A Single Nucleotide Polymorphism Chip-Based Method for Combined Genetic and Epigenetic Profiling: Validation in Decitabine Therapy and Tumor/Normal Comparisons

Eric Yuan,[1] Fatemeh Haghighi,[2] Susan White,[3] Ramiro Costa,[4] Julie McMinn,[1] Kathy Chun,[5] Mark Minden,[6] and Benjamin Tycko[1,3]

[1]Institute for Cancer Genetics, [2]Genome Center, and [3]Department of Pathology, Columbia University Medical Center; [4]New York State Psychiatric Institute, New York, New York and [5]Department of Laboratory Medicine and Hematology and [6]Department of Medicine and Leukemia Unit, Ontario Cancer Institute and University of Toronto, Toronto, Ontario, Canada

Abstract

Progress on several unresolved issues in cancer epigenetics will benefit from rapid and standardized methods for profiling DNA methylation genome-wide. In the area of epigenetic therapy, the demethylating drug decitabine (5-aza-2′-deoxycytidine) is increasingly used to treat acute myelogenous leukemia and myelodysplastic syndrome, but the mechanisms of its anticancer activity have remained unclear. Given the clinical efficacy of decitabine and the uncertainties about its mode of action, it will be useful to optimize methods for following DNA methylation as a biochemical response in individual patients. Here, we describe a single nucleotide polymorphism (SNP) chip-based method (MSNP) for profiling DNA methylation. Using this procedure, the extent of demethylation in bone marrow aspirates from patients with leukemia receiving decitabine can be assessed genome-wide using commercially available (Affymetrix) SNP chips. We validated the accuracy of MSNP by comparing the results with combined bisulfite restriction analysis and by sequencing cloned PCR products from bisulfite-converted DNA. We further validated MSNP in a Wilms' tumor/normal kidney comparison, comparing the results with methylation-sensitive Southern blotting. MSNP simultaneously detects aberrations in DNA copy number and loss of heterozygosity, making it a generally useful approach for combined genetic and epigenetic profiling in tissue samples from cancer patients. (Cancer Res 2006; 66(7): 3443-51)

Introduction

A large body of work has shown that epigenetic aberrations in cancer cells, including both gains and losses of DNA methylation, contribute importantly to tumor initiation and progression (1, 2). However, the correlative and functional data obtained thus far have raised several basic and clinical issues that will benefit from rapid and standardized high-resolution methods for genome-wide epigenetic profiling. The deoxycytidine analogue decitabine (5-aza-2′-deoxycytidine) and related compounds have a long history of use as experimental reagents to study the effects of DNA methylation on gene expression and cell differentiation, and these drugs have been given to patients in various clinical settings, including cancer and hemoglobinopathies (3) However, decitabine has only been accepted as a component of standard therapy recently, with documentation of clinical responses to low-dose regimens with low toxicity in patients with acute myelogenous leukemia (AML) and myelodysplastic syndrome as well as in chronic myelogenous leukemia (CML; refs. 4–8). These important studies, together with the ongoing development of orally active demethylating agents and the possibility of using combinations of DNA methyltransferase and histone deacetylase inhibitors (9, 10), have led to new optimism about the potential for "epigenetic therapies" in clinical oncology. In one of the prior clinical studies, the combined bisulfite restriction analysis (COBRA) method was used to measure DNA methylation of the *p15* gene promoter region in peripheral blood leukocytes, and no consistent correlations with clinical response, or indeed with decitabine treatment itself, were found (6). In a second study, involving patients with CML, a bisulfite conversion/PCR/pyrosequencing assay measuring DNA methylation of LINE1 repetitive sequences was used to monitor the effects of drug treatment, and this procedure revealed a definite demethylating effect in peripheral blood leukocytes, albeit inversely correlated with clinical response (7). Because the mode of action of decitabine as an anticancer agent remains unclear, additional investigation of DNA demethylation as a biochemical response in patients seems warranted. This is particularly true because prior studies have examined only a few types of sequences.

There are many other unresolved issues in cancer epigenetics, arising from gaps in our current knowledge of the genomic location, timing, and extent of gains and losses of DNA methylation in tumor initiation and progression, which would benefit from new analytic methods. Each method for assessing DNA methylation has different advantages and disadvantages in terms of genomic coverage, types of sequences examined, and ease of data interpretation, cost, and technical difficulty (11–13). Microarray-based methods have advantages of broad genomic coverage, commercial availability and standardization of protocols, and ease of data sharing. Here, we describe results from a pilot study in which we have adapted Affymetrix (Santa Clara, CA) single nucleotide polymorphism (SNP) chips to profile the demethylating response to decitabine in patients with AML. The SNP chip-based method (MSNP) method requires small amounts of starting genomic DNA, making it applicable to clinical samples, and it simultaneously detects DNA methylation and DNA copy number, eliminating ambiguities in interpreting the methylation data and allowing an assessment of epigenetic-genetic interactions. Among many other potential uses, the MSNP method is applicable to uncovering tumor-associated changes in DNA methylation, which we illustrate in a Wilms' tumor/normal comparison.

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

Requests for reprints: Benjamin Tycko, Institute for Cancer Genetics, Columbia University Medical Center, 1150 St. Nicholas Avenue, New York, NY 10032. Phone: 212-305-1149; Fax: 212-305-5489; E-mail: bt12@columbia.edu.

doi:10.1158/0008-5472.CAN-05-3739

Materials and Methods

AML patient characteristics and decitabine dosage. Patients with a diagnosis of relapsed or refractory AML were seen in the Leukemia Unit of the Princess Margaret Hospital. In each patient, decitabine was infused i.v. over a period of 3 hours twice daily for 5 days, with the doses as follows: patient 1 (7.5 mg/m^2), patient 2 (10 mg/m^2), patient 3 (17.5 mg/m^2), and patient 4 (20 mg/m^2). Bone marrow aspirates were collected before the first dose of decitabine and on day 8. The cellular composition of these samples was assessed by morphology and, by this criterion, did not reveal a substantial reduction of the malignant cells (percentage of blasts) in this short time frame in any of the patients. At both time points, bone marrow was placed into a PAXGene tube for isolation of RNA and into a sodium heparin tube for isolation of DNA. DNA was isolated by separating mononuclear cells from red cells by Ficoll-Hypaque centrifugation. The mononuclear cells were washed in PBS lacking calcium and magnesium. Following cell lysis in SDS and proteinase K digestion, the DNA was isolated by phenol/chloroform extraction and ethanol precipitation.

Wilms' tumors and matched normal kidneys. Wilms' tumors were untreated sporadic cases, with nephrectomy done at Columbia University Medical Center. Samples of Wilms' tumor and adjacent nonneoplastic kidney cortex were cryopreserved at surgery. The tissues were pulverized under liquid nitrogen and genomic DNA was prepared by a standard protocol using proteinase K/SDS lysis followed by phenol/chloroform extraction and ethanol precipitation.

