Biochemistry (Moscow), Vol. 66, No. 3, 2001, pp. 235-255. Translated from Biokhimiya, Vol. 66, No. 3, 2001, pp. 293-317.

REVIEW

DNA Methylation and Carcinogenesis

A. V. Lichtenstein* and N. P. Kisseljova

Institute of Carcinogenesis, Blokhin Russian Cancer Research Center, Kashirskoe Shosse 24, Moscow, 115478 Russia; fax: (095) 323-1205; E-mail: alicht@online.ru

Received July 20, 2000
Revision received September 21, 2000

Abstract—The hypothesis of the exclusively genetic origin of cancer ("cancer is a disease of genes, a tumor without any damage to the genome does not exist") dominated in the oncology until recently. A considerable amount of data confirming this hypothesis was accumulated during the last quarter of the last century. It was demonstrated that the accumulation of damage of specific genes lies at the origin of a tumor and its following progression. The damage gives rise to structural changes in the respective proteins and, consequently, to inappropriate mitogenic stimulation of cells (activation of oncogenes) or to the inactivation of tumor suppressor genes that inhibit cell division, or to the combination of both (in most cases). According to an alternative (epigenetic) hypothesis that was extremely unpopular until recently, a tumor is caused not by a gene damage, but by an inappropriate function of genes ("cancer is a disease of gene regulation and differentiation"). However, recent studies led to the convergence of these hypotheses that initially seemed to be contradictory. It was established that both factors—genetic and epigenetic—lie at the origin of carcinogenesis. The relative contribution of each varies significantly in different human tumors. Suppressor genes and genes of repair are inactivated in tumors due to their damage or methylation of their promoters (in the latter case an "epimutation", an epigenetic equivalent of a mutation, occurs, producing the same functional consequences). It is becoming evident that not only the mutagens, but various factors influencing cell metabolism, notably methylation, should be considered as carcinogens.

Key words: DNA methylation, DNA-methyltransferases, oncogenes, suppressor genes, mutations, gene activity, silencing, imprinting, histone acetylation

In the world of easy things truth is opposed to lie;
in the world of complicated things one profound truth
is opposed to another not less profound than the first.

Neils Bohr

Methylated residues in DNA were discovered more then 50 years ago [1]. Prokaryotic DNA acquires modified residues N6-methyladenine and 5-methylcytosine, while DNA of higher eukaryotes aquires mainly 5-methylcytosine [2-6]. Methylation of cytosine residues in DNA occurs in bacteria, plants, animals, notably mammals (including human), but was not found in yeast, nematodes, and *Drosophila* [7-10] [1]. This enzymatic reaction occurs within the first minutes after DNA replication, i.e., post-transcriptionally. Since DNA methylation does not lead to an alteration of the nucleotide sequence of DNA, it is considered as an epigenetic event [11]. Although this modification is stable and inherited, it can be reversed by demethylating agents or enzymes, thus differing in principle from mutations of DNA.

In general, the phenomenon of DNA methylation plays a role in the system of "self–foreign"recognition. Bacterial cells can distinguish between their own genetic material and heterogeneous foreign molecules invading the cell due to the methylation-recognition system. Degradation of foreign molecules ensures the genetic stability of the species. Sometimes a single enzyme possesses both methylase and endonuclease activities. Usually a specific DNA site is recognized by two enzymes; one of them methylates it, the other cleaves. In general, the system functions in a way that methylases "label" specific sequences of its "own" DNA and methylation-sensitive restriction endonucleases recognize and cleave only the

* To whom correspondence should be addressed.

[1] During the preparation of this paper, it was demonstrated that *Drosophila melanogaster* DNA contains 5-methylcytosine (Growher, H., Leismann, O., and Jeitsch, A. (2000) *EMBO J.*, **19**, 6918-6923).

0006-2979/01/6603-0235$25.00 ©2001 MAIK "Nauka/Interperiodica"

non-labeled sequences. Thus, a bacterial cell protects itself from invasion by foreign molecules. For example, DNA of a phage that has penetrated into a bacterial cell is cleaved in the target sites by specific endonucleases, whereas the same sequences in the "self" DNA are protected from the restriction by methylation [12].

The role of methylation as a component of the cell "immune system" directed towards the elimination of foreign or "excessive" DNA (or for the suppression of its functions) is apparently conserved in evolution, but the exact mechanisms responsible for the realization of this program can differ significantly [13]. For example, in *Neurospora* cells additional repetitive sequences are eliminated by intensive methylation, followed by the accumulation of point mutations due to the high mutability of 5-methylcytosine residues [14]. Integrated viral sequences in rodent and human cells can be similarly subjected to a stable silencing of transcription due to methylation [15]. Also, it was demonstrated that transgenes in mice are also frequently inactivated by the same mechanism [16].

Transitions from prokaryotes to eukaryotes and from invertebrates to vertebrates during evolution were apparently accompanied by a great increase in the number of genes [10]. These dramatic changes apparently gave rise to new mechanisms of suppression of the unnecessary activity of "excessive" genes—the formation of a nuclear membrane and nucleosomal organization of chromatin in the first case, and conversion of the functional orientation of methylation system in the second. In invertebrates, this system was directed towards the suppression of activity of potentially dangerous DNA sequences (viruses and transposons), while in vertebrates it acquired an additional function of stable repression of endogenous genes (genes of the inactivated X chromosome, imprinted genes, some tissue-specific genes).

The pattern of methylation, which influences significantly the functional state of a gene, is inherited through cell generations. From this point of view, a reliable system of epigenetic heredity (e.g., DNA methylation) is of vital importance for organisms with long lifespan and intensive regeneration of tissue (vertebrates, plants). In contrast, this system is not so essential for animals of small size with short lifespan, i.e., in organisms without intensive cell renewal. This observation can explain the absence of DNA methylation system in nematodes and *Drosophila* [17].

Whatever forces might direct evolution, in mammals DNA methylation serves for the stable functional repression of part of the genetic material (silencing). A relatively small fraction of the genome (~5%) functions during the whole life of a mammal, while the majority of the genome represents genetic ballast (junk DNA), which is subjected to intensive methylation. Also, activation of methylation was observed to accompany cellular processes that lead to the appearance of new DNA; DNA transfection, different types of duplication initiate this modification of new copies. In particular, inverted repeats possess an extremely strong capacity for stimulation of methylation [18]. Hence, eukaryotic genomes are characterized by the presence of two closely related properties: repeated sequences on one hand, and the presence of 5-methylcytosine on the other. Some repeats that have been accumulated in the human genome during evolution are supposed to be functionally useless, whereas some others (transposons) are potentially dangerous. It is evident that they are strongly methylated and, consequently, not active, since their transcription can lead to metabolic chaos in the cell. Methylated regions are also otherwise"restricted", since in the S-phase they are replicated later than the active sequences [19-21]. Methylation-induced inhibition of transcription is mediated by the alteration of chromatin structure (see below).

DNA METHYLATION: PHYSIOLOGY

DNA methylation lies at the base of such prominent biological events as imprinting and inactivation of the X chromosome. During embryogenesis, "waves" of methylation–demethylation are occurring in an orchestrated manner, resulting in a strongly determined methylation pattern of distinct genes and of the genome in general. This pattern is stably conserved in the population of somatic cells. At the stage of the first two cell divisions, total demethylation of the cellular genome occurs, erasing the methylation pattern of the initial germinal cells. The demethylation continues until the implantation of the blastocyst [22]. Two points of view exist concerning the origin of this process. According to the hypothesis of "active demethylation", 5-methylcytosine residues are erased from DNA independently of replication by the recently discovered glycosylase [23-25] or by a DNA-demethylase [26]. The latter is considered to be a highly probable candidate for the general demethylation of the human genome. The enzyme recognizes methylated CpG-dinucleotides and converts 5-methylcytosine to cytosine without damage to the integrity of DNA by initiating demethylation in one point and spreading it across a relatively long DNA segment without interruption [26]. The opposite point of view suggests that genome demethylation during embryonic development is a passive process that consists of repression of so-called "maintenance methylation" coupled to DNA replication [27].

After the implantation of the blastocyst the embryo undergoes active methylation involving the entire genome (*de novo* methylation). Its pattern is generally conserved in somatic cells of the adult organism.

Imprinting

Imprinting is a particular event occurring within a small group (more than 20) of mammalian genes that dif-

fer from the others by the functional inequality of their alleles. Usually the copies of a gene are either both inactivated or are equally active. In the case of imprinted genes, one allele is active, and the status is determined by the origin (from a paternal or a maternal cell). Not all imprinted genes are functionally identified. However, it was established that most of them determine the development of an embryo.

Differential activity of alleles of imprinted genes is caused by variations of methylation of regulatory sequences of these genes, enhancers and silencers, which have opposite functions (Fig. 1a). The former transactivate surrounding genes, while the latter repress their activity. Methylation of regulatory sequences results in the functional inhibition of regulatory sequences—the enhancers no longer activate and the silencers no longer repress the genes. The mosaic of methylated and unmethylated (intact) regulatory sequences determines the imprinting state of a gene. In cells with mutated DNA-methyltransferase the differences of methylation pattern of imprinting genes are less pronounced, and consequently, the differences in their activity disappear (Fig. 1b) [28, 29].

Inactivation of the X Chromosome

Only one X chromosome is active in mammalian cells (activation of the second is lethal). Thus, one of the two X chromosomes in female cells is inactivated by a complicated multi-step process that occurs when the embryo consists only of several cells. The locus *Xist* is activated, a huge quantity of corresponding RNA is synthesized, leading to the reorganization of chromatin structure and, finally, numerous genes (probably not all) on the X chromosome are methylated (a more detailed description of this process is beyond the scope of this review). Methylation attaches stability to the inactivation over cell generations [22, 30].

The process occurs at random in every single cell—in one cell of a female embryo the maternal X chromosome is inactivated, while in an other the paternal one, leading to a mosaic female organism. It consists of two cell clones—in one the paternally inherited X chromosome is "working", in the other the maternally inherited one is active. The mosaic is invisible in a normal organism, since both X chromosomes are intact, but if one of them contains a mutant gene, the corresponding pathological

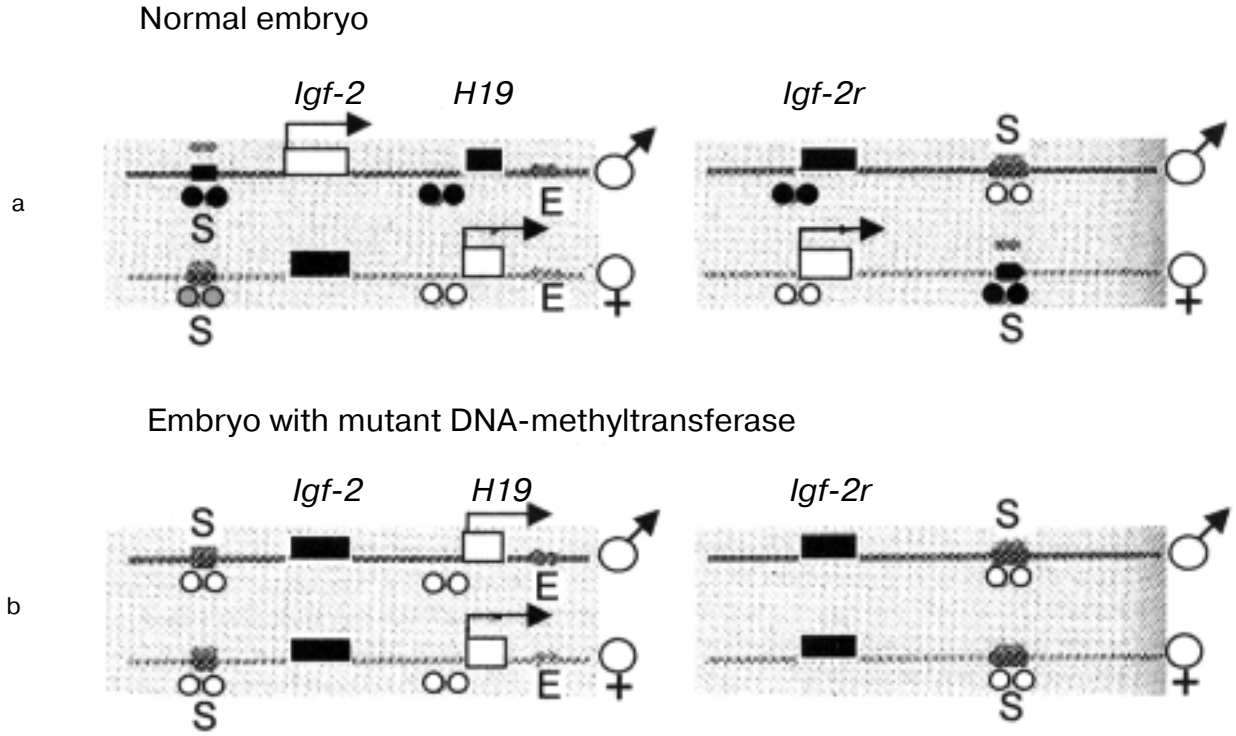

Fig. 1. Differential activity of *Igf-2* (normally the parental allele is active), *Igf-2r*, and *H19* (normally the maternal alleles are active). a) In wild-type cells discrepancies in the expression of imprinted genes are caused by the differences in the methylation of flanking regulatory sequences—enhancers (E) and silencers (S); black circles, completely methylated; gray circles, partly methylated; white circles, unmethylated. b) In embryonic cells with mutated DNA-methyltransferase gene the differences of methylation of regulatory sequences between the alleles are erased, and consequently the differences in the transcriptional rates also disappear; the both alleles are active (*H19*) or not active (*Ifg-2* and *Ifg-2r*) (reprinted by permission from Macmillan Magazines Ltd.).

changes are present in clusters (only in the cells where the chromosome containing a mutated gene is activated). Anhydrotic ectodermal dysplasia is a good example of this phenomenon. In this disease, a gene responsible for the development of sweat glands is mutated. The skin regions without sweat glands form clusters that can be easily demonstrated.

