and the coding portion of exon 1 (*thicker black box*). The thin line represents genomic sequence upstream and downstream of exon 1 with the scale represented with a black bar and 200 bp (*RPIB9*) and 300 bp (*PCDHGA12*). The thick numbered black lines below each gene represent the genomic regions for which we collected bisulfite sequencing data. Two regions were sequenced from the *RPIB9/ABCB1* CGI and one region was sequenced from the *PCDHGA12* CGI. The small black line below *RPIB9* fragment one represents the location of the predicted promoter sequence with AP2, SP1, and TATA binding motifs indicated in bold letters. *B,* bisulfite genomic sequencing. One normal bone marrow sample, one *Sss*I-treated normal PBL DNA sample, three ALL cell lines (SD1, NALM6, and Jurkat), and four primary tumor DNA samples (*RPIB9*: B5, B10, T2, and T5; *PCDHGA12*: B1, B3, B4, and T3) were bisulfite converted and sequenced for two regions in *RPIB9/ABCB1* and one region in *PCDHGA12*. Primary tumor samples beginning with a B are from patients with B-cell ALL and primary tumor samples beginning with a T are from patients with T-cell ALL. Cytosines present within the AP2 (*, *arrow*) and SP1 (**, *arrow*) transcription factor binding sites.

microarray. The nonparametric test for chromosomal clustering of loci methylated in ALL was significant for chromosomes 11 (P = 0.043), 18 (P = 0.044), and 19 (P = 0.009).

Discussion

Despite the inherent biases associated with the methods used in this study, we present an initial global view of the ALL epigenome by providing a low-resolution map of the DNA methylation present within the ALL methylome. Further studies, using whole-genome technologies, will provide a more comprehensive view of the ALL methylome and should be considered. We have identified 262 aberrantly methylated loci, which include genes involved in a variety of important cellular processes, including transcription, cell growth, nucleotide binding, transport, apoptosis, and cell signaling (27). Many of these genes are novel discoveries that have not been associated previously with aberrant methylation in ALL or in other malignancies. Aberrant methylation was confirmed in *ABCB1, DCC, DDX51, DLC-1, potassium channel, subfamily K, member 2* (*KCNK2*), *LRP1B, NKX6-1*, likely orthologue of mouse *neighbor of Punc E11* (*NOPE*), *PCDHGA12, RPIB9*, and *SLC2A14*. The methylation patterns detected in these genes suggest their usefulness as potential biomarkers for ALL. Of particular interest is the usefulness of *DDX51* for distinguishing between B- and T-cell ALL cases (P = 0.011,

Table 2. Characterization of CGI clones

Clones meeting criteria	5,756
Unique CHIP loci	3,435
Known gene	2,267
Within 2,500 bp of TSS	2,090
Within 500 bp of CGI	2,365
Overlaps CGI	1,904
Within 2,500 bp of TSS and 500 bp of CGI	1,781
Hypermethylated in at least 25% of ALL patients	262
Hypermethylated in ALL and within 2,500 bp of TSS	148
Hypermethylated in ALL, within 2,500 bp of TSS and 500 bp of CGI	131

NOTE: Clones not meeting criteria included those (*a*) without *Bst*UI or *Hpa*II restriction sites; (*b*) <50 bp in length; (*c*) aligning to multiple regions in the genome; or (*d*) with homology to the mitochondrial genome.
Abbreviation: TSS, transcription start site.

Fisher's exact test). Treatment of ALL cell lines with a demethylating agent alone or in combination with a histone deacetylase inhibitor *in vitro* resulted in an increase in the expression of these genes, supporting their biological relevance in ALL. Although the explicit roles of these genes in the pathobiology of ALL remain unknown, the epigenetic profiles generated in this study may provide insights toward an improved understanding of ALL and should stimulate further investigations into the roles of these genes in cancer.

