Epigenetic Variation and Human Disease¹,²

Jean-Pierre Issa³

Department of Leukemia, The University of Texas at M.D. Anderson Cancer Center, Houston, TX 77030

ABSTRACT Cytosine guanine dinucleotide (CpG) island methylation is a known mechanism of epigenetic inheritance in postmeiotic cells. Through associated chromatin changes and silencing, such epigenetic states can influence cellular physiology and affect disease risk and severity. Our studies of CpG island methylation in normal colorectal mucosa revealed progressive age-related increases at multiple gene loci, suggesting genome-wide molecular alterations with potential to silence gene expression. However, there was considerable variation in the degree of methylation among individuals of comparable ages. Such variation could be related to genetic factors, lifestyle, or environmental exposures. Studies in ulcerative colitis and hepatocellular cirrhosis and neoplasia revealed that chronic inflammatory states are accompanied by marked increases in CpG island methylation in normal-appearing tissues, confirming the hypothesis that proinflammatory exposures could account for part of the epigenetic variation in human populations. Preliminary data also suggest potential influences of lifestyle and exposure factors on CpG island methylation. It is suggested that epigenetic variation related to aging, lifestyle, exposures and possibly genetic factors, is one of the modulators of acquired, age-related human diseases, including neoplasia. J. Nutr. 132: 2388S–2392S, 2002.

KEY WORDS: • epigenetics • DNA methylation • aging • CpG islands • CpG island methylator phenotype

Epigenetic silencing is a recently uncovered process that refers to clonal, irreversible changes in gene expression that cannot be accounted for by changes in DNA coding sequence (1). Silencing has been described in most multicellular organisms studied and seems to be essential for developmental processes. In humans, silencing is best recognized in inactivation of the X-chromosome in females, as well as inactivation of the silent allele at imprinted loci (2).

The molecular basis of silencing is increasingly well understood. Epigenetic silencing can be accomplished at the level of inhibited transcription or mRNA degradation. The latter mechanism, often referred to as RNA interference, has been studied in plant cells but is likely to play a role in mammalian cells as well (3). Transcriptional repression, however, is thought to be the major mechanism by which mammalian cells achieve silencing (1). Such repression is accompanied by changes in chromatin structure that result in exclusion of transcription factors. Chromatin changes, in turn, are mediated by changes in histone modification, including histone acetylation and methylation (4). In humans, DNA methylation within promoter-rich cytosine guanine dinucleotide (CpG)⁴ islands provides an additional level of epigenetic control, which interacts closely with histone modifications to achieve stable molecular silencing (2,5). Promoter-associated CpG island methylation, then, is a good and sensitive indicator of epigenetic silencing in human cells, although clearly there can be methylation-independent silencing as well.

Promoter-CpG island methylation has emerged as a common molecular defect in cancer cells (6). Evidence suggests that, for some genes, methylation provides a similar selective advantage as mutations or deletions, an argument that has led to the suggestion that hypermethylation should be considered as one of the two steps in Knudson’s model of tumor-suppressor gene inactivation (6,7). The control of CpG island methylation has been thought to be established early in embryogenesis (8). Current models suggest that DNA methylation patterns are erased early in embryogenesis and reestablished soon after implantation, with methylation limited to non-CpG islands.

¹ Presented at the “Trans-HHS Workshop: Diet, DNA Methylation Processes and Health” held on August 6–8, 2001, in Bethesda, MD. This meeting was sponsored by the National Center for Toxicological Research, Food and Drug Administration; Center for Cancer Research, National Cancer Institute; Division of Cancer Prevention, National Cancer Institute, National Heart, Lung, and Blood Institute; National Institute of Child Health and Human Development; National Institute of Diabetes and Digestive and Kidney Diseases; National Institute of Environmental Health Sciences; Division of Nutrition Research Coordination, National Institutes of Health; Office of Dietary Supplements, National Institutes of Health; American Society for Nutritional Sciences; and the International Life Sciences Institute of North America. Workshop proceedings are published as a supplement to The Journal of Nutrition. Guest editors for the supplement were Lionel A. Poirier, National Center for Toxicological Research, Food and Drug Administration, Jefferson, AR, and Sharon A. Ross, Nutritional Science Research Group, Division of Cancer Prevention, National Cancer Institute, Bethesda, MD.

