THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 280, No. 15, Issue of April 15, pp. 14413–14419, 2005
Printed in U.S.A.

Hemimethylation and Non-CpG Methylation Levels in a Promoter Region of Human LINE-1 (L1) Repeated Elements*

Received for publication, December 8, 2004, and in revised form, February 1, 2005
Published, JBC Papers in Press, February 14, 2005, DOI 10.1074/jbc.M413836200

Alice F. Burden‡§, Nathan C. Manley‡¶, Aaron D. Clark‡, Stanley M. Gartler‖, Charles D. Laird‡‖§, and R. Scott Hansen****

*From the Departments of ‡Biology, ‖Genome Sciences, and **Medicine University of Washington, Seattle, Washington 98195*

DNA methylation within the promoter region of human LINE1 (L1) transposable elements is important for maintaining transcriptional inactivation and for inhibiting L1 transposition. Determining methylation patterns on the complementary strands of repeated sequences is difficult using standard bisulfite methylation analysis. Evolutionary changes in each repeat and the variations between cells or alleles of the same repeat lead to a heterogeneous population of sequences. Potential sequence biases can arise during analyses that are different for the converted sense and antisense strands. These problems can be avoided with hairpin-bisulfite PCR, a double-stranded PCR method in which complementary strands of individual molecules are attached by a hairpin linker ligated to genomic DNA. Using human L1 elements to study methylation of repeated sequences, (i) we distinguish valid L1 sequences from redundant and contaminant sequences by applying the powerful new method of molecular barcodes, (ii) we resolve a controversy on the level of hemimethylation of L1 sequences in fetal fibroblasts in favor of relatively little hemimethylation, (iii) we report that human L1 sequences in different cell types also have primarily concordant CpG methylation patterns on complementary strands, and (iv) we provide evidence that non-CpG cytosines within the regions analyzed are rarely methylated.

Methylation of CpG dyads within the promoter region of human LINE1 (L1)[1] transposable elements is important for maintaining transcriptional inactivation and for inhibiting L1 transposition (1–3). Identifying the various states of cytosine methylation in human L1 promoter sequences is thus potentially useful in understanding L1 stability. Methylation states of CpG/CpG dyads are expected usually to be concordant, *i.e.* fully methylated or fully unmethylated. Methylated cytosines on the parent strand usually give rise to methylated cytosines on the daughter strand, and unmethylated cytosines on the parent strand usually give rise to unmethylated cytosines on the daughter strand because of the properties of DNA methyltransferases during DNA replication (4–6). CpG dyads that are discordant in methylation, *i.e.* hemimethylated, are expected to occur at low to moderate frequencies (5, 7–9) with the exception of regions undergoing transcriptional silencing or reactivation and transiently during DNA replication (10–13). Contrary to this expectation, L1 sequences in human fetal fibroblast cultures were reported to be much more methylated on the sense strand than on the antisense strand (11), which would imply high levels of hemimethylation. Estimating methylation patterns on the complementary strands of repeated sequences can, however, be difficult using conventional bisulfite methylation analysis because few if any of the sampled sense and antisense strand sequences are likely to be derived from the same repeat in the same cell. Even sequences that have an apparent match are likely to be derived from different loci that either had identical sequence in the region examined prior to conversion or a match because of the sequence degeneration caused by bisulfite conversion. Estimating methylation patterns on the complementary strands of L1s using conventional analysis of sense and antisense strands also may be confounded by differential biases during PCR amplification and bacterial cloning. For example, the primers used for the sense strand might preferentially amplify methylated sequences from a small subclass of L1 sequences, whereas the antisense primers might amplify a different subclass of L1s that are more hypomethylated.

To overcome these inherent limitations of conventional bisulfite methylation analysis for determining methylation patterns on both strands of L1 repeat elements, we have used a double-strand PCR method, "hairpin-bisulfite PCR" (9). With this approach, the limitations of single-strand PCR can be avoided. Strand biases do not occur because both strands are linked prior to PCR amplification. Bisulfite-converted cytosines can be distinguished from evolutionary C-to-T transitions in L1 sequences by the presence, in the former class, of a guanine on the complementary strand, and information from the exact complementary strands of an individual allele is preserved, allowing allele and cell variation to be studied (9).

Our initial hairpin-bisulfite PCR data on L1 sequences using DNAs from two adult human fibroblast lines revealed levels of hemimethylated CpG dyads that were somewhat higher (14.1 and 10.0%, respectively) than levels observed for a single-copy gene, *FMR1*, in lymphocytes (6.4%) (9) but considerably lower than the 50–80% hemimethylation levels inferred from other published data (11). We now present additional data on the concordance of methylation at CpG dyads within the promoter region of human L1 sequences. Our data (i) distinguish valid from redundant and contaminant sequences by introducing, for repeated sequences, a powerful new method using molecular barcodes, (ii) resolve the controversy on the level of hemimethylation of L1 sequences in fetal fibroblasts, (iii) present data on

* This work was supported by National Institutes of Health Grants GM 53805, HD 02274, and HD 16659. 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.

§ To whom correspondence may be addressed: Dept. of Biology, University of Washington, Box 351800, Seattle, WA 98195. Tel.: 206-616-9385; Fax: 206-543-3041; E-mail: afb4@u.washington.edu (A. F. B.) or cdlaird@u.washington.edu (C. D. L.).

¶ Present address: Dept. of Biological Sciences, Stanford University, Stanford, CA 94305.

[1] The abbreviation used is: L1, LINE1.

the concordance of CpG methylation on complementary strands in L1 sequences from other human cell types, and (iv) assess the methylation status of non-CpG cytosines within the regions analyzed.