Southern blotting. Genomic DNA was digested with the indicated restriction enzymes, electrophoresed through 1% agarose gels, denatured, neutralized, and transferred to Nytran membranes (Whatman, Brentford, Middlesex, United Kingdom). The probes were synthesized by genomic PCR with primers flanking the SNP loci (sequences available on request) and ^{32}P-labeled by random priming (2) and hybridized to the membranes under high stringency conditions in Ultrahyb solution (Ambion, Austin, TX).

MSNP probe preparation. For analyses using the 10K SNP chips, genomic DNA (600 ng) was digested with *Xba*I (New England Biolabs, Ipswich, MA) in a total volume of 40 μL for 3 hours at 37°C followed by heat inactivation for 30 minutes at 65°C. Half of the digestion product was further digested with the methylation-sensitive restriction endonuclease *Hpa*II (New England Biolabs) in a total volume of 100 μL for an additional 3 hours at 37°C and then digested for 30 more minutes with 1 μL additional enzyme to ensure complete digestion followed by heat inactivation for 30 minutes at 65°C. The aliquot of DNA, which was not cut with *Hpa*II, was diluted to the same volume in 1× enzyme buffer and incubated in parallel. Of the resulting digestion products, 20 μL (60 ng) were ligated with *Xba*I adaptors as described in the Affymetrix SNP chip protocol. PCR was carried out essentially as described in the Affymetrix protocol; however, to compensate for the smaller amount of template, PCR was carried out for 38, rather than 35, cycles. The resulting PCR products (*XbaI*-only and *Xba*I/*Hpa*II genomic representations) were analyzed by agarose gel electophoresis and ethidium bromide staining and had the characteristic band pattern described in the Affymetrix protocol. PCR purification was carried out using the MiniElute 96 UF PCR purification kit (Qiagen, Valencia, CA), and probe fragmentation, labeling, and hybridization were carried out as described in the Affymetrix protocol. For analyses using the 50K SNP chips, 3 μg starting DNA was digested first with *Xba*I in a total volume of 60 μL, and 20 μL were further digested in either *Hpa*II or *Msp*I (New England Biolabs) in a total volume of 100 μL as described above. Linker ligation and PCR for preparing the probes was then carried out as described in the Affymetrix protocol. Probes were hybridized to 10K or 50K SNP chips according to the Affymetrix protocol.

DNA copy number analysis. The *Xba*I-only SNP chip data from genomic DNA of AML bone marrow aspirates as well as for two normal human genomic DNA as euploid controls were loaded to dChip2004 after conversion of the *.cel files from binary to text format using the Cel File Conversion Tool (Affymetrix). SNP chip signal intensity values were calculated in dChip2004 using the perfect match/mismatch difference model (14). Briefly, each SNP is represented on the chip array by 20 perfect match and 20 corresponding mismatch targets. For the 20 perfect match/mismatch target pairs, the mismatch signal for a given SNP detected by the scanner was subtracted from its corresponding perfect match signal. This was followed by between-chip normalization and within-chip normalization of the signals using model-based expression. Invariant set normalization conducts pair-wise comparison between chips and creates a normalization curve based on a subset of probes with small within-subset rank difference between the two chips. MBE calculates a signal value for each SNP based on the 20 perfect match/mismatch differences for that given SNP. Gains and reductions in DNA copy number were determined by visual inspection of inferred copy number generated by dChip using analysis of signal intensity values by the hidden Markov model. Loss of heterozygosity (LOH) was scored according to the SNP chip genotype calls; as we have shown previously, even in unpaired tumor samples, stretches of >100 physically contiguous homozygous genotypes in the 10K SNP chip data identify regions of LOH with high sensitivity and specificity (15).

MSNP data analysis. Genotype calls were generated by the Affymetrix software packages GCOS and GDAS, and signal intensity values were generated by dChip2004 as described above. SNP classes were determined for the 10K and 50K SNP chips by a computerized search of the May 2004 Build of the human genome, including the SNP database, in which every *Xba*I restriction fragment containing a SNP represented on the chips was queried for internal *Hpa*II sites and also queried for additional internal SNPs creating or eliminating *Hpa*II sites (see Supplementary Table S1A-F for complete lists of SNPs sorted by class). Class 1 SNPs loci (non-informative for DNA methylation but useful as internal controls for the MSNP analysis and as markers for DNA copy number analysis) were defined as those which do not contain a *Hpa*II site between the two flanking *Xba*I sites. Class 2 SNP loci (informative for DNA methylation) were defined as those situated within *Xba*I fragments that do contain a (nonpolymorphic) *Hpa*II site and therefore would be affected in their amplification and representation in the SNP chip probe preparation by whether the *Hpa*II site is methylated. Class 3 SNP loci (not reliably informative for DNA methylation and also not reliable as internal controls) were defined as those in which a *Hpa*II site in the *Xba*I fragment was potentially created or eliminated by a SNP. To assess the demethylating response to decitabine in the AML bone marrow using the 10K SNP chips, we first separated the SNPs into their respective classes and then eliminated those SNPs that had a raw signal <400 in the *Xba*I-only genomic representation in one or more of the experimental samples. This procedure eliminated <10% of the SNPs, often with Affymetrix "absence calls." Next, we normalized the SNP intensity values from the *Xba*I/*Hpa*II representations such that the average signal for the class 1 SNPs was equal for all samples. Normalized SNP intensities and the demethylating response to decitabine (post-decitabine intensity as percentage of pre-decitabine intensity) were color-coded and displayed using dChip.

Bisulfite conversion of DNA and validations by COBRA and cloning/sequencing. The methylation data for selected SNPs were validated by bisulfite conversion using the CpGenome DNA Modification kit (Chemicon, Temecula, CA) followed by PCR amplification and either restriction analysis (COBRA) or cloning and sequencing. PCR was done using PlatinumTaq (Invitrogen, Carlsbad, CA) and with locus-specific primers matching the bisulfite-converted sequences flanking the CpG dinucleotides to be assayed. PCR primers were selected using MethPrimer (sequences in Supplementary Table S1).[7] COBRA restriction enzymes were selected using Snake-charmer.[8] Bisulfite-converted/PCR-amplified DNA was cloned using the TOPO-TA Cloning kit (Invitrogen) and multiple clones were sequenced.

Results

The MSNP method. SNP Chips are microarrays printed with oligonucleotides spanning SNP at defined positions throughout the genome. Oligonucleotides querying both alleles of each SNP are

[7] http://www.urogene.org/methprimer/index1.html.
[8] http://195.83.84.240/cgrunau/methods/snake_charmer.html.

present on the chip, allowing the genotype to be determined simultaneously at thousands of SNPs. Each SNP is linked to a National Center for Biotechnology Information identifier, with the sequence and chromosomal position of that marker. With the current generation of Affymetrix SNP chips, genomic DNAs to be analyzed are first digested with *Xba*I (or *Hin*dIII), generic linkers are added by ligation, and the hybridization probe is prepared by PCR using universal primers matching sequences included in the linkers. As shown in Fig. 1*A*, the only modification of the Affymetrix protocol needed to adapt these chips for methylation analysis is a digestion of the genomic DNA with a methylation-sensitive restriction enzyme (*Hpa*II in these experiments; but other methylation-sensitive restriction enzymes, such as *Cfo*I, can also be used) after *Xba*I digestion followed by addition of linkers and PCR. Sequences that are unmethylated and therefore cut with *Hpa*II cannot be amplified by PCR and therefore drop out from the final probe. Probes from these *Xba*I/*Hpa*II genomic representations and probes from the *Xba*I-only genomic representations, which serve as controls and are useful for DNA copy number analysis, are hybridized to SNP chips using the Affymetrix protocol. With this procedure, SNPs that convert from a strong signal to a weak or absent signal in comparing sample A with sample B (e.g., a tumor/normal comparison or serial samples from patients treated with decitabine) in the *Xba*I/*Hpa*II representation are losses of methylation (LOM); conversely SNPs that convert from weak or absent signal to strong signal are gains of methylation (GOM). The control *Xba*I-only representations should not show such differences in comparing sample A with sample B.

