Proc. Natl. Acad. Sci. USA
Vol. 94, pp. 2545–2550, March 1997
Medical Sciences

DNA methylation and genetic instability in colorectal cancer cells

CHRISTOPH LENGAUER, KENNETH W. KINZLER, AND BERT VOGELSTEIN*

Howard Hughes Medical Institute at The Johns Hopkins Oncology Center, 424 North Bond Street, Baltimore, MD 21231

Contributed by Bert Vogelstein, December 31, 1996

ABSTRACT **Apparent alterations in DNA methylation have been observed in many cancers, but whether such alterations represent a persistent alteration in the normal methylation process is not known. In this study, we report a striking difference in the expression of exogenously introduced retroviral genes in various colorectal cancer cell lines. Extinguished expression was associated with DNA methylation and could be reversed by treatment with the demethylating agent 5-azacytidine. A striking correlation between genetic instability and methylation capacity suggested that methylation abnormalities may play a role in chromosome segregation processes in cancer cells.**

DNA methylation is one of the most intriguing chemical phenomena affecting the genome. It is essential in prokaryotes, dispensable in lower eukaryotes such as *Saccharomyces cerevisiae*, yet present and presumably important in mammals (1, 2). As gene expression is abnormal in cancer and DNA methylation plays a role in controlling gene expression (3), it was supposed that aberrant DNA methylation might be responsible for expression differences (4). Accordingly, many cancers were shown to have a global hypomethylation of DNA compared with normal tissues (5–7). Treatment of cells or animals with 5-azacytidine (5-aza-C), a demethylating agent that irreversibly inactivates the methyltransferase (8–10), was shown to be oncogenic *in vitro* and *in vivo* (11). Conversely, other studies have shown that hypermethylation of specific sequences is found in some tumors and can be associated with the inactivation of tumor suppressor gene expression (12–14). Mice genetically deficient in methyltransferase were found to be resistant to colorectal tumorigenesis initiated by mutation of the *APC* tumor suppressor gene, and treatment of these mice with 5-aza-C enhanced the resistance (15).

One of the impediments to the interpretation of any of these somewhat conflicting studies is the absence of knowledge about DNA methylation in the precursor cells of the studied cancers. Colorectal tumors, for example, arise from stem cells that form only a minor component of the total epithelial population. Comparison of DNA methylation patterns in colorectal tumors to those of total colorectal epithelium may therefore yield misleading results, as there is no guarantee that the methylation pattern in the stem cells is identical to that of their much more numerous progeny. Thus, statements about "hypo" or "hyper" methylation must be viewed cautiously. Moreover, the mechanisms responsible for the putative altered methylation in tumors, and the time at which such aberrant methylation occurs, are obscure. In particular, whether the putative alteration in DNA methylation represents a historical event that occurred during tumorigenesis rather than a persistent defect in the process of methylation is unknown.

In the present study, we have made some unexpected observations which shed light on these issues. While studying retrovirus-mediated gene expression, we noted a qualitative difference among colorectal cancer cell lines which we later tied to DNA methylation. Because all colorectal cancer cell lines are presumably derived from the same stem cell precursor, the methylation differences observed were likely to be related to the tumorigenic process rather than to differences in cellular origin. Moreover, because *de novo* methylation was assessed, we could conclude that the differences represented alterations that persisted throughout the lifetime of the tumor cell. Furthermore, a striking correlation between genetic instability and methylation capacity suggested that methylation abnormalities might play a role in chromosomal segregation processes in cancer.

MATERIALS AND METHODS

Cell Culture. Cells were cultured in McCoy's 5A medium (modified) supplemented with 10% fetal bovine serum (HyClone), 100 units/ml penicillin, and 0.1 mg/ml streptomycin. Monolayer cultures were grown at 37°C in a 5% CO_2 atmosphere. The cell lines Caco2, Colo205, DLD1, HCT116, HT29, LoVo, SW480, SW48, and SW837 were obtained from the American Type Culture Collection (Rockville, MD), and RKO was generously provided by M. Brattain (Baylor College of Medicine, Houston, TX). The status of these lines with respect to mismatch repair capacity has been analyzed previously (16–21).

Vector Construction and Retrovirus Production. The hygromycin-resistance gene (*HygR*) was PCR amplified from the pCEP4 plasmid (CLONTECH) using the primers 5′-CGA TCA AGC TTC CGT GTT TCA GTT AGC-3′ and 5′-CAT GCG ATA TCC ACC ATG AAA AAG CCT GAA CTC AC-3′ (DNAgency, Malvern, PA) at an annealing temperature of 63°C. The PCR fragment was gel purified using a Spin-X centrifugal filter (Costar), digested with *Hin*dIII and *Eco*RV, and inserted into the *Hin*dIII and *Stu*I sites of the plasmid pGal/Neo (a derivative of G1BgSVNa kindly provided by Genetic Therapy Inc., Bethesda, MD) to yield plasmid pGal/Hyg. Retroviral vector DNA was prepared by cesium chloride banding and transfected into the ecotropic retrovirus packaging cell line BOSC 23 (22). The virus supernatant was harvested 48 hr after transfection and used to infect the amphotropic packaging line PA317 (23). Colonies were selected for 21 days with hygromycin C (Calbiochem). Stable producer clones were cultivated, and virus supernatants were titered (10^6 colony-forming units/ml) and used for infection.

Transfection and Virus Infection. Plasmid DNA was transfected using Lipofectin (GIBCO/BRL), and colonies were selected for 23 days with the neomycin analog G418 (GIBCO/BRL). For virus infection, subconfluent cells were treated with 8 μg/ml Polybrene and infected with the retrovirus Gal/Neo (derived from the retroviral construct G1BgSvNa, Genetic

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

Copyright © 1997 by THE NATIONAL ACADEMY OF SCIENCES OF THE USA
0027-8424/97/942545-6$2.00/0
PNAS is available online at **http://www.pnas.org**.

