

# Both Hypomethylation and Hypermethylation in a 0.2-kb Region of a DNA Repeat in Cancer

Rie Nishiyama,<sup>1</sup> Lixin Qi,<sup>1</sup> Michelle Lacey,<sup>2</sup> and Melanie Ehrlich<sup>1</sup>

<sup>1</sup>Human Genetics Program, Department of Biochemistry, and Tulane Cancer Center, Tulane Medical School and <sup>2</sup>Department of Mathematics, Tulane University, New Orleans, Louisiana

## Abstract

***NBL2* is a tandem 1.4-kb DNA repeat, whose hypomethylation in hepatocellular carcinomas was shown previously to be an independent predictor of disease progression. Here, we examined methylation of all cytosine residues in a 0.2-kb subregion of *NBL2* in ovarian carcinomas, Wilms' tumors, and diverse control tissues by hairpin-bisulfite PCR. This new genomic sequencing method detects 5-methylcytosine on covalently linked complementary strands of a DNA fragment. All DNA clones from normal somatic tissues displayed symmetrical methylation at seven CpG positions and no methylation or only hemimethylation at two others. Unexpectedly, 56% of cancer DNA clones had decreased methylation at some normally methylated CpG sites as well as increased methylation at one or both of the normally unmethylated sites. All 146 DNA clones from 10 cancers could be distinguished from all 91 somatic control clones by assessing methylation changes at three of these CpG sites. The special involvement of DNA methyltransferase 3B in *NBL2* methylation was indicated by analysis of cells from immunodeficiency, centromeric region instability, and facial anomalies syndrome patients who have mutations in the gene encoding DNA methyltransferase 3B. Blot hybridization of 33 cancer DNAs digested with CpG methylation-sensitive enzymes confirmed that *NBL2* arrays are unusually susceptible to cancer-linked hypermethylation and hypomethylation, consistent with our novel genomic sequencing findings. The combined Southern blot and genomic sequencing data indicate that some of the cancer-linked alterations in CpG methylation are occurring with considerable sequence specificity. *NBL2* is an attractive candidate for an epigenetic cancer marker and for elucidating the nature of epigenetic changes in cancer. (Mol Cancer Res 2005;3(11):617–26)**

## Introduction

Both hypermethylation (1, 2) and hypomethylation (3) of DNA have been observed in most tested cancers but in different sequences (4-7). Many specific gene regions become hypermethylated, and some other gene regions and many noncoding DNA repeats become hypomethylated during carcinogenesis (8-10). Nonetheless, hypomethylation and hypermethylation in different parts of the genome in various cancers have been found not to be significantly associated with each other (4, 6, 11). Therefore, cancer-linked DNA hypomethylation is not simply a response to cancer-linked hypermethylation or vice versa.

Here, we show for the first time that, relative to normal somatic tissues, cancers can display both hypomethylation and hypermethylation within the same small region on the same DNA molecule. We analyzed *NBL2*, a tandem 1.4-kb repeat with a complex sequence. We found recently by Southern blot analysis that *NBL2* exhibits either predominant hypermethylation or hypomethylation at *HhaI* sites in 89% of 18 studied ovarian carcinomas and 84% of 51 Wilms' tumors (12). We also observed hypomethylation at *NotI* sites in some ovarian carcinomas. Itano et al. (13) showed that hypomethylation at *NotI* sites was an independent prognostic indicator in hepatocellular carcinoma patients. *NBL2* is often hypomethylated at *NotI* sites in neuroblastomas (hence, the name *NBL2*) and hepatocellular carcinomas (14, 15). This primate-specific sequence (15) is CpG rich (61% C + G; 5.7% CpG). It is present in ~200 to 400 copies per haploid human genome, mostly in the vicinity of the centromeres of four of the five acrocentric chromosomes (12), repeat-rich regions for which only little sequence information is available.

The powerful method that we used to analyze *NBL2* methylation in ovarian carcinomas, Wilms' tumors, and various control cell populations is the hairpin-bisulfite PCR variant of genomic sequencing (hairpin genomic sequencing) developed by Laird et al. (16). In bisulfite-based genomic sequencing, bisulfite causes deamination of unmethylated C residues, whereas methylated C residues (usually only in CpG sequences) are resistant (17). Hairpin genomic sequencing allows analysis of methylation of every CpG dinucleotide pair (dyad) in a given region on covalently linked DNA strands of a restriction fragment. Analysis of both CpGs of a dyad is of special interest because CpG dyads are generally assumed to be symmetrically methylated or symmetrically unmethylated, except during DNA replication. Hairpin genomic sequencing also unambiguously differentiates naturally occurring sequence variation from bisulfite- and PCR-mediated C-to-T conversions at unmethylated cytosines. Our results show that *NBL2*, which

Received 8/22/05; revised 9/28/05; accepted 10/12/05.

**Grant support:** NIH grant CA 81506 and Louisiana Cancer Research Consortium grant.

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.

**Requests for reprints:** Melanie Ehrlich, Human Genetics Program SL31, Tulane Medical School, 1430 Tulane Avenue, New Orleans, LA 70112. Phone: 504-584-2449; Fax: 504-584-1763. E-mail: ehrlich@tulane.edu

Copyright © 2005 American Association for Cancer Research.

doi:10.1158/1541-7786.MCR-05-0146

does not seem to be a gene (12), is especially sensitive to multiple diverse DNA methylation changes during oncogenesis, even within the same 0.2-kb region on the same DNA molecule.

## Results

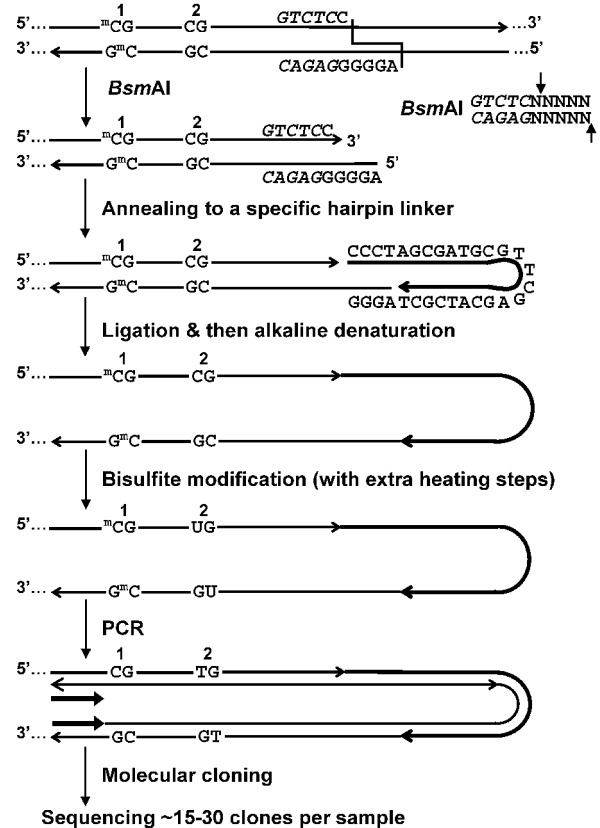
### Hairpin-Bisulfite Genomic Sequencing Strategy and Validation