The SNPs on the chip fall into three categories based on the nucleotide sequences of their corresponding *Xba*I fragments: class 1 SNPs lie within *Xba*I fragments that lack internal *Hpa*II sites, class 2 SNPs lie within *Xba*I fragments that contain *Hpa*II sites at positions other than the SNP itself, and class 3 SNPs (the rarest category; see Tables 1 and 2) fall within a *Hpa*II site or are in *Xba*I fragments that contain additional SNPs affecting *Hpa*II sites. In MSNP, the class 2 SNPs are informative and reliable for assessing DNA methylation. In contrast, the class 1 SNPs are noninformative for DNA methylation but serve as useful internal controls: these SNPs should not show a loss of signal intensity in comparing the microarray data from the *XbaI*-only and *Xba*I/*Hpa*II representations for a given tissue sample and also should not show a loss of signal intensity comparing the *Xba*I/*Hpa*II representations for tissue samples before and after treatment with a demethylating drug. Because class 1 SNPs are in *Xba*I fragments that do not contain *Hpa*II sites, variation from one chip to another for these SNPs should reflect only technical factors like small variations in probe concentration and in the amount of PCR product corresponding to each amplified *Xba*I fragment. Such variations have been minimized by Affymetrix in developing the standard SNP chip protocol. Thus, the class 1 SNPs should be useful for normalizing the SNP Chip data across samples in both *XbaI*-only and *Xba*I/*Hpa*II representations. The rare class 3 SNPs are not reliable for scoring DNA methylation. However, epigenetic and genetic information emerge simultaneous from MSNP, and all three SNP classes are useful for measuring DNA copy number and, when matched tumor/normal samples are available, for scoring LOH.

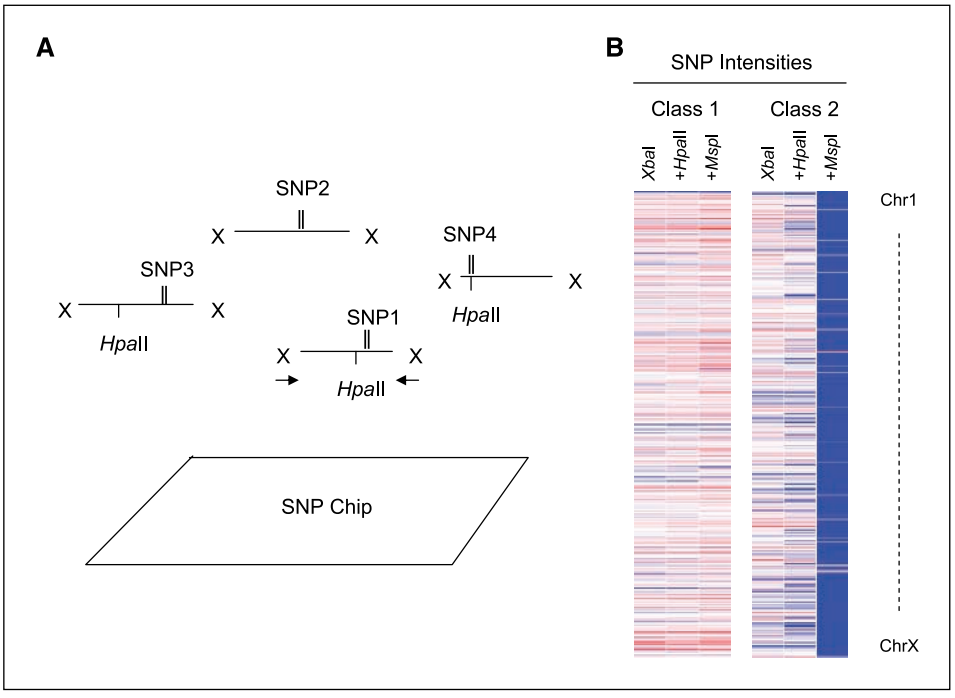

Figure 1. Scoring DNA methylation by MSNP. *A,* linkered *Xba*I genomic fragments. *Arrows,* generic PCR primers. The class 2 SNPs (*SNP1* and *SNP3*), which contain nonpolymorphic *Hpa*II sites in their flanking DNA between the two *Xba*I sites, can be scored for the methylation status of this DNA. The class 1 SNPs (*SNP2*), which lack *Hpa*II sites in the flanking DNA, serve as internal controls. The class 3 SNPs (*SNP4*) map within *Hpa*II sites or are flanked by polymorphic *Hpa*II sites and are therefore not reliably informative for DNA methylation. All three SNP classes are useful for evaluating DNA copy number and LOH. *X, Xba*I. *B,* validation of MSNP by predigesting genomic DNA from a Wilms' tumor with *Xba*I-only, *Xba*I + *Hpa*II, or *Xba*I + *Msp*I before probe synthesis. *Y axis,* SNPs in order of physical position in the genome from chromosome 1 (*Chr1*) to the X chromosome (*ChrX*). The color scale for normalized SNP intensities, from 0 (*blue*) to 5 (*red*), was established using the dChip algorithm (18) based on comparison with 50K SNP chip data obtained with *Xba*I-only genomic representations from a control panel of diploid tissues from nine individuals, setting the mean intensity value for each SNP in this control tissue panel equal to 2 (*white*). Although there is some variation in the individual SNP intensities, the class 1 SNP intensities are insensitive to addition of *Hpa*II or *Msp*I in the predigestion. In contrast, the class 2 SNP intensities are widely reduced with *Msp*I predigestion and detectably but less widely reduced with *Hpa*II predigestion.

Table 1. Signal intensities for the three classes of SNPs in *Xba*I compared with *Xba*I/*Hpa*II and *Xba*I/*Msp*I from a Wilms' tumor

SNP class	No. SNPs	Av. signal (*Xba*I)	Median (*Xba*I)	Av. signal (*Xba*I + *Hpa*II)	Median (*Xba*I + *Hpa*II)	Av. signal (*Xba*I + *Msp*I)	Median (*Xba*I + *Msp*I)
1	38,051	7,254	6,631	7,523	6,957	8,032	7,539
2	14,791	6,474	5,879	5,525	5,069	703	500
3	3,513	6,422	5,792	5,716	5,232	2,525	895

NOTE: Data are raw signal intensities from Affymetrix *Xba*I 50K SNP chips. To increase reliability, only SNPs with signal >400 in the *Xba*I-only genomic representation were considered.