Abbreviations: 5-aza-C, 5-azacytidine; β-gal, β-galactosidase; X-Gal, 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside; LTR, long terminal repeat, MMR, mismatch repair.

*To whom reprint requests should be addressed.

Downloaded from https://www.pnas.org by 122.171.142.147 on January 29, 2025 from IP address 122.171.142.147.

Therapy, Bethesda, MD) or the retrovirus Gal/Hyg at a multiplicity of <0.01. Clones containing retroviral integrations were identified by selection with G418 or hygromycin, respectively. After 3 weeks of selection, individual clones were isolated by limiting dilution and assayed for β-galactosidase (β-gal) expression *in situ*. Cells were fixed and stained with 5-bromo-4-chloro-3 indolyl-β-D-galactopyranoside (X-Gal; Sigma). All infection and staining procedures were performed in parallel on the various lines and repeated at least three times.

Drug Treatment of Cells. Cells were grown in medium containing different concentrations (5–20 μM) of 1-β-D-arabinofuranosylcytosine or (1–10 μM) of 5-aza-C (Sigma) for 72 hr. Drug-treated cells for DNA and RNA isolation were grown in selective media containing 5 μM 5-aza-C.

PCR and Transcription–Translation Reactions. Primer pairs used for genomic amplifications of β-gal sequences were 5′-GGA TCC TAA TAC GAC TCA CTA TAG GGA GAC CAC CAT GGT TAC GGA TTC GGA TCC-3′ with 5′-CTG TTT ACC TTG TGG AGC G-3′ and 5′-GGA TCC TAA TAC GAC TCA CTA TAG GGA GAC CAC CAT GGC TGT GCC GAA ATG GTC C-3′ with 5′-CCG GAG AAC CTG CGT GC-3′, yielding products 2030 bp and 1995 bp in length, respectively. PCR products were used as templates in coupled transcription–translation reactions (Promega). The resultant ^{35}S-labeled proteins were separated by SDS/polyacrylamide gel electrophoresis (Tris–glycine gel) as described by Powell *et al.* (24).

Southern Blot Analysis. DNA samples (5 μg) were separated by electrophoresis on a 0.7% agarose gel and transferred to a Zeta-Probe membrane (Bio-Rad). The blot was probed with a gel-purified 3594-bp *Bam*HI fragment of the β-gal gene derived from the plasmid pCMVβ (CLONTECH), labeled by random priming with [^{32}P]dCTP (25).

RNA Isolation and Northern Blot Analysis. Total RNA was prepared by CsCl gradient ultracentrifugation of guanidine isothiocyanate-lysed cells. Northern blot analysis was performed as described (26), except that QuickHyb (Stratagene) was used for hybridization. As a quantitation control, a gel-purified 300-bp *Pst*I/*Hin*dIII fragment of the glyceraldehyde-3-phosphate dehydrogenase (GAPDH) gene derived from the plasmid pHcGAP (ATCC) (27) was used for rehybridization after stripping filters with 0.1× SSC/0.1% SDS at 100°C for 20 min.

Bisulfite Genomic Sequencing. Bisulfite treatment of genomic DNA was performed as described by Frommer *et al.* (28) and modified by Clark *et al.* (29). For PCR amplification of the bisulfite-modified integrated viral 5′ long terminal repeat (LTR), the primers 5′-GAA TAG AAA AGT TTA GAT TAA GG-3′ and 5′-CAA AAC TAA TTA ACT AAC TAA TAC-3′ (DNAgency) were used, resulting in a 668-bp product. The original sequences corresponding to these primers did not contain 5′-CpG-3′ dinucleotides. Amplified DNA was gel purified and blunt-end-ligated into the *Eco*RV-digested plasmid pZErO-1 (Invitrogen). After transformation, individual colonies were picked, and genomic DNA was prepared and sequenced using the internal primer 5′-CCC ART AAT AAA TCA ATA ATC C-3′. The analysis was performed on HCT116–1 cells before and after 5-aza-C treatment, and on HT29–6 cells. At least 12 DNA clones, each of whose deoxycytosines (when not in a CpG context) were completely converted to uridines, were sequenced for each line.

RESULTS

Expression of Retroviral Sequences. Previous studies of DNA methylation in cancer have largely evaluated sequences already present in the genome of cancer cells. Because methylation is heritable, such studies cannot be used to determine if the process of *de novo* methylation is abnormal at the time of testing. To develop an assay for this purpose, we chose to examine sequences introduced into the cancer cell by retroviral infection. Once integrated, such viral sequences can be methylated (30–32), and preliminary experiments in our laboratory suggested that the expression of such sequences in human cancer cell lines was variable, perhaps reflecting differences in *de novo* methylation (unpublished data). The virus used for these experiments contained the β-gal reporter driven by a retroviral LTR and the G418 resistance gene (*NeoR*) under the control of the simian virus 40 early promoter.

Ten colorectal cancer cell lines were infected at low multiplicity, and G418-resistant clones were selected and stained *in situ* with X-Gal to evaluate the expression of β-gal. HCT116 and HT29 cells exemplify the two distinct patterns of expression observed. Virtually all HT29 clones stained intensely with X-Gal, indicating that the reporter was expressed in the selected cells. In contrast, many HCT116 clones were not stained at all, and even in those colonies which did exhibit staining, only a small fraction of the cells in each clone was stained (Fig. 1). This observation was repeated six times in the two lines, with virtually identical results. All clones contained the retroviral genome, including the *Neo* and β-gal sequences, as indicated by the G418 resistance and confirmed by Southern blots (see later).

