Epigenetic gene silencing in cancer

Benjamin Tycko

Perspective

The classical theory of recessive oncogenesis predicted a mutational mechanism for the inactivation of tumor suppressor (TS) genes. This prediction has been amply confirmed, but an alternative, nonmutational, pathway for loss of TS gene activity has also come into focus. For some TS genes, this epigenetic pathway is more frequent than the mutational one. The best-studied DNA modification that correlates with epigenetic gene silencing is methylation of cytosine residues in CpG sequences, and CpG methylation has recently been linked to an even more general mechanism of epigenetic silencing, histone deacetylation. From a combination of descriptive studies and manipulative experiments, some hints of mechanisms for epigenetic silencing of TS genes in cancer cells are beginning to emerge. Here, I discuss several well-documented examples of epigenetic gene silencing in human cancers. I then consider potential mechanisms for de novo methylation of TS genes in cancer. These include spreading of DNA methylation from repetitive sequences into promoter-associated CpG islands secondary to loss of transcriptional activator proteins, gain of methylation secondary to hyperexpression of transcriptional repressors, primary hypermethylation due to hyperexpression of methyltransferases, and interallelic transfer of methylation via gene pairing. In some cancers, environmental pressures that select for a hypermethylating cellular phenotype may drive these processes. DNA methylation and heritable gene silencing Epigenetic gene silencing refers to nonmutational gene inactivation that can be […]

Find the latest version:

Epigenetic gene silencing in cancer

Benjamin Tycko

Institute of Cancer Genetics and Department of Pathology, Columbia University, College of Physicians and Surgeons, New York, New York 10032, USA

Address correspondence to: Columbia University, Institute of Cancer Genetics, Russ Berrie Research Pavilion, 1150 St. Nicholas Avenue, New York, New York 10032, USA. Phone: (212) 304-7165; Fax: (212) 304-7336; E-mail: bt12@columbia.edu.

The classical theory of recessive oncogenesis predicted a mutational mechanism for the inactivation of tumor suppressor (TS) genes. This prediction has been amply confirmed, but an alternative, nonmutational, pathway for loss of TS gene activity has also come into focus. For some TS genes, this epigenetic pathway is more frequent than the mutational one. The best-studied DNA modification that correlates with epigenetic gene silencing is methylation of cytosine residues in CpG sequences, and CpG methylation has recently been linked to an even more general mechanism of epigenetic silencing, histone deacetylation. From a combination of descriptive studies and manipulative experiments, some hints of mechanisms for epigenetic silencing of TS genes in cancer cells are beginning to emerge. Here, I discuss several well-documented examples of epigenetic gene silencing in human cancers. I then consider potential mechanisms for de novo methylation of TS genes in cancer. These include spreading of DNA methylation from repetitive sequences into promoter-associated CpG islands secondary to loss of transcriptional activator proteins, gain of methylation secondary to hyperexpression of transcriptional repressors, primary hypermethylation due to hyperexpression of methyltransferases, and interallelic transfer of methylation via gene pairing. In some cancers, environmental pressures that select for a hypermethylating cellular phenotype may drive these processes.

DNA methylation and heritable gene silencing

Epigenetic gene silencing refers to nonmutational gene inactivation that can be faithfully propagated from precursor cells to clones of daughter cells. The addition of methyl groups to cytosine residues in CpG dinucleotides in DNA is a biochemical modification that meets this requirement. A very striking example of the potential for CpG methylation to be heritable is provided by epimutations in plants. Genes carrying epimutations cause morphological phenotypes to be transmitted from generation to generation, not based on any alteration in the coding sequence of the relevant genes, but instead caused by CpG (or CpNpG) hypermethylation of their promoter sequences (for example, see ref. 1).

The prototypical mammalian cytosine DNA methyltransferase, encoded by the *DNMT1* gene, is a large enzyme with strong sequence homology to bacterial methyltransferases over its conserved COOH-terminal catalytic domain. More than half of the DNMT1 protein (the NH_2-terminal portion) is not involved in catalysis, but instead is essential for protein-protein interac-

tions and fine tuning of substrate recognition. One critical aspect of this substrate recognition is a preference for DNA substrates that are methylated on 1 strand but not the other. This preference for hemimethylated DNA partly explains the ability of this enzyme to propagate DNA methylation patterns to daughter strands at each S-phase. A second function of the regulatory domain is targeting of DNMT1 to foci of DNA replication in the nucleus. The preference for hemimethylated DNA and physical proximity of the enzyme to newly synthesized DNA combine to account for the observed ability of the enzyme to copy methylation patterns faithfully (2).

Although patterns of DNA methylation are generally passed faithfully to daughter cells, these patterns can also be created or erased. Creation of methylation patterns de novo is critical to epigenetic gene silencing in cancer, and the DNMT1 enzyme does have some de novo methylating activity in vitro. But de novo, as opposed to maintenance, methylation is also catalyzed by other cytosine methyltransferases, distinct from DNMT1. This was first suggested by data from cells homozygous for a null allele of the mouse *Dnmt1* gene. These cells suffered a severe reduction in overall DNA methylation, but retained their ability to methylate invading retroviral DNA. The prediction that additional cytosine methyltransferases exist in mammalian cells was borne out when database searches for conserved methyltransferase motifs led to the cloning of mouse and human versions of 2 additional functional methyltransferases, DNMT3A and DNMT3B. These enzymes exhibit de novo methylating activity both in vitro and in vivo (3). Interestingly, DNMT3B is hyperexpressed, at least at the mRNA level, in some human cancers (4).