As a crude technical validation of this approach, using 50K SNP chips, we compared the genome-wide SNP chip intensities using *Xba*I-only, *Xba*I + *Hpa*II, and *Xba*I + *Msp*I genomic representations from Wilms' tumor DNA as probes. In this analysis, the *Msp*I restriction enzyme is used as the methylation-insensitive isoschizomer of *Hpa*II. The prediction is that the intensities of the class 2 SNPs should be substantially reduced by the predigestion of the genomic DNA with *Msp*I, whereas the intensities of the class 1 SNPs, lacking adjacent *Hpa*II/*Msp*I sites, should not be affected. As shown numerically by the data in Table 1 and visually by the color-coded SNP intensities in Fig. 1*B*, the class 1 SNP intensities were in fact stable with both *Hpa*II and *Msp*I predigestion. Some of the class 2 SNP intensities were reduced with *Hpa*II predigestion, consistent with CpG methylation of many, although not all, of the *Hpa*II sites in these *Xba*I genomic fragments. As predicted, the class 2 SNP intensities were widely and strongly reduced with the *Msp*I predigestion.

MSNP detects DNA copy number aberrations and LOH in bone marrow from patients with AML receiving decitabine. Our first experimental objective was to profile the demethylating response to decitabine genome-wide in bone marrow aspirates from patients with AML. In this series of analyses, we first asked whether DNA copy number aberrations could be found in these leukemia samples. Our series consisted of four patients with AML, with bone marrow samples obtained immediately before treatment and at 8 days after beginning treatment with low-dose decitabine (eight samples total). Figure 2 shows the results obtained with these cases on 10K SNP chips using the *XbaI*-only representations as probes. The SNP intensities are color-coded with blue indicating reduced signal and red indicating increased signal relative to the values obtained on the chips using normal diploid control DNAs. This procedure revealed a hemizygous interstitial deletion of chromosome 5, spanning a 23-Mb region delimited by SNPs rs1898550 and rs1366236, corresponding to bands 5q23-q33, in AML case 3. The copy number analysis also revealed a hemizygous loss of the entire chromosome 7 in this same case and trisomies of chromosomes 8, 10, 13, and 21 in AML case 2 (Fig. 2*A*; data not shown). These findings in AML case 2 were precisely concordant with the standard G-banded karyotype obtained in this case, which showed 50, XX, +8, +10, +13, and +21. Karyotypes were not available from the other patients. No other DNA copy number abnormalities were found, but we detected LOH for the entire chromosome 13 in AML case 1 (Fig. 2*B*). Importantly for interpreting the subsequent DNA methylation data, these chromosomal gains and losses, which are quite characteristic of AML, were seen in the bone marrow aspirates from these three cases both pre-decitabine and post-decitabine (Fig. 2*A* and *B*), thus providing molecular support (bolstering standard morphologic data from examination of the aspirates) for persistence of the leukemia cells during this 8-day time interval.

MSNP detects DNA demethylation in response to decitabine. As a first test of MSNP for analyzing DNA methylation, we asked if the above predictions for the class 1 versus class 2 SNPs on the

Table 2. Signal intensities for the three classes of SNPs in *Xba*I-only compared with *Xba*I/*Hpa*II representations averaged across all four AML patients

SNP class	No. SNPs	Av. signal (*Xba*I)	Median (*Xba*I)	Av. signal (*Xba*I + *Hpa*II)	Median (*Xba*I + *Hpa*II)	ΔAv. signal (%)	% SNPs with signal reduced >30% by *Hpa*II
Before decitabine							
1	7,379	2,235	1,968	2,186	1,971	−2.18	4.08
2	1,525	2,509	2,226	2,345	2,073	−6.54	10.61
3	457	2,475	2,271	2,350	2,101	−5.04	7.82
After decitabine							
1	7,379	2,241	1,912	2,207	1,924	−1.48	4.28
2	1,525	2,487	2,203	2,171	1,907	−12.72	17.16
3	457	2,469	2,218	2,260	2,016	−8.46	12.69

NOTE: Data are from 10K SNP chips. To increase reliability, only SNPs with signal >400 in the *Xba*I-only genomic representations of all four cases were considered.

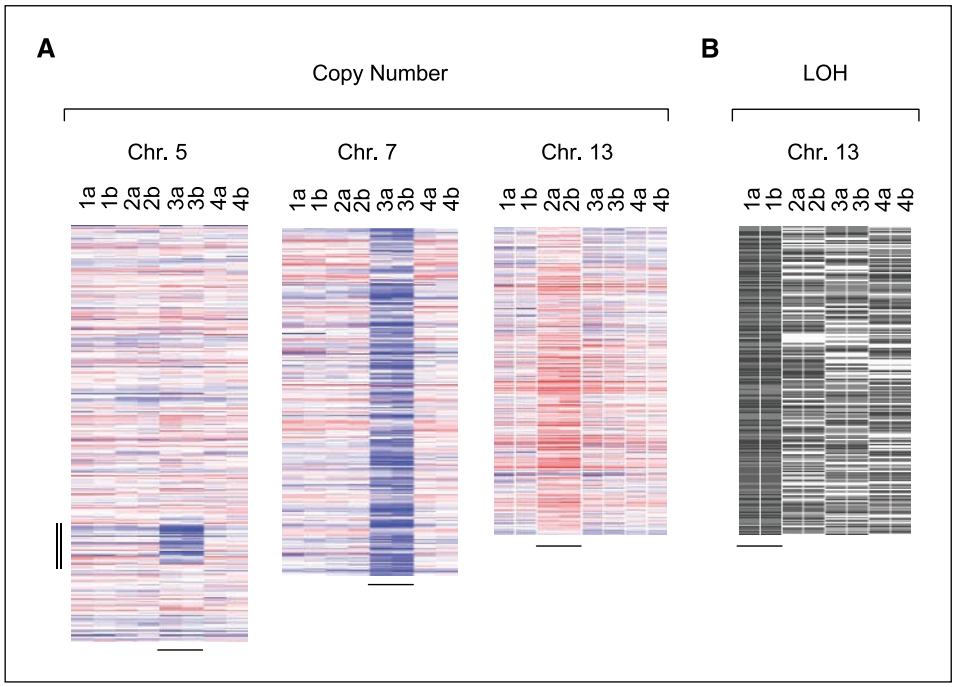

Figure 2. SNP chip data showing chromosomal and subchromosomal DNA gains and losses and LOH in AML bone marrow aspirates. *A, Xba*I-only genomic representations. *Y axis,* SNPs ordered by genomic position (p-tel to q-tel); *X axis,* four AML cases. In these representations, regions of white correspond to SNPs with copy number equal to 2 based on analysis with DNA from a panel of kidney controls known to be diploid from previous cytogenetic and SNP chip data (16). As in Fig. 1, *red,* increase in copy number to a maximum of 5; *blue,* decrease in copy number to a minimum value of 0. The interstitial 5q loss in AML case 3 spans a region of chromosome band 5q23-q33 of ~23 Mb from rs1898550 to rs1366236. In addition to the gain of chromosome 13, gains of chromosomes 8, 10, and 21 were also detected in AML case 2. *B,* AML case 1 does not show DNA copy number alterations, but this leukemia shows LOH for chromosome 13 as indicated by absence of heterozygosity (*gray,* AA SNP genotype; *black,* BB; *white,* AB); heterozygous SNPs are not seen in AML case 1 but are present at a high frequency, randomly distributed along the chromosome, in the other three cases. In each diagram, samples labeled *a* and *b* are pre-decitabine and post-decitabine, respectively. These genetic data indicate that the malignant cell clones persisted in at least three of the four cases (AML cases 1-3) during the 8-day course of decitabine treatment.