MATERIALS AND METHODS

DNA was isolated from GM04522 fetal fibroblasts that originated from a 16-week-old female fetus and were provided to us as passage 8 cells (Coriell Cell Repositories, Camden, NJ). Additional DNA samples from the Laird laboratory included those isolated from male peripheral blood leukocytes (A102) after red blood cell lysis (14), from sperm (A47) using dithiothreitol, proteinase K, phenol, and chloroform (15), and from a female lymphoblastoid line (A54) established by Epstein-Barr transformation of peripheral lymphocytes. DNA methylation data presented previously (9) are from a normal adult female fibroblast line (81-58A) obtained from Dr. George Martin (Department of Pathology, University of Washington) and from a normal male fibroblast line transformed with human telomerase reverse transcription (82-6) obtained from Dr. Peter Rabinovitch (Department of Pathology, University of Washington). Fibroblast tissue cultures were maintained in AmnioMAX™-C100 (Invitrogen) at 37 °C in a 5% CO_2 atmosphere. The transfer of cells was accomplished with trypsin-EDTA (Invitrogen). Purification of tissue culture DNA was performed using Qiagen genomic-tip 20/G. Approximately 2–3 cell doublings occurred between successive cell passages.

We examined L1 CpG and non-CpG methylation within a 104-bp promoter region of human L1 sequences, a region that contains nine CpG sites corresponding to those between nucleotides 219 and 323 as reported by Woodcock *et al.* (11). A more limited set of data from a variety of cell types was obtained for a 118-bp promoter region of L1, focusing on the possibility of non-CpG methylation between nucleotides 328 and 446.

We used hairpin-bisulfite PCR to assess methylation status on both strands of individual DNA molecules (9) with the following modifications. 2–5 μg of genomic DNA were cleaved by 25 units of restriction endonuclease BsmA1 for 1 h at 55 °C. For samples of fetal fibroblasts, a molecular-barcoded hairpin linker, either 5′-PO4-ACCAAGCGATGCGTTDDDDDDDCGAGCATCGCT-3′ or 5′-PO4-ACCAAGCGATGCDDDDDDDGCATCGCT-3′ in which D represents random incorporation of A, G, or T (for review see Ref. 16) was ligated to BsmAI-cleaved genomic DNA for 15 min at 20 °C using 400 units of New England Biolabs T4 DNA ligase in 20 μl with 1× ligation buffer. For DNA samples other than fetal fibroblasts, ligation of hairpin linkers 5′-PO4-ACCAAGCGATGCGTTCGAGCATCGCT-3′ or 5′-PO4-TGCTAGCGATGCGGTTGAGCATCGCT-3′ for the 104- and 118-bp regions, respectively, to BsmAI-cleaved genomic DNA was for 15 min at 20 °C using 400 units of New England Biolabs T4 DNA ligase in 20 μl with 1× ligation buffer.

Bisulfite-conversion cycling conditions were as follows. Hairpin-ligated DNA was denatured in 0.3 M NaOH for 20 min and then heated to 100 °C for 2 min before the addition of 3.4 M sodium bisulfite and 1 mM hydroquinone. The reaction mixture was incubated at 99.9 °C for 1 min followed by 10 cycles of 55 °C for 15 min and 99 °C for 1 min 30 s. Cycling was completed at 55 °C for 1 h and 10 °C for 9 h. A detailed schematic representation of the hairpin-bisulfite PCR method is shown by Laird *et al.* (see Fig. 2 and supporting Fig. 6 in Ref. 9). Post-bisulfite PCR conditions, TOPO TA cloning reactions, and sequencing of all of the samples were carried out as described previously (9) with the exception of the 118-bp region of the promoter for which PCR primers 5′-CTCCACCCAATTC(A/G)AACTTCCC-3′ and 5′-CTAACTTTGGGTAATGG(C/T)GGG(C/T)G-3′ were used and optimized with Failsafe buffer F (Epicenter). We also routinely clean the underside of the PCR machine lid with DNA Zap (Ambion). We have confirmed through barcoding that this procedure reduces contamination substantially.

The cytosine to uracil conversion efficiency based on non-CpG sites was generally high. Only sequences with >98.2% conversion (≤1 unconverted non-CpG cytosine per sequence) were included in our analysis except as described in the next paragraph. The average frequencies of converted non-CpG cytosines in the sequences used for our analysis were 99.4 and 99.8% for the 104- and 118-bp regions, respectively. Sequences with one or more hemimethylated CpG dyads from the 104-bp region had an equally high level of non-CpG conversion efficiency (99.5%) compared with the mean, indicating that hemimethylated dyads are intrinsic to the DNA analyzed and are not a consequence of incomplete conversion of unmethylated CpG dyads. To provide approximately equal representation of data from each of the four passages of fetal fibroblasts analyzed, here we include data from the first 15 well-converted sequences obtained from each passage with the exception of passage 22 from which only 12 well-converted sequences were recovered.

To search for possible non-CpG cytosine methylation in the 118-bp region, we also analyzed 14 L1 sequences from human sperm DNA that were chosen because of their lower conversion efficiency. These lower levels of conversion are not often observed with our normal protocol but were generated in some experiments by using fewer denaturation steps or by moving the denaturation steps to the end of the bisulfite procedure.

RESULTS

The analysis presented here for methylation patterns of L1 sequences in fibroblasts expands on our previous study (9) in several important ways. To provide the highest level of confidence in methylation data, we labeled hairpin-bisulfite PCR products from fetal fibroblasts with molecular barcodes (16). Molecular barcoding allows us to tag genomic sequences at the time of hairpin-linker ligation. The complementary strands of individual DNA molecules, thus, carry this identification through bisulfite conversion, PCR, cloning, and sequencing. We were also able to detect PCR redundancy and contamination, problems often associated with bisulfite conversion (17) and other PCR applications using limiting template (16). The fetal fibroblast data presented here illustrate such barcodes (Fig. 1). For example, only one nucleotide is different between any two pairs of hypermethylated sequences AB257, AB261, and AB287. Such small differences can arise from subtle differences in methylation patterns among different repeats or from errors during PCR amplification of a single allele. In this example, we can conclude that all three sequences are of independent genomic origins because each sequence has a different seven-nucleotide molecular barcode in the hairpin region (Fig. 1). Unmethylated sequences AB307 and AB313 are identical in having no CpG methylation. They, thus, lack the distinctive information provided by partial methylation patterns. We can conclude that these unmethylated sequences also have independent genomic origins because of their different barcodes within the hairpins.