Biochemistry of Methylation

CpG distribution in the human genome. Cytosine is the only DNA residue that is subjected to enzymatic methylation in the human genome. The CpG sequence is the major target for methylation in DNA chains. The dinucleotide CpG in one DNA strand corresponds to the same dinucleotide in the opposite strand (due to the complementary and antiparallel structure of the DNA helix). This facilitates the maintenance function of methylation, which is activated right after DNA replication, and reconstitutes the methylation pattern of the daughter strand according to that of the maternal one (Fig. 2) [31].

The analysis of quantity and distribution of CpG dinucleotides in the genome is essential for understanding of the function of DNA methylation. 5-Methylcytosine is hypermutable (since the amino group in position 6 is extremely unstable, see below) and can undergo accidental deamination giving rise to thymine, which leads to frequent substitutions of G-C pairs by A-T pairs in evolution. This observation explains why the human genome contains ~5-fold less CpG (1 dinucleotide out of 80) than expected (1 out of 16) [32]. This effect is referred to as CpG-suppression [33, 34].

The CpG dinucleotides are distributed in the genome in a very particular way. Two types of distribution exist: a) some of the CpG dinucleotides (~80% of the total number) are dispersed in the genome as individual dinucleotides; b) CpG clusters, usually designated as "CpG-islands" (~20%). The distribution of single CpG seems to be random (they are found mainly in spacer sequences between genes and less often within transcribed sequences), while the CpG-islands are characterized by localization close to structural genes, mainly in their 5′-regions (regulatory sequences, promoters, first exons) [11, 32, 35]. The analysis of completely sequenced human chromosomes 21 and 22 [36, 37] confirms this general statement.

The formal characteristics of CpG-islands, which are present in the promoter regions of ~60% of human genes, are the following: length from 0.5 to 5 kbp (the mean is 1 kbp); average frequency of CpG island is ~1 out of 100 kbp; standard contents of CpG dinucleotide (i.e., corresponding to the calculated 1 : 16 ratio); C+G content exceeds 60%; CpG/GpC ratio not less then 0.6 [38]. Apart from a couple of well known exceptions (see above), the CpG-islands are not methylated in all tissue types independently of the expression of a relevant gene.

The question about the relative spatial localization of structural genes and the CpG-islands seems to be important from this point of view. Approximate estimations provide evidence that more then half of the genes that form the functioning human genome contains CpG-islands. Apparently, all the house-keeping genes and

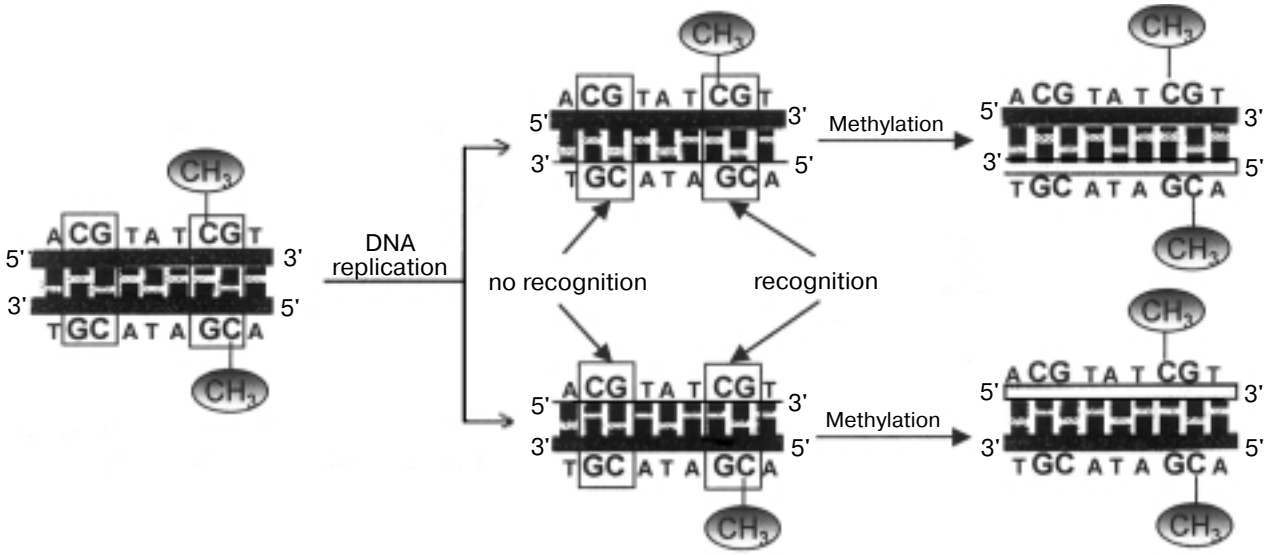

Fig. 2. Post-replicative maintenance methylation of CpG [31]. Methylated CpG dinucleotides are present only in the maternal DNA strand after replication. DNA-methyltransferase recognizes the methylated CpG and restores the same methylation pattern in the daughter DNA strand (reprinted by permission from Academic Press).

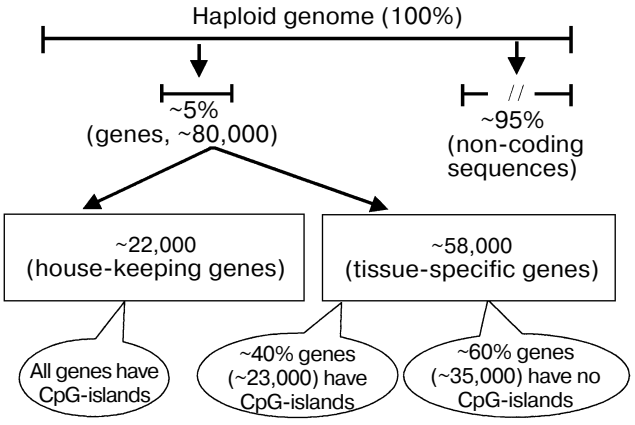

Fig. 3. Distribution of CpG-islands in the human genome. The haploid genome ($3 \cdot 10^9$ bp) contains ~80,000 genes, part of which (~22,000) consists of "house-keeping genes", the other (~58,000) includes genes that have specialized functions (tissue-specific genes). The majority of the former and ~40% of latter contain CpG-islands close to the regulatory sequences. Therefore, ~45,000 CpG-islands of the human genome are approximately equally distributed between "basal" and "specialized" genes.

~40% of genes that carry out specialized functions contain CpG-islands (Fig. 3).

A significant part of the tissue-specific genes does not harbor CpG-islands in their promoters. Instead, they may contain individual CpG dinucleotides. Their methylation level varies significantly (they may be different in different cells and tissues). In this case, methylation does not have such a large-scale influence on the chromatin structure, as methylation of CpG-islands does, but has affects relatively locally, preventing the binding of transcriptional factors. It may be tissue-specific in specialized cells. The promoter of GK-interferon [39] and the CREB (cyclic AMP response element binding-protein) binding region in the promoter of β-globin [40] provide examples of site-specific methylation that leads to the inhibition of transcription. The differences between the two types of promoters are summarized in Table 1 [11].

However, it is evident that the repressive role of methylation of single CpG dinucleotides is limited only by those tissue-specific genes that are regulated by transcriptional factors sensitive to methylation. This statement was supported by the observation that in mouse embryos deficient in DNA-methyltransferase (Dnmt1) DNA demethylation did not lead to untimely or ectopic activation of studied tissue-specific genes [41]. Apparently, methylation is not playing a significant role in the expression of a considerable part of the tissue-specific genes, and this is controlled by other mechanisms.

Methylation patterns. Three possible patterns of methylation of CpG sites are distinguished according to the analysis of the methylation level of a significant number of alleles [34]: completely methylated (almost 100%), unmethylated (about 0%) and partially methylated (between the first two). For the imprinted genes and the genes of the inactivated X chromosome in female cells, the intermediate methylation level (~50%) reflects the alternative methylation of the two alleles.

The genome is thought to be methylated in two steps at the stage of the post-implanted blastocyst [35]: *de novo* methylation, implicating relatively local genome regions, and the spreading of methylation "spots" on surrounding regions (Fig. 4). During the first step, the interaction occurs between the enzyme capable of *de novo* DNA methylation (*trans*-acting factor) and a *cis*-acting element (target DNA, which binds the enzyme due to some yet unclear properties of its structure: enrichment with CG bp, presence of repeats, unusual secondary structure). In the second step the methylation "spot" spreads in the 3′-direction. The resulting methylation pattern is

Table 1. Methylation status and functional peculiarities of promoters that do or do not contain CpG-islands [11]

Promoters containing CpG-islands	Unmethylated in almost all normal cells (independently of transcriptional activity); exceptions are genes of the inactivated X chromosome in female cells and inactive alleles of imprinted genes in autosomes Unmethylated state is associated with an opened nucleosomal conformation that is typical for transcribed genes, with a decrease in the amount of histone H1, with histone hyperacetylation, with replication in the early S-phase of the cell cycle Methylation of CpG-islands is associated with large-scale changes of the chromatin structure, with nucleosomal compactization, with replication in the late S-phase of the cell cycle, and with transcriptional inhibition
Promoters not containing CpG-islands	Variability of methylation pattern Methylation of dispersed CpG dinucleotides close to and within a promoter does not significantly influence the chromatin structure but may restrain binding of transcriptional factors, thus modulating gene activity tissue-specifically

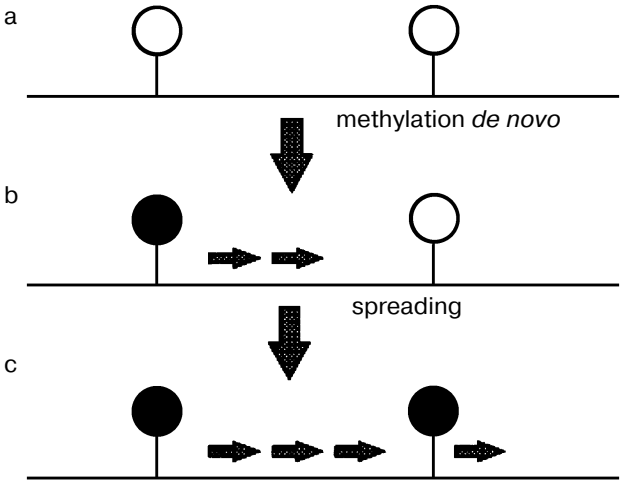

Fig. 4. Two-step methylation process of the CpG sites in the genome of a blastocyst after implantation [35]. a) Initial (unmethylated) locus; b) *de novo* methylation due to a specific interaction between DNA-methyltransferase capable of this reaction and a *cis*-acting methylation signal; c) spreading of the methylation "spot" on the distal, 3′-regions (apparently DNA-methyltransferase 1 is active in this step) (reprinted by permission from Academic Press).

then maintained over many cell generations (maintenance methylation).

Thus, at least two methylation systems are functioning in mammalian cells—*de novo* methylation providing elements of variability to the methylation pattern and maintenance methylation ensuring conservation of an already formed pattern. At the early stages of embryogenesis, the first system is dominating, but later (especially in post-embryonic development) the cells lose to a great extent the ability for *de novo* methylation, while the spreading is maintained [42, 43]. The center of gravity of methylation is then directed towards the system of maintenance methylation that is activated at every cell division (Fig. 2).

The mechanism of DNA methylation is significantly altered in carcinogenesis (see below). Massive methylation of CpG-islands observed in cancer (but not methylated in normal cells) is apparently caused not by *de novo* methylation, but, more likely, by the spreading mechanism, which is fixed by maintenance methylation over cell generations.

Enzymes of methylation. Four DNA-methyltransferases have been identified in mammals: Dnmt1, Dnmt2, Dnmt3a, and Dnmt3b. Different opinions exist about the nature of the DNA methyltransferase methylating *de novo*. Perhaps, it is the DNA-methyltransferase 1 encoded by the *DNMT1* gene that is responsible both for the methyltransferase activity, dominating in mammalian cells, and for the maintenance methylation. To date, Dnmt1 is the best-studied enzyme of the DNA methylation system in vertebrates [11, 44]. Further evidence for the functional importance of this enzyme comes from the observation that a homozygous deletion of the *dnmt1* in mice is lethal at the embryonic stage of development [45]. This fact provides the determining evidence for the necessity of DNA methylation in higher eukaryotes. The same model confirmed the long ago suggested association of DNA methylation with silencing of a related gene [46, 47].