Of the validated candidate genes, *ABCB1, RPIB9*, and *PCDHGA12* have functions that may be related to patient response to chemotherapy or to the etiology of ALL. It has been shown recently that hypomethylation of the *ABCB1* promoter leads to multidrug resistance (28), although methylation of this promoter is associated with the down-regulation of gene expression in ALL (29). The function of *RPIB9* has yet to be confirmed but it likely acts as an activator of *Rap*, which allows B cells to participate in cell-cell interactions and contributes to the ability of B-lineage cells to bind to bone marrow stromal cells, which is required for B-cell maturation (30) and is thought to be important in B-cell ALL (31). Likewise, the function of *PCDHGA12* is presently unknown;

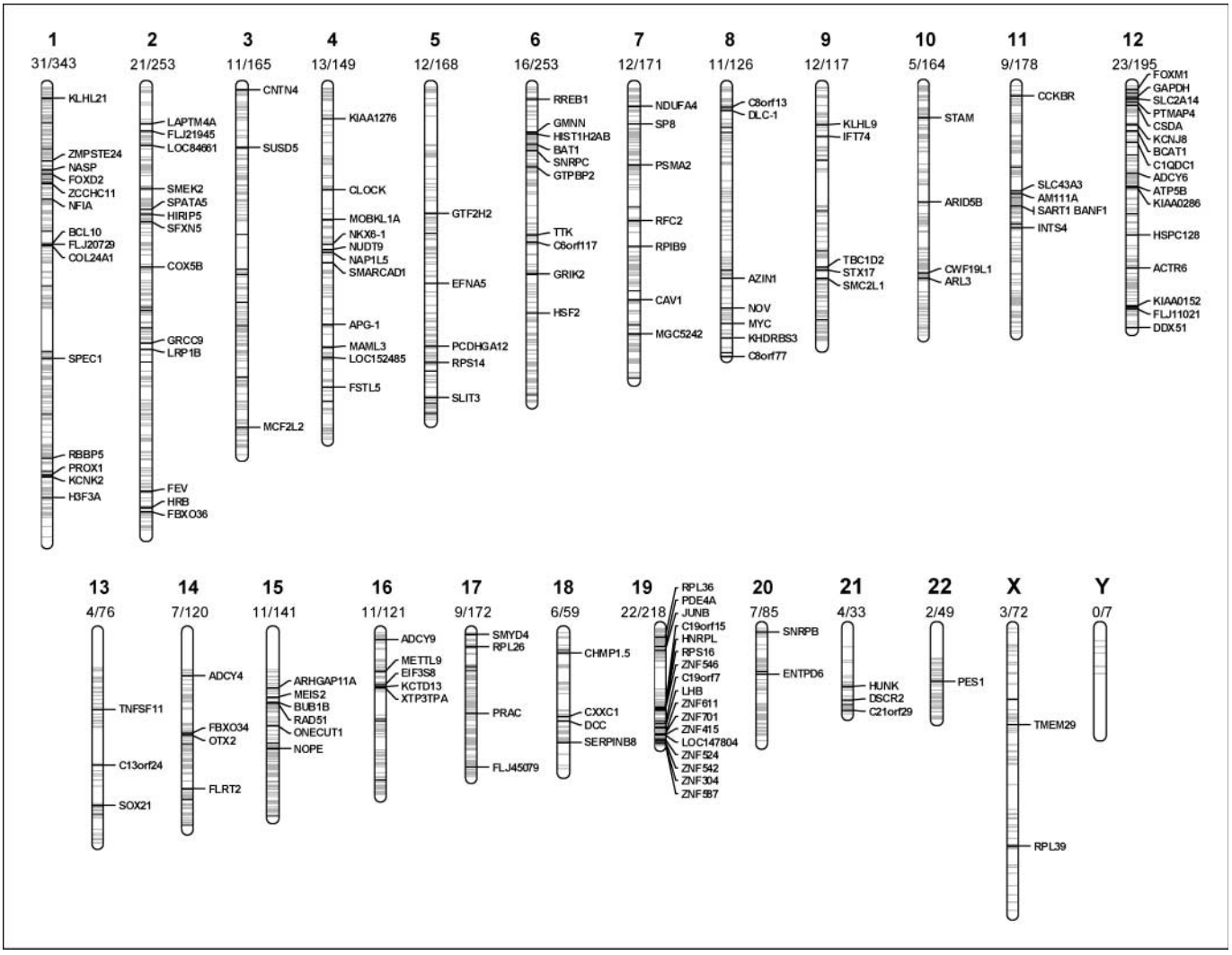

Figure 4. Ideogram representing the chromosomal localization of methylated and unmethylated loci represented on the CGI microarray. Each chromosome is numbered at the top. Immediately below each chromosome number is the number of methylated loci/total number of loci on the microarray for each chromosome. The lighter bands on each chromosome represent the locations of unmethylated loci and the black bands represent locations of loci methylated in ALL. Gene names or the clone designations for unknown loci are indicated beside each methylated locus.