The use of DNAs from passages of fetal rather than adult fibroblasts permits a more direct comparison of our data with those of Woodcock *et al.* (11). To investigate whether or not cell passage alters methylation patterns, we sampled fetal fibroblasts at several points throughout 14 passages beyond the initial passage 8 cells. In this data set of methylation patterns from fetal fibroblast L1s, complementary CpG dinucleotides on sense and antisense strands were usually concordant in their methylation status. For example, the ten illustrated sequences have a total of 81 CpG dyads that are conserved relative to the consensus human sequence (11). Of these 81 dyads, 47 (58.0%) were fully methylated, 27 (33.3%) were fully unmethylated, and 7 (8.6%) were hemimethylated. Thus, 91.3% CpG dyads were concordant in their methylation status (Fig. 1). Our larger data set of 57 L1 sequences from fetal fibroblast DNA, representing 471 consensus CpG dyads, showed a similar high level of methylation concordance (91.7%) and low level of hemimethylation (8.3%).

We analyzed nine of the 22 CpG consensus sites within the L1 promoter (11). For these nine sites, the methylation percentages ranged from 69 to 90% in fetal fibroblast sequences and the two complementary strands had approximately equal levels of methylation (Fig. 2*A*). The levels of sense-strand and antisense-strand methylation were highly correlated at these nine sites ($r^2 = 0.78$). In contrast to our data presented here, methylation levels of L1 sequences from fetal fibroblasts were reported by Woodcock *et al.* (11) to be very different between strands. Our analysis of their data indicates that sense-strand methylation averaged 99.4% and that antisense-strand methylation averaged 40% for these CpG sites in their sequences