Experiments with 5-azacytidine provide another tool for analysis of the association of transcriptional repression with DNA methylation. 5-Azacytidine incorporates in replicating DNA and cannot be methylated due to its chemical structure (N instead of C in position 5). However, its demethylating effect is ensured to a greater extent by repression of DNA-methyltrasferase activity [48]. Transcriptional reactivation of methylated genes and genetic constructions after treatment of cells with 5-azacytidine was demonstrated in many cases. Furthermore, in certain situations this test provides absolute evidence for the involvement of methylation in a phenotypic feature. However, the application of this test is limited by the possibility of side effects.

DNA-methyltransferase 1 is a protein of 190 kD consisting of two domains: a catalytic domain located in the C-terminal part of the protein, which is structurally homologous to bacterial cytosine-methyltransferases, and the regulatory domain located at the N-terminus, which contains a signal sequence that directs the enzyme towards active replicative complexes in dividing cells [49]. The enzymatic activity of Dnmt1 is significantly increased when DNA synthesis starts. This effect is possibly mediated by the activation of the *dnmt1* promoter by the *H-ras* gene product, which is implicated in one of the main pathways of mitogenic signal transfer [50].

Since a replication complex also includes the DNA-methyltransferase 1, the methylation pattern of the daughter DNA strand is restored within the first minutes after replication, using the maternal strand as a template [51, 52].

However, a weak ability of DNA-methyltransferase 1 to methylate DNA *de novo* (the enzyme has a strong preference for hemi-methylated templates, i.e., for DNA with one strand already methylated, see Fig. 2) and the fact that in cells with "knock-out" of the *dnmt1* gene the *de novo* methylation still occurs [53] causes doubt whether Dnmt1 in addition to maintenance methylation is capable of *de novo* methylation.

Recently, a number of genes whose products posess high ability for *de novo* DNA methylation and that are apparently responsible for methylation *in vivo* have been cloned [54]. Two DNA-methyltransferases, Dnmt3a and Dnmt3b, were shown to be essential for the *de novo* methylation and for the embryonic development in mice. Inactivation of appropriate genes is incompatible with the

development of an embryo, arrests the *de novo* reaction, but, at the same time, has no effect on the maintenance methylation of imprinted genes. Different regions of the genome are supposed to be targets for Dnmt3a and Dnmt3b [55, 56]. The role of the forth enzyme, Dnmt2, is not yet clarified [57].

Genome methylation as a dynamic process. The presented picture is descriptive and does not elucidate which factors determine the complicated mosaic of methylation of genomic DNA, and, in particular, what protects the CpG-islands from methylation expansion (it is known that transcribed genes and their promoters are often surrounded by intensively methylated sequences like *Alu* repeats) [36, 37]. Besides, the existence of hemi-methylated CpG sites cannot be explained either (it is not clear what factors protect them from total methylation).

From this point of view, a hypothesis considering the existence of stable genome methylation pattern as an "illusion" is highly attractive. The methylation pattern seems to be a result of the dynamic equilibrium between two constantly active and opposing processes: "expansion" of methylation and protection from it [35]. This hypothesis emerged after the analysis of methylation of 5′-sequences of gene *Aprt*, which is constitutively expressed in tissues and cultured mouse cells (Fig. 5). At the early stages of embryonic development, this sequence is not methylated. Methylation is initiated after the implantation of the blastocyst in a so-called methylation center, the active part of which consists of two copies of the B1 element (repetitive sequences in rodents analogous to the *Alu* repeats of primates). An enzyme capable of *de novo* activity methylates the H1 site, and then methylation spreads downstream towards the gene promoter (the Dnmt1 methylase is supposed to be responsible for the spreading).

The site of binding of the Sp1 transcriptional factor protects a promoter from methylation expansion, arresting the spreading close to the H2 site. The methylation of the H2 site is dynamically unstable (less than 50% methylation) since it is subjected to competition between methylation impulses released from the methylation center and demethylation impulses from the promoter structures (Fig. 5).

Thus, according to this hypothesis, stable methylation seems to be an illusion. This suggests the possibility of modification of the established methylation pattern due to a shift of the equilibrium by induction or suppression of one of the competitive processes that develops through several cell generations. According to [35], exactly the same process occurs during aging and malignant cellular transformation.

Methods of studying of DNA methylation. Various methods of quantitative estimation, as well as detection of methylation state of specific genomic regions, have been applied during the 50 years of study of DNA methylation [9].

The total content of methylated cytosine in DNA is estimated after acidic or enzymatic hydrolysis of DNA. 5-Methylcytosine is detected by different methods (one- or two-dimensional chromatography or electrophoresis, gas chromatography, mass-spectrometry, HPLC), and its relative contents is calculated.

Conjugates of 5-methylcytosine with serum albumin can be used for production of antibodies that effectively (but not absolutely specifically) detect small quantities of this modified residue.

The 5-methylcytosine residing in certain DNA sequences can be detected by the Maxam–Gilbert

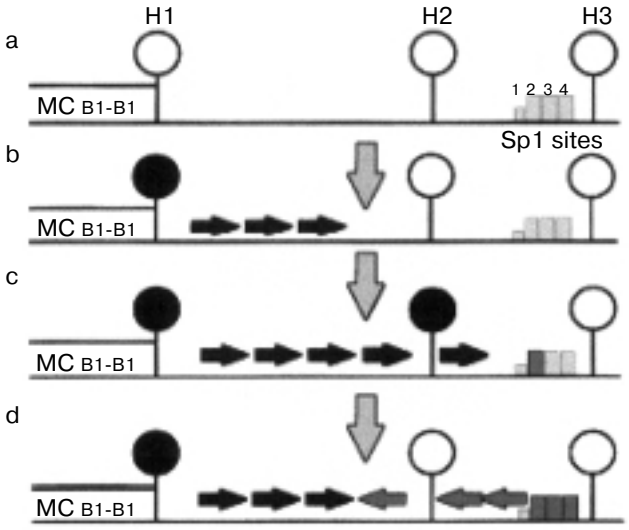

Fig. 5. Model of formation and maintenance of methylation pattern in the 5′-region of the mouse *Aprt* gene [35]. MC, *cis*-acting methylation center (a 838 bp-fragment essential for methylation of flanking regions); B1, repeated elements analogous to the human *Alu* sequences (2 copies), which are located on the 3′-end of the methylation center and determine its activity; H1, H2, and H3, methylation sites (the latter is located within the CpG-island); Sp1, SP1 transcriptional factor binding sites. *De novo* methylation is initiated in the methylation center (*not* by the Dnmt1 methylase) and affects the H1 site, followed by the spreading (arrow to the right) in the distal 3′-direction (by the Dnmt1 methylase). The *cis*-activity of Sp1 sites prevents the H2 and H3 sites from methylation spreading (arrow to the left), which can lead to the inactivation of the *Aprt* gene (the deletion of Sp1 sites leads to uncontrolled methylation of H2 and H3). The methylation is "exhausted" not far from H2 (down arrow) due to the interaction of the Sp1 transcriptional factor (or related proteins) with appropriate binding sites (from 2 to 4). a) The unmethylated allele at the early stages of development (before implantation of the blastocyst); b, c) after the implantation of the blastocyst the H1 site is methylated *de novo* followed by the spreading of the process in the 3′-direction with occasional involvement of the H2 site; d) arrest of the process due to the interaction of the Sp1 transcriptional factor with binding site 2. Competition of these two opposed processes leads to only partial methylation of the H2 site (less than 50%) (reprinted by permission from Academic Press).

method. Since the modified residue reacts with hydrazine to a lesser extent than cytosine or thymine, the corresponding band is absent on the sequencing gel. The confirmation can be obtained by sequencing of the complementary DNA strand.

Restriction endonucleases sensitive or not to methylation of the cytosine residue (for example, the pairs like *Hpa*II and *Msp*I, *Sma*I and *Xma*I) are often used for analysis of the methylation of specific sequences. Together with Southern blot hybridization, this method can be used to establish whether the recognition site of the corresponding enzyme is methylated. This approach is limited by the number of corresponding isoschizomers, which does not cover all CpG-dinucleotides.

MSP (methylation-specific PCR) is a new method that is suitable for estimation of the methylation state of individual CpG-islands. High sensitivity that allows analysis of modified (methylated) alleles in an abundance of wild-type alleles is considered to be a significant advantage of this method [58]. The method consists of treatment of tested DNA by sodium bisulfite that in certain conditions leads to cytosine deamination with the formation of uracil, while the methylated cytosine residues are resistant to this treatment. After a following PCR, the uracil is replaced by thymine. Thus, it is possible to construct primers that will selectively amplify sequences containing (or not containing) methylated cytosine residues in CpG-islands. The method can be used to detect 1 methylated allele in the presence of 1000 unmethylated.

Mapping of all methylated cytosine residues in a known DNA sequence [59] is based, like the previous method, on the replacement of all non-methylated cytosine residues by uracil after treatment if the DNA with sodium bisulfite (after the treatment, the DNA strands become non-complementary in the sites of modification). Amplification of the studied sequence by PCR, followed by the standard sequencing, detects only these cytosine residues that were methylated in the initial strand. The sensitivity of this method corresponds to that of PCR. This method permitted to establish the methylation state of individual genes in early embryonic development (at the stage of two cells), and besides, to detect the CpG-dinucleotides within CpG-islands whose methylation is essential and sufficient for the inhibition of transcription of a corresponding gene [41, 60, 61].

MCA (methylated CpG island amplification) is a PCR-based method that allows for specific amplification of CpG-islands that are differentially methylated in normal and tumor cells [62]. Initially, DNA is digested by *Sma*I restrictase (the recognition site CCCGGG is not cleaved if it acquires 5-methylcytosine), producing fragments with blunt ends. Methylated CCCGGG sites are subsequently cleaved by *Xma*I restrictase, which gives rise to fragments with cohesive ends. Only the latter can interact with adapters and be amplified in the following PCR. Together with the method of representation difference analysis (RDA) [63], this approach allows for the selective amplification of CpG-islands aberrantly methylated in tumor cells. This approach was used for identification and cloning of 33 sequences, including the fragments of known genes (*PAX6*, *Versican*, *α-tubulin*, *CSX*, *OPT*, and others) in colon carcinoma cells.

RLGS (restriction landmark genomic scanning) makes possible the simultaneous analysis of the methylation status of several thousands of CpG-islands in the genome [64, 65]. The method consists of separation of restriction fragments, obtained after cleavade of genomic DNA by a couple of restrictases, by two-dimensional gel electrophoresis. First, the DNA is subjected to digestion by the infrequently cutting restrictase *Not*I (cleaves only non-methylated CpG). After end labeling with radioactive label, the produced fragments are digested by a second enzyme (for example, *Eco*RV) in order to decrease their sizes. The products of hydrolysis are separated in the first dimension (in a capillar agarose gel). Separated fragments *Not*I/*Eco*RV are treated *in situ* (i.e., in the gel) with another enzyme (for example, *Hinf*I), and the electrophoresis in polyacrylamide gel is performed in the second dimension. After autoradiography of the gel, numerous spots are detected. The intensity of each spot reflects the methylation status of the corresponding CpG-island (the position of each spot is strictly determined). The intensity of a spot considered as normal indicates the absence of methylation of both alleles, hemi-reduction of intensity indicates methylation of one allele, and disappearance of a spot indicates methylation of both alleles (the corresponding site is not digested by *Not*I).

Influence on Chromatin Structure

There are two major forms of chromatin—euchromatin (decondensed, potentially active, is replicated in the early S-phase of the cell cycle) and heterochromatin (condensed, not active, is replicated in the late S-phase). The structure and function of chromatin are determined to a significant extent by modification of its components—histone acetylation (leading to decondensation and activation) and DNA methylation (conversely, leading to condensation and inactivation) [66]. Therefore, the active chromatin is characterized by hyperacetylated histones and the absence of 5-methylsytosine, while inactive by hypermethylated cytosine and deacetylated histones. The link between the repressive effect of DNA methylation on the transcription and the histone deacetylation will be discussed below.

Influence of DNA methylation on chromatin structure. Several proteins, repressors of transcription that bind to methyl-CpG-dinucleotides in DNA, have been identified [67-69]. The best-studied are MeCP1 and MeCP2. One of them, MeCP2 (essential for normal

embryonic development), binds tightly to chromatin. It consists of two domains—one domain binds methyl-CpG, the other functions as a transcriptional repressor. The ability of MeCP2 to repress at a distance from the site of binding is explained by the spatial separation of the domains. The flexibility of the chromatin chains is thought to favor this ability (meaning formation of an intermediate loop). In addition, cooperation between MeCP2 molecules is possible, providing the spreading of the MeCP2 binding site outside of the site of the methylated CpG due to the interaction with new MeCP2 molecules. Finally, steric hindrances for the interaction of condensed chromatin with the transcriptional machinery arise.

Histone acetylation. Acetylation of N-terminal tails of lysine residues of H2A, H2B, H3, and H4 histones neutralizes their positive charge and consequently blocks association with nucleosomal DNA. This, in its turn renders the nucleosomal structure decondensed, as well as that of the chromatin in general, and in addition releases the external surface of the nucleosomal DNA for interaction with regulatory factors [70].