Gene silencing by DNA methylation coupled to histone deacetylation

Data for an active role of methylation in gene silencing are both correlative and functional. More than half of all mammalian genes contain CpG-rich sequences (CpG islands) overlapping with their promoter sequences. For the majority of transcriptionally active and potentially active genes in normal cells, these CpG islands are maintained in a fully unmethylated state. In normal non-neoplastic cells, 2 categories of genes illustrate the correlation between CpG methylation and gene silencing: a subset of genes that are subject to parental imprinting and most of those that reside on the inactive X-chromosome. In the upstream regions of prototypical imprinted genes, such as the paternally silenced *H19* gene and the maternally

silenced *SNRPN* gene, the active allele is unmethylated and the silenced allele massively methylated at numerous CpG dinucleotides (see the Perspective by Maher and Reik and the Perspective by Nicholls, this series). This allelic asymmetry is established during gametogenesis and persists in postzygotic development. The resulting correlation of hypermethylation with silencing is so reliable that it can be used in routine clinical diagnosis of syndromes caused by abnormal expression of imprinted genes (see Maher and Reik, this series, and Nicholls, this series). Similarly, the extensive methylation of CpG islands of silent genes on the inactive X-chromosome, and the absence of methylation of these sequences on the active X, has allowed the design of widely used assays for clonality in human tumors and non-neoplastic syndromes.

Functional data supporting a causative role of methylation in gene silencing in non-neoplastic cells are of at least 3 types. In vitro methylation of promoter-reporter constructs inhibits their subsequent expression in transfected cells. Conversely, erasure of DNA methylation using the covalent methyltransferase inhibitor 5-aza-deoxycytidine (AzaC) leads to re-expression of previously methylated genes. Last, homozygous embryos with a germline deletion of the *Dnmt1* gene re-express a number of genes, including the normally silent alleles of several imprinted genes and the abundant but normally repressed endogenous retroviral sequences, which are methylated and silent in heterozygous littermates (5).

In principle, DNA methylation can silence genes by interfering with sequence specific binding of positive transcription factors or by producing more general effects on chromatin. Good evidence exists for the first possibility, but the second has received even more attention. Hindrance of transcription by DNA methylation was long thought to be mediated by general methyl-C binding proteins, and a group of proteins with this property have been characterized. Recent work on the biochemistry of these proteins, exemplified by MeCP2, has led to an important advance in understanding epigenetic gene silencing–forging a link between CpG methylation and a second, synergistic epigenetic modification, histone deacetylation on lysine residues. Deacetylation of histones causes an increase in the positive charge of these proteins. In the simplest models, this modification increases the histone's avidity for DNA or for other histones. The resulting compaction of the chromatin may block access of transcription factors to the DNA or constrain the movement of RNA polymerase. The link between histone deacetylation and DNA methylation was the finding that MeCP2 physically interacts with the transcriptional corepressor protein Sin3A, and in so doing recruits a histone deacetylase (HDAC) to chromatin that contains methylated DNA (6). This biochemical link explains older observations that methylated DNA templates microinjected into cell nuclei can be actively transcribed until the DNA is packaged into chromatin.

Extensive evidence, discussed later here, indicates that genes can be silenced by CpG methylation in cancer cells, but data for pathological gene silencing by histone deacetylation are still sparse. However, the gene for the multifunctional actin-binding protein gelsolin, which is epigenetically silenced in most human breast cancers, can be reactivated either by AzaC or by treatment with the histone deacetylase inhibitor trichostatin-A (7). For certain other genes that are silenced by DNA methylation, inhibition of histone deacetylase is insufficient to cause reactivation. Rather, to reactivate these genes, it is necessary first to alleviate partially the hypermethylation of their promoters with a brief AzaC treatment, after which TSA becomes effective and boosts expression strongly (8).

TS genes silenced by DNA methylation in human cancers

TS genes silenced by DNA methylation in human cancers provide the impetus for this area of research, and a selection of such genes is considered here. From the large number of examples, it should not be concluded that all proved or candidate TS genes undergo frequent methylation in cancer cells; some TS genes are detectably but infrequently methylated. The *RB1* gene is silenced by promoter hypermethylation in some sporadic retinoblastomas, and historically this TS gene provided the first example of this pathway, but the incidence of its epigenetic silencing is only about 10% (reviewed in ref. 9).