We analyzed the 1.4-kb *NBL2* repeat by hairpin genomic sequencing (16) as shown in Fig. 1. In DNA clones resulting from bisulfite treatment and PCR, a genomic m5CpG, the predominant site of vertebrate DNA methylation, will appear as CpG because it escaped bisulfite deamination, and an unmethylated CpG will become TpG due to cytosine deamination followed by amplification. In hairpin genomic sequencing, strand ligation results in the sequence information from both genomic strands of a DNA fragment being present in each strand of the resulting DNA clone (Fig. 1A). Corresponding CpG positions in the two halves of one strand of a DNA clone are compared to determine the methylation status of CpG dyads in the template DNA molecule. For simplicity, we will use the following terms for the DNA clones, which describe the CpG dyad methylation status of the molecule that gave rise to the clone (Fig. 1B): M/M (symmetrically methylated), U/U (symmetrically unmethylated), M/U (hemimethylated with the lower strand unmethylated), and U/M (hemimethylated with the upper strand unmethylated). Not only does hairpin genomic sequencing resolve a symmetrical methylation pattern at a CpG dyad from hemimethylation, but also it allows an unmethylated CpG to be unambiguously distinguished from germ-line C-to-T changes (Fig. 1B). This is especially useful for DNA repeats because of their appreciable sequence variation (16).

From hairpin genomic sequencing of each normal tissue or cancer DNA, we obtained a single or predominant PCR band of the expected size (508 bp) from which 12 to 32 clones were generated and sequenced (Fig. 2). Given the originally self-complementary nature of the ligated DNA for bisulfite treatment and the specificity of bisulfite for denatured DNA, various controls were done to make sure that hairpin genomic sequencing did not yield artifacts. First, we should see essentially only CpG methylation (16) because we were examining postnatal tissues (18). As expected, we found only 0% to 0.3% of non-CpG C residues per tissue sample (0.1% overall; data not shown). In addition, only 0.6% of C residues persisted as C in hairpin genomic sequencing clones from a *NBL2*-containing *Escherichia coli* plasmid (Fig. 2B). In addition, we checked the completeness of bisulfite modification by digesting all hairpin PCR products with *Tsp509I* (recognizing AATT). Gel electrophoresis of the digests indicated complete digestion due to bisulfite-mediated C deamination at genomic AACC or AACT in *NBL2* (data not shown). Hairpin genomic sequencing of a *NBL2* plasmid methylated *in vitro* at most CpG sites by *M.SssI* showed that most of the CpG C residues were retained in the clones.

In addition, we also verified that we were not amplifying only a few template molecules by using a 1:20 dilution of the bisulfite-treated DNA instead of the undiluted sample for PCR.

### A Hairpin-bisulfite PCR sequencing scheme



### B Hairpin-bisulfite PCR sequencing products

Symmetrically methylated (1) and unmethylated (2) CpG's	C-to-T sequence variation (3) and hemimethylation (4)
5' - A <u>C</u> G . . . <u>T</u> G T A A - -	5' - A <u>C</u> G . . . <u>C</u> G T A A - -
3' - T <u>G</u> C . . . <u>G</u> T A T T - -	3' - T <u>A</u> T . . . <u>G</u> T A T T - -
1                      2	3                      4
M/M                      U/U	M/U

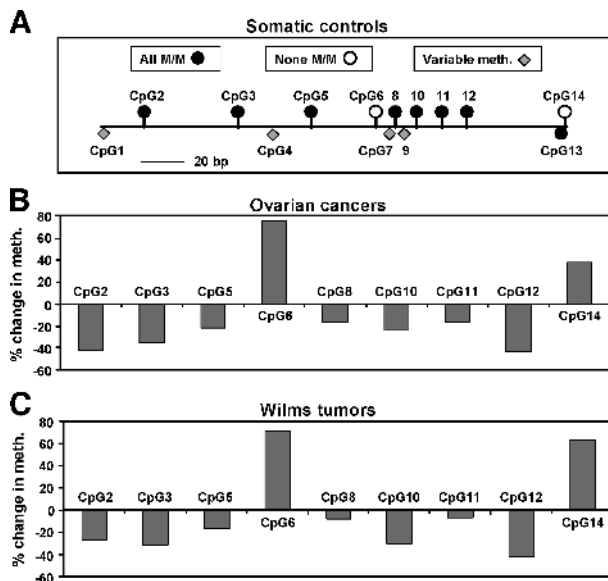
**FIGURE 1.** Hairpin genomic sequencing methodology. **A.** Hairpin genomic sequencing of the *NBL2* repeat with the covalently linked upper and lower strands (not to scale) diagrammed as a hairpin to illustrate their complementarity before bisulfite deamination of all unmethylated C residues. *mC*, 5-methylcytosine. The recognition site for *BsmAI* is in *italics* and its cleavage specificity is shown on the right. **B.** Examples of discrimination between different methylated configurations of CpG dyads and C-to-T changes in a genomic sequence. Part of one strand of each molecular clone is depicted in the hairpin configuration to align CpG positions that were part of a genomic dyad.

A strong PCR product band was obtained with or without dilution from each sample. We conclude that the sequenced molecular clones represent the heterogeneity in the sample DNA, which is consistent with their epigenetic and genetic sequence diversity (Fig. 2). Lastly, we did a mixing experiment with hairpin PCR products amplified as 1:0, 1:3, 1:1, 3:1, and 0:1 mixtures of *M.SssI*-methylated and unmethylated *NBL2* plasmid. We obtained approximately the expected ratios of *Bst*UI-sensitive (CGCG) to *Bst*UI-resistant sites in the PCR product mixtures. Therefore, there was no



*NBL2* DNA clones from somatic controls subject to hairpin genomic sequencing, 7 of the 14 CpG sites were always symmetrically methylated (CpG2, CpG3, CpG5, CpG8, CpG10, CpG11, and CpG12). Two nonadjacent CpGs were never symmetrically methylated (CpG6 and CpG14). One of these, CpG14, was always U/U, and the other, CpG6, was usually U/U but occasionally U/M or M/U. CpG13, which is exactly adjacent to always-unmethylated CpG14 was often replaced by GpG and hence could not be methylated. However, whenever it was not replaced, it was always M/M despite its immediate U/U neighbor (Figs. 2 and 3A). Normal sperm had no M/M sites in this subregion (Fig. 2B), consistent with previous results from various tandem DNA repeats (9).

