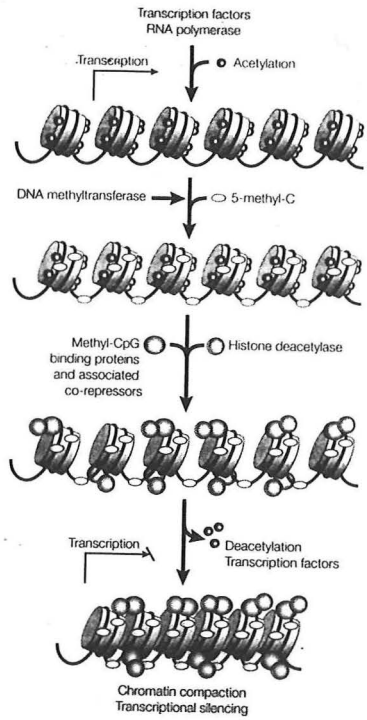
University of Texas Southwestern Medical Center
Internal Medicine Grand Rounds

DNA Methylation: A Two-Edged Sword for Embryogenesis and Cancer

Rody P. Cox, M.D.

February 8, 2001



This is to acknowledged that Rody P. Cox, M.D. has disclosed no financial interests or other relationships with commercial concerns related directly or indirectly to this program. Dr. Cox will be discussing "off-label" uses in his presentation.

Rody P. Cox, M.D.
Professor of Internal Medicine
Clinical and Research Interests:
General Internal Medicine
Medical Education
Biochemical Genetics
Maple Syrup Urine Disease

Cover Illustration: The mechanism whereby DNA methylation and histone deacetylation cooperate to repress transcription. From reference 15.

I acknowledge the outstanding organization and presentation skills of Jan Zerwekh in the preparation of this Grand Rounds lecture.

Introduction

DNA methylation has recently received attention because of its importance in the regulation of gene expression. Most of the genes in the human genome have been sequenced and we are now faced with the challenge of understanding the molecular mechanisms that allow genes to be selectively expressed. Restrictions on gene function are required for the differentiation of specialized cells and a mechanism must exist to turn genes on and also to inactivate them at appropriate times. Failure to repress genes appropriately may result in human diseases and I would like to review certain ones of these today.

Roles for DNA Methylation

1. Embryonic development requires the participation of DNA methylation in order for normal organogenesis and for successful viable birth.
2. Transposons are segments of DNA that move or "hop" from one chromosome to another. They insert themselves randomly into mammalian DNA and potentially disrupt the function of genes. DNA methylation defends the integrity of the genome against these endogenous parasitic elements.
3. DNA methylation has a number of physiological functions including the inactivation of genes on one of the X chromosome, that is silenced in every female cell.
4. The process of DNA methylation is intimately concerned with genetic imprinting, a process that inactivates one of the parental genes so that monoallelic expression of the maternal or paternal occurs in certain normal tissues and some cancers.
5. There are several diseases involving the DNA methylation protein complex where deficiencies in one or another of these proteins leads to inherited syndromes in humans.
6. One of the major roles for DNA methylation is in the origin of tumors.
7. Pharmacogenomics promises that controlling the DNA methylation process may lead to new therapies, not only for cancer, but for other disease processes.

I will discuss each of these areas to provide an overview of the importance of this epigenetic phenomenon in gene regulation.

The concept that epigenetic as well as genetic events may be central to the evolution of gene control and gene silencing was proposed about 20 years ago (1-4). However, until recently, epigenetics has had little impact on medicine. The changes in DNA that are produced by DNA methylation and the attendant alterations in chromatin constitute a heritable state that is not mediated by altered nucleotide sequences in a gene (mutation) or chromosomal aberrations, but appears to be tightly linked to the formation of transcriptionally regulated chromatin. Alterations in methylation patterns in gene promoters or regulatory sequences may disrupt gene function by hypermethylation or, in other instances, restore the function of genes that have previously been inactivated by removing the methylated sequence and reversing the alterations in chromatin. The consequences of epigenetics as exemplified by DNA methylation will enhance our understanding of chromatin modeling and gene regulation. This understanding may present novel possibilities for therapy as well as produce potential sensitive molecular markers to define risk of diseases, monitor prevention strategies, achieve early diagnosis and track the progress of diseases that are dependent on abnormal DNA methylation.

Gene Silencing and Activation

DNA methylation with the accompanying chromatin modifications are two global mechanisms that regulate gene expression. DNA methylation in mammalian cells occurs at the 5-position of cytosines within the CpG dinucleotide. Approximately 70 percent of all CpGs are methylated, but neither 5-methylcytosine distribution nor the special distribution of the CpG dinucleotide is random (5). CpG repeats often occur in clusters called "islands" and are important in cell regulation. These CpG islands account for only about 1 percent of the genome and for 15 percent of the total genomic CpG sites. However, these CpG islands are frequently located within or near promoters of genes and approximately half of them are unmethylated in normal cells (6,7). It is noteworthy that these CpG sites are also hot spots for germ-lined mutations and contribute about 30 percent of all single nucleotide substitution mutations in the germ-line (8). They also are involved in acquired somatic

mutations that may lead to cancer. The reason for the high mutation rate in CpG sites is that during DNA replication, 5-methylcytosine is deaminated to thymine and the thymine substitution represents a point mutation when thymine pairs with adenine. This may lead to an amino-acid substitution in the protein product of that gene affecting its function, i.e. a mutation.

Chromatin Structure

Before discussing the details of DNA methylation, it is necessary to digress and describe the material which comprises chromosomes. This is chromatin which consists of fibers that contain protein and DNA in approximately equal mass plus a small amount of RNA. The DNA in the chromatin is very tightly associated with proteins that are called histones which package and order the DNA into structural units called nucleosomes (9). Chromatin also contains many non-histone proteins, some of which regulate the expression of specific genes. Prior to the experiments of Mendel on sweet peas, it was believed that heredity was a blending of the traits of two parents, giving rise to such expressions as "blood brothers" or "blood lines". Mendel showed that traits were inherited as particles and geneticists conceptually thought for many years that the genes were like beads on a string. It has been found that, indeed, the DNA is wound around a core consisting of the basic protein, histones. The various nucleosomes are then linked together very much like beads on a string (10). The nucleosomes are, therefore, the fundamental unit of organization from which the higher order of chromatin into chromosomes is built. Each nucleosome contains four dimers of histone molecules, designated H2A, H2B, H3 and H4 with 146 DNA nucleotides wound around the histone tetramers (11). The histones that comprise the core of the nucleosomes are not inert balls of protein, but rather they may exist in different configurations because they can be enzymatically modified by methylation, phosphorylation, or acetylation. Such modifications change the histone molecule's shape and other properties. Hyperacetylation of the histones tend to loosen the nucleosome structure so that the DNA surrounding the histone core is extended and becomes available for transcription. In other words, the gene is activated. Contrarily, if the histone molecules are deacetylated, they tend to condense and to pack more tightly, thereby compressing and inactivating the DNA which surrounds them, silencing the genes that are so enclosed (12-14).

DNA Methylation Complex

DNA methylation patterns are established by a complex interplay of three independent DNA methyltransferases. They are designated DNMT-1, DNMT-3A and DNMT-3B as shown in Figure 1 (15). The predominant methyltransferase is DNMT-1, but all three of these transferases are required for normal embryonic development in mammals (16,17). DNMT-1 is the most abundant methyltransferase and reacts with the proliferating cell nuclear antigen. This enzyme is responsible for "maintenance of the methylation patterns" because of its high affinity for hemimethylated DNA during and after DNA replication, but it also possesses limited *de novo* methylation capabilities. DNMT-3A and DNMT-3B are primarily involved in the *de novo* methylation of CpG islands in the DNA.