The level of histone acetylation is determined by the activity of two enzyme types—histone acetyltransferases (HAT) and histone deacetylases (HDAC). A number of transcriptional activators and coactivators (in particular, the very important CBP/300 that participates in the regulation of cellular growth, differentiation, DNA repair, and apoptosis) as well as certain subunits of the basal transcriptional machinery ($TAF_{11}250$) possess histone acetyltransferase activity. In contrast, transcriptional repressors (like Mad and nuclear receptors) are associated with deacetylase activity [71].

Link between DNA methylation and histone acetylation. The inhibitory effect of methylation is mediated by histone deacetylation, which induces the compactization of chromatin [72]. For example, trichostatin A, a specific inhibitor of HDAC, can reverse the effect of methylation on the blocked ribosomal genes (without affecting their methylation state) [73]. The same effect was described for a myeloid-specific gene of mouse M-lysozyme that is methylated and inhibited in non-myeloid cells [74].

The mechanism of histone deacetylation consists of the recruitment of the corepressor complex mSin3A/HDAC by the MeCP2 bound to methylated residues. This complex is responsible for the deacetylation of the N-terminal tails of histones, which results in the appearance of a positive charge on the histones, rendering them capable of tight interaction with nucleosomal DNA. Nucleosomes become compact and lose the interaction with transcriptional factors. In contrast, the recruitment of proteins possessing HAT activity in the complex with transcriptional factors erases the repressive effect of methylation [75, 76].

Competition between transcriptional activators and repressors. A hypothesis of competitive relationships between factors favoring and preventing methylation spreading was discussed above. Perhaps, it should be spread further—on the partnership of the repressive methylation system (including DNA-methyltransferase 1 and MeCP2) and transcriptional activators [70].

Indeed, some data indicate that the transcriptional activity of a gene, DNA methylation status, and histone acetylation mutually influence each other, ensuring a self-maintaining process. It is suggested that an inactivated gene and the hypoacetylated state of its histones attract the methylation system, which fixes this state. And, conversely, an active gene and the hyperacetylated state of its histones protect the corresponding locus from methylation [72].

In this regard, the observation that a lag-phase exists in DNA replication between methylation (~1 min after the synthesis of a corresponding region) and the formation of mature chromatin (10-20 min after the synthesis) seems to be interesting. Apparently, a competition between counteracting factors plays a crucial role during this period. The factors that bind first determine one of the alternative states of chromatin in the future. Thus, as demonstrated in model experiments, the powerful transcriptional activator Gal4-VP16, when present *during* the reconstitution of chromatin on methylated DNA *in vitro*, overcomes the repression and ensures the transcriptional activity of the template, while its addition to the medium *after* chromatin reconstitution fails to restore the transcriptional activity. Further, the presence in the medium of a transcriptional activator that binds the promoter of a gene directly after its replication and *before* the binding of the DNA-methyltransferase 1 may progressively (after several cellular divisions) lead to the demethylation of a specific locus. This hypothesis was confirmed by studying the embryogenesis of *Xenopus laevis*: it was established that binding of transcriptional factors to replicating DNA leads to demethylation [77].

DNA METHYLATION: CARCINOGENESIS AND AGING

The role of DNA methylation in pathology can be most clearly revealed in malignant transformation, where aberrant and stable inactivation of suppressor(s) gene(s) leads to the same consequences as their "physical" damage. This effect was referred to as epimutation [78], i.e., an epigenetic equivalent of a genetic mutation. The abbreviation MAGI (methylation-associated gene inactivation) is also used [79].

It is important to mention that DNA methylation contributes to carcinogenesis not only by an epigenetic mechanism. 5-Methylcytosine can induce mutations due to its considerable instability, and, consequently, it favors genetic carcinogenesis. In addition, methylation suppresses homologous recombination [80]. Therefore, gen-

eral genome demethylation that accompanies malignant transformation may contribute to some extent to genetic instability of a tumor cell by removing the suppression of recombination.

Some data that characterize the contribution of DNA methylation in both components (genetic and epigenetic) of carcinogenesis are presented below.

Genetic Component of Carcinogenesis

Induction of mutations. The close positioning of methyl and amino groups after methylation of cytosine (Fig. 6a) leads to the instability of the latter to such extent that 5-methylcytosine can undergo spontaneous deamination even under normal conditions due to temperature fluctuations [10, 81, 82]. As a result thymine is formed (Fig. 6b), giving rise to a DNA defect (mismatched residues G-T), which induces the reparation system. Since thymine is a normal residue for DNA, and, consequently, it is not recognized as a foreign residue by the repair enzymes, the situation can be resolved in two ways: the restoration of the wild-type sequence (G-C pair) or the appearance of a mutation (replacement of G-C base pair by A-T). Although the repair enzymes erase thymine and restore the initial sequence with higher probability, mutations occur very frequently. Essentially, the instability of 5-methylcytosine resulting in multiple substitutions of C-G by A-T pairs over evolution explains the relatively low presence of CpG dinucleotide in the human genome.

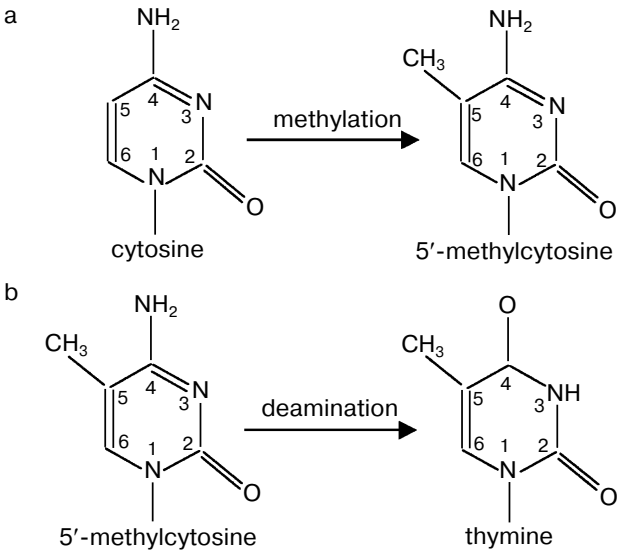

Fig. 6. Reactions of cytosine methylation (a) and 5′-methylcytosine deamination (b). A high frequency of mutations is caused by the instability of 5′-methylcytosine and its transformation to thymine. About 100 reactions of deamination of methylated cytosine occur in each human cell during every 24 h due to temperature fluctuations.

The fact that the instability of 5-methylcytosine gives rise to multiple mutations is important in oncology. The estimations indicate that ~100 reactions of 5-methylcytosine deamination occur in each human cell every 24 h, and some of them lead to the substitution of G-C by A-T pairs. In particular, the observation that 25-30% of ~300 mutations detected in the *p53* suppressor gene in human tumors of different origin (*p53* is considered as the main guardian of genome integrity) are of the type described above provides evidence for the importance of this phenomenon. Therefore, spontaneous mutagenesis (i.e., occurring in the absence of any exogenous or endogenous agents that damage DNA) occurs in human cells very actively and damages many genes, including suppressor genes.

Genome instability. General DNA demethylation is one of the major characteristics of a tumor cell (see below). This process of epigenetic origin may apparently have a large-scale destabilizing effect on the genome. Hence, DNA demethylation in mouse embryonic cells (ES) without the *dnmt1* gene leads to a 10-fold increase in the frequency of deletions and insertions in unique genes [80]. Artificial methylation of DNA regions corresponding to the hot points of recombination in fungi inhibits homologous recombination in meiosis [83]. With regard to these data obtained in experiments with normal cells, genome instability seems to be a highly probable consequence of defects of the methylation system that emerge in carcinogenesis, although direct indications of causal-consecutive links between DNA hypomethylation and chromosomal anomalies in tumor cells have not yet been established.

Epigenetic Component of Carcinogenesis

The tumor cells are characterized by considerable modifications of DNA methylation system: general genome demethylation, increase in DNA-methyltransferase activity, and local hypermethylation [11, 84, 85]. These modifications seem to be contradictory. For example, it should be difficult to combine the increase in DNA-methyltransferase activity and local hypermethylation of the genome on one hand with its overall demethylation on the other hand. The contradictions disappear to a certain extent after more detailed study of the nature of these modifications. It was estimated that the demethylation involves almost exclusively the "dispersed" CpG (their relative contents is high, ~80% of total content, and they are methylated in normal cells), while the local hypermethylation spreads on the CpG-islands (they represent about 20%, and except for certain exceptions are not methylated in normal cells). Thus, it is evident that

methylation of CpG-islands does not compensate demethylation of "dispersed" CpG, so the result is the general genome demethylation. Therefore, the reciprocal relations between the levels of methylation of the CpG-islands and of "dispersed" CpG, which are present in normal cells, are preserved in tumor cells, but in the latter case the equilibrium is shifted in the opposite direction—the islands that were "free" before get methylated, while the "dispersed" CpG dinucleotides become demethylated (Fig. 7). The increased activity of DNA methyltransferase that is typical for tumor cells can be one of the factors, which favor the aberrant methylation of CpG-islands.

General genome demethylation. The decrease in the number of modified DNA residues is one of the earliest marks of transformation (often before the appearance of a tumor) [86, 87]. The observation that the feeding of rodents with a diet without methionine, which leads to a deficit of donors of methyl groups, gives rise to DNA hypomethylation and then to liver tumors provides evidence for the causal relations between DNA hypomethylation and cellular transformation [80, 87].

The causes of genome demethylation and the mechanisms of its carcinogenic effect are not yet clarified. Some observations indicate that this process in human and animal tumor cells can involve several oncogenes, in particular, *K-RAS* [88, 89]. In the case of total demethylation, these local gene-specific alterations emerge relatively early, sometimes long before tumor formation, particularly in benign polyps, the precursors of tumors [90].

Two types of effects of genome demethylation can be suggested taking into account two types of distribution of methylated residues in mammalian DNA (individual dispersed CpG dinucleotides and CpG clusters): local modification, influencing the activity of single genes, and general, which alters the chromosome structure. The former apparently does not play an important role. The promoters of a significant fraction of genes harbor CpG-islands, which are predominantly not methylated, and, consequently, should not be considered as demethylation targets. Another approach can be applied for genes that contain a single CpG in their promoters. Their methylation level varies and may have a tissue-specific effect (methylation and transcription are negatively correlated). Genome demethylation that involves these promoters may indeed favor the activation of appropriate genes (see above). However, the necessity of a *trans*-activating protein factor for the gene activation in addition to the promoter demethylation should also be considered. In the absence of this factor the transcription is impossible, which apparently explains a relatively limited number of genes that are activated in tumors due to genome demethylation.

On the other hand, genome demethylation as a general phenomenon may also influence gene expression indirectly by changing chromatin structure. As mentioned above, the sequences of inactive, compact chromatin (heterochromatin) are intensively methylated, which puts them functionally in a suppressed position (inhibition of transcription, late replication). Total demethylation may dramatically influence chromatin structure, its condensation state, and the timetable of replication. Several observations provide evidence for the reasonability of these suggestions. In particular, it was demonstrated that genome demethylation by 5-azacytidine treatment gives rise to transformation in certain cell cultures [91, 92] and to abnormalities of chromosome segregation in mitosis [11].

Increase in DNA-methyltransferase 1 activity. Two facts suggest the importance of DNA-methyltransferase 1 in cell metabolism: knock-out of *dnmt1* in transgenic mice is lethal at the embryonic stage of development [45], and an increased expression of this gene leads to cell transformation. An increase in DNA-methyltransferase

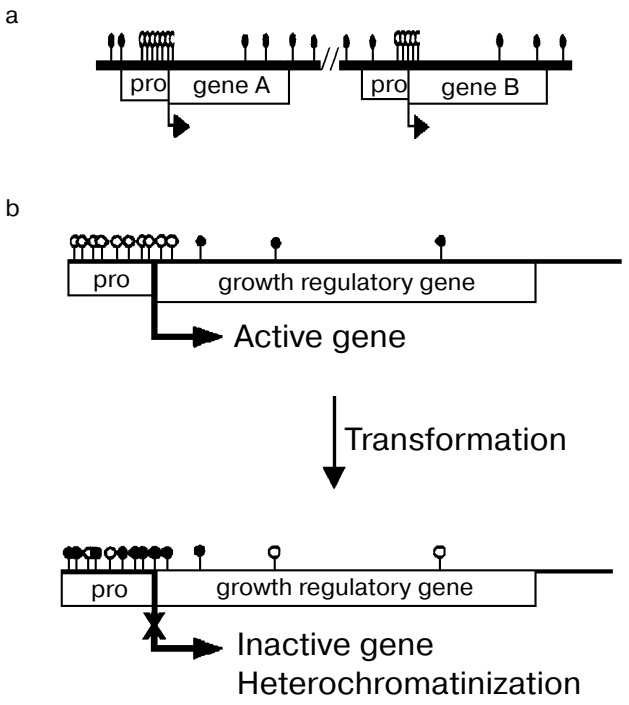

Fig. 7. Alterations of DNA methylation in tumor cells in comparison to normal cells [31]. a) Normal cells are characterized by reciprocal methylation of CpG-islands, which usually are not methylated (white circles), and of dispersed CpG (are usually methylated, black circles). Pro, promoter region of a gene; b) tumor cells are characterized by an opposite tendency—demethylation of dispersed CpG and methylation of CpG-islands. The MAGI phenomenon (methylation-associated gene inactivation) lies in the basis of cellular transformation. Heterochromatinization and inactivation of a repressor gene may favor uncontrolled cell proliferation (reprinted by permission from Academic Press).

activity is one of the well-established properties of transformed cells [11]. It was demonstrated in cell culture studies that the increase in DNA-methyltransferase activity often precedes transformation [11]. In addition, the application of RT-PCR for the analysis of *DNMT1* in colon cells demonstrated progressive increase in *DNMT1* gene expression through tumor progression—from normal mucosa through adenoma to the stage of a mature carcinoma [93].