None of the 146 DNA clones from the 10 cancers had the conserved methylation pattern of normal somatic controls (Fig. 2). Moreover, 56% of the cancer clones had a mixture of both hypomethylated and hypermethylated CpG sites. These methylation changes were defined by the loss of the normally conserved M/M status at CpG2, CpG3, CpG5, CpG8, CpG10, CpG11, or CpG12 and the gain of M/M status at CpG6 or CpG14, sites never normally symmetrically methylated (Figs. 2 and 3). The overall methylation status at each of these nine CpG sites in the cancers was significantly different from that in the somatic controls ( $P < 0.005$ , adjusted for multiple comparisons).



**FIGURE 3.** Comparison of methylation in somatic control tissues (brain, spleen, and lung), ovarian carcinomas, and Wilms' tumors. **A.** Scale of the positions of the CpG sites in the hairpin genomic sequencing region and their methylation status in the somatic controls. The seven CpGs that were always M/M and the two CpGs that were either always U/U (CpG14) or usually U/U and occasionally hemimethylated (CpG6) are above the horizontal line. The variably methylated CpGs are diamonds below the horizontal line. The filled circle below the horizontal line represents CpG13, which was always methylated when present but often not present due to germ-line sequence variation. **B** and **C.** Overall change in methylation in five ovarian carcinomas and five Wilms' tumors at CpGs that were either always M/M or never M/M in somatic controls. The % change in methylation at CpG2, CpG3, CpG5, CpG8, CpG10, CpG11, and CpG12 is shown as the negative of the percentage of cancer clones with hypomethylation (loss of M/M status) at that position; for CpG6 and CpG14, it is the percentage of cancer clones with hypermethylation (gain of M/M status).

### Contribution of Spreading of Methylation or Demethylation to DNA Methylation Changes in Cancer

Spreading of *de novo* methylation along a DNA region has been observed (24-27). We did not see evidence of predominant spreading of *de novo* methylation or demethylation because a pairwise comparison of neighboring CpG sites in *NBL2* in the cancers indicated that there was no statistically significant bias toward adjacent sites having the same methylation status. Furthermore, at CpG6 and CpG8, which are separated by only 6 bp, there were seven clones from four cancers (OvCaN and WT4, WT9, WT21, and WT67) that exhibited opposite methylation changes [i.e., hypermethylation at CpG6 (M/M) and hypomethylation at CpG8 (M/U); Fig. 2]. Overall, the methylation changes in many of the clones suggest multiple discontinuous hits of demethylation and *de novo* methylation within a 0.2-kb region during carcinogenesis.

Nonetheless, there were some DNA clones whose methylation patterns suggested spreading of methylation or demethylation. Some had all 14 CpG dyads unmethylated or all methylated (Fig. 2). Others had the first five or six CpG sites unmethylated on at least one strand. The frequency of clones with no methylation in the first five CpG sites was significantly higher than expected if the methylation at each site was independent, as was the combined frequency of fully methylated or fully unmethylated clones (both  $P$ s  $< 0.0001$ ). In summary, there seems to be spreading of altered DNA methylation patterns in some, but not most, of the copies of *NBL2* in the examined cancers.

### Hemimethylation in the Cancers

In the examined *NBL2* subregion in the somatic controls, 1.6% of the CpG sites and 15% of the somatic control clones displayed hemimethylation. The hemimethylation frequencies rose in the cancers to 3.4% of the CpG sites in 47% of the ovarian cancer clones and 6.6% of the CpG sites in 71% of the Wilms' tumor clones. Because incomplete bisulfite modification was observed at only  $\sim 0.1\%$  of the non-CpG C residues, we are truly describing hemimethylation at CpGs. There was also a change in the distribution of hemimethylation in the cancer DNAs. Many more of the 14 CpG positions were subject to hemimethylation in the cancers than in the somatic controls. In addition, some of the hemimethylated CpGs at a given position in the cancer clones displayed a strong bias for demethylation of the top or the bottom strand (Table 1). Hemimethylated CpG dyads in cancer and control clones usually did not occur as runs but rather had the closest CpG on either side as a M/M or U/U dyad. Furthermore, of the 27 cancer clones containing more than one hemimethylated CpG site, 15 had hemimethylated dyads of opposite polarity with respect to which strand was unmethylated.

### CpG Sites with Preferred Methylation Changes in Cancers

Some normally M/M CpG sites seemed to be more likely to become demethylated in both ovarian carcinomas and Wilms' tumors than others (Fig. 3). To test the significance of this finding, we did a pairwise comparison of methylation changes in cancer clones at the seven normally M/M sites and also at the

**Table 1. Examples of asymmetry in hemimethylation of *NBL2* sites in cancers**

Dyad methylation status	No. clones with indicated type of demethylation		
	CpG8	CpG10	CpG12
Symmetrically demethylated (demethylated in both strands)	9	22	49
Hemimethylated (demethylated in only one strand)*	9	18	11
Hemimethylated as U/M (demethylated only in top strand)	1	18	0
Hemimethylated as M/U (demethylated only in bottom strand)	8	0	11

\*U/M status at CpG10 was seen in five cancers. M/U at CpG8 was seen in five cancers and a CpG12 in four cancers.

two normally unmethylated CpG dyads. In both Wilms' tumor and ovarian carcinoma groups, the following significant differences were observed: demethylation at CpG12 was more frequent than at CpG8 or CpG11, demethylation at CpG2 was more frequent than at CpG5, and demethylation at CpG3 was more frequent than at CpG11 ( $P < 0.05$ , adjusted for multiple comparisons). With respect to the two positions that were never normally symmetrically methylated, CpG6 was significantly more prone to cancer-associated hypermethylation (conversion to M/M) than CpG14 ( $P < 0.00001$ ) in ovarian carcinomas although not in the Wilms' tumors.