Each of the ten lines studied showed one of these two easily distinguishable patterns of β-gal expression (Table 1, top 10 lines). In five lines (HT29, SW480, SW837, Colo 205, and Caco 2), the great majority of cells resistant to G418 were stained. In the other five lines (HCT116, LoVo, SW48, RKO, and DLD1) most cells were unstained. A striking correlation was noted between the propensity of the cells to express exogenous sequences and their proficiency in mismatch repair. All five non-expressing lines were mismatch repair (MMR) deficient, while the five β-gal-expressing lines were normal in this regard. Previous studies have demonstrated that LoVo has an inactivating mutation of the *mutS* homolog *hMSH2* (19, 21), DLD1 has an inactivating mutation of the *hMSH2* partner *GTBP* (17), SW48 and HCT116 have mutations of the *mutL* homolog *hMLH1* (16, 21), and RKO is defective in either *hMLH1* or its

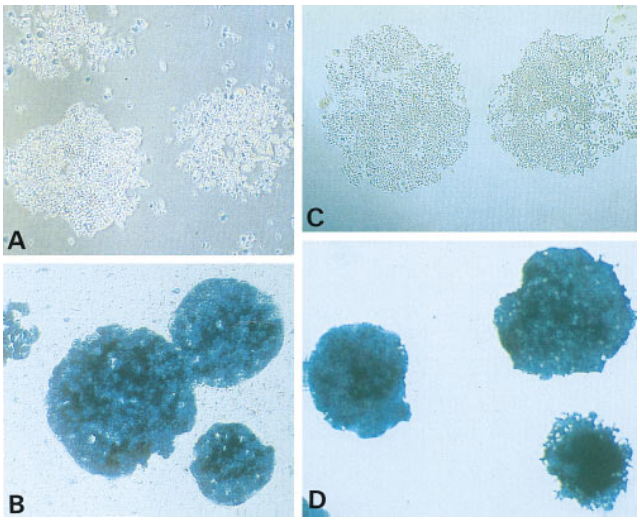

FIG. 1. Histochemical detection of β-gal expression. (*A* and *B*) After infection with the retrovirus Gal/Neo and selection with G418, cells were stained *in situ* with X-Gal to confirm the expression of β-gal. Whereas the majority of cells from individual clones of cell line HCT116 was not stained (*A*), virtually all cells of HT29 stained intensely blue (*B*). (*C* and *D*) Cells were stained *in situ* with X-Gal after retroviral infection with Gal/Hyg and selection with hygromycin. Whereas virtually all HT29 cells stained blue (*D*), the majority of cells from individual clones of HCT116+ch3 were not stained (*C*). (All ×10.)

Downloaded from https://www.pnas.org by 122.171.142.147 on January 29, 2025 from IP address 122.171.142.147.

Table 1. Expression of β-gal gene

Gene transfer agent	Cell line	Expression, % of clones: With 80–100% blue cells	With 50–80% blue cells	With 20–50% blue cells	With <20% blue cells	No. of clones evaluated	MMR status
Gal/Neo virus	HT29	93	4	1	2	190	Proficient
	SW480	91	7	0	2	122	Proficient
	SW837	95	2	0	3	500	Proficient
	Colo205	45	20	20	15	285	Proficient
	Caco2	60	19	5	16	74	Proficient
	HCT116	4	7	8	81	84	Deficient
	LoVo	2	2	6	90	214	Deficient
	SW48	10	6	12	72	405	Deficient
	RKO	10	8	12	70	318	Deficient
	DLD1	26	20	20	34	328	Deficient
Gal/Neo plasmid	HT29	12	7	8	73	464	Proficient
	HCT116	0	0	1	99	254	Deficient
Gal/Hyg virus	HT29	95	1	1	3	2602	Proficient
	HCT116	13	13	27	47	557	Deficient
	HCT116+ch3	11	11	20	58	670	Proficient

MMR-proficient and MMR-deficient cell lines were infected with the retrovirus Gal/Neo (lines 1–10) or were transfected with the plasmid used for Gal/Neo virus production (lines 11 and 12) and were selected with G418. Alternatively, cell lines were infected with the retrovirus Gal/Hyg (lines 13–15) and selected with hygromycin. Individual clones were stained with X-Gal to detect β-gal expression. For each clone the percentage of cells expressing β-gal (blue cells) was determined.

heterodimeric partner *hPMS2* (J. Drummond and P. Modrich, personal communication).

To examine the time course of β-gal expression, HT29 and HCT116 cells were examined at various intervals after retroviral infection. A similar number of cells in both lines (1.3–1.6%) were stained 22 hr after infection, demonstrating similar efficiencies of viral infection. Differences became apparent as early as 2 days after infection, and they were magnified over the time course of the experiment. After selection, virtually all surviving HT29 cells expressed β-gal, while only 4–7% of HCT116 cells expressed the gene (data not shown).

The decreased expression of the exogenously introduced β-gal gene in MMR-deficient lines was not restricted to genes introduced by retroviral transduction. It was also observed after DNA transfection with the plasmid used for virus production. Staining of transfected clones with X-Gal 3 weeks after G418 selection demonstrated β-gal expression in only 1 of 254 HCT116 clones analyzed, while 27% of HT29 clones contained a substantial number of blue cells (Table 1, lines 11 and 12).

Analysis of Retroviral Integrates. To test the possibility that the differences in expression of the exogenously introduced sequences between the MMR-proficient and -deficient cell lines were related to mutational processes affected by mismatch repair, we examined the organization and sequence of the retroviral integrates. Upon Southern blotting, each of 8 clones analyzed from four MMR-deficient lines (nonexpressors) and each of 14 clones from three MMR-proficient lines (expressors) was found to contain only one or two β-gal gene fragments, each of the intensity expected for single-copy genes. The absence of expression was therefore not due to a gross alteration of β-gal sequences in MMR-deficient cell lines.