² Work in the author’s laboratory is supported by grants from the National Cancer Institute, the American Cancer Society, and the George and Barbara Bush Foundation for Innovative Cancer Research.

³ To whom correspondence should be addressed. E-mail: jjissa@mdanderson.org.

⁴ Abbreviations used: CIMP, CpG island methylator phenotype; CpG, cytosine guanine dinucleotide; ER, estrogen receptor; MTHFR, methylene tetrahydrofolate reductase.
land areas, except for the rare genes silenced in normal cells. After birth, methylation levels have been thought to be largely stable throughout life. Hypermethylation in cancer, then, was initially seen as a random molecular event that is selected for in neoplastic cells, thus explaining its relative prominence in this situation (6). Evidence is emerging, however, in favor of an alternate model proposed years ago (9) whereby hypermethylation is initially a property of aging cells in normal tissues. In this model, it was proposed that normal individuals have substantial acquired variability in epigenetic silencing and that this variability is associated with risk of age-related diseases (10). In this review, the extent, sources, and implications of epigenetic variability will be discussed in the context of cancer risk.

**DISCUSSION**

**Epigenetic variation in humans**

DNA methylation within CpG-rich promoters is an excellent (but not exclusive) mark of silencing in human cells (2). Is there, then, variation in DNA methylation between different tissues, individuals, or disease conditions? CpG island methylation was until now thought to be uniformly absent in adult cells, except for the rare normally silenced loci such as the X-chromosome and imprinted loci (11). Several studies have now shown that, while this model is generally true soon after embryogenesis, aging tissues do, in fact, show some degree of CpG island methylation in somatic genes (10). This age-related methylation is likely to be the precursor for most aberrant methylation seen in cancer tissues (12). Thus, one observes at least two types of epigenetic variability: methylation in aging tissues and methylation in neoplastic tissues.

**Age-related methylation.** Promoter methylation of the estrogen receptor (ER) gene CpG island in breast cancer was one of the first reports associating methylation-related gene silencing with a potential physiological consequence in neoplastic tissues (13). The ER gene is essentially unmethylated in normal breast epithelium. In the colon, however, ER is highly methylated in tumor tissues and shows partial methylation in normal tissues that increases linearly with age (9). This finding of age-related methylation preceding neoplasia-related methylation has since been shown to be applicable to most genes hypermethylated in colon cancer (14).

Given that age-related methylation is usually incomplete (9), one can envision two distinct possibilities: a low (partial) degree of methylation in every target cell that then increases uniformly with age, or a patchy methylation whereby some colonic crypts have high degrees of methylation—and presumed silencing, whereas other crypts are essentially unmethylated and have not been affected by methylation-associated silencing. Two studies support the latter situation. A transgene under the control of a colon promoter shows age-related silencing in the mouse that is strictly crypt specific (15). Given that transgene silencing is often mediated by methylation (16), this finding supports the idea of crypt autonomy in the process of silencing/methylation and also suggests that the crypt stem cells are involved in that process. A recent study using microdissected crypts showed that age-related methylation is similarly patchy, with considerable crypt-to-crypt variability (17). These observations demonstrate that aging in human epithelium is accompanied by progressively increasing epigenetic mosaicism in colonic stem cells, such that there likely is considerable crypt-to-crypt variability in gene expression patterns.

Age-related methylation and epigenetic mosaicism in colonic epithelium appear to generate clonal molecular diversity at the level of crypt stem cells. Given that molecular diversity is one of the essential engines of neoplastic transformation, it is reasonable to speculate that age-related methylation is a fundamental predisposing event to spontaneous, age-related neoplastic transformation. Several lines of evidence point in this direction. First, most neoplasms of the colon have very high levels of methylation at many of these loci (12,18), suggesting either that neoplasia started in methylated cells or that the process accelerated in the tumors. Second, this type of methylation is also high in early abnormal lesions of colonic epithelium, including very small adenomas (14,19). Third, ulcerative colitis, a condition associated with a marked increased risk of colonic neoplasia, also has a higher rate of age-related methylation (20), specifically in patients with dysplastic lesions who have a high rate of neoplastic transformation.