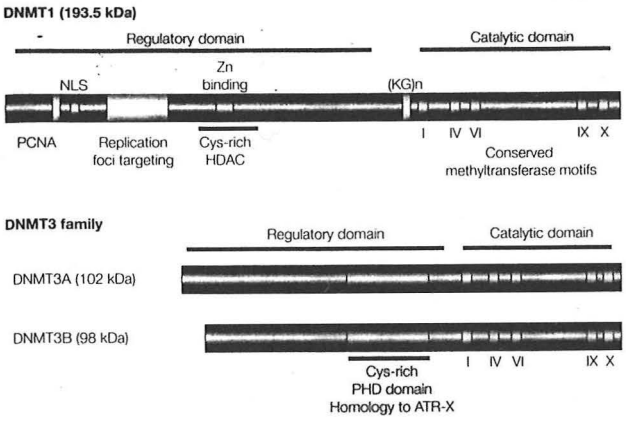


Figure 1. **Schematic structure of the three catalytically active DNA methyltransferases in mammals showing the N-terminal regulatory and C-terminal catalytic domains, and other regions with known or proposed functions.** The catalytic domains of the three enzymes are conserved, but there is little similarity between their N-terminal regulatory domains. PCNA, domain that interacts with proliferating cell nuclear antigen; NLS, nuclear localization signal; (KG)n, lysine-glycine repeat hinge region; HDAC, histone deacetylase interaction domain; PHD, plant homeodomain motif that shows homology to the *ATR-X* (α-thalassaemia, mental retardation, X-linked) gene. (15)

The effects of DNA methylation on chromatin structure was an early observation. It was recognized that this process caused chromatin and its constituent nucleosomes to condense into heterochromatin which is inactive and transcriptionally silent. Unmethylated DNA in the promoter and regulatory regions of genes resulted in the chromatin being in an extended form that was readily available to the RNA polymerases necessary for gene

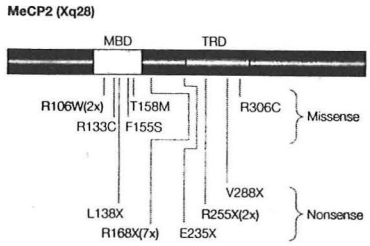


Figure 2. **Summary of mutations identified in the MECP2 protein in people with Rett syndrome.** The methyl-CpG binding domain (MBD) and the transcriptional repression domains (TRD) are indicated. All missense mutations involve highly conserved amino acids within defined functional domains. Nonsense mutations would result in truncated proteins (if stable) that would delete some or all of the conserved domains. The functional consequences of all mutations have yet to be investigated *in vitro*. Most *MECP2* mutations arise from C-to-T transitions within CpG dinucleotides owing to the increased mutability of 5-methylcytosine, implying that these sites are methylated *in vivo*. Several mutational hot-spots have also been identified (R106W, R168X and R255X), The bracketed numbers indicate how many times these mutations have been found (15).

transcription and microscopically this region appears as euchromatin. The mechanism whereby DNA methylation was responsible for and coordinated the chromatin structure has recently been partially illuminated. DNA methylation and histone deacetylation are intimately linked by chromosomal proteins that bind specifically to methylated DNA (18-20). One of these is the well characterized methyl-CpG binding protein-2 (MeCP-2). Figure 2 shows the methyl CpG binding domains designated MBD and the transcriptional repression domain, TRD, of the MeCP-2 protein (15). It is of interest that the gene for MeCP-2 is located on the X chromosome and mutations in this gene are responsible for Rett's Syndrome, a major cause of mental retardation in females, which will be discussed later. MeCP-2 associates with co-repressor protein complexes that include an enzyme, the histone deacetylase. The recruitment of histone deacetylases by MeCP-2 occurs indirectly through its association with an adaptor protein called Sin 3A and perhaps several other proteins (21,22). The subsequent deacetylation of histones results in condensation of the nucleosomes and results in transcriptional silencing.

The role of histone acetylation in gene expression was suggested by studies from my laboratory at NYU in the early 1970s. We found that sodium butyrate could cause the condense chromatin in tumor cells to become extended and to be transcribed. We found that the synthesis of certain oncofetal proteins were then induced in these tumor cells. These proteins include the fetal or placental form of alkaline phosphatase, human chorionic gonadotropin and follicle stimulating protein (23-25). Vernon Ingram at MIT, showed that butyrate was a potent inhibitor of the histone deacetylase and resulted in hyperacetylation of histones (26). My laboratory then showed that butyrate also had an effect on cell cycling and synchronizing cells by inhibiting the late G1 phase

of the cell cycle just before DNA synthesis when chromatin is decondensed and available for transcription (27,28). It is of interest that the new molecular techniques have substantiated these studies from our laboratory. Figure 3 shows a transcriptively active region being targeted for silencing by DNA methylation (15). The subsequent deacetylation of the histones by the recruited histone deacetylase produces the eventual condensation and compaction of the chromatin and transcriptional silencing of that gene.

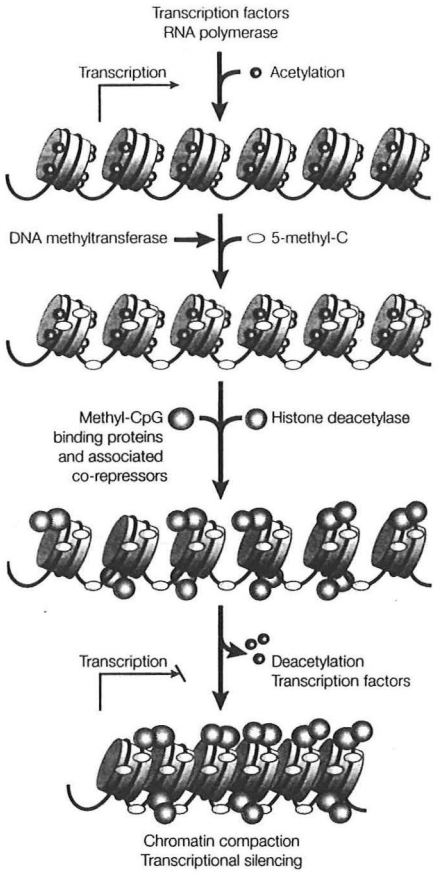


Figure 3. **The mechanism whereby DNA methylation and histone deacetylation cooperate to repress transcription.** A transcriptionally active region targeted for silencing is proposed to acquire DNA methylation first, which then recruits the methyl-CpG binding proteins and their associated co-repressors and histone deacetylases (HDACs). As DNA methyltransferase 1 (DNMT 1) can interact directly with histone deacetylase, it is also possible that transcription is first silenced by deacetylation by other tethering factors, after which the methylation machinery and the methyl-CPG binding proteins are recruited to "cement" the promoter in the silent state. In either case, the deacetylated nucleosomes adopt a more rightly packed structure that inhibits the access of transcription factors to their binding sites (15).

Effect of DNA Methylation on Embryonic Development

The DNA of vertebrates contains tissue specific patterns of methylated cytosine residues. These methylation patterns are transmitted by clonal inheritance through the strong preference of the DNA methyltransferase-1 (DNMT-1) for demimethylated DNA (29,30). This property propogates the methylation pattern from one cell generation to another. In mice, a disruption of the methyltransferase gene by homologous recombination using embryonic stems cells (ES

cells) was carried out to obtain homozygous mutant DNMT-1 deficient ES cells (16). The homozygous mutant ES cells showed a normal morphology and growth rates in tissue culture with no discernible unusual phenotype. Two independent lines of ES cells which were heterozygous for the mutation in the methyltransferase gene were injected into mouse blastocysts which were transferred to the uteri of pseudopregnant females. Mice heterozygous for the DNMT-1 mutation were indistinguishable from the wild-type litter mates. Male heterozygotes were intercrossed with heterozygous females and the progeny were genotyped by Southern blot hybridization. All viable mice that were born were either homozygous normal or heterozygotes for the methyltransferase mutation. However, the litter size was small and it was determined that approximately one-quarter of the embryos had died *in utero* (16).