The low fraction of β-galactosidase-expressing cells within G418-resistant clones seemed inconsistent with inactivation by point mutation. Even in MMR-deficient cells, inactivating mutations occur at a rate of only 10^{-4} to 10^{-5} per gene per generation (18, 33, 34), while over 80% of the selected MMR-deficient cells did not express β-gal (Table 1, top 10 lines). To formally exclude very high rates of mutation within β-gal sequences, we searched for truncating mutations. The β-gal genes in 13 clones from MMR-deficient cells were amplified with PCR, then transcribed and translated *in vitro*. Twelve of 13 clones generated polypeptides of the expected full-length size. Only one clone, LoVo-1, exhibited a truncated β-gal protein (data not shown). Interestingly, this clone exhibited absolutely no X-Gal-staining cells, while in each of the other 12 clones, a small fraction of cells exhibited such staining. Thus, nonsense mutations of β-gal were not likely the reason that some lines failed to express β-gal, suggesting an epigenetic basis.

Methylation and Transcription of β-gal Genes. To determine whether the integrated retroviral sequences were methylated differentially, Southern blots of *Sma*I-digested genomic DNA from the various clones were tested. *Sma*I is a methylation-sensitive restriction endonuclease that does not cleave DNA when the central cytosine of its recognition sequence, 5′-CCCGGG-3′, is methylated (35). DNA was digested with *Hin*dIII alone to reveal the restriction fragment containing the β-gal gene, or with *Hin*dIII plus *Sma*I to test methylation of this fragment. Comparison of the intensity of the *Hin*dIII fragment with that of the *Hin*dIII/*Sma*I fragment revealed that few, if any, of the HT29–6 cells were methylated at the *Sma*I site. In contrast, the *Sma*I digest of HCT116–1 DNA revealed methylation of the *Sma*I site, as demonstrated by its partial resistance to the restriction endonuclease (Fig. 2*A*).

Though methylation of the *Sma*I site was different in HCT116 cells than in the HT29 cells, the relatively low level of *Sma*I resistance did not seem to explain the dramatic difference in β-gal expression between these lines. However, as observed in previous studies of rodent cells (36, 37), the *Sma*I site is only one of several potential methylation sites within the viral 5′LTR. To get a more comprehensive view of the methylation of sequences surrounding the *Sma*I site, we sequenced genomic templates treated with bisulfite. The 214 bp surrounding the *Sma*I site contain 76% of the CpG dinucleotides in the 5′LTR, and we therefore concentrated on this region. The bisulfite genomic sequencing technique allows methylated cytosines to be distinguished through their resistance to bisulfite-mediated deamination; unmethylated cytosines are converted to uridine by bisulfite (28). There were 16 CpG sites within the 214 bp surrounding the *Sma*I site. In HT29 cells, no methylation at any of these CpG sites was found. In contrast, the 5′LTRs in HCT116 cells were associated with an average of 2.33 methylated CpG sites (range 0–9 sites per LTR; $P < 0.025$ by Student's t test). The distribution of methylated CpG residues in the 12 genomic 5′LTR's analyzed

Downloaded from https://www.pnas.org by 122.171.142.147 on January 29, 2025 from IP address 122.171.142.147.

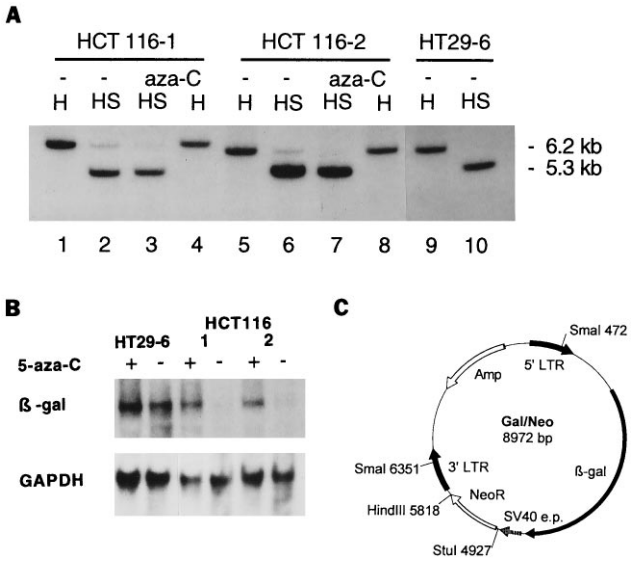

FIG. 2. Methylation of β-gal gene after retroviral infection. (*A*) Analysis of 5′LTR promoter by methylation-sensitive restriction enzyme. DNA from untreated (−) or 5-aza-C-treated cells was digested with *Hin*dIII alone (H) or with *Hin*dIII plus *Sma*I (HS). The blotted gel was probed with a [^{32}P]dCTP-labeled fragment of the β-gal gene. The *Hin*dIII fragment containing β-gal sequences was 6.2 kb in clone HT29–6 (lane 9), and *Sma*I reduced this fragment to 5.35 kb (lane 10), reflecting the invariant size separating the *Hin*dIII and *Sma*I sites in the provirus Gal/Neo. The *Sma*I digest of the HCT116 clones (lane 2 and 6) revealed two bands, indicating methylation of the *Sma*I site in these cells. Treatment of the HCT116 clones with 5-aza-C resulted in the removal of methyl groups, so that only a 5.35-kb fragment was detected after *Sma*I plus *Hin*dIII digestion (lanes 3 and 7). (*B*) Transcription of β-gal gene. Total RNA of the clones HCT116–1, HCT116–2, and HT29–6 was probed with a [^{32}P]dCTP-labeled fragment of the β-gal gene. The level of expression of β-gal mRNA in clones HCT116–1 and 2 was low (lanes 4 and 6) but increased substantially after 5-aza-C treatment (lanes 3 and 5). The 5-aza-C treatment had no effect on β-gal mRNA of clone HT29–6 (lanes 1 and 2). Rehybridization of the filter with a glyceraldehyde-3-phosphate dehydrogenase (GAPDH) probe confirmed that equal amounts of RNA were loaded. (*C*) Scheme of retroviral vector Gal/Neo. The Gal/Neo vector contains the β-gal reporter driven by a retroviral LTR and the G418-resistance gene (*NeoR*) under the control of the simian virus 40 early promoter (SV40 e.p.). Positions of the two *Sma*I and the single *Hin*dIII restriction sites are indicated. The restriction sites *Hin*dIII and *Stu*I were used for the construction of the provirus Gal/Hyg.

is shown in Fig. 3*B*, and examples of the sequencing data are in Fig. 3*A*. As a technical control for the efficiency of bisulfite treatment, it was noted that each of 52 deoxycytosines not in CpG contexts was completely converted to uridines in all of the HCT116 and HT29 clones evaluated.