The causes of age-related methylation in aging cells and tissues are incompletely defined. Fundamentally, one needs to make a distinction between “normal” events embedded in the structure of DNA and “abnormal” events related to specific disturbances that ultimately affect the methylation machinery. One can argue that the process is so ubiquitous in the population, regardless of the presence of neoplasia, that it becomes “normal” by default. Indeed, at a population level, the process seems to be strictly linear, and extrapolation suggests that it starts soon after birth and evolves at a very slow rate (21). Given that methylation patterns are erased early in embryogenesis and reestablished soon after implantation (8), it is possible to propose that the “wave” of methylation that initially establishes tissue-specific patterns does not stop at birth but actually evolves throughout life, albeit at a very slow rate.

**The CpG island methylator phenotype (CIMP).** Another important source of epigenetic variation can be found in the cancers themselves. Aberrant methylation in neoplastic cells varies between different tissues and also varies considerably between different individual patients of a given tumor type (22). One view of this variability is that, by analogy to mutations, aberrant methylation occurs randomly and stochastically and is selected for during the clonal evolution of neoplastic cells. This view, however, does not explain the striking finding in colonic neoplasia that selected genes (which do not show age-related methylation) are concordantly methylated in a subset of tumors (14). This process, termed “CpG island methylator phenotype,” was found to involve all genes studied in colorectal cancer so far (14,23,24) and imparts affected tumors a three- to fivefold higher rate of methylation at any given (predisposed) locus.

Genes affected by CIMP do not appear to share common features beside that of being predisposed to methylation-associated silencing. Functionally, these genes include tumor-suppressor genes, DNA repair genes, as well as potential oncogenes such as cyclooxygenase 2 (COX2) and telomerase reverse transcriptase (TERT) (25). Thus, CIMP seems to be a process that accelerates gene-specific susceptibility to hypermethylation, leading, again, to molecular diversity, this time in cancer cells. Clonal selection further shapes this molecular diversity, selecting for tumor-suppressor gene methylation events and against oncogene methylation events. A gene such as COX2, then, may be under the opposing pressures of methylation related to CIMP and expression related to clonal growth advantage, such that a minority of malignant tumors actually silence the gene, and some display prominent heterogeneity likely related to these opposing pressures (24). By contrast, P16 INK4A methylation is relatively infrequent in primary colonic tumors (26), but very frequent in cultured colon...
cancer cell lines (27), likely reflecting growth advantage imparted by cell cycle deregulation. CIMP, then, is broadly similar to other molecular diversity syndromes such as microsatellite instability and chromosomal instability (28), where multiple random (but predisposed) loci are affected, and a few are selected for (or against).

The causes of CIMP are unknown. Given the dichotomous nature of the phenomenon, it is possible to speculate that a genetic event could lead to this type of aberrant methylation. Candidate genes include the DNA-methyltransferase [DNA (cytosine-5)-methyltransferase; EC 2.1.1.37] enzymes themselves, as well as the elusive proteins that appear to protect against CpG island methylation during development (29). It is possible, however, to consider alternative explanations. For example, CIMP-positive tumors could be arising from a different cellular compartment than CIMP-negative cases. In the colon, for example, the recently described hyperplastic polyp/serrated adenoma/microsatellite-positive carcinoma pathway appears to evolve primarily along a methylator route (30). Given the distinct morphological features of these tumors, which are evident from their earliest stage, it is possible that they arise from a cell distinct from the common colon cancer precursor and with a heightened propensity for CpG island methylation. Another possibility is that CIMP is caused by specific environmental exposures that directly or indirectly lead to methylation (see below), without requiring specific molecular defects in the methylation machinery. In a recent study (31), patients with hyperplastic polyposis of the colon, a condition associated with multiple polypos and colon cancer predisposition, were found to have highly concordant methylation in their separate lesions. These findings support the concept of patient-specific methylation defects related to genetics, lifestyle or exposures.

A model of epithelial carcinogenesis. Combining the two processes described above, one can propose a model of epithelial carcinogenesis that melds the epigenetic and genetic changes that lead to age-related carcinogenesis (32). It is proposed that age-related epigenetic mosaicism subtly alters the expression profiles of individual stem cells, such that some of these become predisposed to transformation through defects in proliferation, DNA repair, differentiation, or apoptosis. Most tumors in adults would arise from such “primed” cells, explaining the high incidence of aberrant methylation in cancer. Neoplastic transformation, however, clearly requires a number of specific molecular alterations, some of which are rate limiting (33). These subsequent steps, then, could be genetic or epigenetic, largely dependent on the molecular instability pathway chosen by the individual cancer. In CIMP-negative cases, epigenetic defects would predominate, with associated genetic defects related, for example, to inactivation of DNA repair genes by hypermethylation. In CIMP-negative cases, genetic defects would predominate instead, including a high preponderance of chromosomal instability. In the colon, genetic analysis preliminarily confirms these predictions (34).