It was suggested that the increase in DNA-methyltransferase activity is not a cause but a consequence of carcinogenesis, namely that it reflects the increased number of proliferating cells in the population (the enzyme activity correlates with the rate of DNA synthesis). Indeed, normalizing the enzyme activity relative to histone H1 synthesis, which reflects the S-phase activity of the cell, to some extent corrects the striking differences between normal and tumor cells observed in early studies [11]. However, in the majority of analyzed cases, such a simple explanation appears to be invalid—the increase in DNA-methyltransferase activity is more considerable than the increase in the number of proliferating cells.

The molecular mechanisms of increased expression of *DNMT1* in tumor and pre-tumor cells actually remain obscure. It is suggested to be a compensatory reaction of a cell to the total genome demethylation. On the other hand, as demonstrated in experiments on cultured cells, DNA methylation and *ras*-mediated pathways of mitogenic signal transduction are closely related. Thus, introduction of an exogenous negative regulator Ras (Gap) in mouse adrenocortical tumor cells Y1 led, on one hand, to the decrease in mRNA level and activity of Dnmt1, and on the other, it restored the normal phenotype. The transfection of exogenous H-*ras* in the resulting revertants increased the level of mRNA expression and the enzyme activity and restored the transformed phenotype [50]. It was demonstrated in experiments with cells transformed by constitutively expressed c-*fos* (Fos is situated after Ras in the signal transduction pathway) that the suppression of the expression or the activity of Dnmt1 restores the normal phenotype without variation of expression of exogenous c-*fos* [94]. Hence, it was established that at least several phenotypic features of Ras-mediated signal pathway, the activation of which has been observed in many tumors, are mediated by the increase in Dnmt1 activity.

In any case, the increase in Dnmt1 activity considerably influences the DNA methylation pattern in tumor cells and apparently contributes to the local hypermethylation phenomenon.

Local hypermethylation. General genome demethylation in tumor cells is frequently associated with local hypermethylation. This involves a relatively small fraction (~20%) of CpG dinucleotides, namely those that form clusters (CpG-islands), which are not methylated in normal cells and are located close to regulatory regions. This type of alteration is typical for tumor cells; it is thought to be one of the main mechanisms of inactivation of suppressor genes [11, 95].

As it will be shown below, hypermethylation of certain suppressor genes is observed in many human tumors, leading to their inactivation and thus favoring tumorogenesis. The mechanism of hypermethylation remains obscure. Apparently, the increase in DNA-methyltransferase activity plays an important role (see above). Thus, the transfection of a cloned human *DNMT1* gene in immortalized human fibroblasts leads to the aberrant methylation of CpG-islands in promoter regions of several genes, including *E-cad* and *HIC1*. At the same time, CpG-islands associated with other genes (for example, with $p16^{INK4A}$) are not altered in methylation state, although *DNMT1* is expressed constantly [96]. Therefore, it is evident that the increase in Dnmt1 activity plays a certain role in the selective aberrant methylation of CpG-islands, but not in total CpG-island methylation. Apparently the mechanisms that ensure the specificity of the enzyme towards hemi-methylated DNA are damaged in transformed cells together with the mechanisms of the protection of the CpG-islands from methylation.

From this point of view, the analysis of kinetics of local hypermethylation of CpG-islands (the suppressor gene *E-cad* is taken as an example) seems to be interesting for clarification of the mechanism of this process [97]. The promoter structure of this gene is analogous to that of mouse *Aprt* (Fig. 5): it contains a CpG-island (its methylation leads to gene inactivation), an Sp1-binding site, and intensively methylated *Alu* repeats (analogous to the mouse B1 elements). Hyperexpression of the *DNMT1* gene gave rise to the spreading of methylation in human transformed fibroblasts, which gradually increased through several cell generations. The *Alu* repeats served as initial points for this spreading, which finally resulted in the involvement of the promoter and gene inactivation. In general, the events described above are in agreement with the hypothesis about the dynamic equilibrium between methylation expansion and protection from it (Fig. 5) that provides the basis for the formation of the methylation pattern [35]. Consequently, the hypothesis is proposed that hypermethylation of promoters could be caused by a shift of this equilibrium due to activation of the spreading (increase in DNA-methyltransferase 1 level), attenuation of the protection (deletion or damage of Sp1 sites or of other sites of the same type, decrease in the level of Sp1-binding proteins), and a combination of both.

Whatever is the distinct mechanism of CpG-island methylation, it provides a stable inactivation of a related gene (i.e., MAGI phenomenon—methylation-associated gene inactivation) due to the appearance of steric hindrance for the binding of methyl-sensitive transcriptional factors to the promoter sites, or to heterochomatinization caused by methylcytosine-binding proteins, which recruit histone deacetylases to the repressor protein complex.

The biological consequences of the MAGI phenomenon differ significantly depending on the type of the gene involved. If it is one of the house-keeping genes, it would be most probably a lethal event for the cell, but it would not have any dramatic consequence for the organism. Inactivation of one of the tissue-specific genes will apparently alter to some extent differentiated cellular phenotype without influence on the cell viability.

Inactivation of any suppressor gene (or gene of the reparation system) can have much more dramatic consequences since it favors uncontrolled cell proliferation. Thus, the distinct mechanism of inactivation of a suppressor gene, namely, mutation (alteration of nucleotide sequence) or epimutation (methylation of nucleotide sequence), makes no difference regarding the final biological consequences.

Much data confirming the contribution of epigenetic events (DNA methylation) in the appearance and progression of human tumors have been accumulated during recent years (Table 2). It was demonstrated that such suppressor genes as *p16*^{ink4a}, *p15*^{ink4b}, *HIC* (hypermethylated in cancer), *WAF*, and others are inactivated in many tumors by this pathway.

Brief characterization of some of the above-mentioned genes that are often inactivated by the epigenetic mechanism in tumor cells is presented below.

Calcitonin gene. The promoter of this gene was the first found to contain a CpG-island that is not methylated in all examined normal tissues but is intensively methylated in many solid human tumors [98], in leukemia, and in cell lines transformed by various viruses [99]. The role of this gene in carcinogenesis is still not clarified.

Rb1 gene. The retinoblastoma gene was the first classical suppressor gene for which the MAGI phenomenon was established. It is suggested that all (or almost all)

Table 2. Genes hypermethylated in tumors for which suppressor function is established or suggested [11]

Gene	Locus	Tumor
Genes with established suppressor function		
RB1	13q14.2	retinoblastoma
VHL	3p25	renal cancer
CDKN2A	9p21	many solid tumors and lymphomas
CDKN2B	9p21	acute leukemias, Burkitt's lymphoma
CDH1 (E-cadherin)	16q22.1	bladder, breast, colon, and liver cancers
MLH1	3p	colon and rectal cancers
PTEN	10q23.3	prostate cancer
Genes with suggested suppressor function		
ERS1 (estrogen receptor)	6q25	breast, colon, lung cancers, leukemias
HIC1	17p13.3	brain tumors, breast, colon, renal cancers
MDG1	1p35	breast cancer
TP73	1p36	acute lympholeukemia, Burkitt's lymphoma
GSTP1	11q13	prostate cancer
MGMT	10q24	brain tumors
CALCA (calcitonin)	11p15	many carcinomas, leukemias
CACNA1G (T-type calcium channel)	17q21	colon, rectal, and stomach cancers; acute myeloid leukemia
THBS1 (angiogenesis inhibitor)	—	gliomas

antiproliferative signals in the cell are mediated by Rb or the related proteins p107 and p130 [100]. Hypermethylation of a CpG-island in the *Rb1* promoter occurs in approximately 10-15% of cases of sporadic (unilateral) retinoblastoma. Inactivation of *Rb1* is mediated by the epigenetic mechanism only in this tumor type. In all other cases (*Rb1* inactivation is observed in many forms of cancer) this gene is inactivated due to mutations or deletions in it. Therefore, it could be concluded that mutations and epimutations of *Rb1* are mutually exclusive, which confirms the same functional result of these alterations.

VHL (von Hippel-Lindau) gene. Alterations of this suppressor gene are frequently detected in clear-cell renal carcinoma and in other tumor types. The deletion of one allele of the *VHL* gene is associated with hypermethylation of a CpG-island in the promoter of the second allele in approximately 20% of sporadic clear-cell renal carcinomas [31].

The p16 gene (CDKN2A, *MTS1*, *INK4A)* encodes for a constitutively expressed inhibitor of cyclin-dependent kinase. The p16 protein plays a key role in cell cycle control by arresting the cyclin D-Rb signal pathway. The damage or inactivation of this protein leads to the loss of control of the cell cycle: cyclin-dependent kinase phosphorylates Rb, resulting in the release of E2F transcriptional factors from an inactive complex. This leads to the activation of genes responsible for the entrance to the S-phase of the cell cycle. Since lesions or deletions of the *p16* gene are very frequently detected in tumors of different origins, a hypothesis emerged that *p16* is a suppressor gene that is frequently damaged in carcinogenesis. Hypermethylation of a CpG-island in the 5′-region of this gene is found, according to different estimations, in 20-67% of solid human tumors of various localization [11].

p53 gene family. The *p53* gene plays an essential role of "guardian of the genome"; lesions in it have been identified in more than half of all human cancers. *p53* does not contain a CpG-island in the promoter region. Consequently, the MAGI phenomenon cannot be applied to it. However, the contribution of methylation to the *p53* inactivation is considerable, but it is not mediated by the epigenetic (as in the majority of cases), but by the genetic mechanism. Dispersed methylated CpG dinucleotides, which are present in large quantity within this gene, favor the replacement of G-C pairs with A-T pairs due to their hypermutability (approximately one third out of more than 300 mutations detected in this gene is of this type). Mutations that are so common for *p53* are rarely detected in the *p73* gene (a *p53* homolog), which harbors a large and typical CpG-island in the promoter. Hypermethylation of this gene was observed in acute lymphoblastic leukemia and lymphoma [31].

The INK4A/ARF locus encodes for two suppressor proteins in a manner that is unique for mammalian cells [101]. Two leader exons (different for each protein and each possessing its own promoter) are associated by splicing with common exons 2 and 3, forming different open reading frames. This results in the synthesis of two completely different proteins (p16 and ARF), each playing a significant role as a suppressor in cell cycle regulation. The former participates as a crucial element in the Rb-mediated signal pathway (see above), the latter increases the level of functional p53 by interaction with MDM2 (which binds to p53 and induces its degradation). The *ARF* promoter contains a CpG-island that is hypermethylated in cell lines derived from colon carcinoma. In addition, the *ARF* promoter contains a p53-responsive element that ensures the inhibition of transcription of this gene when the p53 level increases. Hence, a self-regulatory "loop" that functions by a feed-back mechanism and coordinates the expression of several genes (*p53*, *ARF*, and *MDM2*) is present in the cells. Methylation of the *ARF* promoter can alter the fine mechanism of maintenance of normal p53 level.

The p15 gene (CDKN2B) is located close to *p16*, has a strong homology with it, and also functions through the inhibition of cyclin-dependent kinase. In contrast to constitutively expressed *p16*, the *p15* gene is expressed in response to the growth-inhibiting effect of TGF-β. Hypermethylation of a CpG-island in the *p15* promoter was observed in several cases of glioblastoma. In contrast to solid tumors, where hypermethylation of the CpG-island of *p15* is rarely observed, it becomes the main pathway of the inactivation of this suppressor gene in tumors of the blood system [11].

E-cad gene. The cadherin gene family encodes for the cell surface proteins implicated in cell–cell adhesion. Their role of principal regulators of cell proliferation and differentiation consists of alteration of activity of certain genes through the β-catenin–transcriptional factor Lef/Tcf signal pathway. Association of closely located epithelial cells by E-cadherin bridges initiates an anti-growth signal directed into the cell [102]. *E-cad* is a suppressor gene whose inactivation favors the acquirement of invasive and metastasizing properties by cells. Transcriptional inhibition of this gene is observed in many human tumors, being in some cases mediated by the MAGI phenomenon, particularly, in breast, stomach, prostate, liver, and thyroid cancers [103].

Estrogen receptor gene (ER). Functional inactivation of the *ER* gene is considered to be the main cause of hormonal resistance of breast cancer cells. The conclusion that *ER* is a suppressor gene that modulates cell growth and differentiation was recently made. Its functions are altered in different forms of malignant neoplasia. In addition to mammary gland, *ER* expression is observed in various other tissues, in particular, in brain. A CpG-island resides in the *ER* promoter and its hypermethylation in many transformed cell lines is combined with inhibition of expression, demethylation by 5-azacytidine giving rise

to reactivation of the gene [104]. A clear association of hypermethylation of the *ER* CpG-island with many types of human tumors, in particular, with colon carcinoma, has been established. Introduction of the *ER* gene in cells derived from these tumors inhibited their growth [105, 106].