Examination of the homozygous DNMT-1 mutant embryos isolated in mid-gestation at Day 9½-10 were stunted in their developmental stages. Analysis of the homozygous embryos revealed that major organ rudiments were present, but they were smaller than wild-type litter mates. Close inspection revealed significantly increased cell death and considerably reduced mitosis in the homozygous as compared with wild-type or heterozygous embryos. Immuno-blot analysis and quantitation indicated a marked reduction in the amount of methyltransferase protein in homozygous mutant embryos. The amount of radioactive methyl groups incorporated into cytosine was approximately 30 percent of the wild-type in homozygous embryos. The conclusions of this study were that DNA methylation was required for normal embryonic development and that the DNMT-1 enzyme was essential for normal embryonic development beyond mid-gestation. In this study, there was some confusion regarding the 30 percent of residual methylation of DNA observed in homozygous mutants. However, later, it was determined that there were two other isozymes of the methyltransferase, DNMT-3A and DNMT-3B.

To determine the roles of DNMT-3A and DNMT-3B for *de novo* methylation of DNA during embryonic development, DNMT-3A and DNMT-3B deficient mice were generated from embryonic stem cells (17). Analysis of the knock-out mice for each of the DNMT-3 isozymes indicated that the heterozygous mice were

grossly normal and fertile while the homozygous mutant mice, which had been derived from intercrosses of heterozygous mice survived to term, but the mutant mice were runted and died at about four weeks of age. The major organ system affected in the mutants was the nervous system, particularly the rostral regions of the nervous system. Although the DNMT-3A and 3B enzymes exhibit overlapping functions, they each carry out *de novo* methylation of different sequences in developing mice. However, the two DNMT-3 isozymes have subtle but distinct functional properties. DNMT-3B mice are defective in methylation of tandem repeats in the centromeric region of chromosomes. Recently, a mutant human DNMT-3B gene was shown to be associated with a syndrome in humans called ICF (31-33). Children with this defect show Immunodeficiencies, Centromeric instabilities, and Facial anomalies. DNA methylation studies showed that the satellites which are major components of heterochromatin adjacent to centromeres are hypomethylated in this syndrome, similar to the hypomethylation of centromeric minor satellite repeats in the homozygous DNMT-3B mutant mice.

Transposons and the Integrity of the Human Genome

Transposable elements are mobile sequences that occur throughout the genome. There are two major classes of these elements, DNA transposons and retrotransposons (34). Retrotransposons are self-replicating RNA templates that are probably derived from ancient retroviruses (35,36). These unstable ribose-based polymers are converted to stable DNA by reverse transcription. The first class are DNA transposons which are maintained in chromosomal sites where they are transcribed. On the other hand, retrotransposons are reversely transcribed into the DNA prior to their integration into chromatin. These mobile elements influence genomic structure and function through several mechanisms, including insertion into exons, transduction of 3′ flanking sequences and recombination of repetitive sequences with unequal crossovers. It has been estimated that transposon elements may account for as much as 40 percent of the DNA in the human genome, while exons that are the coding sequences for proteins comprise as little as 5 percent (34,35). It has been postulated that DNA methylation may have originated as a genome defense system to silence expressions of elements such as transposons and retrotransposons and to limit their spread through the genome (34,37). These elements may be considered as parasitic DNA which represents a significant threat to the structural integrity of the genome

because they can mediate recombination between non-allelic repeats and also can cause chromosome rearrangements or translocations. The retrotransposons may integrate in genes and disrupt their expression. On the other hand, others have postulated that the transposons and other parasitic elements provide a flexibility to the genome and their shuffling into new sites may have created new genes during human evolution. That retrotransposons of the long interspersed nuclear elements (LINE or L-1) subclass may contribute to disease was established by demonstrating their presence in factor VIII gene of a hemophilia A patient (38,39), their insertion into the gene for Duchenne muscular dystrophy (40,41), their involvement with type II retinitis pigmentosa (42), the β globulin gene of the β thaloassemia (43) and chronic granulomatosis disease (44). The L1 element in these five families were either germ-line or early embryonic insertions and were inherited in these families. A rare cause of a somatic mutation in the APC gene of colon cancer was the insertion of an L1 element in the last exon (45). Although there are estimated to be over 100,000 L1 elements in the human genome, most are defective and are incapable of retrotransposition. Thus, only small numbers of L1s are able to be mobile and insert themselves into genes. The number of patients reported whose mutation are the result of L1 insertions is relatively small, but this may represent a major underestimate since detection of L1 insertion has technical difficulties (34).

A second sub-class of mobile elements are the non-autonomous retrotransposons, primarily *Alu* elements, which require the reverse transcriptions encoded by L1 retrotransposons for their conversion to DNA transposons. *Alu* retrotransposition also has been associated with human disease. Eleven insertions of full-length Alu elements have been identified as the molecular basis of disease. These include insertions in the factor IX gene in hemophilia B (46), the NF1 gene in neurofibromatosis (47), the FGFR2 gene in Aperts syndrome (37), the DPC gene in desmoid tumors (37), the XSCID gene in X-linked immunodeficiency disease (37) and BRCA2 gene in breast cancer (48). All of these Alu insertions are in germ-line cells and are inherited.

Since most retrotransposons have been discovered in genes of germ-line or undifferentiated cell, it is suggested that undermethylation at CpG residues and the concomitant hyperacetylation of histones in the nucleosomes

facilitate the transposon insertion. This supports the hypothesis that one of the major functions of CpG methylation is to reduce the expression of mobile elements in differentiated cells and protect the integrity of our genome from these endogenous parasitic DNA elements.

X Chromosome Inactivation

X chromosome inactivation is the mechanism used in mammals for dosage compensation of X-linked genes between chromosomally XX females and XY males. One of the two X chromosomes of females becomes transcriptionally inactive in every cell of the early embryo and remains so in all somatic cells throughout life. This represents a highly unusual form of gene regulation in that the whole or almost all of the chromosome is silenced. Chromatin is heterochromatic, condensed, late replicating, heavily methylated CpG's and histones are hypoacetylated. X chromosome inactivation is initiated from an X inactivation center on the X chromosome located at Xq13 region (49). The X inactivation center in man is called XIST and is expressed from the inactive X chromosome and is necessary for *cis* X inactivation. This region of the X encodes a non-translated RNA. The mechanism whereby the RNA spreads throughout the inactive X chromosome and how it induces transcriptional silencing are unknown. Recently, Mary F. Lyon has suggested that interspersed nuclear elements, of the LINE-1 class, act as promoters of spreading. This hypothesis would imply that the retrotransposons actually have a function physiologically to silence one of the X chromosomes in female cells (49). Recent evidence indicates that the X chromosome is, indeed, rich in L1 elements, there being about 26 percent of L1s on the inactive X chromosome and only approximately 13 percent of L1s on autosomal DNA (50). Further support for this hypothesis is that approximately 10 percent of genes on the human X chromosome escape inactivation being expressed from both the active X chromosome as well as the putative inactive X and the L1 content of these active segments is markedly reduced.

Irrespective of the role of transposons in X chromosome activation, it is clear that the inactive X has the properties that are typical of heterochromatin. It is condensed and replicates its DNA late in the S phase of the cell cycle. Its histones are hypoacetylated and the cytosines in the CpG islands are heavily methylated. The

acquisition of heterochromatic properties by the inactive X is believed to be the result of coating with the RNA produced by the X inactivation center. It may be that the L1 elements serve as binding sites or as boosting stations for the RNA, perhaps acting through RNA and protein complexes. However, it is clear that the methylation of the CpGs is involved in stabilizing the inactive state of the X chromosome. This is an extension of the concept that gene silencing is a result of the cell defense mechanism that senses the presence of invading genomic parasites such as transposable elements and silences them. It would appear that the cell makes use of this mechanism to provide dosage compensation by silencing one X chromosome in every female cell with the L1 transposons acting as the instigating influence. This suggestion, however, remains speculative.