Is this model applicable to other tissues? Both published and unpublished evidence supports the presence of age-related methylation in other epithelial tissues such as liver, stomach, bladder, and prostate (35,36) (Issa, J.-P., unpublished results). The CIMP phenotype has also been described in other epithelial cancers (37–40). It has been difficult to demonstrate age-related methylation in nonepithelial tissues, however. It is not clear whether this reflects the choice of genes analyzed or the properties of the methylation machinery in these different tissues. Nevertheless, cultured fibroblasts do show passage-dependent increased methylation of ER (41) and insulin-like growth factor 2 (IGF2) (44), whereas cultured vascular smooth muscle cells show high levels of ER methylation (43). It is likely that some variation of these processes will be found in all tissues and cell types.

Implications for nonmalignant age-related diseases. Epigenetic mosaicism and the model developed above have potential implications for the pathogenesis of several age-related diseases. Atherosclerotic vascular disease, for example, is characterized by discrete foci of abnormal proliferation (44), and it is easy to see how epigenetic variability could result in the acquisition of a gene expression profile that favors the development of this process. Indeed, hypermethylation of the ER gene has been observed in nonneoplastic vascular cells (43,45). Similar dynamics could be at work in other aging diseases, including neurodegenerative diseases.

Sources of epigenetic variation

An essential feature of age-related methylation is the high degree of patient-to-patient variability one observes within a relatively homogeneous group of patients (21). In fact, aging alone accounts for a minor degree of the variability in a population (20–40%). Sources of variation, then, include both technical issues as well as biological issues. Gene-specific methylation measurement continues to be an imperfect art, and until assays are improved, this will always be a source of variability. Nevertheless, there is emerging evidence suggesting that exposures and lifestyle factors, including diet, could affect the process of epigenetic variability.

Tissue-specific factors. Different normal tissues and cell types show distinct patterns of promotor methylation both in normal-appearing tissues and neoplastic tissues (12). Of course, the latter observation may well be related to the former in that age-related methylation was found to be the predominant explanation for hypermethylation in cancer (14). Examples of these tissue-specific differences include the fact that ER methylation is high in normal liver, intermediate in normal colon, and very low in normal breast epithelium (12). The frequency of ER methylation in the respective tumor types follows this pattern (9,13). These differences correlate with differences in basal gene expression levels and suggest that active transcription may be part of the mechanism that protects CpG islands from age-related and neoplasia-related methylation. A second important source of tissue variability in methylation changes, particularly in neoplasia, relates to differential selective advantages for gene-specific epigenetic inactivation. For example, retinoblastoma 1 (RB1) methylation is almost exclusively found in retinoblastomas (46), whereas von Hippel-Lindau (VHL) methylation is almost exclusively found in renal cancers (47). In both of these cases, the tissue type of methylation is the same as the tumor type in patients who carry germ line mutations in these genes, suggesting that the methylation tissue distribution relates to strong selective pressure for inactivation in those tissues (but not others). Other likely sources of tissue-specific epigenetic variability are differential exposure to carcinogenic insults (48), differential rates of proliferation, as well as, possibly, varying intrinsic susceptibility to aberrant methylation.

Genetic factors. There are few studies that address the potential for genetic changes to cause or influence epigenetic patterns. At a gene-specific level, it has been shown that expansion of a triplet repeat in the fragile X mental retardation gene (FMR1) promoter leads to hypermethylation and gene silencing (49). Similarly, the 9:22 chromosomal translocation typical of chronic myelogenous leukemia is associated with hypermethylation of the translocated Abelson murine leukemia oncogene (ABL) promoter (50,51). Other than these two
observations, no specific genetic anomalies have been described in association with hypermethylation in normal tissues or cancer. In fact, most genes hypermethylated in cancer have no associated nucleotide changes, including areas in and around the promoter (6). In addition, no specific gene mutations in the methylation machinery have been described in cancer, to date, although such mutations could conceivably explain simultaneous hypermethylation of multiple genes in the subset of cancers affected by CIMP (18).