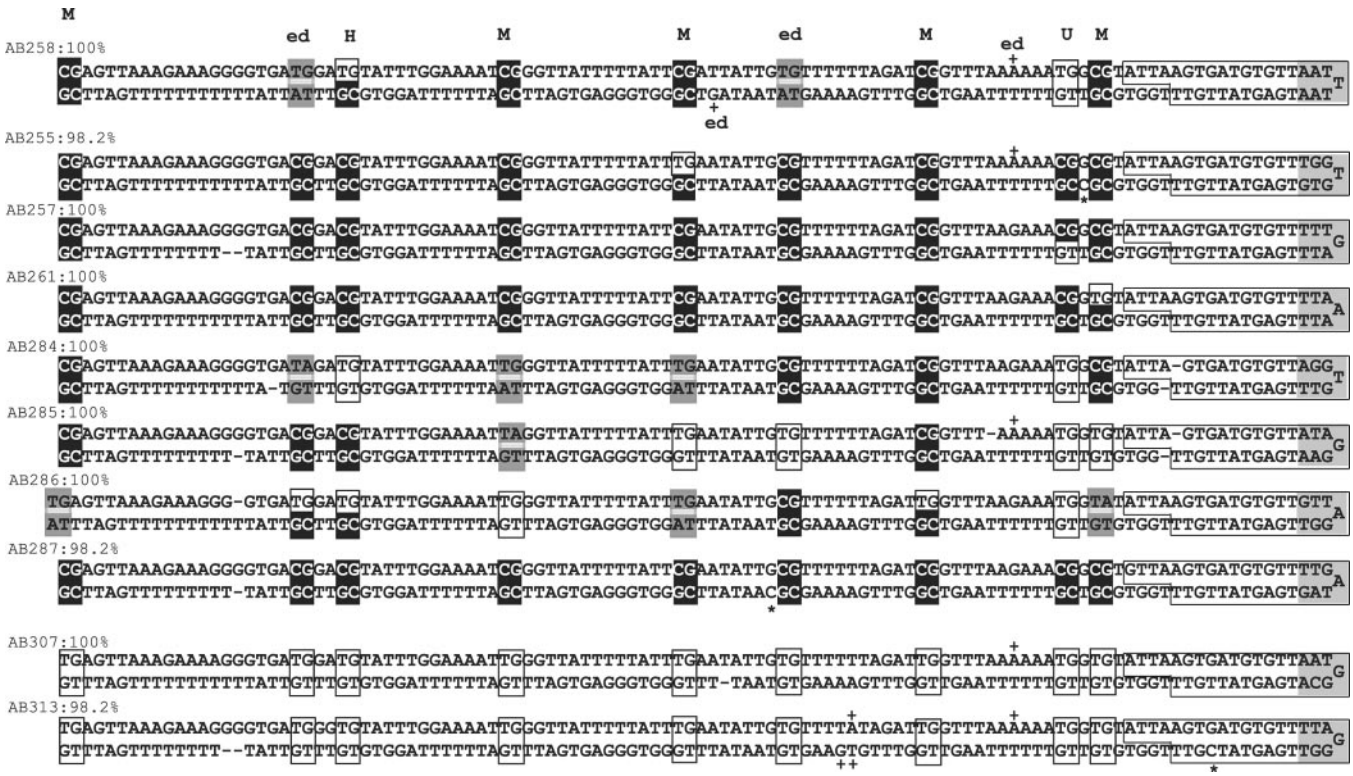

FIG. 1. **Methylation of a 104-bp L1 promoter sequence in DNAs from fetal fibroblast passages 10, 14, and 22 using barcoded hairpin-bisulfite PCR.** Methylation states were inferred using barcoded hairpin-bisulfite PCR as described under "Materials and Methods." As illustrated and labeled for the top sequence AB258, methylated CpG dyads (*M*) are indicated in *black*, unmethylated CpG dyads (*U*) are *boxed*, hemimethylated CpG dyads (*H*) are *black* on the methylated strand and *boxed* on the unmethylated complementary strand, evolutionary divergence (*ed*) within consensus CpG dyads is indicated in *gray*, and evolutionary divergence not within CpG dyads is indicated by "+." The hairpin linker on the *right* is *boxed*, and bases *shaded* within the linker represent the seven-nucleotide barcode. For the first eight sequences, evolutionary divergences within consensus CpG dyads are G/C→A/T transitions. In examples of divergence, information on the complementary strands of individual DNA duplexes distinguishes evolutionary changes from PCR errors (9). Identification number and conversion efficiency of each sequence are listed on the *left*. Only two non-converted non-CpG cytosines (*asterisk*) were observed in the first eight sequences. Only sequences with ≥98.2% conversion of non-CpG cytosines (representing zero or one non-conversion event per sequence) were used for analysis. *Dashes* within the sequence indicate sequencing errors resulting from uncalled bases. The first four analyzed sequences from passage 10 (*top four*) and passage 22 (*next four*) are shown here. Additionally, two fully unmethylated sequences from passage 14 were observed (*bottom two*). Distinctive barcodes indicate that these 10 sequences were derived from different genomic L1s. For example, sequence AB307 had a barcode of 5′-AATGGCA-3′ and sequence AB313 had a barcode of 5′-TTAGGGT-3′. For the former barcode, the C is either a PCR error (because cytosines were not included in the barcodes (16)) or the C represents a low level of error during synthesis of the barcoded linker. One unconverted non-CpG cytosine (*asterisk*) was observed in the last sequence.

(Fig. 2*B*). Such discordance in the levels of strand methylation would, if correct, imply high levels of hemimethylation. To search for possible explanations for these differences, we asked whether or not the passaging of fetal fibroblasts in cell culture significantly altered the level of hemimethylation. No systematic differences in the levels of hemimethylation were observed for DNA from cell passages 10, 14, 17, and 22 (Fig. 3). The percentages of CpG dyads hemimethylated in L1 sequences from these four passages were 9.5, 8.5, 7.4, and 7.5%, respectively, values that are not significantly different from the average of 8.3% ($p > 0.50$).

Site-specific differences in symmetry of methylation do, however, exist in L1 sequences from human fibroblasts. The highest average level of hemimethylation that we observed in fetal fibroblasts was 16% for site 3, whereas the data in Ref. 11 indicate a value at this site that was 3-fold higher or 50% (Fig. 3). The levels of hemimethylation were markedly lower than those reported in Ref. 11 with the exception of site 6 for which a low level of hemimethylation is consistent between both studies (Fig. 3).

We also analyzed DNAs from peripheral blood leukocytes, lymphoblastoid cells, sperm, and previously established adult fibroblasts (9). Unlike fetal fibroblast DNAs, these other DNAs did not have molecular barcodes incorporated within the hairpin linker. Each sequence included in our analysis did, however, have a unique methylation pattern (for example, see Fig. 3 in Ref 9), providing confidence that contamination and limited-template redundancy did not bias these data. DNAs from these other tissues and cell types had levels of hemimethylation ranging from 4.8 to 20.8% (Table I).

Our hairpin-bisulfite PCR data on L1 sequences also provide an opportunity to assess the methylation status of cytosines that are not within CpG dinucleotides. For the 104-bp region, we confirm the observation of Woodcock *et al.* (11) that non-CpG cytosines are well-converted and therefore unmethylated in DNA from fetal fibroblasts (11). In this region, we observed seven unconverted non-CpG cytosines in 57 well-converted sequences (>98.2% conversion of non-CpG cytosines) with no site represented more than once (Fig. 4*A*). We conclude that methylation of non-CpG cytosines in this region is rare.

We were also interested in the possibility of non-CpG cytosine methylation in the 118-bp promoter region. This region includes two non-CpG cytosines on the sense strand at nucleotide positions 432 and 436 that were reported to be non-converted and hence possibly methylated in most clones analyzed by conventional bisulfite conversion and PCR (11). We analyzed 64 well-converted sequences (>98.6% conversion of non-CpG cytosines) that include these two nucleotides from a variety of cell types. We observed eight unconverted non-CpG cytosines distributed throughout the region with only one site