Affymetrix 10K SNP chip were met in data from the four patients with AML before and after decitabine. First considering the pretreatment bone marrow samples and comparing the results for the *Xba*I-only and *Xba*I/*Hpa*II genomic representations, the data in Table 2 indicate that the class 2 SNPs in fact lost signal intensity to a greater extent than the class 1 SNPs. The results, averaged across all four patients, showed that 6% of the class 2 SNPs lost a substantial amount of their signal intensity in the *Xba*I/*Hpa*II representations compared with the *XbaI*-only representations before decitabine, whereas a somewhat greater number (10%) of these SNPs lost signal in this comparison after decitabine. These findings suggested that decitabine did have a net demethylating effect even when the results were averaged over all four patients. This observation is compatible with the prior literature reporting partial demethylation of certain gene sequences (notably repetitive elements) after decitabine, with variable responses

Table 3. Normalized signal intensities for the class 2 SNPs in each AML patient

Patient	Av. signal (*Xba*I)	Av. signal (*Xba*I + *Hpa*II)	ΔAv. signal (%)	% SNPs with signal reduced >30% by *Hpa*II
Before decitabine				
1	2,526	2,345	−7.16	10.62
2	2,514	2,363	−5.99	10.62
3	2,510	2,360	−5.97	9.84
4	2,487	2,312	−7.03	11.34
After decitabine				
1	2,507	1,976	−21.18	27.54
2	2,492	2,297	−7.79	10.30
3	2,466	2,098	−14.93	17.97
4	2,484	2,312	−6.91	12.85

NOTE: Data are from 10K SNP chips. For pre-decitabine, the class 2 SNP intensities were normalized by setting the mean intensity of the class 1 SNPs in each case equal across the experiment.

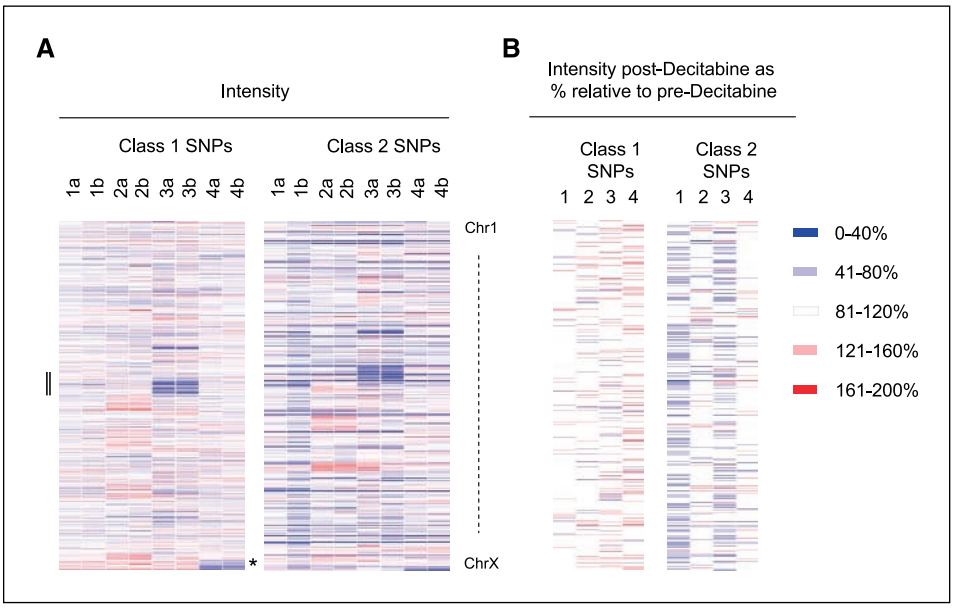

Figure 3. MSNP shows the demethylating response to decitabine in bone marrow aspirates from patients with AML. *A,* color-coded MSNP intensity values from the *Xba*I/*Hpa*II genomic representations, after normalization as described in the text, for AML cases 1 to 4, with *a* for pre-decitabine and *b* for post-decitabine bone marrow aspirates. The SNPs are arranged by physical position in the genome from chromosome 1 to the X chromosome. As a useful internal validation, the reduced copy number of the X chromosome in patient 4 (*asterisk*) is consistent with the fact that this patient is a male, whereas the other three patients are females. As expected, the control class 1 SNPs do not show a major alteration in signal in any patient as a function of decitabine treatment (*white*), whereas the class 2 SNPs show partial but widespread demethylation (*blue*) most prominent in patients 1 and 3. The genetic alterations in the AML cells, for example, the hemizygous losses of chromosomes 5 and 7 in AML case 3 (*double line*), are seen superimposed on the methylation data. *B,* display of the post-decitabine SNP intensity as a percentage of the pre-decitabine value, emphasizing the strong demethylating responses in AML cases 1 and 3. For clarity in this display, the values have been binned as indicated in the scale (*right*).

between patients. Accordingly, when we broke the data down by individual patients, the results became more striking: as shown in Table 3, before decitabine, all four patients showed similar DNA methylation by MSNP, whereas after decitabine treatment the AML bone marrow from patient 1 showed more than a third of all class 2 SNPs with an intensity decrease of >30% in the *Xba*I/*Hpa*II compared with *Xba*I-only genomic representation, suggesting a strong demethylating response to the drug in this patient.