Reversal of Methylation. The documentation of methylation of 5′LTR sequences reveals a correlation with β-gal expression that could not necessarily be regarded as causative. To obtain functional evidence that differences in methylation were responsible for the expression differences, we treated cells with 5-aza-C. This compound inactivates methyltransferase activity, reverses methylation, and restores retroviral expression in retrovirally infected rodent cells (8–10, 38). In 8 of 9 nonexpressor clones tested, 5-aza-C significantly increased the fraction of cells expressing β-gal in a dose-dependent manner (Fig. 4*A*). In contrast, in three expressor clones, the high level of β-gal expression was not affected by 5-aza-C (example in Fig. 4*A*). The single nonexpressor clone in which 5-aza-C had no effect was LoVo-1 (Fig. 4*A*). As mentioned above, LoVo-1 had a truncating mutation in the β-gal gene that should not be reversible by drug treatment.

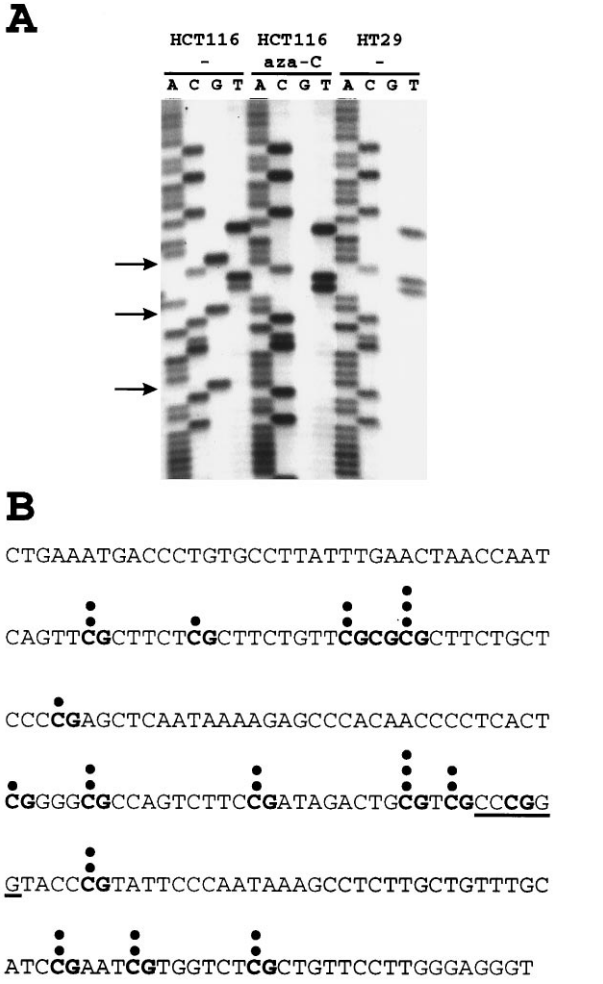

FIG. 3. Bisulfite genomic sequencing of 5′LTR. Genomic DNAs of infected cell lines were treated with bisulfite, converting deoxycytosine but not 5-methylcytosine residues into uracil through deamination. After treatment, the viral 5′LTR region was PCR-amplified, the PCR products were cloned, and a 214-bp region surrounding the *Sma*I site (see Fig. 2*C*) was sequenced. (*A*) Examples of clones HCT116–1 (before and after 5-aza-C treatment) and HT29–6 are shown. Arrows mark the locations of methylated CpG dinucleotides in HCT116–1. (*B*) Distribution of methylated CpG dinucleotides within the 5′LTR promotor of HCT116–1. The DNA sequence of the 214-bp region surrounding the *Sma*I site (underlined) is shown. Its 16 CpG dinucleotides are typed in boldface. The shaded area represents the region that is shown on the sequencing gel in *A*. Each dot represents a 5-methylcytosine residue that was not converted into uracil in one of the 12 cloned PCR products.

Treatment with 5-aza-C resulted in the removal of methyl groups, so that only a 5.3-kb fragment was found after *Sma*I plus *Hin*dIII digestion in clone HCT116–1 (Fig. 2*A*). Similar degrees of methylation of the *Sma*I site were observed in three different HCT116 clones tested, all reversible with 5-aza-C. No resistance to *Sma*I, either before or after 5-aza-C treatment, was observed in three HT29 clones similarly tested (example in Fig. 2*A*). This reduction in methylation was confirmed and quantitated by bisulfite genomic sequencing. The average number of methylated CpG dinucleotides was reduced by 84% after 5-aza-C treatment of HCT116 cells (0.37 methylated CpG sites per 5′LTR, range 0–3, $P < 0.020$ by Student's t test; example in Fig. 3*A*).

To determine whether the reversal of inactivation of β-gal expression by 5-aza-C was permanent, we removed the drug from the medium and measured β-gal expression at subsequent times. As shown in Fig. 4*B*, the clones reverted to the

Downloaded from https://www.pnas.org by 122.171.142.147 on January 29, 2025 from IP address 122.171.142.147.