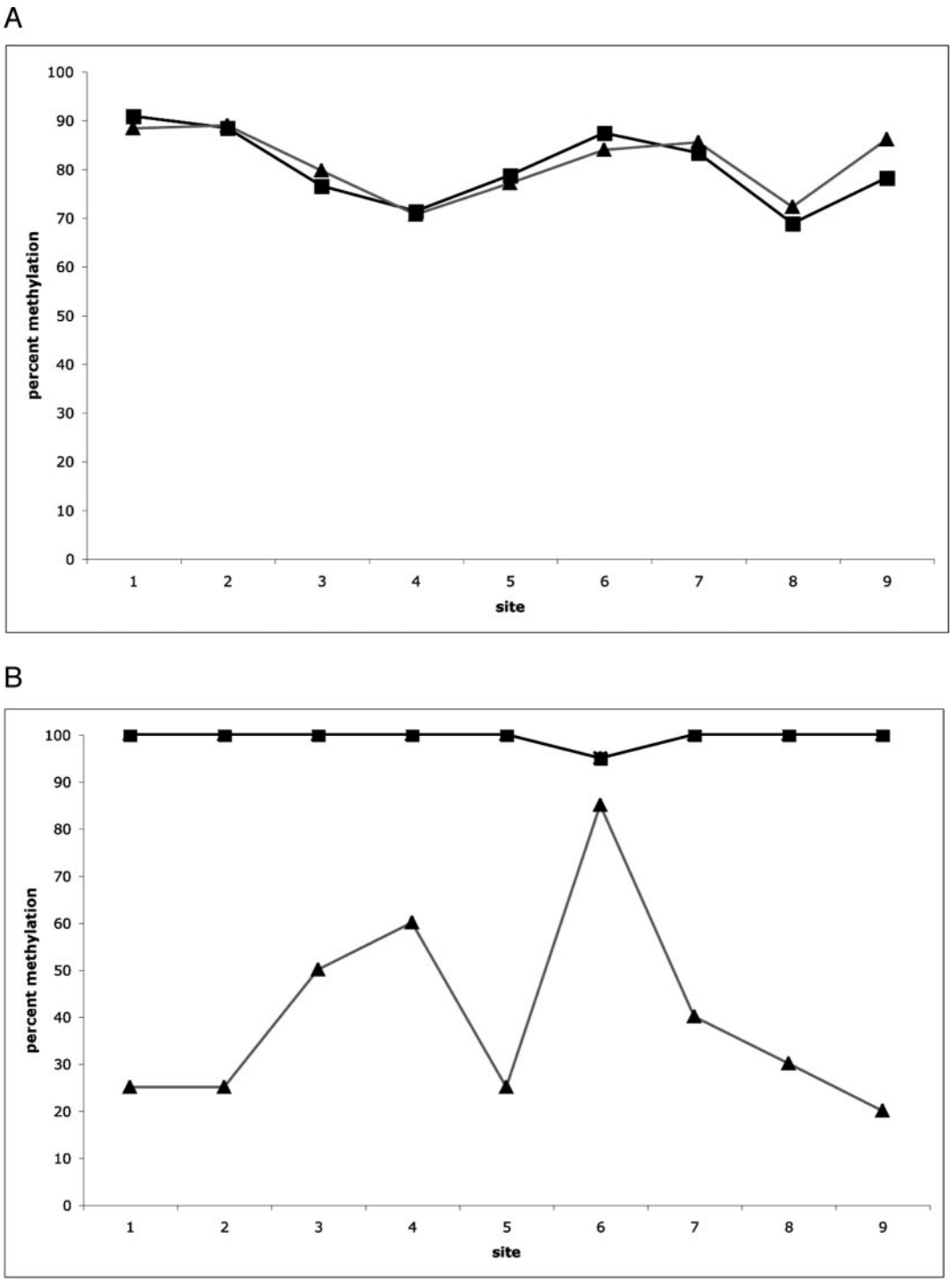

FIG. 2. **Average percent methylation on sense and antisense strands at nine different sites within a 104-bp region of the L1 promoter.** *A*, DNA from four passages of human fetal fibroblasts amplified by barcoded hairpin-bisulfite PCR. 15 sequences from passages 10, 14, and 17 were analyzed. 12 sequences were analyzed for passage 22. In these sequences, the number of CpG dyads that are evolutionarily conserved relative to the consensus from Ref. 11 are 126, 129, 123, and 93, respectively. Average methylation values for the nine sites on sense-strand (■) and antisense-strand (▲) sequences were 80.4 and 81.3%, respectively. *B*, methylation levels reported by Woodcock *et al.* (11) for these nine sites. These data were obtained by independent PCR of sense and antisense strands and were estimated from Fig. 3 of Ref. 11.

(nucleotide 401) represented more than once. None of these eight unconverted cytosines was at site 432 or 436 (Fig. 4*B*). We also examined 14 L1 sequences of this region from human sperm DNA that were chosen for analysis because of their lower levels of conversion efficiency (90–98% conversion, see "Materials and Methods" for the basis of these lower conversion efficiencies). 50 unconverted non-CpG cytosines were observed in these 14 sequences, occurring most often in the region with three closely spaced CpGs. Only one non-conversion event was observed at site 432. None was observed at site 436 (Fig. 4*C*). We conclude that methylation of non-CpG cytosines is rare in this region of the human L1 promoter, including at nucleotides 432 and 436.

DISCUSSION

In our initial study of methylation of human L1 sequences using hairpin-bisulfite PCR, we analyzed CpG methylation in the published human L1 consensus sequence (11), focusing on

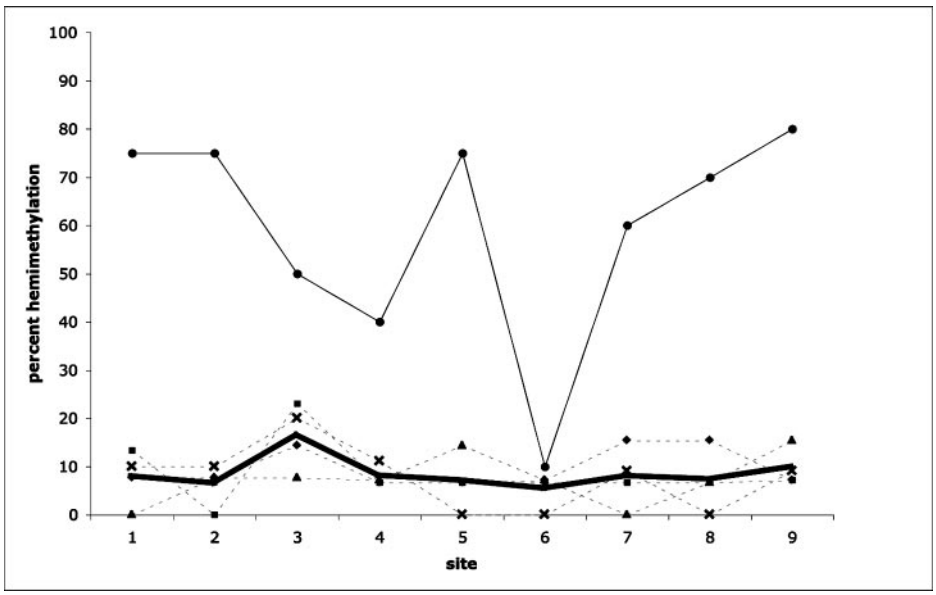

FIG. 3. **Percent hemimethylation of L1 promoter sites within a 104-bp region in DNA from four passages of human fetal fibroblasts.** Fibroblasts from a 16-week female fetus obtained from Coriell Cell Repository were cultured as described under "Materials and Methods." An estimated 2–3 cell doublings occurred between successive passages with DNA from passage 10 (♦) representing the earliest sample and DNA from passage 22 (×) representing the last sample from these cells. Hemimethylation data for DNA from intermediate passages 14 (■) and 17 (▲) are also shown as well as the average hemimethylation value (*boldface line*) for the combined data. The percent of hemimethylation at the same nine CpG sites estimated from Fig. 3 of Woodcock *et al.* (11) is displayed for comparison (●).