however, an association between promoter methylation and gene silencing has been shown recently in astrocytomas for another member of this family, *PCDHGA11* (32). Bisulfite sequence analysis of these genes, including the putative promoter and first exon of *RPIB9* (this region also overlaps the first intron of *ABCB1*) and the first exon of *PCDHGA12*, provided insight into mechanisms that might be responsible for the low levels of mRNA expression observed in the NALM-6 and Jurkat cell lines. Methylation of transcription factor binding sites (SP1 and AP2) located in the promoter of *RPIB9* may inhibit the binding of the SP1 and/or the AP2 transcription factor(s), thereby controlling its transcriptional regulation. In addition, the dense methylation observed in both fragments may contribute to the down-regulation of both *RPIB9* and *PCDHGA12* by rendering the chromatin structure inaccessible to the transcription machinery.

An intriguing but unanswered question is why gene promoter methylation seems to manifest in a nonrandom fashion in cancers. Methylation patterns are specific to cell or tumor types and even to subtypes within the same tumor category. In addition to revealing methylation patterns specific to ALL, we identified genomic hotspots, which harbor an overabundance of methylated loci. In particular, we show that the methylated loci on chromosome 19 tend to be aggregated in a methylation hotspot toward the telomere of the q arm. Others have reported the aberrant methylation of *NES-1* in ALL (33), which is also located within this aggregate of methylated loci on chromosome 19 (but not present on the CGI microarray used in this study). Given that the statistical analysis used herein considered only the loci present on the chip versus those loci that were differentially methylated in ALL, the hotspot detected is not a reflection of the extraordinarily high density of genes present on chromosome 19. Notably, the zinc-finger genes look as if they are overrepresented; however, this is likely an artifact due to the clustering of gene family members within a region of HSA19 that seems to be a methylation target (34, 35). Interestingly, this region also harbors the maternally imprinted genes *IMP01, PEG3,* and *ZIM2*. Given that it has been shown that *de novo* methylation targets the promoters of imprinted genes in stressful situations, such as tumorigenesis (36), we cannot help but speculate that this imprinted region of chromosome 19 might be a target for *de novo* methylation in ALL.

It is apparent that certain regions of the genome are targeted for methylation and that these targets vary with cancer type. What is not clear is the mechanism that is responsible for the disease-specific targeting of methylated loci. We hypothesize that just as unique chromosomal translocations are associated with specific malignancies, unique targeted chromosomal methylation hotspots are also associated with specific malignancies, such as ALL. However, because translocations are likely a direct result of genomic instability, which is associated with hypomethylation and not hypermethylation, we would expect alternate genomic regions to be prone to the phenomenon of targeted methylation. This hypothesis is corroborated in this study given that the genomic regions associated with the most commonly reported translocations in ALL (*ABL,* 9q34; *BCR,* 22q11; *TEL,* 12p13; *AML,* 21q22; *ENL,* 19p13; *MLL,* 11q23; and *AF-4,* 4q21) are not densely methylated with the exception of 12p, which is also associated with allelic loss in ALL. Other regions of the genome that commonly experience allelic loss in ALL include 9p, 20q, 5p, 10p, 6q, and 17p, none of which was identified to harbor an overabundance of methylated loci in this study.

Finally, the concept of intranuclear chromosomal territories and loops of chromatin that may interact functionally (reviewed in ref. 37) is noteworthy and suggests that the proximity of chromosomal regions within the nucleus might facilitate the spreading of methylation. This is particularly interesting because hypermethylated gene promoters tend to be aggregated within the genome of cancer cells (38), the spreading of epigenetic events and of gene silencing are known to occur (39), chromosomes colocalize temporally in resting and activated B cells (40), and in some instances, nuclear architecture is thought to affect several processes, including epigenetic mechanisms (41, 42). Although this mechanism is purely speculative, the existence of nonrandom patterns of methylation in ALL is intriguing and warrants further investigation. We propose that it is time to begin the study of the epigenetic geography of tissues, cells, chromosomes, and genes to better understand their contribution to the aberrant epigenetic events observed in cancer and disease.