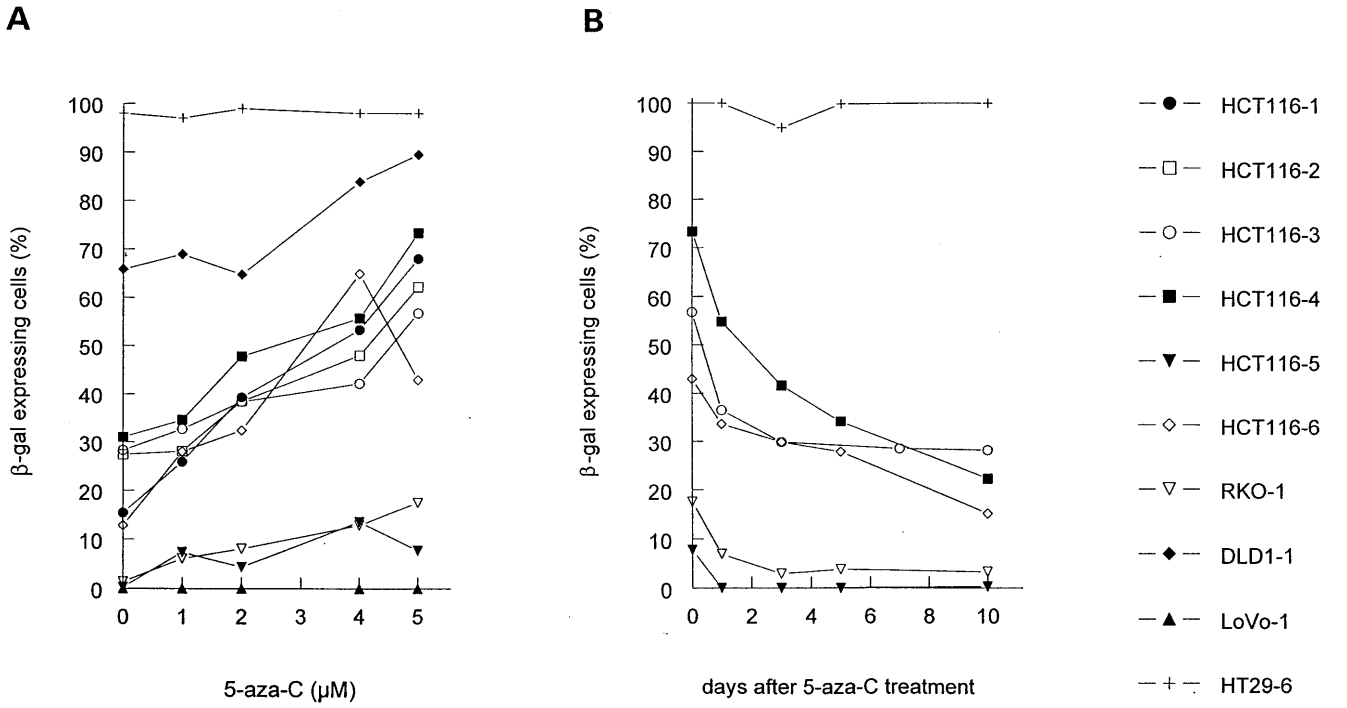

FIG. 4. Reversal of β-gal inactivation by 5-aza-C. (*A*) Nine MMR-deficient clones (HCT116–1, HCT116–2, HCT116–3, HCT116–4, HCT116–5, HCT116–6, RKO-1, DLD1–1, and LoVo-1) and the MMR-proficient clone HT29–6 were treated with different concentrations (0, 1, 2, 4, or 5 μM) of 5-aza-C for 72 hr and stained with X-Gal. In 8 of 9 MMR-deficient clones, 5-aza-C significantly increased the fraction of cells expressing β-gal in a dose-dependent manner. The single MMR-deficient clone in which 5-aza-C had no effect was LoVo-1. LoVo-1 had a truncating mutation in the β-gal gene (see text). (*B*) To determine whether the reversal of inactivation of β-gal expression by 5-aza-C was permanent, the drug (5 μM) was removed from the medium, and β-gal expression was measured at subsequent times. The clones reverted to the same basal level of β-gal expression found before 5-aza-C treatment.

same basal level of β-gal expression found before 5-aza-C treatment. To ensure that the reversion was not simply the result of cellular toxicity due to 5-aza-C treatment, we treated cells with another cytidine analog, the antimetabolite 1-β-D-arabinofuranosylcytosine. This drug exhibited toxicity at 10 μM as high as that observed with 6 μM 5-aza-C, but treatment with 1-β-D-arabinofuranosylcytosine at concentrations of 5, 10, or 20 μM for 4 days did not increase the fraction of cells expressing β-gal (data not shown).

Finally, to ensure that the drug affected expression at the RNA level, we performed Northern blot analysis. The level of expression of β-gal mRNA in each of three HCT116 clones tested was low but increased substantially after 5-aza-C treatment. The 5-aza-C treatment had no effect on β-gal mRNA in three HT29 clones tested (examples in Fig. 2*B*).

Methylation and Failure of MMR. The observations recorded above suggested that the absence of a MMR system might in some way lead to the methylation of exogenously introduced sequences. To test this hypothesis, we tested the HCT116 derivative (HCT116+ch3) in which the MMR deficiency had been corrected by transfer of a normal chromosome 3 containing the *hMLH1* gene missing in HCT116 cells (39). Western blot analysis showed, as expected, that the hMLH1 protein was expressed in the HCT116+ch3 cell line but missing in HCT116 (data not shown). Because the transferred chromosome 3 was tagged with a G418-resistance marker, the retrovirus Gal/Neo (Fig. 2*C*) could not be used and a new retroviral vector, Gal/Hyg, was constructed by replacing the *NeoR* gene of Gal/Neo with a hygromycin-resistance gene (*HygR*). The expression of β-gal in HT29 and HCT116 cells after infection with this virus mimicked that resulting from Gal/Neo infection (Fig. 1*D* and the last three lines of Table 1). The pattern of β-gal expression in the HCT116+ch3 line was very low, similar to the pattern obtained by infection of HCT116 and the other MMR-deficient lines (Fig. 1*C* and Table 1). These data demonstrate that *de novo* methylation of exogenously introduced sequences in the MMR-deficient lines is not due to the absence of MMR activity.