TABLE I
Percent CpG hemimethylation of a 104-bp L1 promoter region in human DNAs

Percent hemimethylation in human DNAs from different cell types using hairpin-bisulfite PCR of a 104-bp CpG-rich region within the L1 promoter. Fetal fibroblast samples had molecular barcodes. 12–17 sequences were analyzed for each sample. Percent hemimethylated dyads was calculated by dividing the number of hemimethylated CpG dyads by the total number of consensus sequence dyads (Ref. 11) among the sequences from an individual sample.

DNA sample	Number informative CpG dyads (sequences)	Percent hemimethylated CpG dyads
Male adult fibroblast line (82-6 hTERT)	120 (17)	10.0
Female adult fibroblast line (81-58A)	128 (16)	14.1
Male leukocyte DNA (A102)	120 (17)	20.8
Sperm (A47)	125 (13)	4.8
Female lymphoblastoid DNA (A54)	99 (15)	6.1
Fetal fibroblast passage 10	126 (15)	9.5
Fetal fibroblast passage 14	129 (15)	8.5
Fetal fibroblast passage 17	123 (15)	7.4
Fetal fibroblast passage 22	93 (12)	7.5

a 104-bp promoter region (9). We reported that hemimethylation averaged 12.1% of CpG dyads for two different adult fibroblast lines. Here we apply new methods and present more extensive data expanding on questions of hemimethylation within the 104-bp promoter region of human L1s from fetal fibroblasts. We also address the question of non-CpG methylation within the 104-bp region and within an adjacent 118-bp region of the human L1 promoter.

Technical Aspects of Hairpin-Bisulfite PCR and Distinguishing Valid L1 Sequences from Redundant and Contaminant Sequences Using Molecular Barcodes—We can be confident that the sense- and antisense-strand methylation patterns presented here represent information from a single allele and locus, even when a sequence is repeated many times in the human genome as is the case for L1 sequences. The hairpin-bisulfite PCR technology provides information on methylation patterns on the complementary strands of individual DNA molecules. The complementary strands remain attached via the hairpin linker throughout the bisulfite conversion and PCR amplification, cloning, and sequencing steps. Thus, this method eliminates the problems of biases in sense and antisense methylation analysis because both strands are amplified in a single step (Fig. 1).

Two technical considerations reinforce the conclusion that the inferred methylation patterns from our hairpin-bisulfite methylation analyses, including the hemimethylated dyads, are an accurate reflection of patterns on both strands of individual L1 sequences. As described under "Materials and Methods," the frequency of hemimethylated dyads is considerably higher than expected from inefficient bisulfite conversion. Additionally, sequences that are the likely result of cross-over PCR are extremely rare. If template switching were to occur during PCR amplification, we would observe sequences in which the sense and antisense strands were not precisely complementary because of the sequence divergence among members of the L1 family. In our initial publication (9) on hairpin-bisulfite PCR, we reported that no case of detectable cross-over PCR had been observed in >200 analyzed L1 sequences. Our most recent analysis of 636 L1 sequences that met our criteria for efficient bisulfite conversion revealed four sequences that were probably the result of cross-over PCR. These four sequences were not included in further analyses. We conclude that the rate of cross-over PCR among analyzed L1 sequences is somewhat <1/100 sequences. Thus, the frequencies of hemimethylated dyads reported here are unlikely to be significantly inflated by problems of inefficient bisulfite conversion or cross-over PCR.

The fetal fibroblast sequences presented here (Fig. 1) have unique barcodes, thus verifying that they are appropriately amplified sequences rather than re-amplified products or contaminants. We introduced molecular barcodes for hairpin-bisulfite PCR of a single-copy gene, *FMR1*, to distinguish valid sequences from redundant examples of an allele that had previously been obtained as well as from contaminant sequences (16). Similar use of molecular barcodes is reported here for the family of repeated L1 elements. The molecular barcodes have in most cases identified valid sequences that also had distinctive methylation patterns. Contamination from previous PCR