The Hic-1 gene (hypermethylated in cancer), which is supposed to be a suppressor gene, encodes for a protein referred to the family of transcriptional factors with a "zinc finger" sequence. Hypermethylation of the CpG-island of this gene has been observed in many human tumors and is associated with inhibition of expression. The *Hic-1* promoter contains a p53-binding site, indicating that *Hic-1* is possibly implicated in this signal pathway [11].

BRCA1 gene. Alterations of the *BRCA1* suppressor gene favor the development of breast cancer. Unexpectedly, the absence of mutations of this gene was observed in sporadic cancers. This observation allowed the suggestion of another, epigenetic mechanism of its inactivation. Indeed, *BRCA1* (but not *BRCA2*) is hypermethylated in some cases of breast cancer, in the cell lines derived from breast cancer, and in ovary cancer cells [31].

GST-π gene. The gene of glutathione-S-transferase-π is hypermethylated and its transcription is repressed in almost all studied cases of primary prostate carcinoma [107]. Methylation of the regulatory sequence of *GST-π* is a molecular "border" between benign and malignant forms of prostate carcinoma and the most constant marker of this tumor form. Besides, methylation of *GST-π* has been found in breast and renal carcinomas, but extremely rarely in other tumor types. The contribution of the hormonal factor in carcinogenesis is considered to be a common feature of malignant neoplasia with inactivated *GST-π*.

MLH1 gene. The product of this gene participates in the reparation of mismatched base pairs arising due to DNA replication errors. Mutations of this gene were found in patients with hereditary non-polyposis colorectal cancer (HNPCC), and aslo with the sporadic form of cancer of the same localization. *MLH1* inactivation leads to the accumulation of mutations in short repetitive (microsatellite) sequences. The loss of functional activity of this gene is often caused by the methylation of its CpG-island. It was demonstrated in recent publications that microsatellite instability in sporadic forms of colon and rectal cancers and cancers of endometrium and stomach is predominantly caused by the same reasons [31].

CACNA1G gene. A gene encoding for the T-type protein of calcium channels has been identified recently [108]. It is located on the 17q21 chromosome, where the phenomenon of loss of heterozygosity is often observed in various forms of human malignant neoplasia. The gene is expressed in normal colon mucosa and in bone marrow cells, while its transcription is repressed due to the methylation of the CpG-island in colon and rectal carcinomas, in stomach cancer, and in acute myeloid leukemia. *CACNA1G* is supposed to regulate the flow of calcium in the cell, a process that is directly associated with cell proliferation and apoptosis.

THBS1 gene. The gene of thrombospondin-1 that encodes for an inhibitor of angiogenesis is expressed in many tissues and is regulated by the *p53* and *RB* gene products. A decrease in expression of *THBS1* has been observed in many tissues. The re-expression of this gene in breast cancer cells inhibits the growth of a tumor, angiogenesis, and metastasizing. To date, mutations, deletions, or translocation of this gene have not been detected in human tumors. Hypermethylation of a CpG-island associated with inactivation of *THBS1* was observed in gliomas [109].

A general conclusion that can be drawn from this overview is that both types of events (genetic and epigenetic) contribute to the appearance and progression of human tumors, the relative role of every event varying significantly in each case. This conclusion is illustrated by the suppressor genes *p16* and *p15* shown as examples in Fig. 8.

Acquired features of tumor cells. A relatively small group of molecular, biochemical, and cellular features (acquired features) that are present in the majority (if not all) of tumor cell types has been established during the investigation of cancer. The acquiring by a cell of all of these properties is an essential and apparently sufficient condition for the development of a malignant (metastasizing) tumor. Six principal features of malignant neoplasia (Table 3), which appear in random order, are distinguished [100].

Only transformed cells, which possess all of these features, are capable of evoluting into a metastasizing tumor. The destabilization of the genome is the only situation in which the accumulation of several independent genetic events (4-7) during a relatively short period of time (life span of an individual) is possible.

From this point of view the contribution of epigenetic events in carcinogenesis seems to be most evident for the appearance of such properties as insensitivity to growth signals ("epimutation" of *Rb*), invasion, and metastasizing (inactivation of *E-cad*). However, taking into account that the mechanisms of apoptosis and angiogenesis are under multiple controls, which include positive (activating) as well as negative (inhibiting) factors, it is possible to suggest that the contribution of epigenetic factors to these alterations in a tumor cells is highly significant. Furthermore, available data about the influence of aberrant methylation on the signal pathways of the p53 family and about the destabilizing effect of DNA demethylation indicate that the epigenetic component may considerably contribute to the principal feature of a tumor cell—the instability of the genome.

Aberrant methylation of CpG-islands in a tumor as a large-scale phenomenon. The majority of studies of the methylation of CpG-islands in tumors have been directed towards distinct genes. However, the possibility to analyze

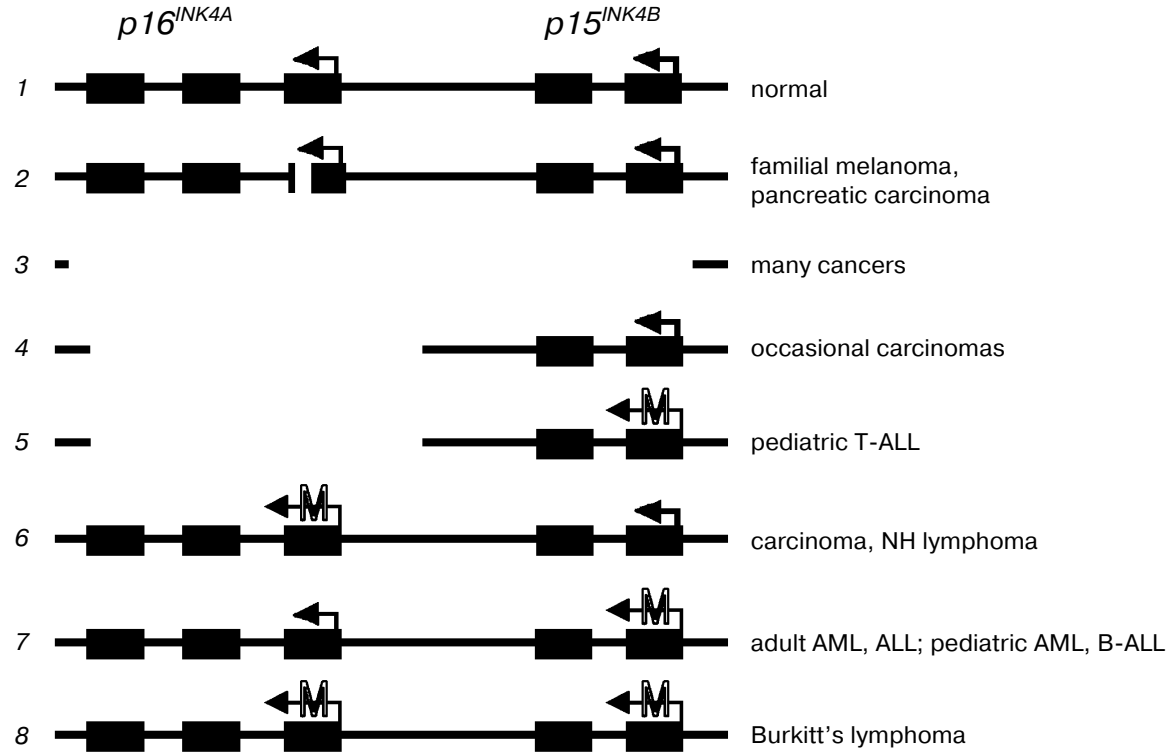

Fig. 8. Contribution of genetic and epigenetic events to the appearance of different malignant human neoplasia (the suppressor genes $p16^{INK4A}$ and $p15^{INK4B}$ are taken as examples) [11]: *1*) norm; *2-4*) genetic events (mutations, deletions); *5*) combination of genetic and epigenetic events (MAGI); *6-8*) MAGI; box, exon; opened box, mutation in the exon; arrow to the left, active gene; arrow with "M", MAGI (methylation-associated gene inactivation) (reprinted by permission from Academic Press).

simultaneously the methylation state of thousands of CpG-islands by the RLGS method (see above) appeared recently. This approach was used to compare ~1200 CpG-islands in 98 human tumors of different origins with corresponding normal tissues [63]. This "bird flight sight" on the general methylation of the genome of normal and tumor cells allowed important conclusions to be drawn. In particular, it was established that methylation of the CpG-islands on one hand is tumor-specific (i.e., some sites are methylated in numerous tumors of various localization and are not methylated normally), and on the other hand has some specificity with regard to different tumor types (i.e., some sites are methylated in some tumors but are not methylated in others).

Table 3. Acquired features of tumor cells [100]

Feature	Molecular mechanism (example)
Self-maintenance with growth signals	activation of H-*ras* oncogene
Insensitivity to anti-growth signals	inactivation of Rb suppressor
Arrest of apoptosis	activation of *Bcl-2*
Unlimited proliferative potency	activation of telomerase
Stimulation of angiogenesis	production of VEGF inducer
Invasion and metastasizing	inactivation of E-cadherin

Estimations show that the average number of CpG-islands that are hypermethylated in the genome of a tumor cell is ~600 (from 0 to 4500) out of the total in the human genome (~45,000), thus a tumor cell should acquire approximately the same number of aberrantly inhibited genes. The absence of any specific difference in the level of methylation between different stages of tumor progression indicates that these alterations appeared in the early stages of tumor development. The scale of detected alterations suggests the existence of a large number of genes that are crucially important for the tumor growth and progression (in addition to already detected ones). The detection of CpG-islands that are hypermethylated in tumors is the first step in identification of the related genes, most of which are not yet identified.

Aging and Cancer

Old age is one of the major risk factors in the development of malignant neoplasia ("cancer is a disease of advanced age"). This conclusion is still true, although the number of cases of the disease in young people is increasing, which is explained by constantly increasing loading with carcinogens. Until recently, this evident dependence has been explained by the necessity of accumulation of a critical amount of genetic defects (sufficient for the appearance of a tumor), which evidently takes a lot of time. Therefore, the latent period for the appearance of the clinical indications of a malignant neoplasia may take years. This explanation is still completely relevant, but, in addition, another fundamental link between aging and cancer is revealed by the analysis of DNA methylation phenomenon.

Physiological aging is accompanied in many tissues by changes in the expression of several genes that cannot be caused by the accumulation of mutations [78, 110]. Since methylation lies at the origin of silencing, is was suggested long ago that methylation participates in the transcriptional changes caused by aging as well as in the appearance of tumors [111-113]. Massive hypermethylation, which was detected in the promoters of suppressor genes recently, confirmed the importance of the epigenetic mechanisms in carcinogenesis and evoked interest in them with regard to aging. The results of relevant investigations can be summarized in several basic statements. First, total genome hypomethylation is observed in aging cells (cultured human fibroblasts) [114]. Second, local hypermethylation occurs in certain cases [115]. Furthermore, the CpG-islands in the promoter regions of several genes are hypermethylated in tumor cells as well as in aging normal cells [95]. Thus, it was established that hypermethylation of the *ER* (estrogen receptor) gene promoter, which is strongly expressed in colon cancer cells, is also observed in cells of normal mucosa surrounding the tumor [105]. Further investigation demonstrated that methylation of the *ER* promoter in the colon mucosa is related to aging, since it is not detected in young persons but it is found and is increasing progressively with the age of an individual [105]. Taking into account that the introduction of the *ER* gene in colon cancer cells leads to significant inhibition of their growth, the authors [105] suggest that aging-related methylation of the promoter of this gene provides a selective advantage to the cell subpopulation with this defect and a prerequsite for the appearance of a tumor.

The situation with an imprinted *IGF2* gene (that encodes for a fetal growth factor) is more complicated and interesting. Only the parental allele of this gene is active in mouse and human cells. This gene is actively expressed in tumors of different origin, including leukemia and colon, breast, and lung cancer. *IGF2* has four different promoters (P1, P2, P3, and P4), and, consequently, four different leader exons. Closely situated P2, P3, and P4 promoters are located within a CpG-island and are coordinately imprinted (i.e., only the maternal allele is selectively methylated). The P1 promoter is located at a significant distance from them (20 kbp upstream from P2) outside of the CpG-island, consequently is not imprinted, and therefore, both its alleles can be expressed. This occurs in adult human liver, where IGF2 is produced most intensively. The methylation pattern of the *IGF2* promoters changes considerably during aging and malignant growth. Thus, it was established that the P2, P3, and P4 promoters, only one allele of which is normally methylated (the maternal one), are progressively methylated with aging on the second allele (i.e., the initially intact paternal allele starts to be subjected to this modification) [116]. As a result, the imprinting is in fact abolished in some tissues during aging, as well as during malignant growth (since both alleles of P2, P3, and P4 promoters are methylated), while the gene is nevertheless actively expressed through the non-imprinted P1 promoter. Thus, a rather particular situation arises: increase in the gene expression as a result of the switching of the active promoter due to hypermethylation.