The von Hippel-Lindau gene in renal cell carcinomas. Germline mutations in the *von Hippel-Lindau (VHL)* gene, on distal chromosome 3p, predispose to renal cell carcinomas and hemangioblastomas, and somatic mutations of *VHL* are characteristic of sporadic renal cell carcinomas. An important early study of *VHL* in sporadic renal cell carcinomas showed that the promoter was hypermethylated in about 20% of the tumors, establishing a precedent for this route to TS gene inactivation (10). There were no coding mutations in these cases, and, in a renal cell carcinoma cell line, the *VHL* locus was reactivated by AzaC. A second comprehensive study of *VHL*-associated tumors occurring in the setting of the inherited syndrome showed that promoter hypermethylation is a frequent pathway, alternative to loss of heterozygosity (LOH), for the somatic "second hit" in tumorigenesis (11). In that study, more than 30% of the tumors (renal cell carcinomas and hemangioblastomas) that had not undergone LOH had lost expression of the nonmutated *VHL* allele via promoter hypermethylation. An indirect indication of the specificity and biologic relevance of *VHL* promoter hypermethylation is that it has not been found in cancer types other than those that are associated with the von Hippel-Lindau syndrome.

The hMLH1 gene in colonic, gastric, and endometrial cancer. The *hMLH1* gene, on proximal chromosome 3p, encodes a protein essential for DNA mismatch repair. This gene is mutated in the germline of a high percentage of people with hereditary nonpolyposis colon cancer, and it is also inactivated in about 15–25% of sporadic colon cancers. These cancers show a high rate of replicative errors in repetitive (microsatellite) DNA tracts, a phenotype often described as microsatellite instability (MIN); germline mutation of a second mismatch repair gene, *hMSH2*, represent a rarer cause of MIN. Data from several studies have indicated that more than half of spo-

radic cases of colon cancer with MIN are free of mutations in *hMLH1* but exhibit substantial hypermethylation of the *hMLH1* promoter (for example, see ref. 12). Gene reactivation by AzaC showed that hypermethylation causes silencing, and this treatment proved sufficient to restore mismatch repair competence to the cancer cells. Subsequent studies have uncovered a high frequency of *hMLH1* promoter hypermethylation in other common epithelial malignancies, including endometrial and gastric cancers. In all of these series, strong correlations emerged between promoter hypermethylation, loss of hMLH1 protein, and the MIN phenotype. Gene specificity of the epigenetic silencing was observed in each of these tumor types, as hypermethylation of *MSH2* was not detected in any of several studies. As appears to be true for colon cancers, epigenetic inactivation of *hMLH1* is also more frequent than mutational inactivation in gastric and endometrial cancers.

The p16/INK4A locus in multiple cancers. The *p16* gene resides on chromosome 9p in a region that is subject to frequent deletion or LOH in various human cancers. This gene encodes a cyclin-dependent kinase inhibitor that restrains the cell cycle by preventing phosphorylation and inactivation of Rb. p16 is expressed from a complex locus, with an alternative promoter and an alternative reading frame encoding the p19[ARF] protein, a positive regulator of the p53 tumor suppressor. Adding to this complexity, *p16* lies immediately adjacent to a second CDK inhibitor gene, encoding the p15 protein. There is evidence for epigenetic silencing of all 3 of these genes in human cancers, but for brevity only the *p16* gene is discussed here. After this gene was cloned, it became apparent that, despite a very high frequency of deletions and mutations in cancer cell lines, a large group of primary cancers with LOH for this region lacked deletions or coding alterations. Because the case for *p16* as a TS gene was otherwise strong, several laboratories sought and found evidence for its inactivation in primary cancers via an epigenetic pathway (reviewed in ref. 13). These and numerous subsequent studies showed that promoter hypermethylation is the most frequent pathway for *p16* inactivation in human carcinomas, including those that arise in the lung, oropharynx, bladder, cervix, liver, colon, pancreas, and other sites (13). Several of these studies also showed that p16 mRNA is re-expressed after exposure of cancer cell lines to AzaC.

Detailed analysis of a colon cancer with retention of heterozygosity showed that 1 allele of p16 had sustained a coding mutation and was expressed, whereas the other allele was nonmutated, hypermethylated, and silent. In that study, the *p16* promoter regions were sequenced, and no mutations were found (14). Similarly, in a study of pancreatic cancers, Schutte et al. found that the *p16* promoter was hypermethylated only on alleles of *p16* where no coding mutation could be found (15). By all of these criteria, promoter hypermethylation is the proximate cause of *p16* silencing and can qualify as an alternative form of "second hit" leading to biallelic inactivation.

Importantly, in many tumors, *p16* is inactivated solely by DNA hypermethylation (i.e., both "hits" are epigenetic), and evidence is accumulating that this process begins quite early in tumor evolution. Assays performed by methylation-sensitive PCR on small tissue biopsies showed that *p16* hypermethylation is detectable in preinvasive bronchial mucosal lesions occurring in smokers (16).