DISCUSSION

The methylation of exogenously introduced retroviral sequences is not a new observation (30–32, 36). The novelty of the observations reported here lies in the differences in methylation and expression of foreign genes in MMR-proficient versus MMR-deficient cell lines. This observation was completely unexpected and highly statistically significant ($P < 0.00001$ by χ^2). Some previous studies have suggested that methylation can be affected in various ways by DNA repair abnormalities, though the results were somewhat contradictory and no specific effect of MMR has been delineated (40, 41). Our experiments showed that MMR could not be directly responsible for the methylation differences, as correction of the MMR deficiency in HCT116 cells did not alter its ability to methylate exogenously introduced sequences.

Previous studies showing that the methylation of specific DNA sequences was different in tumor cells than in bulk preparations of "normal cells" could be interpreted in many ways. The tumor cells could have originated from stem cells that had different patterns of methylation than the bulk of the normal, differentiated cells used for comparison. Additionally, the tumor cells could have passed through some bottleneck in which methylation was altered, and this pattern could then have been transmitted to daughter cells through normal maintenance methylation. In contrast, the fact that methylation of exogenous sequences in colorectal cancer lines differed dramatically among the cell lines we tested shows that methylation abnormalities are not just historical events, but represent continuing physiologic differences that persist throughout the lifetime of the tumor cells.

Downloaded from https://www.pnas.org by 122.171.142.147 on January 29, 2025 from IP address 122.171.142.147.

Which of the two methylation patterns observed here represents the "normal" type likely to be found in nonneoplastic stem cell precursors? Though this question cannot be answered definitively at present, we feel it is likely that the methylation competence of the MMR-deficient tumor cells represents the physiologically normal state. This judgment is based on prior observations demonstrating that retroviral gene expression is often extinguished over time in normal tissues and that this extinction is associated with a methylation of the retroviral LTR similar to that described here (30, 31, 36, 38). According to this hypothesis, the MMR-deficient lines are methylation proficient (MMR^-, MET^+), while the MMR-proficient lines are methylation deficient (MMR^+, MET^-).

What is the basis upon which cells with a MMR^+, MET^- phenotype are selected during tumorigenesis? We propose the following speculative explanation. It has long been suspected that some sort of genetic instability is necessary for a tumor to accumulate the numerous genetic alterations that accompany its development (42–44). There appear to exist two pathways of genetic instability in colorectal cancer. The first is found in about 15% of colorectal tumors and involves point mutations, microdeletions, and microinsertions associated with MMR deficiency (45, 46). The second is found in MMR-proficient cells and involves gains and losses of whole chromosomes (47). We suggest that methylation abnormalities are intrinsically and directly involved in the generation of the second type of instability, thus allowing for the selection of MET^- cells during the clonal evolution of tumors. This hypothesis is supported by the observation that demethylation is associated with chromosomal aberrations, including mitotic dysfunction and translocation (48, 49), and is consistent with the innovative hypothesis relating methylation and aneuploidy recently put forward by Thomas (50).

The authors thank S. Baylin for critical reading of the manuscript and Genetic Therapy Inc. for providing the retrovirus G1BgSVNa. This work was supported by grants from the Fonds zur Förderung der wissenschaftlichen Forschung (C.L.), the Clayton Foundation, and National Institutes of Health Grants CA 35494 and CA 62924. B.V. is an Investigator of the Howard Hughes Medical Institute.