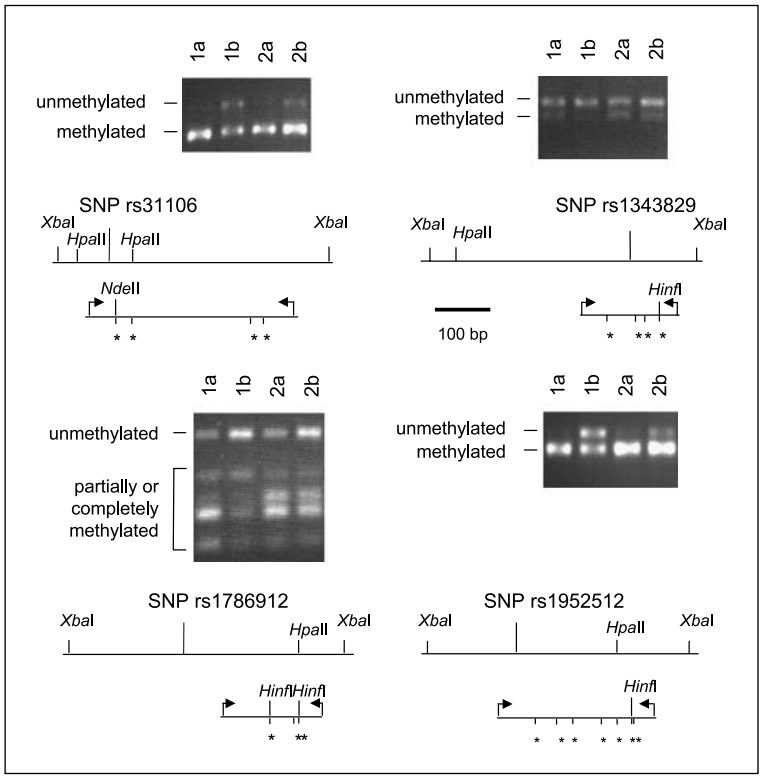

Figure 4. Validation of the MSNP data by COBRA at four SNP loci. Restriction maps of the *Xba*I fragments containing each SNP, with the COBRA strategies indicated below each map. *Arrows,* PCR primers for COBRA; *asterisks,* CpG dinucleotides present in the unconverted DNA sequence. Lanes labeled *a* are pre-decitabine bone marrow aspirates and *b* are post-decitabine bone marrow aspirates. COBRA results confirm the greater demethylating response to decitabine in AML case 1 versus AML case 2 (compare ratios of methylated and unmethylated bands in each lane). SNP rs31106 is in the intergenic region between *IRX6* and *MMP2* genes in chromosome band 16p12.2; SNP rs1786912 is in an intron of the *CTNNA3* gene in chromosome band 10q21; and SNP rs1343829 is in an intron of the *EXTL1* gene in chromosome band 1p36.

Also evident from Table 3, patient 3 showed some demethylation, whereas patients 2 and 4 showed little demethylation by this simple numerical analysis.

Having seen that the strategy would work in principle, we next asked whether the demethylating response to decitabine could be visualized in the MSNP data. To increase reliability, we first eliminated the 10% of SNPs that gave the weakest signals in the *Xba*I-only representations, most which were associated with Affymetrix absence calls in the genotype data (see Materials and Methods). Next, we normalized the SNP intensity values from the *Xba*I/*Hpa*II representations such that the average signal for the class 1 SNPs was equal for all samples. We then calculated percent differences in signal intensity post-decitabine versus pre-decitabine for each SNP. These procedures yielded numerical data sets formally similar to those from microarray-based expression studies and therefore amenable to the same types of visual displays. To identify AML cases that responded to decitabine with significant DNA demethylation, we displayed the normalized SNP intensities for each sample before and after decitabine (Fig. 3*A*). We also displayed the response to decitabine in each case at each SNP locus by expressing the post-decitabine SNP intensity as a percentage of the pre-decitabine intensity (Fig. 3*B*). In our color-coding scheme, blue represents low signals (Fig. 3*A*) or loss of DNA methylation (Fig. 3*B*), white represents intermediate signals (Fig. 3*A*) or no change in methylation (Fig. 3*B*), and red represents high signals (Fig. 3*A*) or GOM (Fig. 3*B*). We displayed the data according to the genomic positions of the SNPs. As expected from the numerical data (Tables 1 and 2), this display showed no systematic decrease in the intensities of the class 1 SNPs after decitabine in any of the AML cases. In contrast, two of the AML cases (1 and 3) showed clear decreases in the intensity of many of the class 2 SNPs after decitabine. These demethylating responses were evident for case 1 in the simple display of the normalized SNP intensities before and after treatment (Fig. 3*A*) and were highlighted in this case and revealed more clearly in case 3 in the display of the post-decitabine intensity as percentage of the pre-decitabine intensity (Fig. 3*B*). In contrast, the two other cases (2 and 4) showed much less demethylation (Fig. 3*A* and *B*).

A potential complicating factor in interpreting DNA methylation data during decitabine or other anticancer therapies is replacement of neoplastic cells by normal cells. MSNP gives both genetic and epigenetic information; in our series of AML cases, the genetic data showing chromosomal or subchromosomal gains, losses, or LOH indicated persistence of the neoplastic clones in the bone marrow of patients 1, to 3. The SNP chips did not reveal LOH or DNA copy number changes in patient 4, although standard histopathology indicated persistence of abnormal cells in the bone marrow of all four patients. Thus, MSNP can provide data to increase confidence that any demethylation being observed is taking place in the cancer cells and is not an artifact of replacement by a different cell population.

Validation of the SNP chip data by COBRA and bisulfite sequencing. We next wished to test the demethylation implied by the MSNP data using an independent assay of CpG methylation. We picked five SNP loci for validation and employed a standard method of bisulfite conversion of the genomic DNA followed by PCR with locus-specific primers matching the bisulfite-converted sequences followed by restriction digestion to query the methylation status of internal CpG sites (COBRA). We assessed CpG methylation in the DNA adjacent to these five SNPs in AML case 1, indicated by MSNP to have a strong demethylating response to decitabine, and AML case 2, indicated by MSNP to have a weak response. The COBRA results were consistent with the MSNP data, with comparison of the methylated and unmethylated band intensities confirming more extensive demethylation in case 1 compared with case 2 at all five SNP loci (Figs. 4 and 5). We further tested one of these SNPs, rs963172, by sequencing sets of cloned bisulfite PCR products. This procedure queried the methylation status of four CpG dinucleotides in the immediate vicinity of this SNP. We cloned and sequenced multiple bisulfite/PCR products from the pre-decitabine and post-decitabine bone marrow aspirates from these two patients. As shown by the diagrams of the bisulfite sequencing data (Fig. 5), the inferences from MSNP were validated: substantial demethylation post-decitabine was seen in AML case 1, with less demethylation seen in AML case 2.