The E-cadherin gene in carcinomas. The *E-cadherin* gene encodes a cell-surface adhesion protein that is thought to play an important role in homotypic cell-cell adhesion and maintenance of epithelial morphology, and it may also help control cell growth and differentiation by its interaction with the multifunctional cytoplasmic protein β-catenin. Indeed, E-cadherin protein is absent or mislocalized in certain classes of discohesive and poorly differentiated cancers, notably diffuse gastric carcinomas and lobular breast carcinomas. About half of all diffuse gastric cancers and lobular breast carcinomas carry inactivating coding mutations in the *E-cadherin* gene. An alternative pathway to somatic mutation, epigenetic silencing of *E-cadherin* by promoter hypermethylation occurs frequently in sporadic gastric cancers (17), as well breast, prostate, and colon cancers (18, 19). AzaC reactivates this gene, indicating that promoter hypermethylation actively silences its expression in tumors (18, 19). However, in some cell lines from breast and other cancers, there is evidence for a dominant pathway for extinguishing E-cadherin expression (20). Recent data show that this novel pathway involves the overexpression of specific transcriptional repressor proteins (21). As discussed later here, these findings raise the possibility that the observed promoter hypermethylation in these cancers occurs as a consequence, rather than a cause, of the primary transcriptional silencing. It is not yet clear whether all human cancers with loss of *E-cadherin* have undergone hypermethylation of the promoter of this gene. One study of transformed mouse keratinocytes suggests a methylation-independent pathway for gene silencing, as these cells lose E-cadherin mRNA expression as a result of the loss of a transactivating protein, but apparently without increasing CpG methylation (22).

Imprinted genes: epigenetic silencing of H19 and activation of IGF-2 in pediatric cancers. Chromosome 11p15 is a frequent site of LOH in pediatric solid tumors, and, in Wilms' tumors (WTs), embryonal rhabdomyosarcomas, hepatoblastomas, and adrenal cortical carcinomas, the lost alleles are always maternal in origin. This parent-of-origin effect suggests that imprinted genes contribute to the pathogenesis of these tumors. The *H19* gene, on chromosome 11p15, encodes a nontranslated cytoplasmic RNA. In normal cells, the paternal copy is silenced and heavily CpG methylated, and this gene is therefore expressed monoallelically from the unmethylated maternal copy. Given that the *H19* gene and its immediate upstream DNA are hypermethylated in sperm, but not in ova, and that this paternal hypermethylation persists throughout post-zygotic development, the primary imprint is considered to correspond to paternal silencing; that is, *H19* is paternally imprinted. *H19* lies about 100 kb downstream from another imprinted gene, encoding IGF-2, which is monoallelically expressed from the opposite paternal, allele. The reciprocal expression of these 2 genes is not

coincidental, but reflects a *cis* interaction involving shared regulatory sequences, sometimes referred to as "expression-competition." There is good evidence from manipulative experiments in mice, as well as from the study of human tumors (see later here), that the master regulatory sequences for this process reside within the differentially methylated DNA immediately upstream of *H19* (see Maher and Reik, this series). Hence, the primary imprint on *H19*, methylation of the paternal allele, also accounts for the opposite allele-specific expression of *IGF2*.

The *H19/IGF2* locus has been the subject of intense study in childhood tumors, as well as in the Beckwith-Wiedemann overgrowth syndrome (Maher and Reik, this series). In these neoplasms, there is frequent conversion from the normal monoallelic methylation of the *H19* gene and its upstream sequences to an abnormal pattern of biallelic hypermethylation. As a result of this change, both alleles now resemble the imprinted paternal allele of *H19*. Hence, the epigenetic alteration in neoplastic tissue that causes complete loss of expression of H19 RNA (23 and references therein) is best described as a somatic "gain of imprinting" at this locus. In WTs, the silencing of *H19* invariably correlates, at the whole-tissue level, with a reciprocal biallelic activation of *IGF2*, which is often referred to as a somatic "loss of imprinting" of this gene. The end point of biallelic silencing at *H19* and biallelic activation at *IGF2* is an abnormal bipaternal pattern of gene expression at both loci, with hyperexpression of an antiapoptotic growth factor (IGF-2) (24) and loss of a transformation-suppressing RNA (H19) (25).

An important and well-documented feature of altered functional imprinting of *H19/IGF2* is that it occurs quite early in the multistage evolution of WTs. The abnormal bipaternal epigenetic state is detectable not only in the tumors, but also frequently in the adjacent histologically normal kidney parenchyma, where it exists as a 50–90% tissue mosaicism for the bipaternally programmed cells (23, 26, 27). Mosaicism for the bipaternal epigenotype is a specific precancerous lesion for WT, and it is an alternative pathway to 11p15 LOH. It is not observed in control kidneys from patients with other conditions, and it is also never found in kidneys from WT patients whose tumors have undergone 11p15 LOH. Because epimutation of *H19* often occurs before the earliest stages of overt neoplastic transformation, and because the gain of DNA methylation on the maternal allele is locally restricted to the *H19* gene (23), a gene-specific mechanism, rather than a global disruption in genomic methylation, probably accounts for this phenomenon. One possible mechanism, imprint transfer via gene pairing, is discussed later here.