Acknowledgments

Received 10/30/2006; revised 1/2/2007; accepted 1/16/2007.

Grant support: National Cancer Institute grants CA100055 and CA097880 and the National Library of Medicine grant LM07089 (C.W. Caldwell).

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 Angel Surdin for editorial support.

References

1. Jaffe ES, Harris NL, Stein H, Vardiman JE. World health organization classification of tumors. Pathology and genetics of tumours of haematopoietic and lymphoid tissues. Lyon: IARC Press; 2001.
2. Graux C, Cools J, Michaux L, Vandenberghe P, Hagemeijer A. Cytogenetics and molecular genetics of T-cell acute lymphoblastic leukemia: from thymocyte to lymphoblast. Leukemia 2006;9:1496–510.
3. Pui CH, Crist WM, Look AT. Biology and clinical significance of cytogenetic abnormalities in childhood acute lymphoblastic leukemia. Blood 1990;76:449–63.
4. Wagner W, Saffrich R, Wirkner U, et al. Hematopoietic progenitor cells and cellular microenvironment: behavioral and molecular changes upon interaction. Stem Cells 2005;23:1180–91.
5. Yoder JA, Soman NS, Verdine GL, Bestor TH. DNA (cytosine-5)-methyltransferases in mouse cells and tissues. Studies with a mechanism-based probe. J Mol Biol 1997;270:385–95.
6. Roman-Gomez J, Castillejo JA, Jimenez A, Barrios M, Heiniger A, Torres A. The role of DNA hypermethylation in the pathogenesis and prognosis of acute lymphoblastic leukemia. Leuk Lymphoma 2003;44:1855–64.
7. Shteper PJ, Siegfried Z, Asimakopoulos FA, et al. ABL1 methylation in Ph-positive ALL is exclusively associated with the P210 form of BCR-ABL. Leukemia 2001;15:575–82.
8. Zheng S, Ma X, Zhang L, et al. Hypermethylation of the 5′ CpG island of the FHIT gene is associated with hyperdiploid and translocation-negative subtypes of pediatric leukemia. Cancer Res 2004;64:2000–6.
9. Matsushita C, Yang Y, Takeuchi S, et al. Aberrant methylation in promoter-associated CpG islands of multiple genes in relapsed childhood acute lymphoblastic leukemia. Oncol Rep 2004;12:97–9.
10. Huang TH, Perry MR, Laux DE. Methylation profiling of CpG islands in human breast cancer cells. Hum Mol Genet 1999;8:459–70.
11. Rahmatpanah FB, Carstens S, Guo J, et al. Differential DNA methylation patterns of small B-cell lymphoma subclasses with different clinical behavior. Leukemia 2006;10:1855–62.
12. Shi H, Guo J, Duff DJ, et al. Discovery of novel epigenetic markers in non-Hodgkin's lymphoma. Carcinogenesis 2007;28:60–70.
13. Yan PS, Chen CM, Shi H, Rahmatpanah F, Wei SH, Caldwell CW, Huang TH. Dissecting complex epigenetic alterations in breast cancer using CpG island microarrays. Cancer Res 2001;61:8375–80.
14. Heisler LE, Torti D, Boutros PC, et al. CpG island microarray probe sequences derived from a physical library are representative of CpG Islands annotated on the human genome. Nucleic Acids Res 2005;33:2952–61.
15. Rozen S, Skaletsky H. Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol 2000;132:365–86.
16. Li LC, Dahiya R. MethPrimer: designing primers for methylation PCRs. Bioinformatics 2002;18:1427–31.