Genetic Imprinting

Genomic imprinting is a process similar to that described under X chromosome inactivation. However, only one gene is expressed according to its parental origin. Genetic imprinting is a phenomenon where only one allele of a parent is expressed while the other is silenced and inactivated. Epigenetic events, primarily DNA methylation with histone deacetylation silences one parent's gene or an area of a chromosome in which several parental genes may reside. This arrangement of coordinately imprinting genes within a cluster offers some insight into the mechanisms by which cells establish and maintain appropriate imprints on functionally related genes (51-54).

Beckwith Wiedemann Syndrome (BWS)

As an example, Beckwith Wiedemann Syndrome occurs on human chromosome 11p15.5. The usual phenotype exhibits organomegaly, hypoglycemia, hemihypertrophy, genitourinary abnormalities, and in 5 percent of children, embryonal tumors, most frequently Wilms' tumor. The vast majority of familial clustering and sporadic cases of BWS show no cytogenetic abnormality, but approximately 2-5 percent do have duplications or inversions or translocations affecting the distal 11p portion of the chromosome. It appears that the imprinting of one or more genes on the short arm of chromosome 11 causes this syndrome. In familial cases, the risk to offspring depends on the sex of the transmitting parent with maternal transmission associated with a greatly

increased penetrance. Approximately 20 percent of sporadic cases demonstrate paternal disomy for chromosome 11p15.5. The characteristics of the imprinting is hypermethylation of this region with histone condensation which is observable as heterochromatin and late DNA replication in this region. At least 4-5 genes that contiguously lie in this region have been implicated in contributing to the BWS syndrome (55).

Prader-Willi and Angelman's Syndromes

Other human genetic disorders are attributed to genetic imprinting or imprinting centers that involve a region of a chromosome that contains several genes. Two of these diseases are Prader-Willi Syndrome and Angelman's Syndrome which are neurobehavioral diseases. Prader-Willi Syndrome is characterized by infantile hypotonia, gonadal hypoplasia, obsessive-compulsive behavior, and a hyperphagia leading to profound obesity. The paternal allele is imprinted and transmission is through the maternal locus. Angelman's Syndrome is characterized by mental retardation with microcephaly, a round face, large mouth, absent speech, ataxic gait, and a happy demeanor. Transmission is by the paternal locus while the maternal gene is imprinted. The genetic locus for these diseases encompasses chromosome 15q11-q13. This region may cytologically have deletions or chromosomal rearrangements, although the underlying silencing of the genes appears to be by hypermethylation and consequent histone condensation (56).

Genetic Diseases Due to Mutations in the DNA Methylation Complex

Three diseases have been described that are due to deficiencies in the proteins of the DNA methylation complex. They include the ICF, Rett and Fragile X Syndromes. A common feature of all of these diseases is a variable degree of mental retardation implicating DNA methylation-dependent gene control pathways as being particularly important for brain development and function.

ICF Syndrome

The ICF Syndrome is autosomal recessive inherited disease. The major findings are a variable Immunodeficiency with reduction in serum immunoglobulin levels; instability of Centromeric regions of

chromosomes, particularly numbers 1 and 16; Facial anomalies include telecanthus, epicanthic folds, low set ears, and protrusion of the tongue. Most patients also exhibit a large abdomen with thin arms and legs. They have significant psychomotor retardation and are delayed in their ability to sit and to walk. Most patients are recognized during the first year of life because of severe recurrent respiratory infections which are attributed to combined immunodeficiency. The most frequent infections are severe recurrent respiratory infections, purulent sinusitis and episodes of pneumonia (56,57).

Cytogenetic abnormalities are extensive and include juxtacentromeric heterochromatin which is greatly elongated and thread-like in metaphase chromosomes, these alterations may be associated with the formation of complex multiradiate chromosomes. The centromeres appear embryonic and are hypomethylated. These juxtacentromeric regions are subject to persistent interphase self-association and may be extruded into nuclear blebs or micronuclei. Abnormalities are largely confined to satellites 2 and 3 at the juxtacentrmeric regions of chromosome 1and 16 and, occasionally, chromosome 9 (58). Normally, satellite DNA is heavily methylated at cytosine residues, but in ICF Syndrome, it is almost completely unmethylated in all tissues (59,60).

The basic defect in DNA methylation in ICF Syndrome are mutations in the DNA methyltransferase-3B (DNMT 3B). As shown in Figure 4, these mutations tend to cluster at the 3 ' region of the gene and involve the catalytic domain (15). The DNMT 3B gene has been mapped to chromosome 20q11.2, the ICF susceptibility locus (61). Most of the ICF mutations leave the end-terminal regulatory domain intact and this region of the protein may be critical for embryonic survival. Mutations or loss of the end-terminal domain may be lethal to human embryos. The effect of a defect in centromeric methylation is not known. Targeting of DNMT 3B to the centromeric heterochromatin may be accomplished by its interaction with DNA binding proteins. It is hypothesized that altered centromeric chromosome structure may globally change gene transcription in a subtle way and this may affect the development of the brain more than of other organs because of the complexity of neurological differentiation.

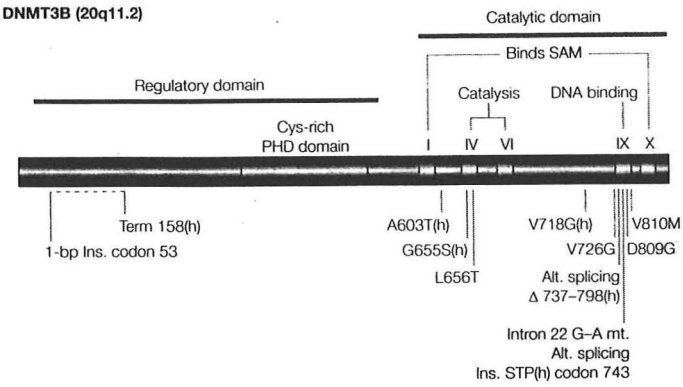


Figure 4. **Summary of mutations in DNA methyltransferase 3B (DNMT3B) identified in ICF syndrome patients.** Many individuals were compound heterozygotes for two independent mutations (denoted with an h). Nearly all mutations affect the catalytic domain. Loss of catalytic activity was confirmed for the D809G mutation in cell line overexpression studies. The conserved methyltransferase motifs within the catalytic domain are indicated with roman numerals. The cysteine-rich PHD (plant homeodomain) region in the N terminus is highly homologous to a PHD motif within the α–thalassaemia, mental retardation, X-linked (ATR-X) protein. The S-adenosyl-$_L$-methionine (SAM) binding domain is indicated, as well as motifs involved in catalysis (IV,V) and DNA binding (IX). Several mutations (methyltransferase.) result in altered splicing patterns (alt.) and the insertion (ins.) or deletion (Δ) of coding sequence (15).