There was also evidence of preferential epigenetic patterning involving multiple CpG positions in the sequenced *NBL2* region from the cancers. Eleven cancer clones derived from four cancers had the following methylation status: CpG4, U/M; CpG12, M/U; CpG1, CpG3, CpG5, CpG6, and CpG10, U/U; CpG7, CpG8, CpG9, CpG11, CpG13, and CpG14, M/M (Fig. 2, first row for WT4, second row for WT9 and OvCaO, and third row for WT67). This methylation pattern constitutes changes from the normally conserved methylation status of the five underlined CpG sites.

To try to explain site preferences for cancer-linked methylation changes or for the conserved methylation patterns in somatic controls, we looked for possible effects of the sequence 1 to 3 bp on either side of each CpG. No rules for predicting the methylation status in the somatic controls or cancers based on adjacent sequences could be deduced just from the region subjected to genomic sequencing. However, Southern blot analysis of *NBL2* arrays gave us further insights as described below.

#### Distinguishing All Cancer and Somatic Control Clones by Methylation Status at Several CpGs

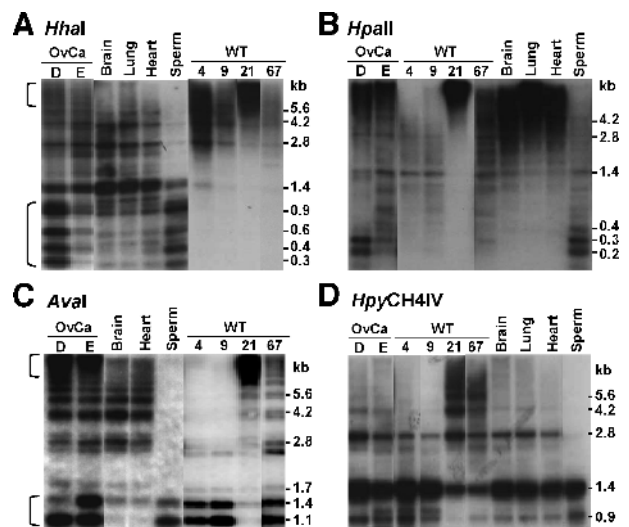
We searched for a few CpG sites whose methylation status could be used to distinguish all the cancer-derived molecular clones from all the somatic control clones. We found such sites with 100% predictive power by generating a classification tree from the data. All but two of the cancer-derived clones displayed symmetrical methylation at CpG6 (M/M) or demethylation at CpG10 (U/U or U/M); none of the somatic control-derived clones had these epigenetic attributes. The two

exceptional tumor clones could not display hypomethylation because CpC or CpT replaced CpG6 (Fig. 2B, WT67 and WT21, hyphen at the CpG6 position). Those two clones exhibited hypermethylation at CpG14, which distinguishes them from all somatic control-derived clones. Our ability to distinguish all *NBL2* cancer clones from all *NBL2* somatic control clones also shows the purity of the cancer DNA samples used for this analysis.

#### Analysis of CpG Methylation by Southern Blotting

All cancers in this study and an additional 13 ovarian carcinomas and 46 Wilms' tumors had been examined by Southern blot analysis for methylation at *HhaI* and *NorI* sites with a 1.4-kb *NBL2* probe (12). *HhaI* digests of DNAs from various postnatal somatic control tissues from 15 individuals gave very similar distributions of intermediate molecular weight hybridizing fragments (e.g., see Fig. 4A), whereas *NorI* digests all gave very high molecular weight hybridizing fragments (12). A comparison of cancers and somatic controls revealed predominant hypermethylation at *HhaI* sites in 81% of the 69 cancers and hypomethylation in 4% of them (e.g., Fig. 4A). Advantages of Southern blot analysis are that it can show long-range methylation patterns not identifiable by genomic sequencing, especially in tandem repeats, and it provides results from the population average of all the copies of the examined sequence.

In this study, we compared the above Southern blot results describing methylation at all *HhaI* sites throughout the *NBL2* arrays and the hairpin genomic sequencing results for methylation at every CpG in a 0.2-kb *NBL2* subregion (Table 2). *HhaI* site methylation scores were approximated from phosphorimager quantitation of Southern blot results



**FIGURE 4.** Representative Southern blot analysis of *NBL2* hypermethylation and hypomethylation in cancer DNAs. Ovarian carcinoma, Wilms' tumor, and control DNAs were digested with the indicated CpG methylation-sensitive enzymes and probed with the 1.4-kb *NBL2* sequence. Brackets in **A** and **C** indicate the separate hypermethylated and hypomethylated fractions of *NBL2* repeats in OvCaD and OvCaE, although the hypomethylated repeats were more prominent, especially for *HhaI* digests. Different exposures from the same blot were used.

(+1 to +3, increasing hypermethylation relative to somatic controls; -1 to -3, increasing hypomethylation). The genomic sequencing data for each cancer clone were quantified as the weighed average of hypermethylation at the two normally unmethylated CpGs and hypomethylation at the seven normally methylated CpGs. There was a significant association between *NBL2* methylation changes in the cancers determined by these two assays ( $P < 0.001$ ). Therefore, both are monitoring similar methylation changes. In addition, a comparison of the overall m5CpG content of the DNA (by high-performance liquid chromatography analysis) and the total proportion of methylated sites in the *NBL2* 0.2-kb subregion indicated a significant association between these two epigenetic variables ( $P = 0.001$ ) as well as between global DNA methylation and *NBL2* *HhaI* site methylation. Therefore, *NBL2* methylation changes in cancer are linked to global DNA methylation changes.

To enable us to analyze methylation changes at *HpaII*, *AvaI*, *HpyCH4IV*, or *BstUI* sites in *NBL2* arrays from cancer samples, we first characterized methylation of normal somatic DNAs digested with these CpG methylation-sensitive enzymes and hybridized to a *NBL2* probe. All somatic controls from various postnatal somatic tissues derived from different individuals gave very similar results with a given enzyme (Fig. 4; data not shown). However, there was much more resistance of *NBL2* arrays in all these controls to cleavage by some of these enzymes than by others. These results could not be explained by the frequency of the restriction sites in *NBL2*. There is an average of ~9 to 10 *HpaII* sites versus 5 to 6 *HhaI*, 2 to 4 *AvaI*, 3 to 5 *HpyCH4IV*, and 6 *BstUI* sites per *NBL2* monomer (Genbank Y10752 and AC0128692). Nonetheless, *NBL2* arrays in somatic controls were much more resistant to digestion by *HpaII* than by the other enzymes, and *HpyCH4IV* gave more cleavage than the other enzymes (Fig. 4). We disproved the possibility that the low extent of digestion of *NBL2* arrays by *HpaII* in somatic control DNAs was due to

sequence variation by showing complete digestion of all tested samples to <0.4-kb fragments by *MspI*, an isoschizomer of *HpaII*. *MspI* is resistant to CpG methylation, except at GGCCGG sites (28). In addition, internal controls for all digests showed that no inhibitors were present. The preferential methylation of *HpaII* sites in *NBL2* from somatic controls observed in these Southern blot assays was consistent with genomic sequencing data (Fig. 2, CpG2, CpG5, and CpG11).