Another important phenomenon is thought to be related to aging: the spreading of methylation from one allele to another that was not methylated before. This phenomenon, referred to as "co-suppression", was shown in plants, when the methylated transgene mediates the aging-dependent methylation and inactivation of the endogenous allele [117]. The methylated transgene associates with the endogenous gene by homologous recombination, and the DNA-methyltransferase considers the formed complex as semi-methylated DNA like during DNA replication. Thus, according to the hypothesis, the spreading of methylation from one allele to another occurs. The fact that many genes have analogs (pseudogenes) that are not active and are intensively methylated may be also applied for the explanation of aging-associated methylation of several genes. Therefore, the spreading of the methylation onto the active genes may also occur in this manner with the participation of pseudogenes [11].

Aging-associated methylation of CpG-islands was also demonstrated for some other genes—*N33*, *MyoD*, *Versican* [95]. Their hypermethylation, which is frequently observed in colon cancer tissue, was also found in cells of normal mucosa surrounding the tumor and also in fibroblasts and other tissues, the methylation level correlating with the age of the patient. It was also demonstrated that the majority of loci that are hypermethylated in colon carcinoma cells and in bladder and prostate carcinomas are also hypermethylated (although to a lesser extent) in corresponding normal tissues [118]. Similar data are available for breast [119] and lung cancer [120]. These data lead to the suggestion that the increase in aging-related methylation of a gene in various tissues is a usual event that favors the appearance of a tumor.

The reasons of the change in gene methylation with aging remain obscure. It is suggested that the level of the expression of a gene may play a certain role (the higher the expression is, lower is the probability of methylation). At the same time, environmental factors, carcinogens, and the activity of the methylation system itself very probably also influence this process since the rates of DNA methylation in tissues of individuals of the same age differ significantly.

The above-mentioned data distinguish two types of gene methylation: aging-related methylation and methylation that is specific for tumor development [95]. When applied to the colon carcinoma model, the first type suggests methylation of genes in cells of normal mucosa, which increases with the age of an individual. This process favors the appearance of a tumor through inactivation of genes inhibiting cell proliferation and stimulating differentiation (like *ER*, *IGF2*, *MyoD*).

Methylation of the second type concerns a group of tumors, which performs the CIMP phenotype (CpG-island methylator phenotype), i.e., simultaneous methylation of many CpG-islands, and spreads on a group of genes (like *VHL*, *p16*, *hMLH1*) that are almost never methylated in normal tissues. CIMP is a phenotype of molecular instability that is responsible for the majority of cases of inactivation of suppressor genes and has important functional links with other cellular processes. Thus, in addition to the several examples given above, the $CIMP^+$ phenotype may be responsible for many cases of microsatellite instability in sporadic forms of colon carcinoma (Fig. 9). It was demonstrated that the CpG-island of *hMLH1* (one of the genes of reparation of mismatched base pairs) is frequently methylated in this form of cancer (mutations of this gene are generally not detected in these cases). As a result, the damages that appear during DNA replication are not repaired properly, leading to a cascade of changes in microsatellite loci and in other genes. The expression of the $CIMP^+$ phenotype (Fig. 9) is also found in colon cancer as well as in many other forms of human malignant neoplasia. The reasons for the appearance of the $CIMP^+$ phenotype remain unknown, but it can be suggested that the exhaustion of cellular capacities for protection from the expansion of methylation (damage of *cis*-acting factors and/or decrease in the level of *trans*-acting factors) may play a crucial role in the appearance of this phenotype (see above).

The phrase of N. Bohr taken as an epigraph to this review, which referred to the problems of physics in the beginning of the century, appears to be applicable to the biology of the end of the century. Every alternative concept that initially seemed to be contradictory ("cancer is a disease of genes" according to the point of view of genetics and "cancer is a disease of gene regulation" according to "epigenetics") was shown later to adequate-

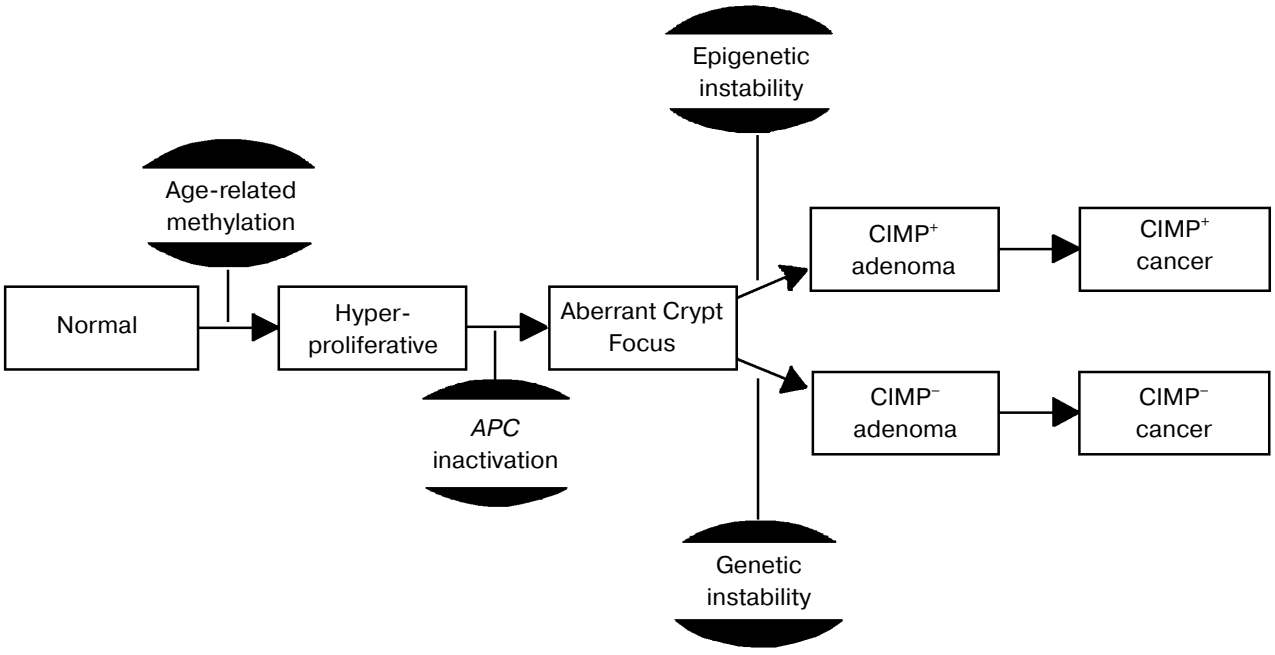

Fig. 9. Succession of genetic and epigenetic events (colon carcinoma is taken as an example) [95]. Several genes undergo methylation in some normal epithelial cells due to aging-related methylation, and these cells thus acquire the capacity for increased proliferation. The gate-keeper gene is inactivated (mutation of *APC* or of *β-catenin*) in some of them, leading to the appearance of an aberrant cryptical focus. These precancerous lesions can afterwards give rise to a tumor by two possible pathways: by acquiring genetic instability mediated by *p53* mutations ($CIMP^-$ phenotype) or through the $CIMP^+$ phenotype, obtaining epigenetic instability (reprinted by permission from Academic Press).

ly reflect only one side of a very complicated process. In contrast to the first, the truthfulness of which became evident more than a quarter of a century ago after the discovery of oncogenes, the understanding of the role of epigenetic mechanisms came only during recent years with the investigation of the DNA methylation phenomenon.

The significance of this shift in the way of thinking cannot be overestimated. On one hand, it became evident that the mechanisms of carcinogenesis are much more complicated than had been suggested before, and that the pathways leading to this pathology are much more variable. On the other hand, it is now possible to answer difficult questions that were apparent before, in particular, about the possible existence of non-genotoxic (i.e., not DNA damaging) carcinogens. Apparently, one can get a positive answer to this question. Indeed, not only direct changes in DNA structure but certain other factors related to cellular metabolism as well can lead to inactivation of suppressor genes and DNA repair genes, thus favoring malignant transformation of a cell. An old observation that feeding experimental animals a diet without methionine (resulting in a deficit of methyl group donors) gives rise to liver tumors supports this conclusion. Taking into account the metabolic links in the cell ("everything is linked with everything"), it can be suggested that many other metabolites may influence the DNA methylation system indirectly.

It is clearly evident that the relative contribution of genetic and epigenetic factors in various types of tumors varies considerably. It should be mentioned in this context that epigenetic regulation consists not only of DNA methylation, but of other mechanisms and chromatin components as well, in particular of the proteins with chromodomain [121]. The possibility that this mechanism can participate in the pathogenesis of malignant transformation of a cell should not be excluded.

The change in DNA methylation contributes significantly to the genetic component of carcinogenesis due to the hypermutability of the 5-methylcytosine (up to a third of *p53* mutations are caused in this way) and due to genome destabilization that is caused by general demethylation. These data provide additional evidence for the role of spontaneous mutagenesis, which occurs in cells constantly and independently of the presence of mutagens in the environment.

In contrast to mutations, which cannot be reversed, the DNA modifications, although rather stable, can be reversed. 5-Azacytidine possesses this capacity, suggesting prospects for clinical applications of its analogs that are without side effects. An effect in therapy should be expected for malignant neoplasia with a significant contribution of the epigenetic component. Such drugs are now being designed.

Intensive studies of DNA methylation show in perspective some other clinical applications of this phenomenon, in particular as genetic markers. Thus, the identification of DNA methylation pattern (for example, using DNA "chips") could be applied for early diagnostics, monitoring, and prognosis of tumors.

REFERENCES

1. Hotchkiss, R. D. (1948) *J. Biol. Chem.*, **175**, 315-332.
2. Wyatt, G. R. (1951) *Biochem. J.*, **48**, 581-584.
3. Shapiro, H. S., and Chargaff, E. (1960) *Biochim. Biophys. Acta*, **39**, 68-82.
4. Doscocil, J., and Sormova, Z. (1965) *Biochim. Biophys. Acta*, **95**, 513-515.
5. Vanyushin, B. F., Tkacheva, S. G., and Belozersky, A. N. (1970) *Nature*, **225**, 948-949.
6. Vanyushin, B. F., Belozersky, A. N., Kokurina, N. A., and Kadirova, D. X. (1968) *Nature*, **218**, 1066-1067.
7. Mazin, A. L., Mukhovatova, L. M., Shuppe, H. G., and Vanyushin, B. F. (1984) *Dokl. AN SSSR*, **276**, 760-762.
8. Mazin, A. L., Gimadutdinov, O. A., Antoniv, T. T., Mukhovatova, L. M., Shuppe, H. G., and Vanyushin, B. F. (1989) *Izv. AN SSSR*, **6**, 922-925.
9. Doerfler, W. (1983) *Ann. Rev. Biochem.*, **52**, 93-124.
10. Bird, A. P. (1995) *Trends Genet.*, **11**, 94-100.
11. Baylin, S. B., Herman, J. G., Graff, J. R., Vertino, P. M., and Issa, J.-P. (1998) *Adv. Cancer Res.*, **72**, 141-196.
12. Noyer-Weidner, M., and Trautner, T. A. (1993) in *"DNA Methylation: Molecular Biology and Biological Significance"* (Jost, J. P., and Saluz, H. P., eds.) Birkhauser Verlag, Basel, pp. 39-108.
13. Bestor, T. H. (1990) *Philos. Trans. Roy. Soc.* (London), **326**, 179-187.
14. Selker, E. U., Fritz, D. Y., and Singer, M. J. (1993) *Science*, **262**, 1724-1728.
15. Doerfler, W. (1993) in *"DNA Methylation: Molecular Biology and Biological Significance"* (Jost, J. P., and Saluz, H. P., eds.) Birkhauser Verlag, Basel, pp. 262-299.
16. Sasaki, H., Allen, N. D., and Surani, M. A. (1993) in *"DNA Methylation: Molecular Biology and Biological Significance"* (Jost, J. P., and Saluz, H. P., eds.) Birkhauser Verlag, Basel, pp. 469-486.
17. Jablonka, E., and Regev, A. (1995) *Trends Genet.*, **11**, 383-384.
18. Selker, E. U. (1999) *Cell*, **97**, 157-160.
19. Goldman, M. A., Holmquist, G. P., Gray, M. C., Caston, L. A., and Nag, A. (1984) *Science*, **224**, 686-692.
20. Hatton, K. S., and Schidkraut, C. L. (1988) *Mol. Cell Biol.*, **8**, 2149-2158.
21. Schmidt, M., and Migeon, B. R. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 3685-3689.
22. Razin, A., and Shemer, R. (1995) *Hum. Mol. Genet.*, **4**, 1751-1755.
23. Jost, J. P., and Jost, Y. C. (1995) *Gene*, **157**, 265-266.
24. Jost, J. P. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 4684-4688.
25. Jost, J. P., Siegmann, M., Sun, L., and Leung, R. (1995) *J. Biol. Chem.*, **270**, 9734-9739.
26. Bhattacharya, S. K., Ramchandani, S., Cervoni, N., and Szyf, M. (1999) *Nature*, **397**, 579-583.
27. Rougier, N. (1998) *Genes Dev.*, **12**, 2108-2113.
28. Li, E., Beard, C., and Jaenisch, R. (1993) *Nature*, **366**, 362-364.
29. Surani, M. A. (1993) *Nature*, **366**, 302-303.
30. Rigs, A. D. (1975) *Cytogenet. Cell Genet.*, **14**, 9-25.