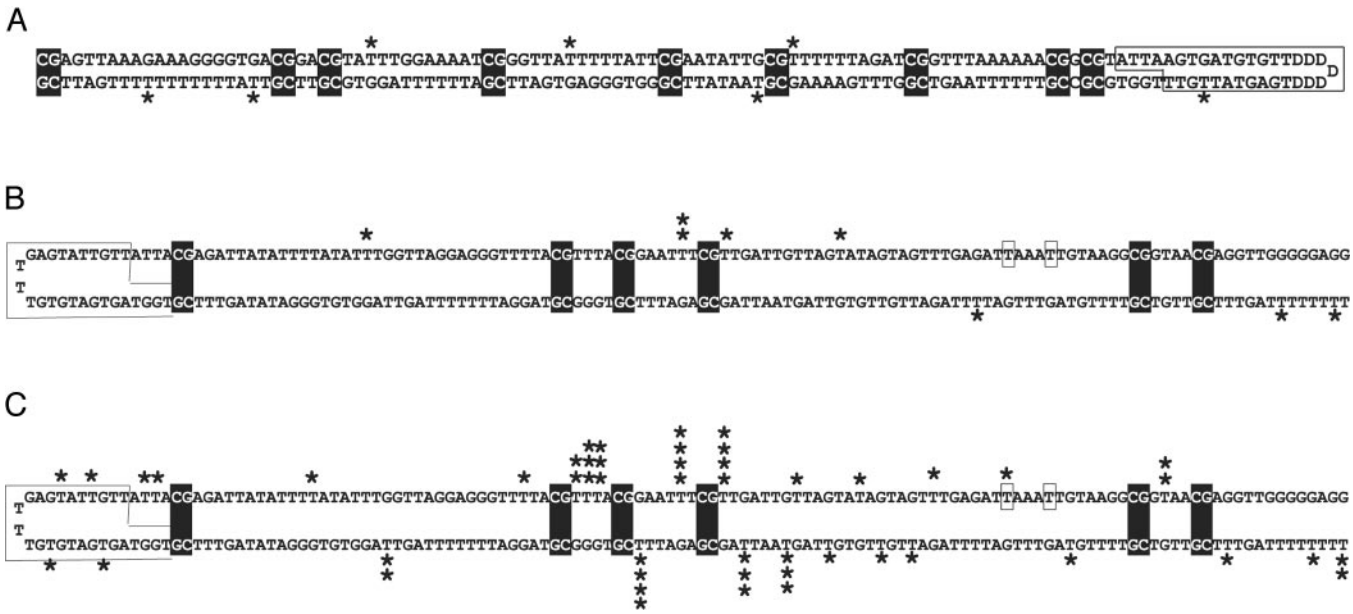

FIG. 4. **Hairpin-bisulfite conversion of non-CpG cytosines in L1 promoter regions.** *A*, conversion of non-CpG cytosines in the 104-bp region from 57 well-converted (≥98.2%) fetal fibroblast DNA sequences. We observed only seven unconverted non-CpG cytosines (*asterisk*) of the possible 3249 opportunities for non-conversion. All seven events were at sites represented only once. Potentially methylated CpG sites from the consensus are *black*, and the hairpin linker is *boxed* on the *right*. *B*, conversion of non-CpG cytosines in the 118-bp region from 64 well converted (≥98.6%) DNA sequences from a variety of human cell types. This region includes two non-CpG cytosines (nucleotides 432 and 436, *boxed* on *upper strand*) that were previously reported to be consistently unconverted and hence possibly methylated (11). We observed only eight unconverted non-CpG cytosines (*asterisk*) of the possible 4736 opportunities for non-conversion. Only one site was represented more than once. None of the eight unconverted cytosines was at site 432 or 436. Of these 64 sequences, 33 are from adult fibroblasts, 16 are from adult blood leukocytes, and 15 are from an established lymphoblastoid line. Potentially methylated CpG sites from the consensus are *black*, and the hairpin linker is *boxed* on the *left*. *C*, conversion of non-CpG cytosines in the 118-bp region from 14 less-well converted (90–98%) human sperm DNA sequences. Lower levels of conversion in these sequences were created experimentally as described under "Materials and Methods." 50 unconverted non-CpG cytosines (*asterisk*) were observed. Only one non-conversion event was observed at site 432. None was observed at site 436.

reactions is identified by redundant matching barcodes in PCR products and, in most cases, by the identity of the inferred methylation pattern. For completely methylated or unmethylated sequences, however, the molecular barcodes are indispensable in determining whether a sequence is from an L1 allele not previously sampled. This information is crucial for accurate quantification of various methylation states among alleles and cells, as exemplified in Fig. 1, and described more fully under "Results."

Levels of Hemimethylation in L1 Sequences from Fetal Fibroblasts in Cell Culture—The most frequently recovered L1 sequences from fetal fibroblasts were densely methylated with primarily concordant methylation patterns at consensus CpG dyads (Fig. 1). From our larger data set of 471 CpG dyads (57 L1 sequences) from fetal fibroblast DNA, 91.7% ($n = 432$) were concordantly methylated and 8.3% ($n = 39$) were hemimethylated. Among these 57 sequences, only two (3.5%) were observed that did not fit the pattern of hypermethylation. Distinctive barcodes indicated that these two sequences were derived from different genomic L1s (Fig. 1). CpG dyads in these two sequences were strikingly concordant in their methylation status in that all of the dyads were fully unmethylated. Site-specific variations in methylation were observed in fetal fibroblast L1s, but the two complementary strands had approximately equal and highly correlated levels of methylation (Fig. 2*A*). Thus, we observe primarily concordant methylation in both hypermethylated and hypomethylated L1 sequences and low levels of hemimethylation at all of the nine CpG sites analyzed.

Our data on levels of hemimethylation in L1 sequences of fetal fibroblasts conflict with data reported by Woodcock *et al.* (11). These authors analyzed the methylation patterns of sense and antisense strands independently using bisulfite conversion and concluded that there were large differences between strand methylation. Specifically, Fig. 3 in Ref. 11 indicates that sense-strand methylation in their analyzed sequences was close to 100%, whereas antisense-strand methylation averaged ~40% (see Fig. 2*B*). If correct, these data would indicate that most of the CpG sites in this region (59%) were hemimethylated. In contrast, our data indicate that sense- and antisense-strand methylation in this region are highly concordant with an average level of hemimethylation of only 8.3% (Fig. 2*A*). The levels of hemimethylation in L1s from these fetal fibroblasts were not significantly different between samples from the four passages. Thus, it seems unlikely that different stages of passaging of fetal fibroblasts in tissue culture would explain these different estimates of L1 hemimethylation (Fig. 3).

The validity of conclusions drawn from the independent analysis of the two complementary strands of DNA relies on the assumption that the independent PCR amplification and cloning of the complementary strands is unbiased and that the separate amplifications of sense and antisense strands draw from the same population of previously double-stranded molecules. This assumption is not always valid. Warnecke *et al.* (18), for example, provided evidence of strand bias for the retinoblastoma tumor suppressor gene when using single-strand PCR. Further studies indicated that unmethylated T-rich DNA had been amplified preferentially, thus leading to a PCR bias and to an inaccurate estimate of methylation (19). Variations in methylation among individual L1 elements, including a low percentage of hypomethylated alleles (Fig. 1), provide ample opportunities for observing apparent strand difference under conditions of strand-biased PCR. Efforts to control for methylated *versus* unmethylated bias using cloned sequences (11) may not be sufficient within the context of single-strand DNA amplification of repeated sequences that vary in sequence as well as methylation status.

Methylation of L1 Sequences from Different Cell Types—We asked whether or not levels of hemimethylation of CpG dyads

in L1s from different cell types are comparable with those that we observe from cultured fetal fibroblasts. L1 sequences from DNAs of other human cells had levels of hemimethylation that spanned the average observed for fetal fibroblasts ranging from 4.8% in sperm to 20.8% in a leukocyte sample. Although these hemimethylation values show considerable variation, none approaches the high level implicit in data from Woodcock *et al.* (11).