All 18 ovarian cancer DNAs and 13 of the 15 Wilms' tumor DNAs examined with at least three of the above enzymes exhibited altered Southern blot patterns of *NBL2* methylation relative to somatic controls (Table 2; data not shown). We compared Southern blot data from cancer DNAs digested with different enzymes (Fig. 2) with the caveats that *HpaII* digests give an underestimate of hypermethylation and *HpyCH4IV* digests give an underestimate of hypomethylation. Importantly, *HhaI* sites seemed to undergo *de novo* methylation during carcinogenesis more frequently than *AvaI*, *HpyCH4IV*, and *BstUI* sites despite all of these enzymes giving mostly intermediate molecular weight *NBL2*-hybridizing bands in somatic controls (Fig. 4; Table 2; data not shown). This suggests some sequence specificity to cancer-linked hypermethylation. In addition, the distribution of *NBL2*-containing restriction fragments in *HhaI* digests and *AvaI* digests of OvCaD and OvCaE indicated that *NBL2* arrays can be bifurcated into two epigenetic components differing in the extent of methylation at a given restriction site in several of the cancers (Fig. 4A and C, brackets). Long tandem regions of hypermethylation at these two kinds of restriction sites were observed as increases in *NBL2* signal in >10-kb fragments, although those tumors also displayed increases in low molecular weight signal relative to the somatic controls. Individual fractions of *NBL2* repeats with respect to long-range methylation patterns might correspond to *NBL2* arrays on different acrocentric chromosomes.

**Table 2. Methylation changes in *NBL2* repeats in Wilms' tumors and ovarian carcinomas relative to controls somatic controls as determined by hairpin sequencing or Southern blot analysis**

Sample	Summary of genomic sequencing results*		<i>NBL2</i> methylation scores from Southern blot with the indicated DNA digests †				Global DNA methylation, ‡ % C methylated
	Hypermethylation (%)	Hypomethylation (%)	<i>HhaI</i>	<i>AvaI</i>	<i>HpaII</i>	<i>BstUI</i>	
OvCaD	22	50	-2	↓	↓	↓	3.31
OvCaE	33	47	-1	↓	↓	↓	2.94
WT9	33	26	+1	↓	↓	↓	3.09
WT4	50	30	+1	↓	↓	NC	2.88
OvCaN	63	11	+2	NC	↓	↑	3.76
WT67	78	9.3	+1	↓	↓	↓	3.45
OvCaO	81	9.3	+2	↓	↓	↑	3.73
WT21	86	4.9	+3	↑	↑	↑	3.90
OvCaQ	87	14	+3	NC	↓	↑	3.57
WT16	89	5.3	+3	↑	↑	↑	3.67
ICF B LCL	19	53	+3	↓	↓	ND	ND
Patient C LCL	63	12	+2	↓	NC	ND	ND

\*Hypermethylation in *NBL2* at the normally unmethylated CpG6 and CpG14 was calculated as the overall percentage of these two sites with symmetrical methylation, and hypomethylation was calculated as the overall loss of symmetrical methylation at the normally methylated CpG2, CpG3, CpG5, CpG8, CpG10, CpG11, and CpG12.

† *HhaI* site methylation scores were from previous Southern blot analyses with phosphorimager quantitation (12). Negative values, overall hypomethylation at *HhaI* sites; positive values, overall hypermethylation at these sites. For the other CpG methylation-sensitive enzymes, downward and upward arrows denote hypomethylation and hypermethylation, respectively. NC, no change in methylation relative to the somatic controls; ND, not determined.

‡ Global genomic methylation levels determined by high-performance liquid chromatography analysis of DNA digested to mononucleosides (11). Depending on the tissue, somatic controls have 3.43% to 4.04% of genomic C residues methylated.



cancer, there was only one cancer DNA clone (in WT67; Fig. 2) that displayed both hypermethylation and hypomethylation at precisely these positions. This observation and the nonrandom nature of many of the other DNA methylation changes observed by hairpin genomic sequencing and Southern blot analyses indicate that losses and gains of methylation in *NBL2* during carcinogenesis are often targeted to specific CpG positions and in specific patterns within the repeat.

The targeting of *NBL2* for nonrandom hypermethylation and hypomethylation cannot be explained by transcription-related binding of sequence-specific DNA binding proteins as is the case for certain promoters (31). *NBL2* underwent extensive cancer-linked alterations in methylation despite its lack of transcription in normal tissues and in most analyzed cancers and the absence of an *in silico*-predicted gene structure (12). Therefore, silencing of transcription is not necessary for all cancer-associated DNA hypermethylation, although it has been implicated in promoter hypermethylation (32). Moreover, an *in silico* search for consensus sites for sequence-specific DNA-binding proteins in *NBL2* (<http://www.cbil.upenn.edu/tess>) did not yield putative sites that could explain the observed methylation patterns. Whether gene regions and other DNA repeats incur the same type of interspersed hypermethylation and hypomethylation, which might occur especially in the early stages of tumorigenesis, remains to be determined.

Cancer-linked demethylation of *NBL2* was often observed in more than one of the seven normally methylated CpG positions with intervening CpGs that retained methylation. With respect to hemimethylation, cancer clones had a higher frequency of hemimethylated CpG sites than somatic control clones, and 15 clones had two hemimethylated sites with opposite strands unmethylated. These results indicate that demethylation by inhibition of maintenance methylation after DNA replication is not the major source of cancer-linked hypomethylation. Instead, they suggest some kind of active demethylation. However, there was a small, but statistically significant, percentage of cancer DNA clones containing runs of unmethylated CpGs on one or both strands. An overall decrease in fidelity of maintenance methylation could contribute to this cancer-associated demethylation, but the finding that most hypomethylation was intermittent along the examined cancer DNA molecules and the considerable sequence specificity observed in the cancer-linked CpG hypomethylation suggests more than just the previously reported inaccurate methylation maintenance (33). The mechanism for demethylation in cancer is uncertain. However, it is clear that mammals have the capacity for active demethylation as seen in the male pronucleus of the mouse zygote (34). Hemimethylated sites in vertebrate DNAs were described previously (16, 35). There is evidence for demethylation specifically of one strand in a transcription regulatory region followed by demethylation of the other during normal vertebrate development (36, 37). Our results suggest that hemimethylated sites are a rather stable intermediate in demethylation of DNA during carcinogenesis and that this demethylation of one strand of a CpG dyad occurs with preferences for certain CpGs and with strand preferences at some of these CpGs.