Pathways for de novo methylation of TS genes

Invasion of CpG islands by methylation: secondary hypermethylation after gain of repressors or loss of activators, or primary hypermethylation after overwhelming assault by methyltransferases? The mechanisms of pathological de novo DNA hypermethylation in cancer precursor cells are currently unknown. In approaching this problem, we first need to consider 3 general models. In the first model (Figure 1a), the primary event leading to locus-specific hypermethylation in cancer precursor cells is loss of expression of 1 or more critical transcription factors. In this "loss-of-protection" model, hypermethylation occurs as a secondary event: somehow, the initial loss of transcription factor occupancy creates a chromatin configuration around the promoter region that is permissive for subsequent de novo methyltransferase activity. The second model (Figure 1b) envisions the converse situation, in which there is hyperexpression of a transcriptional repressor in the cancer precursor cell. Here too, the observed promoter methylation occurs as a secondary event. In contrast, the third model (Figure 1c) posits that de novo methylation is the primary inactivating event. Relevant to this model is the contentious question of whether total cytosine DNA methyltransferase activity is increased in cancer cells. The consensus of the evidence, supported by
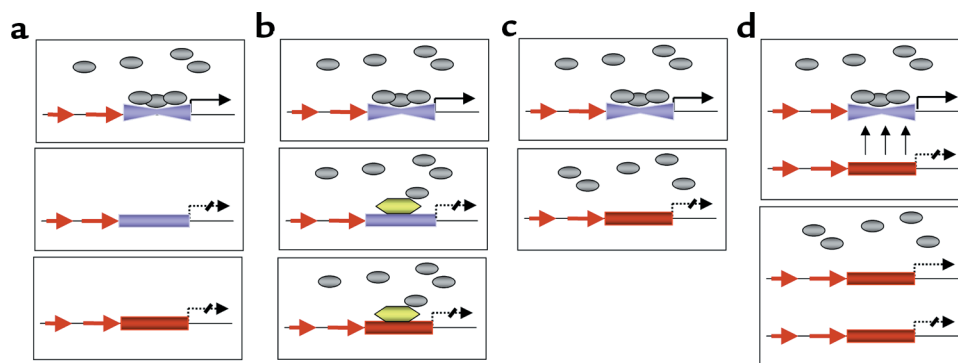
Figure 1
Hypothetical pathways for de novo methylation of gene promoters in cancer precursor cells. (**a**) Primary silencing by loss of activating transcription factors, followed by secondary de novo methylation. (**b**) Primary silencing by overexpression of transcriptional repressors, followed by secondary de novo methylation. (**c**) Primary de novo methylation by hyperexpressed DNA methyltransferase, without loss of transcription factors. (**d**) Interallelic transfer of DNA methylation at a locus with preexisting allele-specific hypermethylation (either parental imprinting or "first-hit" de novo methylation). Methylated DNA is shown in red. The gene promoter is in blue, and changes in chromatin configuration are indicated by a change in shape. Activating transcription factors are in gray, and repressor proteins are in yellow. The red arrows are methylated repetitive elements, and the black arrows indicate gene transcription. In **a**–**c**, DNA methylation is shown as spreading from preexisting methylated sequences within the repetitive elements.

immunohistochemistry, is that DNMT1 protein is indeed elevated in some cancer cells relative to their presumed normal precursors, but that significant cell-to-cell heterogeneity in DNMT1 expression exists within a given neoplasm (28).

Cancer cell lines with hypermethylation of the endogenous *E-cadherin* gene are relatively deficient in their ability to transcribe transfected reporter genes under control of the *E-cadherin* promoter, compared with similar cell lines without such hypermethylation (29, 30). Clearly, this finding is compatible either with the loss-of-protection or gain-of-repression models of gene silencing, but is not easily squared with the hypothesis that hypermethylation at the endogenous locus is the initial event in the pathway. A follow-up study using somatic cell hybrids suggested that a dominant repressing activity might account for *E-cadherin* silencing (20), and the overexpression of a specific transcriptional repressor, the Snail protein, is a strong candidate to be the primary regulatory event that culminates in both silencing and methylation (21). However, in 1 study, carcinoma cell lines with *E-cadherin* promoter hypermethylation could still express reporter genes driven by the *E-cadherin* promoter (19). Also conflicting with the loss-of-protection or the gain-of-repression models as global explanations for gene silencing is the fact that AzaC can reactivate E-cadherin expression in some cancer cell lines (18, 19).

Given the continuing controversies, additional experiments are needed to determine the frequency of transcription factor deficiency or transcriptional repressor overexpression as explanations for *E-cadherin* gene silencing. Similar experiments will be of interest to probe the role of transcription factor deficiency or repressor overexpression in the silencing of other TS genes that are hypermethylated in cancer cells. A recent study used somatic cell genetics and chromosome transfers (31) to test for deficiency of *trans*-acting factors or overexpression of transcriptional repressors in a renal cell carcinoma line with *VHL* silencing and hypermethylation, and the findings did not support either of these mechanisms.

Spreading of methylation from repetitive DNA into CpG islands. Evidence from diverse organisms has suggested that repetitive DNA is a preferred target for epigenetic silencing, and this has been viewed as a host-defense system directed against viruses and transposons (see the Perspective by Bestor, this series). All the TS genes that are pathologically hypermethylated in cancer cells contain CpG islands or island-like CpG-rich sequences in their upstream regions. These sequences are completely or nearly completely unmethylated in the normal tissues in which the cancers arise, but these "protected" sequences are adjacent to other, non–CpG island, sequences that are methylated in normal cells. Often, these adjoining, methylated, sequences consist of repetitive DNA, notably Alu elements.