Methylation Status of Non-CpG Cytosines in the L1 Promoter—Methylated cytosines that are not within CpG contexts are occasionally reported (11). Consistent methylation at such non-CpG cytosines, if verified, would be interesting in part because the mechanism of methylation inheritance would be different from that for CpG methylation, which depends on the symmetry of CpG/CpG dyads (4–6). For the 104-bp region of the human L1 promoter, we confirmed the observation of Woodcock *et al.* (11) that non-CpG cytosines in this region are rarely, if ever, methylated in human fetal fibroblasts (Fig. 4*A*).

These authors (11) did, however, report that two non-CpG cytosines within the 118-bp region at nucleotides 432 and 436 of the human L1 promoter are consistently unconverted and hence are possibly methylated. We find no support for this conclusion in our hairpin-bisulfite PCR data. In 64 well-converted sequences (>98.6% conversion), from two adult fibroblast lines, leukocytes, and lymphoblastoid cells, we found that these two cytosines were always converted and hence unmethylated (Fig. 4*B*). To explore the possibility that we missed non-CpG cytosine methylation by selecting only well-converted sequences, we examined a small number of less-well converted L1 sequences from human sperm DNA. In 14 sequences with 90–98% conversion of non-CpG cytosines, only one instance of non-conversion at site 432 was observed for these two sites (Fig. 4*C*). Other unconverted cytosines in these poorly converted sequences occurred most often in the region with three closely spaced CpGs (Fig. 4*C*). Regions with high CG content, especially when hypermethylated, are refractory to denaturation, which will inhibit bisulfite conversion (17). Although our data on these two cytosines were not from fetal fibroblasts, DNAs from two adult fibroblasts lines and from two other cell types were examined. It is possible that L1s in fetal fibroblasts are different from other L1s, but we consider it more likely that incomplete conversion at these sites and subsequent PCR redundancy explain the previous conclusion of methylation of several specific non-CpG cytosines (11). Therefore, we conclude that non-CpG cytosines in these regions are rarely, if ever, methylated including at nucleotides 432 and 436.

Concluding Remarks—For human L1 promoter sequences, our results resolve a controversy in the literature. Our data are consistent with high levels of concordant CpG methylation on complementary strands, hence with low levels of hemimethylation inferred by Hansen (3) and Laird *et al.* (9), and are inconsistent with high levels of hemimethylation reported by Woodcock *et al.* (11). Data from a variety of cell types, including cultured fetal fibroblasts, support this conclusion. We also report that two non-CpG cytosines within the L1 promoter reported by Woodcock *et al.* (11) to remain unconverted during bisulfite conversion are routinely converted in our hairpin-bisulfite PCR procedure in DNAs from different cell types. We conclude that these and all of the other non-CpG cytosines in the regions analyzed are rarely, if ever, methylated. The methods of analysis used here for L1 sequences are applicable to other repeated sequences for which methylation analyses on complementary strands are especially difficult. The application of these experimental approaches will likely provide insights into the origins and consequences of variations in methylation patterns.

Acknowledgments—We thank Megan McCloskey, Diane Genereux, Brooks Miner, Reinhard Stöger, and Jessica Sneeden for helpful suggestions.

REFERENCES

1. Thayer, R. E., Singer, M. F., and Fanning, T. G. (1993) *Gene* (*Amst.*) **133,** 273–277
2. Hata, K., and Sakaki, Y. (1997) *Gene* (*Amst.*) **189,** 227–234
3. Hansen, R. S. (2003) *Hum. Mol. Genet.* **12,** 2559–2567
4. Bestor, T. H., and Ingram, V. M. (1983) *Proc. Natl. Acad. Sci. U. S. A.* **80,** 5559–5563
5. Pfeifer, G. P., Steigerwald, S. D., Hansen, R. S., Gartler, S. M., and Riggs, A. D. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87,** 8252–8256
6. Hermann, A., Goyal, R., and Jeltsch, A. (2004) *J. Biol. Chem.* **279,** 48350–48359
7. Wigler, M. H. (1981) *Cell* **24,** 285–286
8. Otto, S. P., and Walbot, V. (1990) *Genetics* **124,** 429–437
9. Laird, C. D., Pleasant, N. D., Clark, A. D., Sneeden, J. L., Hassan, K. M., Manley, N. C., Vary, J. C., Jr., Morgan, T., Hansen, R. S., and Stoger, R. (2004) *Proc. Natl. Acad. Sci. U. S. A.* **101,** 204–209
10. Saluz, H. P., Jiricny, J., and Jost, J. P. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83,** 7167–7171
11. Woodcock, D. M., Lawler, C. B., Linsenmeyer, M. E., Doherty, J. P., and Warren, W. D. (1997) *J. Biol. Chem.* **272,** 7810–7816
12. Zhou, Y., Magill, C. W., Magill, J. M., and Newton, R. J. (1998) *Genome* **41,** 23–33
13. Singal, R., and vanWert, J. M. (2001) *Blood* **98,** 3441–3446
14. Poncz, M., Solowiejczyk, D., Harpel, B., Mory, Y., Schwartz, E., and Surrey, S. (1982) *Hemoglobin* **6,** 27–36
15. Edwards, M. J., Wenstrup, R. J., Byers, P. H., and Cohn, D. H. (1992) *Hum. Mutat.* **1,** 47–54
16. Miner, B. E., Stoger, R. J., Burden, A. F., Laird, C. D., and Hansen, R. S. (2004) *Nucleic Acids Res.* **32,** e135
17. Grunau, C., Clark, S. J., and Rosenthal, A. (2001) *Nucleic Acids Res.* **29,** E65
18. Warnecke, P. M., Stirzaker, C., Melki, J. R., Millar, D. S., Paul, C. L., and Clark, S. J. (1997) *Nucleic Acids Res.* **25,** 4422–4426
19. Stirzaker, C., Millar, D. S., Paul, C. L., Warnecke, P. M., Harrison, J., Vincent, P. C., Frommer, M., and Clark, S. J. (1997) *Cancer Res.* **57,** 2229–2237