Rett's Syndrome

Rett Syndrome is a neurodevelopmental disorder characterized by loss of acquired skills after a period of normal development in infant girls (62). At the age of 7-18 months, developmental stagnation occurs followed by rapid deterioration of higher brain functions. Within a year and a half, this deterioration generally leads to severe dementia, autism, loss of purposeful movements in use of the hand described as a sterotypic "hand-washing" movement, jerky truncal ataxia, and acquired microcephaly. Following this profound regressive stage, there may be apparent stability lasting several decades. Later in life, insidious neurological abnormalities supervene, mainly a spastic parapareses, vasomotor disturbances of the lower limbs, and epilepsy. The exclusive involvement of females suggests a dominant mutation on one of the X chromosome. Attempts to map the gene were hindered by the frustratingly small number of familial Rett cases available. However, exclusion mapping based on comparison of X-chromosome haplotypes among affected sisters or half-sisters has focused attention on a region of the X chromosome Xq28. The responsible gene encoded methyl-CpG-binding protein-2 (MeCP-2) was implicated as the cause of this disease (63). Both nonsense and missense mutations were found with multiple recurrences of each of these mutations in unrelated sporadic cases as well as affected

families (see Figure 2). It is of interest that all of the nucleotide substitutions involved a C to T transition at a CpG hot spot, emphasizing that methylcytosine is frequently deaminated into thymine (63,64). In keeping with this prediction, it is estimated that the incidence of Rett's is 1 per 10,000 female births with 99.5 percent of all cases being sporadic. Such a high incidence requires a high rate of new mutations which would be expected in CpG islands. The severity of Rett's phenotype emphasizes that MeCP-2 is important as a molecular link binding 5-methylcytosine and the co-repressor Sin 3a through its transcriptional repression domain (TRD), thus recruiting histone deacetylases and other proteins necessary to silence gene complexes. It is of interest that MeCP-2 is not expressed in embryonic stem cells nor in proliferating embryonic tissue. However, methylation becomes important during the stages of organ differentiation when specific tissue developmental patterns of gene expression are established (65). MeCP-2 is essential for the maintenance of methylation patterns throughout life. Therefore, Rett's Syndrome is a human disorder caused by genetic defects in a component of the epigenetic silencing machinery. The exclusive involvement of neural differentiation in Rett patients suggests there may be related proteins that exist in other tissues which mitigate the effect of mutations in MeCP-2 on non-neural tissues (66). Hendrich and his collaborators have carried out a genomic data base search and have identified and characterized a family of mammalian methyl CpG-binding proteins which may complement the function of MeCP-2 and explain the lack of pleiotropic effects on organs other than the brain in Rett Syndrome (20).

Fragile X Syndrome

The Fragile X Syndrome may be the most frequent cause of inherited mental retardation, the incidence being about 1 in 1,500 males and 1 in 2,500 females. The syndrome is associated with a fragile site on the X chromosome at Xq27.3, which is induced to appear as a gap or break, when cell culture conditions adversely affect deoxynucleotide synthesis. Genetic studies have shown that the mutation is *cis*-acting and located at or very closely to the fragile site. Approximately 20 percent of males who carry the mutation have no clinical or cytogenetic expression of the fragile X site, the remaining 80% are severely retarded. Carrier females have approximately a 30 percent incidence of some mental retardation, that usually is associated with a higher

frequency of fragile X expression in cultured cells than in mentally normal carriers. Penetrance also appears variable in different sibships and even within the same family. Daughters of normal transmitting males rarely express any symptom while penetrance is high in sons and daughters of carrier women. Initially, a very high mutation frequency at the fragile X locus was assumed to account for the high incidence of the disease, particularly in view of the low reproductive fitness of affected males (67).

The involved gene is called FMR-1 (Fragile X Mental Retardation-1) and was discovered to contain CCG-repeats within the first exon. Normal individuals had between 6 and 52 CCG copies with a mean of 29. Among affected individuals, the repeat length was dramatically increased well beyond 230 CCG repeats and usually 600 or more (68,69). Concomitant with the large expansion of the trinucleotide was abnormal methylation within the region rich in CG dinucleotides (a CpG island) immediately upstream of the gene (70,71). This marked expansion of the CCG repeats was associated with the absence of FMR-1 transcription. Nonpenetrant carrier males and many such females had repeat lengths of intermediate size without abnormal methylation. This situation was called a pre-mutation. The intermediate size in the CCG repeats ranged from approximately 50-200 and upon transmission to the offspring, displayed remarkable instability, with offspring usually exhibiting allele sizes different from the transmitting parent and frequently distinct from other siblings. The increase in the CCG repeat length and the larger the size was proportional to the risk of a full expression with respect to mental retardation. The abnormal methylation of CpG islands is consistent with an imprinting mechanism being responsible for the retardation in Fragile X Syndrome patients. In addition to the mental retardation seen primarily in males, there are dysmorphic findings which include a long face, large everted ears, autism, hand-biting, hyperactivity, and enlarged testicles.

The mechanism responsible for *de novo* methylation and consequent histone deacetylation of the CpG islands upstream of the gene and its subsequent silencing are not clear, but may result from the potential of the repeats to form hairpin structures (72). Such hairpin loops in either "odd" DNA structures are good substrates for the DNMT-1 enzyme. An alternative explanation is that the *de novo* methylation may in some way be related to the

genome defense against repetitive elements. The expanded repeats may resemble or adopt one or more of the features of parasitic elements, for example, retrotransposons and this targets them for methylation. An interesting feature is that growing fragile X cells in culture with 5-azadeoxycytidine results in reduced methylation of the CpG promoter islands and re-expression of the FMR-1 protein. The FMR-1 protein is abundant in neurons, and contains several RNA-binding motifs and associates with translating ribosomes in an RNA dependent manner. It has been proposed that the FMR-1 protein may be important in sustaining protein synthesis in neurons and that its deficiency leads to mental retardation.

Tumorogenesis

For several decades it has been recognized that most cancers have genetic alterations and these genetic changes are manifested in the cellular DNA as single nucleotide substitutions, deletions, duplications, and microsatellite repeats, or at the chromosomal level, translocations and chromosomal aberrations. These genetic changes are viewed as critical in the progression of cancer, particularly if one accepts the multihit theory required for tumor formation and the selective advantage of cancer cells during tumor progression. Many tumors exhibit a cascade of mutations involving different genes. Some mutations lead to hyperplasia; subsequent mutations in other genes to neoplasia and others to metastasis (73).

However, recently it has become apparent that the development of some human cancers also may depend on the functional inactivation of tumor suppressor genes, for example, the retinoblastoma (RB) gene and the p53 gene. The inactivation of tumor suppressor genes can be achieved by reversible modifications of chromatin producing transcriptional silencing. These effects may be an additional important mechanism by which cancers arise in nature. The best studied DNA modification that correlates with epigenetic gene silencing is methylation of cytosine residues in the CpG sequences. CpG methylation has recently been linked to an even more general mechanism of epigenetic silencing and that is histone deacetylation as described above (74,75).

It is difficult to envision how information concerning DNA methylation might be stably maintained through multiple rounds of cell division. The heritability of newly acquired traits is a hallmark of clonal expansion in tumors. Enzymatic methylation of the C-5 position of cytosine residues can affect epigenetic inheritance by altering the expression of genes. Transmission of the methylation pattern through cell divisions is conferred by the DNMT-1 methyl transferase as described earlier (16,17). Thus, in addition to the well-known role of mutational hot spots in human DNA where methylated cytosine is deaminated to thymine producing a mutation, the CpG islands may be methylated silencing the gene. Therefore, epigenetic and genetic changes can be viewed as two complimenting mechanisms both leading to cancer.