In *de novo* methylation of *NBL2* in cancer, DNMT3B is likely to be the main enzyme involved as determined by our

analysis of B-cell LCLs from controls and from ICF patients who have inactivating mutations in *DNMT3B* that eliminate most DNMT3B activity (30). The much lower levels of *NBL2* methylation in ICF LCLs than in control LCLs indicate that DNMT3B is necessary for establishing the normal *NBL2* methylation pattern during development. The hypermethylation of *NBL2* at *HhaI* sites in control LCLs relative to somatic control tissues could be explained by overexpression of *DNMT3B* (as well as *DNMT3A* and *DNMT1*) during transformation with EBV (38) and/or by the oncogenic transformation-associated loss of fidelity of DNA methyltransferases (33). *In vitro* transformation of lymphocytes by EBV may provide a model for understanding *NBL2* methylation changes during malignant transformation *in vivo* because both hypomethylation and hypermethylation relative to control somatic tissues were observed in *NBL2* in normal LCLs. In the two ICF LCLs subject to genomic sequencing, despite the overall hypomethylation of *NBL2*, some hypermethylation was observed at CpG6, although not at CpG14, the other site at which we could analyze cancer-linked hypermethylation. In addition, the control LCL displayed more methylation at CpG6 than CpG14. Similarly, CpG6 was hypermethylated significantly more frequently than CpG14 in ovarian carcinomas. Moreover, CpG6 was occasionally hemimethylated in somatic controls, whereas CpG14 was always symmetrically unmethylated. These findings might be related to the dynamic system of normal maintenance methylation and *de novo* methylation proposed by Pfeifer et al. (39). At *NBL2*, there may be infrequent *de novo* methylation of CpG6 in one strand in normal cells, which is not followed by maintenance methylation. In contrast, there may be frequent hemimethylation at this site with subsequent maintenance methylation on oncogenic transformation.

An *in vitro* study of methylation by DNMT3B indicated strong sequence preferences for *de novo* methylation (40). Our findings, especially from Southern blot analysis, support this idea, although the specificities that we found do not match the *in vitro* ones. This enzyme may have its sequence preferences strongly altered *in vivo*. Both our genomic sequencing and Southern blot analysis indicate that *HpaII* sites (CCGG) have an especially high level of methylation in *NBL2* in normal somatic tissues. In addition, our Southern blot analysis suggests that *HhaI* sites (GCGC), which were missing from the bisulfite-sequenced region, were more frequently *de novo* methylated in the cancers than *HpyCH4IV* (ACGT), *AvaI* (CYCGRG), and *BstUI* (CGCG) sites.

Evidence had been provided for cross-talk between demethylation and *de novo* methylation pathways in tumorigenesis (41) and in *Arabidopsis* containing an antisense DNA methyltransferase transgene (42). However, hypermethylation of 5' regions of tumor suppressor genes and hypomethylation of LINE1 interspersed repeats, satellite DNA, and promoter regions of cancer-testes antigen genes (4, 6, 11, 43) have been shown to be statistically independent of each other, although all such changes are linked to cancer. Although hypermethylation at certain DNA sequences and hypomethylation at others in cancer are not associated with one another, this study shows that both cancer-linked hypomethylation and hypermethylation are somehow targeted to *NBL2* repeats. We hypothesize that a

chromatin structure change in *NBL2* arrays occurs during oncogenesis that predisposes to both demethylation and *de novo* methylation in *cis*. Alternatively, *NBL2* arrays, which have a high overall m5CpG content, might first be demethylated during tumorigenesis and the resulting chromatin structure change might favor further demethylation as well as *de novo* methylation.

Previous bisulfite-based genomic sequencing studies of cancer DNA usually involved unmethylated CpG-rich promoter regions that become hypermethylated mostly homogeneously (22, 23, 44). There may be several reasons for *NBL2* displaying surprisingly complex, nonrandom patterns of altered methylation during carcinogenesis. It is apparently not a gene, and its methylation status probably confers no selective advantage to a developing tumor. This is unlike the situation with promoters of tumor suppressor genes whose almost complete methylation can benefit the growing tumor by repressing transcription or stabilizing this repression. In addition, unlike most DNA regions from cancers analyzed by genomic sequencing, *NBL2* normally has very low levels of methylation at some CpGs and complete methylation at many others so that both cancer-linked increases and decreases of DNA methylation can be observed. Furthermore, it seems to be an unusually frequent target for multiple methylation changes during carcinogenesis. As such, it is a good candidate for a cancer marker as well as a source of insight into cancer-linked epigenetic alterations without the skewing of DNA methylation patterns by oncogenic selection pressures.

## Materials and Methods

### Patients and DNA Samples

With institutional review board approval, primary tumor samples were obtained from surgery patients before chemotherapy or radiation therapy. Informed consent was given by all patients or unlinked samples were used. The LCLs were described previously (45–47) or available from the Coriell Institute (Camden, NJ; GM17900, AG14836, AG14953, and AG15022). Control somatic tissues were autopsy specimens of trauma victims (individuals A, B, and C, all males of 56, 19, and 68 years, respectively). DNA was purified as described previously (11).