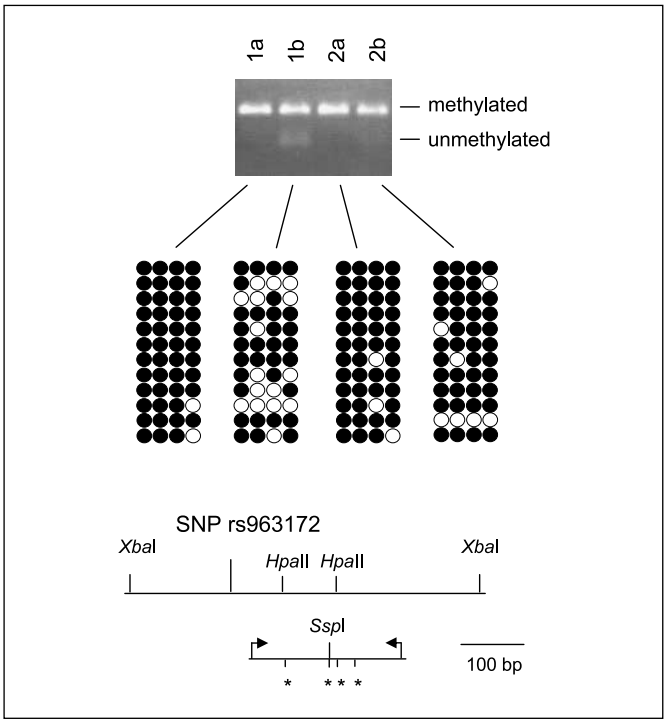

Figure 5. Validation of the MSNP data by COBRA and bisulfite sequencing at a fifth SNP locus. *Hpa*II sites in the *Xba*I fragment containing SNP rs963172 were reported in the MSNP data as significantly demethylated after decitabine treatment in AML case 1 (40% reduction in signal) but not in AML case 2 (14% decrease in signal). This difference in response is confirmed both by COBRA and by sequencing of multiple clones from bisulfite conversion/PCR. There is less overall demethylation in AML case 2, although one clone in this case is demethylated at all four CpG sites. Thus, a small percentage of the malignant cells in patient 2 and a larger percentage of cells in patient 1 have responded to the drug by DNA demethylation. SNP rs963172 is in an intron of the *PCSK6* gene in chromosome band 15q26. The COBRA strategy is indicated below the restriction map. *Arrows,* PCR primers for COBRA; *asterisks,* CpG dinucleotides present in the unconverted DNA sequence. Lanes labeled *a* are pre-decitabine and *b* are post-decitabine.

Validation of MSNP by Southern blotting in a tumor/normal comparison. All of the validations above relied on PCR, and we also wished to validate the MSNP method by a non-PCR-based assay. Southern blotting with methylation-sensitive restriction enzymes is useful for this purpose, but it requires more genomic DNA than was available from the bone marrow aspirates. Accordingly, our second proof-of-principle experiment involved a cancer/normal comparison, with a series of four Wilms' tumor/

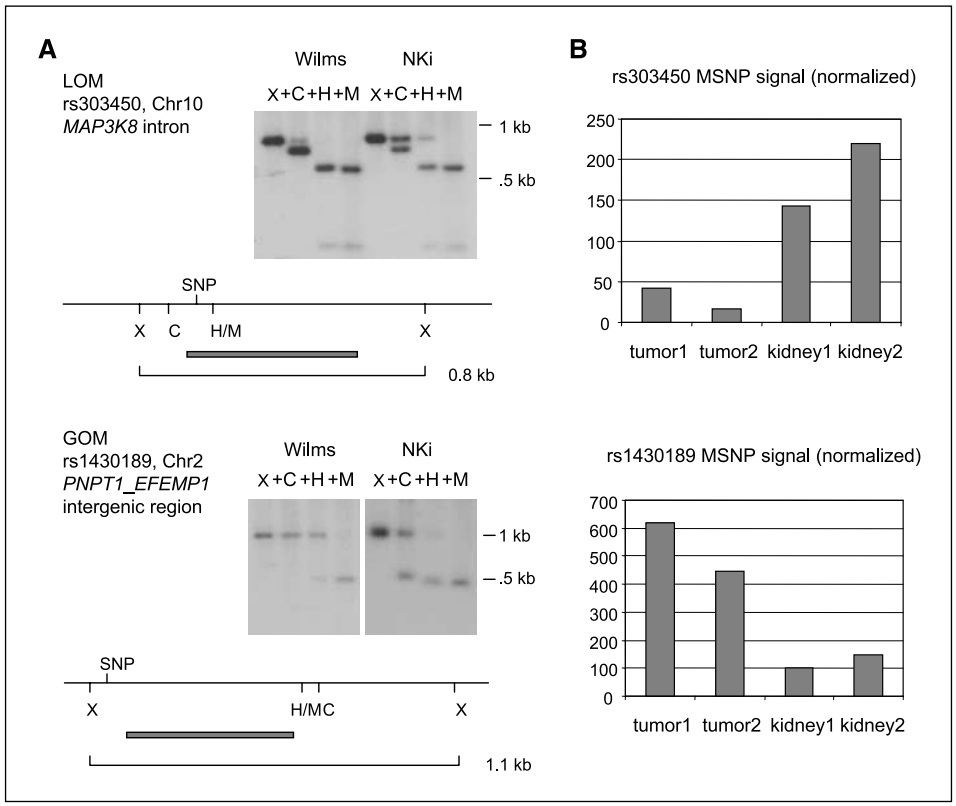

Figure 6. Gains and losses of DNA methylation identified by MSNP and validated by Southern blotting in a Wilms' tumor/normal kidney comparison. Two regions differentially methylated based on the 10K SNP chip readout were validated by digesting genomic DNA from Wilms' tumors and adjacent normal kidney with methylation-sensitive restriction enzymes and hybridizing the Southern blots with the indicated probes. SNP ID numbers, chromosomal location, and gene contexts. *X, Xba*I; *C, Cfo*I; *H, Hpa*II, *M, Msp*I. GOM is validated by increased intensities and LOM by decreased intensities of the uncut (*top*) bands in the Wilms' tumor lanes (*+C*, *+H*).

normal kidney pairs. These tumors, and adjacent nonneoplastic kidney parenchyma, yield ample material for Southern analysis. Using MSNP, again with the 10K Affymetrix SNP chips, we searched for class 2 SNP loci (*Xba*I fragments) with recurrent alterations in DNA methylation, requiring that at least two of the four tumor/normal pairs showed the GOM or LOM for that SNP. SNPs were chosen by calculating percent difference in normalized class 2 SNP *Xba*I/*Hpa*II signals between tumor and matched normal kidney DNA. By defining strong GOM as at least a 50% increase in the signal for both tumors, we identified a group of 18 SNPs, including rs1430189. Similarly, we found 52 LOM SNPs, including rs303450, in which signal was decreased by at least 50% in both tumors relative to their matched normal kidneys. We carried out Southern blotting of tumor and normal genomic DNAs using methylation-sensitive restriction enzymes and hybridized the blots with probes for these two SNP loci. As shown in Fig. 6, the predicted reduction in high molecular weight bands (LOM) or increase in high molecular weight bands (GOM) in Wilms' tumor compared with normal kidney control were clearly seen with the methylation-sensitive enzymes.

Discussion

These results introduce and validate a useful method for concurrent genetic and epigenetic analysis for clinical and basic studies in cancer biology. MSNP has many potential applications, but in this pilot study we have emphasized its use in characterizing and monitoring the biochemical response to a promising "epigenetic drug," decitabine. Several interesting questions relating to epigenetic therapies will likely be amenable to future investigation by MSNP. For example, what is the correlation between decitabine dose and DNA methylation? Because this drug has multiple effects on the cell, including cytotoxicity at high doses, a direct correlation would not necessarily be expected; in fact, in the small series of patients analyzed here by MSNP, the strongest demethylating response occurred in the patient (AML case 1) who received the lowest dose. This observation may reflect the necessity for cell proliferation for genomic demethylation in response to decitabine, a consequence of the fact that this drug works by inactivating methyltransferases, thereby promoting passive demethylation during each S phase. In this report, we only assessed the demethylating response early in treatment, and given this drug mechanism, using MSNP to follow the biochemical response during longer courses of treatment will be of interest in future studies. Second, are there distinct sets of loci (*Xba*I genomic fragments in the MSNP method) that are highly sensitivity versus poorly sensitive to demethylation in response to decitabine? Results of the feasibility study reported here suggest that this may be true, and if this is confirmed in larger series of cases, it might be possible to design a small panel of marker loci, encompassing representative high and low sensitivity sites, to follow the extent of the biochemical response to this drug during treatment. Third, does the extent of demethylation genome-wide or at specific types of loci (highly sensitive or poorly sensitive to demethylation) correlate with the clinical response to decitabine as measured by patient survival? Applying MSNP to a large number of AML and myelodysplastic syndrome cases and normal bone marrow can potentially answer these questions.