A promising area of investigation is the relationship between polymorphisms in common genes and hypermethylation in normal and cancer tissues. Polymorphisms in the methylenetetrahydrofolate reductase (MTHFR; EC 1.5.1.20) gene have been reported in preliminary studies to be associated with variation in global levels of methylation (52). It is not yet known whether such polymorphisms affect gene-specific promoter methylation in normal or neoplastic tissues, although preliminary data suggest that it is not the case in the colon (Issa, J.-P., unpublished results). Other promising genes to study include additional genes in the folate pathway and the DNA [(cytosine-5)-methyltransferase (EC 2.1.1.37)] family of proteins.

**Exposure factors.** DNA methylation has long been suspected to mediate some of the effects of environmental exposures and lifestyle factors on disease risk (53). Early studies demonstrated an effect of both carcinogen exposures (54) and methyl-deficient diets on global methylation levels (55). The recognition of promoter CpG island methylation as a potential mediator of epigenetic alterations in neoplastic cells prompted more recent studies into this process.

Physical carcinogens have been shown to influence the degree of methylation in tumors in both experimental systems and epidemiological studies. It has been demonstrated that cells and tumors exposed to chemotherapeutic agents had a higher frequency of inactivation of specific genes (56), although, in this setting, it is difficult to tease out direct effects on methylation from clonal selection for methylated cells. In animal models of lung carcinogenesis, exposure to different agents resulted in varying frequency of aberrant methylation (48), which occurs quite early in some cases (57). Recently, epidemiological observations have suggested a link between tobacco exposure and aberrant methylation in human lung cancer (58). The mechanisms underlying these observations are unknown. In particular, it is not clear whether carcinogenic exposures affect methylation directly or whether the inflammation and injury responses result in increased cell turnover that is then associated with accelerated methylation, as has been described in ulcerative colitis (20).

Viruses are known to trigger DNA methylation processes on cellular infection, and such methylation can sometimes spread to involve endogenous loci adjacent to the sites of viral integration (59). In cell culture, viral transformation can also trigger aberrant methylation and expression of DNA [cytosine-(5)-methyltransferases (41,60). Epidemiologically, infection-driven tumors such as hepatitis B- and C-related liver tumors (61) and Helicobacter pylori-related gastric cancers (62) seem to have higher levels of gene-specific methylation than the tumors that arise independent of such exposure. Here again, the mechanism of the effect is unknown, although it is tempting to speculate that inflammation and chronic proliferative stimulation, so characteristic of hepatocellular disease associated with viral infection, play a major role in accounting for aberrant methylation.

Dietary factors are attractive candidates to explain some of the variation one observes in age-related methylation, but the large studies needed to analyze this issue remain to be done. In liver cancer, chronic alcohol exposure and associated hepatitis and cirrhosis are accompanied by high levels of methylation of P16 and other genes (63). This effect may well be indirect, through an injury response analogous to viral-driven tumors. In the colon, a small study suggested that high-fiber diets are associated with reduced levels of ER methylation, whereas reduced estrogenic hormone levels (via premature menopause, for example) were associated with higher levels of methylation (Issa, J.-P., unpublished results). This effect, which needs to be reproduced in large studies, may be mediated via direct effects on gene expression or indirect effects on proliferation and inflammation.

Overall, there are few studies that directly test the hypothesis that exogenous exposures are associated with different degrees of aberrant gene-specific methylation and silencing in normal aging tissues and in neoplasms. This is one of the most important questions that will need to be addressed using large studies in the next few years.

**Clinical implications of epigenetic variation**

The processes described above have several implications relevant to the clinical management of age-related diseases. A central prediction of epigenetic variability is that individuals with high degrees of this phenomenon are at increased risk of tumor formation (32). Thus, it is possible that measurement of methylation in normal epithelium could be used to stratify patients for screening and prevention approaches. This will require careful studies of appropriate gene markers and methods of detection with prospective validation of positive results. Another implication of these studies is that preventive approaches could have different efficacy depending on the underlying epigenetic state of the target lesion. For example, colonic tumors with COX2 methylation may not be responsive to COX2-targeted chemopreventive approaches (24).

The most remarkable difference between genetic and epigenetic events in carcinogenesis is that the latter are potentially reversible with pharmacological manipulation (64). Several drugs that manipulate methylation levels through inhibition of DNA methyltransferases are now in clinical trials for patients with cancer (65), and it is possible that, should safe ones be identified, these approaches could be applied to disease prevention.

**LITERATURE CITED**