### Hairpin-Bisulfite PCR and Cloning

Hairpin-bisulfite PCR was done basically as described by Laird et al. (16) using a *NBL2* sequence (Y10752, Genbank) to design primers and the hairpin linker. Human DNA (0.5 µg) or *NBL2*-containing pDMHD-1 (50 ng; ref. 14) plus 450 ng λDNA carrier were digested with 10 units *Bsm*AI and ligated to 5'-CCCTAGCGATGCGTTCGAGCATCGCT-3'. The DNA was denatured with 0.6 mol/L NaOH at 37°C for 15 minutes followed by incubation in boiling water for 1 minute. At hourly intervals during the 5-hour bisulfite treatment, the sample was incubated four times in boiling water for 1 minute. In an ultrafiltration device (Microcon-100, Millipore), bisulfite-modified DNA was washed thrice with water, desulfonated with 0.3 mol/L NaOH at 37°C for 15 minutes, and eluted in 50 µL of 10 mmol/L Tris-HCl, 1 mmol/L EDTA (pH 7.5). The primers for subsequent PCR had a 3' T or A corresponding to deamination

products from a non-CpG C residue or its complement. The primers were F2-1, 5'TTTTTGTGGGTTTGTGTTAGT-3', and R2-2, 5'-CAAAAACATCTTTATTCCTCTA-3'. F2-1 was replaced by F2-2, 5'-AYGTGGTTTGGGTTAGGTAT-3', in the second round of PCR. Only the F2-2 primer had a CpG in the analogous unmodified genomic sequence (at positions 2 and 3). After denaturation at 94°C for 15 minutes, PCR was done (Hotstar, Qiagen) for 30 cycles on 2 µL of the bisulfite-treated DNA (94°C for 15 seconds, 52°C for 15 seconds, 72°C for 1 minute, and a final extension at 72°C for 5 minutes). Then, 1 µL of the product was amplified analogously for an additional 35 cycles. Purified fragments obtained by electrophoresis in a 1.5% agarose gel were used for cloning (TA Cloning kit, Invitrogen), transformation (*E. coli*, Top10F), and sequencing (Translational Genomics Research Institute, Phoenix, AZ). Due to much sequence variation, there were three CpG positions that were much less frequently present in the clones than the 14 CpG positions referred to above. They are omitted for simplicity, but their methylation status did not change any of the conclusions of this study.

### Southern Blot Analysis

For Southern blot analysis, 1.5 µg human DNA was digested with 15 to 30 units restriction endonuclease overnight according to the manufacturer's procedures (New England Biolabs), all with parallel internal controls as described previously (12). At least three diverse somatic control tissues and sperm DNA were included as references in each blot.

### Reverse Transcription-PCR

Using random primers in one set and oligo(dT) in a duplicate set, cDNA was synthesized from 3 µg total RNA that had been treated with 3 units DNase I (amplification grade, Invitrogen) for 45 minutes at room temperature. Real-time PCR (SYBR Green PCR Master Mix, Applied Biosystems) was done with previously described primers and conditions (12). Semiquantitative RT-PCR with evaluation of the product by gel electrophoresis was also done as described previously (12).

### Statistical Methods

Genomic sequencing methylation data were analyzed using R version 2.0.1 (<http://www.rproject.org>).  $\chi^2$  test statistics were used to assess differences of proportions, and strengths of association for continuous and ordinal variables were evaluated using the standard Pearson's correlation and Kendall's  $\tau$  statistics, respectively. Where appropriate, *P*s were adjusted for multiple tests using the Holm procedure. Classification trees were generated using the RPART library (48).

## Acknowledgments

We thank Drs. Louis Dubeau, Murali Chintagumpala, Jeffrey Dome, Lisa Diller, Martin Champagne, and Hisaki Nagai for supplying the cancer samples and the *NBL2* plasmid and Kathryn E. Donahoe for valuable editorial assistance.

## References

1. Baylin SB, Hoppener JW, de Bustros A, Steenbergh PH, Lips CJ, Nelkin BD. DNA methylation patterns of the calcitonin gene in human lung cancers and lymphomas. *Cancer Res* 1986;46:2917–22.
2. Issa JP. CpG island methylator phenotype in cancer. *Nat Rev Cancer* 2004;4: 988–93.