Tumor Suppressor Gene Silenced by DNA Methylation in Human Cancers

The von Hippel-Lindau Gene in Renal Cell Carcinoma

Germ-line mutations in the von Hippel-Lindau (VHL) gene on distal chromosome-3p predisposed to renal cell carcinoma and hemangioblastomas. Somatic mutations of the VHL gene are also characteristic of sporadic renal cell cancers. In sporadic renal cell carcinomas, the promoter regulatory region was hypermethylated in about 20 percent of the tumors, suggesting that this is an important contributor to tumor suppressor gene inactivation (76). When these genes were sequenced, there were no coding mutations and in a renal cell carcinoma line, the VHL locus was reactivated by growing the cancer cells in medium with azadeoxycytosine. With respect to inherited renal cell carcinomas, promoter hypermethylation was responsible for the loss of heterozygosity and was the "second hit" for tumorogenesis. More than 30 percent of inherited renal cell carcinomas and hemangioblastomas had lost expression of a non-mutated VHL gene through promoter hypermethylation in one allele and had a heritable mutation in the other (77,78).

hMLH-1 Gene in Colonic, Gastric and Endometrial Cancer

The hMLH-1 gene on chromosome 3p encodes a protein essential for DNA mismatch repair of simple repetitive elements that comprise microsatellite sequences. Germ-line mutations in this gene are reported in a large percentage of patients with hereditary non-polyposis colon cancer and it is also inactivated in about 25 percent

of sporadic colon cancers. There is a high occurrence of errors in DNA replication characterized by expansion of nucleotide repeats reported in these patients which is described as a microsatellite instability (MIN). Approximately half of the sporadic cases of colon cancer do not have mutations in the hMLH-1 gene, but rather exhibit a high degree of hypermethylation of the promoter region of that gene (79). Growth of colon cancer cell in culture with this phenotype in the presence of azadeoxycytosine show that this treatment reverses the hypermethylation sufficiently to restore DNA mismatch repair competence to the cancer cells. Similar findings have been reported for endometrial and gastric cancers. Thus, it appears that in colon cancer, epigenetic inactivation of the hMLH-1 gene is a frequent occurrence (74).

The p16/INK4A Locus in Cancers

The p16 gene resides on chromosome 9p in a region that is subject to frequent mutations or loss of heterozygosity in various tumors. This gene, which is a tumor suppressor, is a cyclin-dependent kinase inhibitor and is mutated in many solid tumors as well as lymphomas. The p16 gene restrains the cell cycle by preventing phosphorylation and the inactivation of the RB gene, an important tumor suppressor. The INK4A locus encodes a second tumor suppressor gene, p19. The two genes also act as inhibitors of the p53 pathway that regulates cell mortality, cell cycle arrest, and apoptosis. Thus, the p16/INK 4A gene silencing is similar to simultaneous inactivation of both RB and p53 which are extremely critical targets for tumorogenesis (78). The epigenetic silencing of p16 and p19 indicate that hypermethylation of their promoters is a frequent pathway for inactivation of these genes in human carcinomas including those that arise in the lung, oropharynx, bladder, cervix, liver, colon, pancreas, and other sites (74). Moreover, the inactivation of these two genes suggest that an imprinting box or site is near the promoters and attracts the DNA methylation complex (74).

Genetic Imprinting and Wilms Tumor (WTs)

Chromosome 11p15 is a frequent site of loss of heterozygosity in pediatric solid tumors including Wilms tumor. In normal human tissue, the insulin-like growth factor gene (IGF2) is imprinted so that the paternal allele is expressed and the maternal gene is inactivated. The only exception is human liver where expression is biallelic

because of promoter switching after birth (78). Direct evidence linking aberrant imprinting with tumorogenesis was identified when approximately 70 percent of Wilms tumors were found to have a bialleic expression of the IGF2, a gene that encodes for growth factors known to be oncogenic when overexpressed. A second tumor suppressant gene, H19, also was found to be inactivated by imprinting in many Wilms tumors (80). Since the insulin-like growth factor 2 is always paternally produced, loss of imprinting at the maternal locus for this growth factor also may be related to the loss of expression of the H19 gene (55,78). This conclusion is supported by findings in transgenic mice which are H19 null and show bialleic expression of the IGF2 gene. Another imprinted gene involved in Wilms tumor formation is WT1, a tumor suppressor located on human chromosome 11p13. It exhibits a complicated imprinting pattern in which the paternal gene is imprinted in some tissues, (i.e., brain) while in others the maternal gene is imprinted (i.e., connective tissue and lymphocyte). As described, genomic imprinting is an epigenetic form of gene regulation that results in the expression of only one parental allele (78). Imprinted genes not only play an important role in embryogenesis and behavioral development, but as shown by the examples above, they are mechanistically involved in carcinogenesis. Imprinted genes are functionally haploid, imprinted tumor suppressor genes and *proto*-oncogenes are particularly vulnerable to inactivation and activation, respectively. The imprinting of these genes also varies among species, individual tissues, cells, and stages of embryonic development. Therefore, the overall effect of genomic imprinting on cancer susceptibility and its penetrance is potentially great.

Detection of Aberrant DNA Methylation in the Plasma and Serum of Cancer Patients

The p16 tumor suppressor gene is often inactivated by promoter region hypermethylation in many human cancers. Recent evidence suggests that tumor cells may release DNA into the circulation which is enriched in the serum and plasma allowing detection of both mutations and microsatellite alterations as well as aberrant DNA methylation patterns. In a study of serum from patients with non-small cell lung cancers, promoter region hypermethylation was present in 45 percent of the sera analyzed (81). The sera of cancer patients is enriched in DNA, containing on an average four times the amount of free DNA compared to normal controls. Moreover, studies indicate that the aberrant methylation and inactivation of the p16 gene may be an early event in the

genesis of non-small cell lung cancers (82). If these findings are substantiated, they demonstrate the potential use of an epigenetic change as a biomarker to identify persons at high risk for lung cancer. Sensitive PCR methods are available using primers that are specific for either methylated or modified unmethylated DNA. Recently, the hypermethylation of promoter regions of other oncogenetic genes were described. These include the DNA repair gene, MGMT, the detoxifying gene, GSTP1, and the DAP kinase gene. Using all of these markers, 68 percent of non-small cell lung cancer primary tumors exhibited abnormal promoter hypermethylation in at least one of the genes. p16 was hypermethylated in 41 percent of the patients examined. Thirty-three percent showed hypermethylation of the DAP kinase gene, 9 percent of the GSTP1, and 27 percent of the MGMT gene (81).

As described, methylation is the main epigenetic modification in humans and changes in methylation patterns play an important role in tumorogenesis. Many of the tumors examined in this study were combinations of genetic and epigenetic events. That is, one of the tumor suppressor genes was inactivated by hypermethylation while the other allel contained a mutation or deletion. Using PCR methods in which there was a chemical modification of unmethylated CpG islands compared to methylated regions, PCR detection is extremely sensitive ($\cong$ 68 percent). Moreover, the specificity (73%) is also quite high since no abnormal methylations were observed in serum DNA from normal individuals or from cancer patients in which methylation of the DNA was not present in the primary tumor (81). Thus, analysis of serum and plasma holds the potential of detecting molecular markers of tumorogenesis, they may define the risk of disease in patients during the so-called pre-mutation and pre-tumor interval. It also opens the possibility of monitoring prevention strategies for cancers and the potential to achieve an early diagnosis and track the progression of the disease when abnormal DNA methylation is involved in promoting neoplasia.

Pharmacogenomics

In general, chemotherapeutic agents are selected on the basis of tumor type, clinical stage, the patient's age, and other non-molecular considerations. Patients are treated based on a belief that a particular therapy may be beneficial, but there is currently no way of predicting which agents will be more specific for a particular cancer.