Loss of promoter occupancy by activating factors, or altered promoter configuration secondary to occupancy by repressors, may predispose to promoter inactivation by allowing methylation to spread from adjacent densely methylated repetitive DNA. Experiments with transfected cells and transgenic mice suggest that when a single class of transcription factor binding sites, the Sp1 sites commonly found within CpG islands, is mutated, DNA methylation can encroach from adjacent repetitive sequences into the mutated CpG island. However, it appears that genes can be silenced by this mechanism even in cells that still have normal promoter occupancy. Graff et al. showed that in cells engineered to overexpress DNMT1, methylation spread progressively from flanking Alu sequences into the CpG island of a previously active *E-cadherin* gene (32). The term "spreading" is used in these studies to describe the observed progressive accumulation of methylation in the Alu-flanking promoter sequences over several cell generations, but the local interactions between methyltransferases and the replicating DNA that underlie this process are not yet understood.

Gene-specific hypermethylation and natural selection for a cellular "methylator" phenotype. Some of the experiments just described do not precisely mimic the situation in human cancer cells, as, with rare exceptions, mutations have not been found in the promoter sequences of hypermethylated TS genes, and loss of Sp1-like transcription factors may not be common. Nevertheless, spreading of CpG methylation from adjacent repetitive DNA into CpG islands is potentially important for de novo methylation in cancer. This epigenetic change may occur in rare cells that suffer transient loss of transcription factor occupancy or in cells that express methyltransferases at levels sufficient to breach the normal protective barriers.

Baylin and coworkers have proposed that specific TS genes, such as *p16* and *hMLH1*, which have repetitive DNA immediately adjacent to their CpG islands, are inherently susceptible to promoter hypermethylation, particularly in cells with a hyperactive methylation system (13). To test this idea, these authors screened for methylated CpG islands in cancer cells and analyzed methylation status of 30 such sequences in the colon during normal aging and in different classes of colon cancers. They found that the majority of CpG islands methylated in colon cancer are also methylated in a subset of normal colonic cells during aging. In contrast, methylation of several "cancer-specific" CpG islands was found exclusively in a subset of colorectal cancers, which displayed a "CpG island methylator phenotype." These cancers showed hypermethylation of multiple CpG islands, including a high incidence of *p16* promoter methylation, and they also included the majority of sporadic colorectal cancers in which MIN resulted from *hMLH1* methylation (33). The notion of a "methylator" phenotype in a subgroup of human cancers is also supported by a large study of gastric carcinomas, in which all of the tumors with *hMLH1* hypermethylation also showed *p16* and *E-cadherin* methylation (17). Similarly, Lengauer et al. (34) had previously described human colon cancer cell lines that are proficient at methylating exogenous retroviral DNA and others that lack this ability.

These observations may be understood within a coherent model inspired by the striking observation that cancers with the MIN phenotype occur commonly

in the proximal colon but rarely in the distal colon and rectum. Breivik and Gaudernack (35) have proposed that nitrosylated bile acids are major carcinogens in the proximal colon and that fecal mutagens are more important in the distal colon. Because bile acid derivatives cause O6-methylguanine adducts in DNA that cannot be effectively removed by the mismatch repair system, these authors postulated a selective pressure, applying to proximal but not distal tumors, for a hypermethylating phenotype and the loss of mismatch repair. In proximal colonic tissue, this form of DNA damage creates a futile cycle of failed attempts at strand incision and mismatch repair, which in turn induces checkpoint activation and cellular growth arrest. Silencing of *hMLH1* by promoter methylation would abrogate this form of "self-inflicted" DNA damage and would enable the cell to bypass the checkpoint. In this scenario, a specific environmental challenge (bile acid metabolites), together with the structural susceptibility of the *hMLH1* gene to promoter hypermethylation, drives the natural selection for cells with the hypermethylating phenotype. For proximal colon cancers, the inactivation of *p16* and other TS genes (as well as otherwise unrelated genes that carry appropriate promoter and flanking DNA sequences) occurs concurrently and influences the outcome of clonal evolution, but it is viewed as a secondary consequence of the 2 primary driving forces.

Propagation of epigenetic silencing via gene pairing. Homologous gene pairing is known to mediate gene silencing in diverse species, and, in organisms with functional methyltransferases, this process is often associated with extensive de novo methylation. A particularly well-studied phenomenon of this type occurs in the fungus *Ascobolus immersus* and is referred to as MIP, for methylation induced premeiotically. This process is observed when 2 haploid cells, 1 with a methylated gene and the other with a nonmethylated version of the same gene, fuse in preparation for a transient diploid phase. At a low but easily detectable frequency, MIP causes extensive methylation of the previously active and unmethylated allele. The direct relationship between the efficiency of this methylation and the degree of sequence relatedness of the target DNAs, together with the fact that gene conversion accompanies the methylation, strongly implies that MIP proceeds through physical pairing of the repeated sequences. The simplest explanation is that it reflects a direct action of methyltransferase on transient DNA heteroduplexes formed between the unmethylated and methylated alleles (36).