Among many other potential uses, the MSNP method is also applicable to uncovering tumor-associated changes in DNA methylation, which we have illustrated here in a Wilms' tumor/normal comparison. We have studied Wilms' tumor as a paradigm for the combined role of genetic and epigenetic aberrations during tumor initiation and progression most recently in a genomic

profiling study using 10K resolution SNP chips for genetic analysis (15). MSNP can extend these types of studies to cancer epigenetics, potentially achieving a simultaneous genetic and epigenetic analysis with genomic coverage. Other uses, for example, in profiling DNA methylation in the development of normal tissues and in uncovering novel imprinted loci or other loci with allele-specific DNA methylation (which would be detected as conversion from an AB genotype call to AA or BB in the *Xba*I/*Hpa*II genomic representations), are equally obvious. The methodology in this report will likely be improved and streamlined over time probably with input from Affymetrix and other companies. Currently, our protocol requires 600 ng genomic DNA and is therefore applicable to small clinical samples ($\sim 10^5$ cells). A valuable future objective will be to miniaturize the protocol, so that even smaller numbers of cells can be analyzed.

Lastly, MSNP examines preferentially non-CpG island and non-repetitive sequences, and it therefore offers a unique window on a large portion of the epigenome that arguably has received insufficient attention. The *Xba*I restriction fragments queried by the SNP chips were chosen by Affymetrix to be roughly equally spaced along the human chromosomes, to have a size in the 0.5- to 2-kb range, and to contain SNPs with high frequencies of heterozygosity. Accordingly, <1% of these SNPs, on both 10K and 50K chips, are in CpG islands. CpG islands have been heavily scrutinized, nearly to the exclusion of the rest of the genome, in previous studies of DNA methylation in cancer cells, for the legitimate reason that they overlap gene promoters. However, loss of DNA methylation genome-wide in non-CpG island sequences may also contribute to genomic instability (2, 16). In addition, intragenic and intergenic regions other than CpG islands contain abundant evolutionarily conserved sequence blocks, and MSNP in conjunction with gene expression profiling has the potential to reveal methylation-dependent regulatory functions of these non-CpG island sequences.

Addendum

A related article describing the use of Affymetrix filing arrays for methylation profiling of specific genomic regions appeared while this work was under review (17).

Acknowledgments

Received 10/17/2005; revised 1/8/2006; accepted 1/23/2006.

Grant support: This work was supported in part R01CA 102270 from the NIH.

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.

References

1. Baylin SB, Herman JG, Graff JR, Vertino PM, Issa JP. Alterations in DNA methylation: a fundamental aspect of neoplasia. Adv Cancer Res 1998;72:141–96.
2. Feinberg AP, Tycko B. The history of cancer epigenetics. Nat Rev Cancer 2004;4:143–53.
3. Egger G, Liang G, Aparicio A, Jones PA. Epigenetics in human disease and prospects for epigenetic therapy. Nature 2004;429:457–63.
4. van den Bosch J, Lubbert M, Verhoef G, Wijermans PW. The effects of 5-aza-2′-deoxycytidine (decitabine) on the platelet count in patients with intermediate and high-risk myelodysplastic syndromes. Leuk Res 2004;28:785–90.
5. Wijermans P, Lubbert M, Verhoef G, et al. Low-dose 5-aza-2′-deoxycytidine, a DNA hypomethylating agent, for the treatment of high-risk myelodysplastic syndrome: a multicenter phase II study in elderly patients. J Clin Oncol 2000;18:956–62.
6. Issa JP, Garcia-Manero G, Giles FJ, et al. Phase 1 study of low-dose prolonged exposure schedules of the hypomethylating agent 5-aza-2′-deoxycytidine (decitabine) in hematopoietic malignancies. Blood 2004;103:1635–40.
7. Issa JP, Gharibyan V, Cortes J, et al. Phase II study of low-dose decitabine in patients with chronic myelogenous leukemia resistant to imatinib mesylate. J Clin Oncol 2005;23:3948–56.
8. Kantarjian HM, O'Brien S, Cortes J, et al. Results of decitabine (5-aza-2′deoxycytidine) therapy in 130 patients with chronic myelogenous leukemia. Cancer 2003;98:522–8.
9. Cheng JC, Yoo CB, Weisenberger DJ, et al. Preferential response of cancer cells to zebularine. Cancer Cell 2004;6:151–8.
10. Insinga A, Monestiroli S, Ronzoni S, et al. Inhibitors of histone deacetylases induce tumor-selective apoptosis through activation of the death receptor pathway. Nat Med 2005;11:71–6.
11. Rush LJ, Raval A, Funchain P, et al. Epigenetic profiling in chronic lymphocytic leukemia reveals novel methylation targets. Cancer Res 2004;64:2424–33.
12. Laird PW. Cancer epigenetics. Hum Mol Genet 2005;14:R65–76.
13. Ching TT, Maunakea AK, Jun P, et al. Epigenome analyses using BAC microarrays identify evolutionary conservation of tissue-specific methylation of SHANK3. Nat Genet 2005;37:645–51.
14. Li C, Wong WH. Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. Proc Natl Acad Sci U S A 2001;98:31–6.
15. Yuan E, Li CM, Yamashiro DJ, et al. Genomic profiling maps loss of heterozygosity and defines the timing and stage dependence of epigenetic and genetic events in Wilms' tumors. Mol Cancer Res 2005;3:493–502.
16. Yamada Y, Jackson-Grusby L, Linhart H, et al. Opposing effects of DNA hypomethylation on intestinal and liver carcinogenesis. Proc Natl Acad Sci U S A 2005;102:13580–5.
17. Schumacher A, Kapranov P, Kaminsky Z, et al. Microarray-based DNA methylation profiling: technology and applications. Nucleic Acids Res 2006;34:528–42.
18. Zhao X, Li C, Paez JG, et al. An integrated view of copy number and allelic alterations in the cancer genome using single nucleotide polymorphism arrays. Cancer Res 2004;64:3060–71.

A Single Nucleotide Polymorphism Chip-Based Method for Combined Genetic and Epigenetic Profiling: Validation in Decitabine Therapy and Tumor/Normal Comparisons

Eric Yuan, Fatemeh Haghighi, Susan White, et al.

Cancer Res 2006;66:3443-3451.