The emerging field of pharmacogenomics, through the study of genes that influence drug activity, toxicity and metabolism may provide an opportunity to tailor our chemotherapy treatment and to eliminate some of the uncertainties of tumor response. Genetic polymorphisms that influence drug metabolism, for example, enzymes in the cytochrome system affect the metabolism of many drugs. In the *New England Journal of Medicine* lead article on November 9, 2000, a different sort of mechanism was described to predict the clinical response of gliomas to alkylating agents (83). The authors discovered that an epigenetic mechanism explained the resistance of certain gliomas to nitrosourea alkylating agents. Carmustine (BCNU) and other nitrosoureas kill by alkylating the O^6 position of guanine and thereby crosslink adjacent strands of DNA. A DNA repair enzyme O^6-methyguanine-DNA methyltransferase (MGMT) rapidly reverses the alkylation and prevents the formation of crosslinks. Approximately 30 percent of gliomas lack MGMT, but sequencing the gene shows that mutations in the gene for this enzyme are very unusual. It was found that methylation of the promoter region of the gene with consequent transcriptional silencing seemed to account for the variation in tumor response to these chemotherapeutic agents. In the reported study, 12 of 19 glioma patients with methylated promoters in the MGMT gene improved when treated with carmustine. However, only 1 of 28 patients with an unmethylated promoter responded. Overall survival and the progression of the glioma was much delayed in patients whose tumors had methylated promoters. Although the number of patients studied was relatively small, the implications are that patients could be selected for a particular chemotherapeutic agents based on the silencing of genes which interfere with the therapeutic efficiency of the treatment. Pharmacogenomic studies will contribute to the specificity of treatment and are the subject of both clinical research and may better direct pharmacological research in targeting particular tumors.

Summary

DNA methylation represents perhaps the best studied of the epigenetic mechanisms available for gene silencing and altering of gene expression. It represents a mechanism that allows genes to be selectively transcribed or silenced. The role of epigenetics in modulating gene function required for differentiation of specialized cells and the aberrations that can occur which result in tumorogenesis represent the edges of a two-edged sword. One edge is essential for normal development and function. The other edge results in aberrations in embryogenesis and in the cascade that results in the conversion of normal cells to cancer. If we can learn to control the beneficial effects of methylation and reverse or prevent the deleterious effects, our understanding of human physiology will be greatly advanced. The sequencing of the human genome is complete. We are now faced with the challenge to understand the regulation and function of our heredity endowment. Studies on epigenetics and genetic alterations should go hand-in-hand in our approaches to better controlling and understanding human health.

References

1. Razin A, Riggs AD. DNA methylation and gene function. *Science* 210:604, 1980.
2. Stein R, Razin A, Cedar H. *In vitro* methylation of the hamster adenine phosphoribosyltransferase gene inhibits its expression in mouse L cells. *Proc Natl Acad Sci USA* 79:3418, 1982.
3. Jones PA, Taylor SM. Cellular differentiation, cytidine analogs and DNA methylation. *Cell* 20:85, 1980.
4. Antequera F, Bird A. CpG islands. *IN: DNA Methylation: Molecular Biology and Biological Significance.* Jost JP, Saluz HP (Eds.). Basel, Switzerland: Birkhauser Verlag. 169-185, 1993.
5. Cooper DN, Krawczak M. Cytosine methylation and the fate of CpG dinucleotides in vertebrate genomes. *Hum Genet* 83:181, 1989.
6. Bird A, Taggart M, Frommer M, Miller OJ, Macleod D. A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA. *Cell* 10:91, 1985.

7. Antequera F, Bird A. Number of CpG islands and genes in human and mouse. *Proc Natl Acad Sci USA* 90:11995, 1993.

8. Cooper DN, Youssoufian H. The CpG dinucleotide and human genetic disease. *Hum Genet* 78:151, 1988.

9. Filipski J, Leblanc J, Youdale T, Sikorska M, Walker PR. Periodicity of DNA folding in higher order chromatin structures. *EMBO J* 9:1319, 1990.

10. Kornberg RD. Chromatin structure: a repeating unit of histones and DNA. *Science* 184:868, 1974.

11. Richmond TJ, Finch JT, Rushton B, Rhodes D, Klug A. Structure of the nucleosome core particle at 7Å resolution. *Nature* 311:532, 1984.

12. van Holde KE. *Chromatin*. New York: Springer Verlag. 1989.

13. Keshet I, Lieman-Hurwitz J, Cedar H. DNA methylation affects the formation of active chromatin. *Cell* 44:535, 1986.

14. Tazi J, Bird A. Alternative chromatin structure at CpG islands. *Cell* 60:909, 1990.

15. Robertson KD, Wolffe AP. DNA methylation in health and disease. *Nat Rev Genet* 1:11, 2000.

16. Li E, Bestor TH, Jaenisch R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* 69:915, 1992.

17. Okano M, Bell DW, Haber DA, Li W. DNA methyltransferases Dnmt3a and Dnmt3b are essential for *de novo* methylation and mammalian development. *Cell* 99:247, 1999.

18. Boyes J, Bird A. DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. *Cell* 64:1123, 1991.

19. Cross SH, Meehan RR, Nan X, Bird A. A component of the transcriptional repressor MeCP1 shares a motif with DNA methyltransferase and HRX proteins. *Nat Genet* 16:256, 1997.

20. Hendrich B, Bird A. Identification and characterization of a family of mammalian methyl-CpG binding proteins. *Mol Cell Biol* 18:6525, 1998.

21. Laherty CD, Yang WM, Sun JM, Davie JR, Seto E, Eisenman RN. Histone deacetylases associated with the mSin3 corepressor mediate mad transcriptional repression. *Cell* 89:349, 1997.

22. Hassig CA, Fleischer TC, Billin AN, Schreiber SL, Ayer DE. Histone deacetylase activity is required for full transcriptional repression by mSin3A. *Cell* 89:341, 1997.

23. Deutsch ST, Silvers DN, Cox RP, Griffin MJ, Ghosh NK. Ultrastructural and enzymic modulation of HeLa cells induced by sodium butyrate and the effects of cytochalasin B and colcemid. *J Cell Sci* 21:391, 1976.

24. Ghosh NK, Cox RP. Production of human chorionic gonadotropin in HeLa cell cultures. *Nature* 259:416, 1976.

25. Ghosh NK, Cox RP. Induction of human follicle-stimulating hormone in HeLa cells by sodium butyrate. *Nature* 267:435, 1977.

26. Riggs MG, Whittaker RG, Neumann JR, Ingram VM. n-Butyrate causes histone modification in HeLa and Friend erythroleukaemia cells. *Nature* 268:462, 1977.

27. Fallon RJ, Cox RP. Cell cycle analysis of sodium butyrate and hydroxyurea, inducers of ectopic hormone production in HeLa cells. *J Cell Physiol* 100:251, 1979.

28. Lieblich JM, Weintraub BD, Rosen SW, Ghosh NK, Cox RP. Secretion of hCG-α subunit HCG by Hela strains. *Nature* 265:746, 1977.

29. Pradhan S, Bacolla A, Wells RD, Roberts RJ. Recombinant human DNA (cytosine-5) methyltransferase 1. Expression, purification, and comparison of *de novo* and maintenance methylation. *J Biol Chem* 274:33002, 1999.

30. Pradhan S, et al. Baculovirus-mediated expression and characterization of the full-length murine DNA methyltransferase. *Nucleic Acids Res* 25:4666, 1997.

31. Xu G-L, et al. Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. *Nature* 402:187, 1999.

32. Hansen RS, et al. The DNMT3B DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome. *Proc Natl Acad Sci USA* 96:14412, 1999.

33. Xie S, et al. Cloning, expression and chromosome locations of the human DNMT3 gene family. *Gene* 236:87, 1999.

34. Prak ETL, Kazazian HH. Mobile elements and the human genome. *Nat Rev Genet* 1:134, 2000.

35. Yoder JA, Walsh CP, Bestor TH. Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet* 13:335, 1997.

36. Colot V, Rossignol J-L. Eukaryotic DNA methylation as an evolutionary device. *BioEssays* 21:402, 1999.

37. Kazazian JHH, Moran JV. The impact of L1 retrotransposons on the human genome. *Nature Genet* 19:19, 1998.

38. Kazazian HH, et al. Haemophilia A resulting from the *de novo* insertion of L1 sequences represents a novel mechanism for mutation in man. *Nature* 332:164, 1988.

39. Dombrowski BA, Mathias SL, Nanthakumar E, Scott AF, Kazazian HH. Isolation of an active human retrotransposable element. *Science* 254:1805, 1991.