31. Herman, J. G. (1999) *Sem. Cancer Biol.*, **9**, 359-367.
32. Gardiner-Garden, M., and Frommer, M. (1987) *J. Mol. Biol.*, **196**, 261-268.
33. Mazin, A. L., and Vanyushin, B. F. (1987) *Mol. Biol. (Moscow)*, **21**, 543-551.
34. Mazin, A. L., and Vanyushin, B. F. (1987) *Mol. Biol. (Moscow)*, **21**, 5552-5562.
35. Turker, M. S. (1999) *Sem. Cancer Biol.*, **9**, 329-337.
36. Hattori, M., Fujiyama, A., Tailor, T. D., et al. (2000) *Nature*, **405**, 311-328.
37. Dunham, I., Hunt, J. E., Collins, J. E., et al. (1999) *Nature*, **402**, 489-495.
38. Antequera, F., and Bird, A. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 11995-11999.
39. Melwin, M. J., McGurn, M. E., and Bort, S. J. (1995) *Eur. J. Immunol.*, **25**, 426-430.
40. Iguchi-Ariga, S. M., and Schaffner, W. (1989) *Genes Dev.*, **3**, 612-619.
41. Walsh, C. P., and Bestor, T. H. (1999) *Genes Dev.*, **13**, 26-34.
42. Toth, M., Lichtenberg, U., and Doerfler, W. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 3728-3732.
43. Turker, M. S., Mummaneni, P., and Bishop, P. L. (1991) *Somat. Cell Mol. Genet.*, **17**, 151-157.
44. Bestor, T. H., Laudano, A. P., Mattaciano, R., and Ingram, V. M. (1983) *J. Mol. Biol.*, **203**, 971-983.
45. Li, E., Bestor, T. H., and Jaenisch, R. (1992) *Cell*, **69**, 915-926.
46. Bird, A. P. (1992) *Cell*, **70**, 5-8.
47. Bird, A. P. (1986) *Nature*, **321**, 209-213.
48. Creusot, F., Acs, G., and Christman, J. K. (1982) *J. Biol. Chem.*, **257**, 2041-2048.
49. Bestor, T. H., and Verdine, G. L. (1994) *Curr. Opin. Cell Biol.*, **259**, 946-951.
50. MacLeod, A. R., Rouleou, J., and Szyf, M. (1995) *J. Biol. Chem.*, **270**, 11327-11337.
51. Chuang, L. S. H., Ian, H. I., Koh, T. W., Ng, H. H., Xu, G., and Li, B. F. L. (1997) *Science*, **277**, 1996-2000.
52. Vanyushin, B. F. (1984) *Curr. Topics Microbiol. Immunol.*, **108**, 99-114.
53. Lei, H. (1996) *Development*, **122**, 3195-3205.
54. Okano, M., Xie, S., and Li, E. (1998) *Nature Genet.*, **19**, 219-220.
55. Okano, M., Bell, D. W., Haber, D. A., and Li, E. (1999) *Cell*, **99**, 247-257.
56. Xu, G.-L., Bestor, T. H., Bourc'his, D., Hsiehi, C.-L., Tommerup, N., Bugge, M., Hulten, M., Qu, X., Russo, J. J., and Viegas-Pequignot, E. (1999) *Nature*, **402**, 187-191.
57. Okano, M., Xie, S., and Li, E. (1998) *Nucleic Acids Res.*, **26**, 2536-2540.
58. Herman, J. G., Graff, J. R., Myohanen, S., Nelkin, B. D., and Baylin, S. B. (1996) *Proc. Natl. Acad. Sci. USA*, **93**, 9821-9826.
59. Clark, S. J., Harrison, J., Paul, C. L., and Frommer, M. (1994) *Nucleic Acids Res.*, **22**, 2290-2297.
60. Warnecke, P. M., Mann, J. R., Frommer, M., and Clark, S. J. (1998) *Genomics*, **51**, 182-190.
61. Cote, S., Sinnett, D., and Monparler, R. (1998) *Anti-Cancer Drugs*, **9**, 743-759.
62. Toyota, M., Ho, C., Ahuja, N., Jair, K.-W., Li, Q., Ohe-Toyota, M., Baylin, S., and Issa, J.-P. J. (1999) *Cancer Res.*, **59**, 2307-2312.
63. Lisitsyn, N., Lisitsyn, N., and Wigler, M. (1993) *Science*, **259**, 946-951.
64. Eng, C., Herman, J. G., and Baylin, S. B. (2000) *Nature Genet.*, **24**, 101-102.
65. Costello, J. F., Fruhwald, M. C., Smiraglia, D. J., Rush, L. J., Robertson, G. P., Gao, X., Wright, F. A., Feramisco, J. D., Peltomaki, P., Lang, J. C., Schuller, D. E., Yu, L., Bloomfield, C. D., Caligiuri, M. A., Yates, A., Nishikawa, R., Huang, H.-J. S., Petrelli, N. J., Zhang, X., O'Dorisio, M. S., Held, W. A., Cavenee, W. K., and Plass, C. (2000) *Nature Genet.*, **25**, 132-138.
66. Tazi, J., and Bird, A. (1990) *Cell*, **60**, 909-920.
67. Meehan, R. R., Lewis, J. D., McKay, S., Kleiner, E. L., and Bird, A. P. (1989) *Cell*, **58**, 499-507.
68. Meehan, R. R., Lewis, J. D., and Bird, A. P. (1992) *Nucleic Acids Res.*, **20**, 5085-5092.
69. Lewis, J. D., Meehan, R. R., Henzel, W. J., Maurer-Fogy, I., Jeppensen, P., Klein, G., and Bird, A. P. (1992) *Cell*, **69**, 905-914.
70. Jones, P. L., and Wolffe, A. P. (1999) *Sem. Cancer Biol.*, **9**, 339-347.
71. Archer, S. Y., and Hodin, R. A. (1999) *Curr. Opin. Gen. Dev.*, **9**, 171-174.
72. Ng, H.-H., and Bird, A. (1999) *Curr. Opin. Gen. Dev.*, **9**, 158-163.
73. Chen, Z. J., and Pikaard, C. S. (1997) *Genes Dev.*, **11**, 2124-2136.
74. Ammerpohl, A., Schmitz, A., Steinmuller, L., and Renkawitz, R. (1998) *Nucleic Acids Res.*, **26**, 5256-5260.
75. McInemey, E. M., Tsai, M. J., O'Malley, B. W., and Katzenellenbogen, B. S. (1996) *Proc. Natl. Acad. Sci. USA*, **93**, 10069-10073.
76. Spencer, T. E., Jenster, G., Burcin, M. M., Allis, C. D., Zhou, J., Mizzen, C. A., McKenna, N. J., Onate, S. A., Tsai, S. Y., Tsai, M. J., and O'Malley, B. W. (1997) *Nature*, **389**, 194-198.
77. Matsuo, K., Silke, J., Georgiev, O., Marti, P., Giovannini, N., and Rungger, D. (1998) *EMBO J.*, **17**, 1446-1453.
78. Holliday, R. (1991) *Mutat. Res.*, **250**, 351-363.
79. Antequera, F., Boyes, J., and Bird, A. (1990) *Cell*, **62**, 503-514.
80. Chen, R. Z., Pettersson, U., Beard, C., Jackson-Grusby, L., and Jaenisch, R. (1998) *Nature*, **395**, 89-93.
81. Antequera, F., and Bird, A. (1993) in *DNA Methylation: Molecular Biology and Biological Significance* (Jost, J. P., and Saluz, H. P., eds.) Birkhauser Verlag, Basel, pp. 169-185.
82. Mazin, A. L., Gimadutdinov, O. A., Turkin, S. I., Burtseva, N. N., and Vanyushin, B. F. (1985) *Mol. Biol. (Moscow)*, **19**, 903-914.
83. Maloisel, L., and Rossignol, J. L. (1998) *Genes Devel.*, **12**, 1381- 1389.
84. Jones, P. A. (1996) *Cancer Res.*, **56**, 2463-2467.
85. Laird, P. W., and Jaenisch, R. (1994) *Hum. Mol. Genet.*, **3**, 1487-1495.
86. Gama-Sosa, M. A., Slagel, V. A., and Trewyn, R. W. (1983) *Nucleic Acids Res.*, **11**, 6883-6894.
87. Pogribny, I. P., Poirier, L. A., and James, S. J. (1995) *Carcinogenesis*, **16**, 2863-2867.
88. Wainfan, E., and Poirier, L. A. (1992) *Cancer Res.*, **52**, 2071-2077.
89. Feinberg, A. P., and Vogelstein, B. (1983) *Nature*, **301**, 89-92.
90. Goelz, S. E., Vogelstein, B., and Hamilton, S. R. (1985) *Science*, **228**, 187-190.

91. Rainier, S., and Feinberg, A. P. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 6384-6388.
92. Walker, C., and Nettesheim, P. (1986) *Cancer Res.*, **46**, 6433-6437.
93. El-Deiry, W. S., Nelkin, B. D., and Celano, P. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 3470-3474.
94. Bakin, A. V., and Curran, T. (1999) *Science*, **283**, 387-390.
95. Toyota, M., and Issa, J.-P. J. (1999) *Sem. Cancer Biol.*, **9**, 349-357.
96. Vertino, P. M., Yen, R. W. C., Gao, J., and Baylin, S. D. (1996) *Mol. Cell. Biol.*, **16**, 4555-4565.
97. Graff, J. R., Herman, J. G., Myohanen, S., Baylin, S. B., and Vertino, P. M. (1997) *J. Biol. Chem.*, **272**, 22322-22329.
98. Baylin, S. B., Hoppener, J. W., de Bustros, A., Steenbergh, P. H., Lips, C. J., and Nelkin, B. D. (1986) *Cancer Res.*, **46**, 2917-2922.
99. Baylin, S. B., Fearon, E. R., and Vogelstein, B. (1987) *Blood*, **70**, 412-417.
100. Hanahan, D., and Weinberg, R. A. (2000) *Cell*, **100**, 57-70.
101. Robertson, K. D., and Jones, P. A. (1998) *Mol. Cell Biol.*, **18**, 6457-6473.
102. Christofori, G., and Semb, H. (1999) *Trends Biochem. Sci.*, **24**, 73-76.
103. Graff, J. R., Herman, J. G., Lapidus, R. G., Chopra, H., Xu, R., Jarrard, D. F., Isaacs, W. B., Pitha, P. M., and Davidson, N. E. (1995) *Cancer Res.*, **55**, 5195-5199.
104. Ferguson, A. T., Lapidus, R. G., and Baylin, S. B. (1995) *Cancer Res.*, **55**, 2279-2283.
105. Issa, J. P., Ottaviano, Y. L., Celano, P., Hamilton, S. R., Davidson, N. E., and Baylin, S. B. (1994) *Nature Genet.*, **7**, 536-540.
106. Issa, J. P., Zehnbauer, B. A., Civin, C. I., Collector, M. I., Sharkis, S. J., Davidson, N. E., Kaufmann, S. H., and Baylin, S. B. (1996) *Cancer Res.*, **56**, 973-977.
107. Lee, W. H., Morton, R. A., Epstein, J. I., Brooks, J. D., Campbell, P. A., Bova, G. S., Hsieh, W. S., Isaacs, W. B., and Nelson, W. G. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 11733-11737.
108. Toyota, M., Ho, C., Ohe-Toyota, M., Baylin, S., and Issa, J.-P. J. (1999) *Cancer Res.*, **59**, 4535-4541.
109. Li, Q., Ahuja, N., Burger, P. C., and Issa, J. P. (1999) *Oncogene*, **18**, 3284-3289.
110. Smith, J. R., and Pereira-Smith, O. M. (1996) *Science*, **273**, 63-67.
111. Vanyushin, B. F., Nemirovsky, L. E., Klimenko, V. V., Vasiliev, V. K., and Belozersky, A. N. (1973) *Gerontologia* (Basel), **19**, 138-152.
112. Vanyushin, B. F., and Berdyshev, G. D. (1977) *Molecular–Genetic Mechanisms of Aging* [in Russian], Meditsina, Moscow.
113. Holliday, R. (1987) *Science*, **238**, 163-170.
114. Holliday, R. (1987) *Exp. Cell. Res.*, **166**, 543-552.
115. Swisshelm, K., Disteche, C. M., and Thorvaldsen, J. (1990) *Mutat. Res.*, **237**, 131-146.
116. Issa, J. P. J., Vertino, P. M., Boehm, C. D., Newsham, I. F., and Baylin, S. B. (1996) *Proc. Natl. Acad. Sci. USA*, **93**, 11757-11762.
117. Flavell, R. B. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 3490-3496.
118. Liang, G., Salem, C. E., Yu, M. C., Nguyen, H. D., Gonzales, F. A., Nguyen, T. T., Nichols, P. W., and Jones, P. A. (1998) *Genomics*, **53**, 260-268.
119. Huang, T. H. M., Perry, M. R., and Laux, D. E. (1999) *Hum. Mol. Genet.*, **8**, 459-470.
120. Shiraishi, M., Chuu, Y. H., and Sekiya, T. (1999) *Proc. Natl. Acad. Sci. USA*, **96**, 2913-2918.
121. Cavalli, G., and Paro, R. (1998) *Curr. Opin. Cell Biol.*, **10**, 354-360.