Epigenetic silencing via gene pairing (Figure 1d) could be relevant to human cancer in 2 situations. For an imprinted TS gene, such as *H19*, a silent allele exists in every normal cell, and transfer of the silent epigenetic state to the active allele would lead to complete, biallelic silencing. For nonimprinted genes, like *VHL, hMLH1, p16*, or *E-cadherin*, once 1 copy has become methylated through a de novo pathway, biallelic silencing could be achieved via gene pairing and transfer of the silent epigenetic state. A potential example of this pathway was in the first report of *VHL* promoter hypermethylation; in that study, 1 renal cell carcinoma retained both alleles of the gene and showed biallelic hypermethylation (10).

Future directions: a total genomic map of DNA methylation in cancer cells

A number of challenging mechanistic questions, touched on in the preceding sections, may be resolved by additional experiments using existing technologies, but the future also holds promise for completing the descriptive phase of this research. One exciting possibility will be to capitalize on the completion of the human genome sequence, and the parallel development of DNA-microarray ("DNA chip") technologies, by developing methods to assess comprehensively sequence-specific changes in DNA methylation as they occur during cancer initiation and progression. The study by Issa and coworkers (33), in which methylation-dependent PCR and representational difference analysis identified differentially methylated CpG islands in cancerous and normal colonic cells, points the way for other genomic approaches to the study of epigenetic changes. A recent pilot study by Huang et al. (37) provides another example. These authors arrayed DNAs corresponding to several hundred human CpG islands on filters and then probed these filters with PCR products corresponding to heavily methylated and unmethylated DNA sequences from normal tissues and cancer cell lines. The resulting direct visual readout identified CpG islands that were specifically hypermethylated in the cancer cells. About 10% of the CpG islands were hypermethylated, and many of these appeared to be "hot spots" for methylation, appearing repeatedly as positives in screens with probes from different cancer cell lines. It is easy to imagine that application of similar methods to more stringent comparisons — between early preneoplastic lesions and adjacent normal tissues, for example — will advance our understanding of the structural features of genes that target them for hypermethylation and, by pointing to recurrent silencing events, will reveal a biologically meaningful sequence of early epigenetic changes in cancer precursor cells.

1. Cubas, P., Vincent, C., and Coen, E. 1999. An epigenetic mutation responsible for natural variation in floral symmetry. *Nature.* **401**:157–161.
2. Bestor, T.H., and Verdine, G.L. 1994. DNA methyltransferases. *Curr. Opin. Cell Biol.* **6**:380–389.
3. Okano, M., Bell, D.W., Haber, D.A., and Li, E. 1999. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell.* **99**:247–257.
4. Robertson, K.D., et al. 1999. The human DNA methyltransferases (DNMTs) 1, 3a and 3b: coordinate mRNA expression in normal tissues and overexpression in tumors. *Nucleic Acids Res.* **27**:2291–2298.
5. Li, E., Bestor, T.H., and Jaenisch, R. 1992. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell.* **69**:915–926.
6. Nan, X., et al. 1998. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature.* **393**:386–389.
7. Hoshikawa, Y., Kwon, H.J., Yoshida, M., Horinouchi, S., and Beppu, T. 1994. Trichostatin A induces morphological changes and gelsolin expression by inhibiting histone deacetylase in human carcinoma cell

lines. *Exp. Cell Res.* **214**:189–197.

8. Cameron, E.E., Bachman, K.E., Myohanen, S., Herman, J.G., and Baylin, S.B. 1999. Synergy of demethylation and histone deacetylase inhibition in the re-expression of genes silenced in cancer. *Nat. Genet.* **21**:103–107.

9. Ohtani-Fujita, N., et al. 1997. Hypermethylation in the retinoblastoma gene is associated with unilateral, sporadic retinoblastoma. *Cancer Genet. Cytogenet.* **98**:43–49.

10. Herman, J.G., et al. 1994. Silencing of the VHL tumor-suppressor gene by DNA methylation in renal carcinoma. *Proc. Natl. Acad. Sci. USA.* **91**:9700–9704.

11. Prowse, A.H., et al. 1997. Somatic inactivation of the VHL gene in Von Hippel-Lindau disease tumors. *Am. J. Hum. Genet.* **60**:765–771.

12. Herman, J.G., et al. 1998. Incidence and functional consequences of hMLH1 promoter hypermethylation in colorectal carcinoma. *Proc. Natl. Acad. Sci. USA.* **95**:6870–6875.

13. Baylin, S.B., Herman, J.G., Graff, J.R., Vertino, P.M., and Issa, J.P. 1998. Alterations in DNA methylation: a fundamental aspect of neoplasia. *Adv. Cancer Res.* **72**:141–196.

14. Myohanen, S.K., Baylin, S.B., and Herman, J.G. 1998. Hypermethylation can selectively silence individual p16ink4A alleles in neoplasia. *Cancer Res.* **58**:591–593.