40. Holmes SE, Dombroski BA, Krebs CM, Boehm CD, Kazazian HH. A new retrotransposable human L1 element from the LRE2 locus on chromosome 1q produces a chimaeric insertion. *Nature Genet* 67:143, 1994.

41. Narita N, et al. Insertion of a 5' truncated L1 element into the 3' end of exon 44 of the *dystrophin* gene resulted in skipping of the exon during splicing in a case of Duchenne muscular dystrophy. *J Clin Invest* 91:1862, 1993.

42. Schwahn U, et al. Positional cloning of the gene for X-linked retinitis pigmentosa 2. *Nature Genet* 19:327, 1998.

43. Divoky V, et al. A novel mechanism of β-thalassemia. The insertion of L1 retrotransposable element into β globin IVSII. *Blood* 88:148a, 1996.

44. Meischl C, de Boer M, Ahlin A, Roos D. Intronic insertion of a LINE-1 fragment as the cause of chronic granulomatous disease. *Eur J Hum Genet* 8:697, 2000.

45. Kochanek S, Renz D, Doertier W. Transcriptional silencing of human Alu sequences and inhibition of protein binding in the B box regulatory elements by 5'-CG-3' methylation. *FEBS Lett* 360:115, 1995.

46. Vidaud D, et al. Haemophilia B due to a *de novo* insertion of a human-specific Alu subfamily member within the coding region of the factor IX gene. *Eur J Hum Genet* 1:30, 1993.

47. Wallace MR, et al. A *de novo* Alu insertion results in neurofibromatosis type 1. *Nature* 353:864, 1991.

48. Miki Y, Katagiri T, Kasumi F, Yoshimoto T, Nakamura Y. Mutation analysis in the BRCA2 gene in primary breast cancers. *Nature Genet* 13:245, 1996.

49. Lyon MF. Line-1 elements and X chromosome inactivation: A function for "junk" DNA? *Proc Nat Acad Sci* 97:6248, 2000.

50. Bailey JA, Carrel L, Chakravarti A, Eichler EE. Molecular evidence for a relationship between Line elements and X chromosome inactivation: The Lyon repeat hypothesis. *Proc Nat Acad Sci* 97:6634, 2000.

51. Bartolomei MS, Tilghman SM. Genomic imprinting in mammals. *Ann Rev Genet* 31:493, 1997.

52. Reik W, Walter J. Imprinting mechanisms in mammals. *Curr Opin Genet Dev* 8:154, 1998.

53. Surani MA. Imprinting and the initiation of gene silencing in the germ-line. *Cell* 93:309, 1998.

54. Constancia M, Pickard B, Kelsey G, Reik W. Imprinting mechanisms. *Genet Res* 8:881, 1998.

55. Maher ER, Reik W. Beckwith-Wiedemann syndrome: imprinting in clusters revisited. *J Clin Invest* 105:247, 2000.

56. Reis A, Dittrich B, Greger V, Buiting K, Lalande M, Gillessen-Kaesbach G, Anvret M, Horsthemke B. Imprinting mutations suggested by abnormal DNA methylation patterns in familial Angelman and Prader-Willi syndromes. *Am J Hum Genet* 54:741, 1994.

57. Smeets DFCM, et al. ICF syndrome: a new case and review of the literature. *Hum Genet* 94:240, 1994.

58. Franceschini P, et al. Variability of clinical and immunological phenotype in immunodeficiency-centromeric instability-facial anamolies syndrome. *Eur J Pediatr* 154:840, 1995.

59. Jeanpierre M, et al. An embryonic-like methylation pattern of classical satellite DNA is observed in ICF syndrome. *Hum Mol Genet* 2:731, 1993.

60. Kondo T, et al. Whole-genome methylation scan in ICF syndrome: hypomethylation of non-satellite DNA repeats D4Z4 and NBL2. *Hum Mol Genet* 9:597, 2000.

61. Wijmenga C, et al. Localization of the ICF syndrome to chromosome 20 by heterozygosity mapping. *Am J Hum Genet* 63:803, 1998.

62. Hagberg B, Alcardi J, Dias K, Ramos O. A progressive syndrome of autism, dementia, ataxia and loss of purposeful hand use in girls: Rett's syndrome: report of 35 cases. *Ann Neurol* 14:471, 1983.

63. Wan M, et al. Rett syndrome and beyond: recurrent spontaneous familial MECP2 mutations at CpG hotspots. *Am J Hum Genet* 65:1520, 2000.

64. Ballestar E, Yusufzai TM, Wolffe AP. The effects of Rett syndrome mutations of the methyl-CpG binding domain of the transcriptional repressor MeCP2 on selectivity for association with methylated DNA. *Biochemistry* 39:7100, 2000.

65. Tate P, Skarnes W, Bird A. The methyl-CpG binding protein MeCP2 is essential for embryonic development in the mouse. *Nature Genet* 12:205, 1996.

66. Willard HF, Hendrich BD. Breaking the silence in Rett syndrome. *Nature Genet* 23:127, 1999.

67. Warren ST, Nelson DL. Trinucleotide repeat expansions in neurological disease. *Curr Opin Neurobiol* 3:752, 1993.

68. Kremer EJ, et al. Mapping of DNA instability at the fragile X to a trinucleotide repeat sequence p(CCG)n. *Science* 252:1711, 1991.

69. Oberle I, et al. Instability of a 550-base pair segment and abnormal methylation in Fragile X syndrome. *Science* 252:1097, 1991.

70. Coffee B, Zhang F, Warren ST, Reines D. Acetylated histones are associated with FMR1 in normal but not fragile X-syndrome cells. *Nature Genet* 22:98, 1999.

71. Fu YH, et al. Variation of the CCG repeat at the Fragile X site results in genetic instability: Resolution of the Sherman paradox. *Cell* 67:1047, 1991.

72. Smith SS, Laayoun A, Lingeman RG, Baker DJ, Riley J. Hypermethylation of telomere-like foldbacks at codon 12 of the human c-Ha ras gene and the trinucleotide repeat of the FMR-1 gene of fragile X. *J Mol Biol* 243:143, 1994.

73. Janne PA, Mayer RJ. Chemoprevention of colorectal cancer. *N Eng J Med* 342:1960, 2000.

74. Baylin SB, Herman JG, Herman JR, Vertino PM, Issa J-P. Alterations in DNA methylation: a fundamental aspect of neoplasia. *Adv Cancer Res* 72:141, 1998.

75. Jones PA, Laird PW. Cancer epigenetics comes of age. *Nature Genet* 21:163, 1999.

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

77. Prowse AH, et al. Somatic inactivation of the VHL gene in Von Hippel-Lindau disease tumors. *Am J Hum Genet* 60:765, 1997.

78. Tycko B. Epigenetic gene silencing in cancer. *J Clin Invest* 105:401, 2000.

79. Herman JG, et al. Incidence and functional consequences of hMLH1 promoter hypermethylation in colorectal carcinoma. *Proc Natl Acad Sci USA* 95:6870, 1998.

80. Dao D, et al. Multipoint analysis of human chromosome 11p15/mouse distal chromosome 7: inclusion of H19/ICF2 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, 1999.

81. Esteller M, et al. Detection of aberrant promoter hypermethylation of tumor suppressor genes in serum DNA from non-small cell lung cancer patients. *Cancer Res* 59:67, 1999.

82. Belinsky SA, et al. Aberrant methylation of $p16^{INK4a}$ is an early event in lung cancer and a potential biomarker for early diagnosis. *Proc Nat Acad Sci* 95:11891, 1988.

83. Esteller M, et al. Inactivation of the DNA-repair gene MGMT and the clinical response of gliomas to alkylating agents. *N Eng J Med* 343:1350, 2000.