1. Bird, A. (1992) *Cell* **70,** 5–8.
2. Noyer-Weldner, M. & Trautner, T. A. (1993) in *DNA Methylation: Molecular Biology and Biological Significance*, eds. Jost, J. P. & Saluz, H. P. (Birkhauser, Basel), pp. 39–108.
3. Yeivin, A. & Razin, A. (1993) in *DNA Methylation: Molecular Biology and Biological Significance*, eds. Jost, J. P. & Saluz, H. P. (Birkhauser, Basel), pp. 523–568.
4. Bird, A. (1986) *Nature (London)* **321,** 209–213.
5. Gama-Sosa, M. A., Slagel, V. A., Trewyn, R. W. & Ehrlich, M. (1983) *Nucleic Acids Res.* **11,** 6883–6894.
6. Goelz, S. E., Vogelstein, B., Hamilton, S. P. & Feinberg, A. P. (1985) *Science* **228,** 187–190.
7. Feinberg, A. P., Gehrke, C. W., Kuo, K. C. & Ehrlich, M. (1988) *Cancer Res.* **48,** 1159–1161.
8. Santi, D. V., Garrett, C. E. & Barr, P. J. (1983) *Cell* **33,** 9–10.
9. Santi, D. V., Normen, A. & Garrett, C. E. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 6993–6997.
10. Jüttermann, R., Li, E. & Jaenisch, R. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 11797–11801.
11. Landolph, J. R. & Jones, P. A. (1982) *Cancer Res.* **42,** 817–823.
12. Baylin, S. B., Makos, M., Wu, J. J., Yen, R. W., de Bustros, A., Vertino, P. & Nelkin, B. D. (1991) *Cancer Cells* **3,** 383–390.
13. Herman, J. G., Merlo, A., Mao, L., Lapidus, R. G., Issa, J. P., Davidson, N. E., Sidransky, D., Baylin, S. B. (1995) *Cancer Res.* **55,** 4525–4530.
14. Jones, P. A. (1996) *Cancer Res.* **56,** 2463–2467.
15. Laird, P. W., Jackson-Grusby, L., Fazeli, A., Dickinson, S. L., Jung, W. E., Li, E., Weinberg, R. A. & Jaenisch, R. (1995) *Cell* **81,** 197–205.
16. Papadopoulos, N., Nicolaides, N. C., Wei, Y. F., Ruben, S. M., Carter, K. C., Rosen, C. A., Haseltine, W. A., Fleischmann, R. D., Fraser, C. M., Adams, M. D., Venter, J. C., Hamilton, S. R., Peterson, G. M., Watson, P., Lynch, H. T., Peltomaki, P., Mecklin, J. P., de la Chapelle, A., Kinzler, K. W. & Vogelstein, B. (1994) *Science* **263,** 1625–1629.
17. Papadopoulos, N., Nicolaides, N. C., Liu, B., Parsons, R. E., Lengauer, C., Palombo, F., D'Arrigo, A., Markowitz, S., Willson, J. K. V., Kinzler, K. W., Jiricny, J. & Vogelstein, B. (1995) *Science* **268,** 1915–1917.
18. Shibata, D., Peinado, M. A., Ionov, Y., Malkhosyan, S. & Perucho, M. (1994) *Nat. Genet.* **6,** 273–281.
19. Umar, A., Boyer, J. C., Thomas, D. C., Nguyen, D. C., Risinger, J. I., Boyd, J., Ionov, Y., Perucho, M. & Kunkel, T. A. (1994) *J. Biol. Chem.* **269,** 14367–14370.
20. da Costa, L. T., Liu, B., El-Deiry, W. S., Hamilton, S. R., Kinzler, K. W. & Vogelstein, B. (1995) *Nat. Genet.* **9,** 10–11.
21. Liu, B., Nicolaides, N. C., Markowitz, S., Willson, J. K. V., Parsons, R. E., Jen, J., Papadopoulos, N., Peltomaki, P., de la Chapelle, A., Hamilton, S. R., Kinzler, K. W. & Vogelstein, B. (1995) *Nat. Genet.* **9,** 48–55.
22. Pear, W. S., Nolan, G. P., Scott, M. L. & Baltimore, D. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 8392–8396.
23. Miller, A. D. & Buttimore, C. (1986) *Mol. Cell. Biol.* **6,** 2895–2902.
24. Powell, S. M., Petersen, G. M., Krush, A. J., Booker, S., Jen, J., Giardiello, F. M., Hamilton, S. R., Vogelstein, B. & Kinzler, K. W. (1993) *N. Engl. J. Med.* **329,** 1982–1987.
25. Feinberg, A. P. & Vogelstein, B. (1983) *Anal. Biochem.* **132,** 6–13.
26. El-Deiry, W. S., Nelkin, B. D., Celano, P., Yen, R.-W. C., Falco, J. P., Hamilton, S. R. & Baylin, S. B. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 3470–3474.
27. Tso, J. Y., Sun, X. H., Kao, T. H., Reece, K. S. & Wu, R. (1985) *Nucleic Acids Res.* **13,** 2485–2502.
28. Frommer, M., McDonald, L. E., Millar, D. S., Collis, C. M., Watt, F., Grigg, G. W., Molloy, P. L. & Paul, C. L. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 1827–1831.
29. Clark, S. J., Harrison, J., Paul, C. L. & Frommer, M. (1994) *Nucleic Acids Res.* **22,** 2990–2997.
30. Harbers, K., Schnieke, N., Stuhlmann, H., Jahner, D. & Jaenisch, R. (1981) *Proc. Natl., Acad. Sci. USA* **78,** 7609–7613.
31. Jahner, D., Stuhlmann, H., Stewart, C. L., Harbers, K., Lohler, J., Simon, I. & Jaenisch, R. (1982) *Nature (London)* **298,** 623–628.
32. Bednarik, D. P. (1993) in *DNA Methylation: Molecular Biology and Biological Significance*, eds. Jost, J. P. & Saluz, H. P. (Birkhauser, Basel), pp. 300–329.
33. Bhattacharyya, N. P., Skandalis, A., Ganesh, A., Groden, J. & Meuth, M. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 6319–6323.
34. Eshleman, J. R., Lang, E. Z., Bowerfind, G. K., Parsons, R., Vogelstein, B., Willson, J. K., Veigl, M. L., Sedwick, W. D. & Markowitz, S. D. (1995) *Oncogene* **10,** 33–37.
35. Roberts, R. (1990) *Nucleic Acids Res.* **18,** S2331–S2365.
36. Challita, P.-M. & Kohn, D. B. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 2567–2571.
37. Challita, P.-M., Skelton, D., El-Khoueiry, A., Yu, X.-J., Weinberg, K. & Kohn, D. B. (1995) *J. Virol.* **69,** 748–755.
38. Hoeben, R. C., Migchielsen, A. A. J., van der Jagt, R. C. M., van Ormondt, H. & van der Eb, A. J. (1991) *J. Virol.* **65,** 904–912.
39. Koi, M., Umar, A., Chauhan, D. P., Cherian, S. P., Carethers, J. M., Kunkel, T. A. & Boland, C. R. (1994) *Cancer Res.* **54,** 4302–4312.
40. Branch, P., Hampson, R. & Karran, P. (1995) *Cancer Res.* **55,** 2304–2309.
41. de Wind, N., Dekker, M., Berns, A., Radman, M. & te Riele, H. (1995) *Cell* **82,** 321–330.
42. Loeb, L. (1991) *Cancer Res.* **51,** 3075–3079.
43. Hartwell, L. (1992) *Cell* **71,** 543–546.
44. Nowell, P. C. (1976) *Science* **194,** 23–28.
45. Modrich, P. (1994) *Science* **266,** 1959–1960.
46. Marra, G. & Boland, C. R. (1995) *J. Natl. Cancer Inst.* **87,** 1114–1125.
47. Lengauer, C., Kinzler, K. W. & Vogelstein, B. (1997) *Nature (London)*, in press.
48. Schmid, M., Grunert, D., Haaf, T. & Engel, W. (1983) *Cytogenet. Cell Genet.* **36,** 554–561.
49. Schmid, M., Haaf, T. & Grunert, D. (1984) *Hum. Genet.* **67,** 257–263.
50. Thomas, J. H. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 480–482.

Downloaded from https://www.pnas.org by 122.171.142.147 on January 29, 2025 from IP address 122.171.142.147.