15. Schutte, M., et al. 1997. Abrogation of the Rb/p16 tumor-suppressive pathway in virtually all pancreatic carcinomas. *Cancer Res.* **57**:3126–3130.

16. Belinsky, S.A., et al. 1998. Aberrant methylation of p16(INK4a) is an early event in lung cancer and a potential biomarker for early diagnosis. *Proc. Natl. Acad. Sci. USA.* **95**:11891–11896.

17. Suzuki, H., et al. 1999. Distinct methylation pattern and microsatellite instability in sporadic gastric cancer. *Int. J. Cancer.* **83**:309–313.

18. Yoshiura, K., et al. 1995. Silencing of the E-cadherin invasion-suppressor gene by CpG methylation in human carcinomas. *Proc. Natl. Acad. Sci. USA.* **92**:7416–7419.

19. Graff, J.R., et al. 1995. E-cadherin expression is silenced by DNA hypermethylation in human breast and prostate carcinomas. *Cancer Res.* **55**:5195–5199.

20. Hajra, K.M., Ji, X., and Fearon, E.R. 1999. Extinction of E-cadherin expression in breast cancer via a dominant repression pathway acting on proximal promoter elements. *Oncogene.* **18**:7274–7279.

21. Batlle, E., et al. 2000. The transcription factor Snail is a repressor of E-cadherin gene expression in epithelial tumour cells. *Nat. Cell Biol.* **2**:84–89.

22. Rodrigo, I., Cato, A.C., and Cano, A. 1999. Regulation of E-cadherin gene expression during tumor progression: the role of a new Ets-binding site and the E-pal element. *Exp. Cell Res.* **248**:358–371.

23. Dao, D., et al. 1999. Multipoint analysis of human chromosome 11p15/mouse distal chromosome 7: inclusion of H19/IGF2 in the minimal WT2 region, gene specificity of H19 silencing in Wilms' tumorigenesis and methylation hyper-dependence of H19 imprinting. *Hum. Mol. Genet.* **8**:1337–1352.

24. Christofori, G., Naik, P., and Hanahan, D. 1995. Deregulation of both imprinted and expressed alleles of the insulin-like growth factor 2 gene during beta-cell tumorigenesis. *Nat. Genet.* **10**:196–201.

25. Hao, Y., Crenshaw, T., Moulton, T., Newcomb, E., and Tycko, B. 1993. Tumour-suppressor activity of H19 RNA. *Nature.* **365**:764–767.

26. Moulton, T., et al. 1994. Epigenetic lesions at the H19 locus in Wilms' tumour patients. *Nat. Genet.* **7**:440–447.

27. Taniguchi, T., Sullivan, M.J., Ogawa, O., and Reeve, A.E. 1995. Epigenetic changes encompassing the IGF2/H19 locus associated with relaxation of IGF2 imprinting and silencing of H19 in Wilms tumor. *Proc. Natl. Acad. Sci. USA.* **92**:2159–2163.

28. De Marzo, A.M., et al. 1999. Abnormal regulation of DNA methyltransferase expression during colorectal carcinogenesis. *Cancer Res.* **59**:3855–3860.

29. Hennig, G., et al. 1995. Progression of carcinoma cells is associated with alterations in chromatin structure and factor binding at the E-cadherin promoter in vivo. *Oncogene.* **11**:475–484.

30. Ji, X., Woodard, A.S., Rimm, D.L., and Fearon, E.R. 1997. Transcriptional defects underlie loss of E-cadherin expression in breast cancer. *Cell Growth Differ.* **8**:773–778.

31. Kuzmin, I., et al. 1999. Analysis of aberrant methylation of the VHL gene by transgenes, monochromosome transfer, and cell fusion. *Oncogene.* **18**:5672–5679.

32. Graff, J.R., Herman, J.G., Myohanen, S., Baylin, S.B., and Vertino, P.M. 1997. Mapping patterns of CpG island methylation in normal and neoplastic cells implicates both upstream and downstream regions in de novo methylation. *J. Biol. Chem.* **272**:22322–22329.

33. Toyota, M., et al. 1999. CpG island methylator phenotype in colorectal cancer. *Proc. Natl. Acad. Sci. USA.* **96**:8681–8686.

34. Lengauer, C., Kinzler, K.W., and Vogelstein, B. 1997. DNA methylation and genetic instability in colorectal cancer cells. *Proc. Natl. Acad. Sci. USA.* **94**:2545–2550.

35. Breivik, J., and Gaudernack, G. 1999. Carcinogenesis and natural selection: a new perspective to the genetics and epigenetics of colorectal cancer. *Adv. Cancer Res.* **76**:187–212.

36. Colot, V., Maloisel, L., and Rossignol, J.L. 1996. Interchromosomal transfer of epigenetic states in *Ascobolus*: transfer of DNA methylation is mechanistically related to homologous recombination. *Cell.* **86**:855–864.

37. Huang, T.H., Perry, M.R., and Laux, D.E. 1999. Methylation profiling of CpG islands in human breast cancer cells. *Hum. Mol. Genet.* **8**:459–470.