17. Vandesompele J, De PK, Pattyn F, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol 2002;3:research0034.1–11.
18. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(−ΔΔ C(T)) method. Methods 2001;25:402–8.
19. Gebhard C, Schwarzfischer L, Pham TH, et al. Genome-wide profiling of CpG methylation identifies novel targets of aberrant hypermethylation in myeloid leukemia. Cancer Res 2006;66:6118–28.
20. Garcia-Manero G, Jeha S, Daniel J, et al. Aberrant DNA methylation in pediatric patients with acute lymphocytic leukemia. Cancer 2003;97:695–702.
21. 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.
22. Yuan BZ, Jefferson AM, Baldwin KT, Thorgeirsson SS, Popescu NC, Reynolds SH. DLC-1 operates as a tumor suppressor gene in human non-small cell lung carcinomas. Oncogene 2004;23:1405–11.
23. Sonoda I, Imoto I, Inoue J, et al. Frequent silencing of low density lipoprotein receptor-related protein 1B (LRP1B) expression by genetic and epigenetic mechanisms in esophageal squamous cell carcinoma. Cancer Res 2004;64:3741–7.
24. Voorrips RE. MapChart: software for the graphical presentation of linkage maps and QTLs. J Hered 2002;93:77–8.
25. Das R, Dimitrova N, Xuan Z, et al. Computational prediction of methylation status in human genomic sequences. Proc Natl Acad Sci U S A 2006;103:10713–6.
26. Rollins RA, Haghighi F, Edwards JR, et al. Large-scale structure of genomic methylation patterns. Genome Res 2006;16:157–63.
27. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell 2000;100:57–70.
28. Baker EK, Johnstone RW, Zalcberg JR, El-Osta A. Epigenetic changes to the MDR1 locus in response to chemotherapeutic drugs. Oncogene 2005;24:8061–75.
29. Garcia-Manero G, Daniel J, Smith TL, et al. DNA methylation of multiple promoter-associated CpG islands in adult acute lymphocytic leukemia. Clin Cancer Res 2002;8:2217–24.
30. McLeod SJ, Shum AJ, Lee RL, Takei F, Gold MR. The Rap GTPases regulate integrin-mediated adhesion, cell spreading, actin polymerization, and Pyk2 tyrosine phosphorylation in B lymphocytes. J Biol Chem 2004;279:12009–19.
31. Gibson LF. Survival of B lineage leukemic cells: signals from the bone marrow microenvironment. Leuk Lymphoma 2002;43:19–27.
32. Waha A, Guntner S, Huang TH, et al. Epigenetic silencing of the protocadherin family member PCDH-γ-A11 in astrocytomas. Neoplasia 2005;7:193–9.
33. Roman-Gomez J, Jimenez-Velasco A, Castillejo JA, et al. Promoter hypermethylation of cancer-related genes: a strong independent prognostic factor in acute lymphocytic leukemia. Blood 2004;104:2492–8.
34. Ashworth LK, Batzer MA, Brandriff B, et al. An integrated metric physical map of human chromosome 19. Nat Genet 1995;11:422–7.
35. Shannon M, Ashworth LK, Mucenski ML, Lamerdin JE, Branscomb E, Stubbs L. Comparative analysis of a conserved zinc finger gene cluster on human chromosome 19q and mouse chromosome 7. Genomics 1996;33:112–20.
36. Pantoja C, de Los Rios L, Matheu A, Antequera F, Serrano M. Inactivation of imprinted genes induced by cellular stress and tumorigenesis. Cancer Res 2005;65:26–33.
37. Cremer T, Cremer M, Dietzel S, Muller S, Solovei I, Fakan S. Chromosome territories—a functional nuclear landscape. Curr Opin Cell Biol 2006;18:307–16.
38. Keshet I, Schlesinger Y, Farkash S, et al. Evidence for an instructive mechanism of *de novo* methylation in cancer cells. Nat Genet 2006;38:149–53.
39. Di CL, Buschbeck M, Gutierrez A, et al. Altered epigenetic signals in human disease. Cancer Biol Ther 2004;3:831–7.
40. Skok JA, Brown KE, Azuara V, et al. Nonequivalent nuclear location of immunoglobulin alleles in B lymphocytes. Nat Immunol 2001;2:848–54.
41. Fuks F. DNA methylation and histone modifications: teaming up to silence genes. Curr Opin Genet Dev 2005;15:490–5.
42. Baylin SB. DNA methylation and gene silencing in cancer. Nat Clin Pract Oncol 2005;2 Suppl 1:S4–11.