3. Feinberg AP, Vogelstein B. Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature* 1983;301:89–92.
4. Ehrlich M, Woods C, Yu M, et al. Quantitative analysis of association between DNA hypermethylation, hypomethylation, and DNMT RNA levels in ovarian tumors. *Oncogene*. In press 2005.
5. Narayan A, Ji W, Zhang X-Y, et al. Hypomethylation of pericentromeric DNA in breast adenocarcinomas. *Int J Cancer* 1998;77:833–8.
6. Santourlidis S, Florl A, Ackermann R, Wirtz HC, Schulz WA. High frequency of alterations in DNA methylation in adenocarcinoma of the prostate. *Prostate* 1999;39:166–74.
7. Bariol C, Suter C, Cheong K, et al. The relationship between hypomethylation and CpG island methylation in colorectal neoplasia. *Am J Pathol* 2003;162:1361–71.
8. De Smet C, De Backer O, Faraoni I, Lurquin C, Brasseur F, Boon T. The activation of human gene MAGE-1 in tumor cells is correlated with genome-wide demethylation. *Proc Natl Acad Sci U S A* 1996;93:7149–53.
9. Ehrlich M. DNA methylation in cancer: too much, but also too little. *Oncogene* 2002;21:5400–13.
10. Jones PA, Baylin SB. The fundamental role of epigenetic events in cancer. *Nat Rev Genet* 2002;3:415–28.
11. Ehrlich M, Jiang G, Fiala ES, et al. Hypomethylation and hypermethylation of DNA in Wilms tumors. *Oncogene* 2002;21:6694–702.
12. Nishiyama R, Qi L, Tsumagari K, et al. A DNA repeat, NBL2, is hypermethylated in some cancers but hypomethylated in others. *Cancer Biol Ther* 2005;4:440–8.
13. Itano O, Ueda M, Kikuchi K, et al. Correlation of postoperative recurrence in hepatocellular carcinoma with demethylation of repetitive sequences. *Oncogene* 2002;21:789–97.
14. Nagai H, Kim YS, Yasuda T, et al. A novel sperm-specific hypomethylation sequence is a demethylation hotspot in human hepatocellular carcinomas. *Gene* 1999;237:15–20.
15. Thoraval D, Asakawa J, Wimmer K, et al. Demethylation of repetitive DNA sequences in neuroblastoma. *Genes Chromosomes Cancer* 1996;17:234–44.
16. Laird CD, Pleasant ND, Clark AD, et al. Hairpin-bisulfite PCR: assessing epigenetic methylation patterns on complementary strands of individual DNA molecules. *Proc Natl Acad Sci U S A* 2004;101:204–9.
17. Frommer M, McDonald LE, Millar DS, et al. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc Natl Acad Sci U S A* 1992;89:1827–31.
18. Dodge JE, Ramsahoye BH, Wo ZG, Okano M, Li E. *De novo* methylation of MMLV provirus in embryonic stem cells: CpG versus non-CpG methylation. *Gene* 2002;289:41–8.
19. Warnecke PM, Stirzaker C, Melki JR, Millar DS, Paul CL, Clark SJ. Detection and measurement of PCR bias in quantitative methylation analysis of bisulphite-treated DNA. *Nucleic Acids Res* 1997;25:4422–6.
20. Dubeau L. The cell of origin of ovarian epithelial tumors and the ovarian surface epithelium dogma: does the emperor have no clothes? *Gynecol Oncol* 1999;72:437–42.
21. Millar DS, Paul CL, Molloy PL, Clark SJ. A distinct sequence (ATAAA)<sub>n</sub> separates methylated and unmethylated domains at the 5'-end of the GSTP1 CpG island. *J Biol Chem* 2000;275:24893–9.
22. Melki JR, Vincent PC, Clark SJ. Concurrent DNA hypermethylation of multiple genes in acute myeloid leukemia. *Cancer Res* 1999;59:3730–40.
23. Amoreira C, Hindermann W, Grunau C. An improved version of the DNA Methylation database (MethDB). *Nucleic Acids Res* 2003;31:75–7.
24. Toth M, Lichtenberg U, Doerfler W. Genomic sequencing reveals a 5-methylcytosine-free domain in active promoters and the spreading of preimposed methylation patterns. *Proc Natl Acad Sci U S A* 1989;86:3728–32.
25. Turker MS. Gene silencing in mammalian cells and the spread of DNA methylation. *Oncogene* 2002;21:5388–93.
26. Nguyen C, Liang G, Nguyen TT, et al. Susceptibility of nonpromoter CpG islands to *de novo* methylation in normal and neoplastic cells. *J Natl Cancer Inst* 2001;93:1465–72.
27. Yan PS, Shi H, Rahmatpanah F, et al. Differential distribution of DNA methylation within the RASSF1A CpG island in breast cancer. *Cancer Res* 2003;63:6178–86.
28. Busslinger M, deBoer E, Wright S, Grosveld FG, Flavell RA. The sequence GGCMCGG is resistant to MspI cleavage. *Nucleic Acids Res* 1983;11:3559–69.
29. Hansen RS, Wijmenga C, Luo P, et al. The DNMT3B DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome. *Proc Natl Acad Sci U S A* 1999;96:14412–7.
30. Gowher H, Jeltsch A. Molecular enzymology of the catalytic domains of the Dnmt3a and Dnmt3b DNA methyltransferases. *J Biol Chem* 2002;277:20409–14.
31. Hornstra IK, Yang TP. High-resolution methylation analysis of the human hypoxanthine phosphoribosyltransferase gene 5' region on the active and inactive X chromosomes: correlation with binding sites for transcription factors. *Mol Cell Biol* 1994;14:1419–30.
32. Clark SJ, Melki J. DNA methylation and gene silencing in cancer: which is the guilty party? *Oncogene* 2002;21:5380–7.
33. Ushijima T, Watanabe N, Shimizu K, Miyamoto K, Sugimura T, Kaneda A. Decreased fidelity in replicating CpG methylation patterns in cancer cells. *Cancer Res* 2005;65:11–7.
34. Santos F, Hendrich B, Reik W, Dean W. Dynamic reprogramming of DNA methylation in the early mouse embryo. *Dev Biol* 2002;241:172–82.
35. Liang G, Chan MF, Tomigahara Y, et al. Cooperativity between DNA methyltransferases in the maintenance methylation of repetitive elements. *Mol Cell Biol* 2002;22:480–91.
36. Saluz HP, Jiricny J, Jost JP. Genomic sequencing reveals a positive correlation between the kinetics of strand-specific DNA demethylation of the overlapping estradiol/glucocorticoid-receptor binding sites and the rate of avian vitellogenin mRNA synthesis. *Proc Natl Acad Sci U S A* 1986;83:7167–71.
37. Singal R, vanWert JM. *De novo* methylation of an embryonic globin gene during normal development is strand specific and spreads from the proximal transcribed region. *Blood* 2001;98:3441–6.
38. Tsai CN, Tsai CL, Tse KP, Chang HY, Chang YS. The Epstein-Barr virus oncogene product, latent membrane protein 1, induces the downregulation of E-cadherin gene expression via activation of DNA methyltransferases. *Proc Natl Acad Sci U S A* 2002;99:10084–9.
39. Pfeifer GP, Steigerwald SD, Hansen RS, Gartler SM, Riggs AD. Polymerase chain reaction-aided genomic sequencing of an X chromosome-linked CpG island: methylation patterns suggest clonal inheritance, CpG site autonomy, and an explanation of activity state stability. *Proc Natl Acad Sci U S A* 1990;87:8252–6.
40. Handa V, Jeltsch A. Profound flanking sequence preference of Dnmt3a and Dnmt3b mammalian DNA methyltransferases shape the human epigenome. *J Mol Biol* 2005;348:1103–12.
41. Pogribny IP, Miller BJ, James SJ. Alterations in hepatic p53 gene methylation patterns during tumor progression with folate/methyl deficiency in the rat. *Cancer Lett* 1997;115:31–8.
42. Jacobsen SE, Meyerowitz EM. Hypermethylated SUPERMAN epigenetic alleles in *Arabidopsis*. *Science* 1997;277:1100–3.
43. Kaneda A, Tsukamoto T, Takamura-Enya T, et al. Frequent hypomethylation in multiple promoter CpG islands is associated with global hypomethylation, but not with frequent promoter hypermethylation. *Cancer Sci* 2004;95:58–64.
44. Rush LJ, Raval A, Funchain P, et al. Epigenetic profiling in chronic lymphocytic leukemia reveals novel methylation targets. *Cancer Res* 2004;64:2424–33.
45. Ehrlich M, Buchanan K, Tsien F, et al. DNA methyltransferase 3B mutations linked to the ICF syndrome cause dysregulation of lymphocyte migration, activation, and survival genes. *Hum Mol Genet* 2001;10:2917–31.
46. Tuck-Muller CM, Narayan A, Tsien F, et al. DNA hypomethylation and unusual chromosome instability in cell lines from ICF syndrome patients. *Cytogenet Cell Genet* 2000;89:121–8.
47. Gisselsson D, Shao C, Tuck-Muller C, et al. Interphase chromosomal abnormalities and mitotic missegregation of hypomethylated sequences in ICF syndrome cells. *Chromosoma* 2005;114:118–26.
48. Breiman L, Friedman JH, Olshen RA, Stone CJ. Classification and regression trees. Belmont (CA): Wadsworth; 1984.