REVIEW ARTICLE
UDC: 616-006.4:577.27:576.385.5

Arch Oncol 2005;13(3-4):108-14.
DOI:10.2298/AOO0503108D

Cancer epigenome: A review

Bogomir Dimitrijević

Institute for Nuclear Sciences "Vinča", Laboratory for Molecular Genetics and Radiobiology, Belgrade, Serbia & Montenegro, Address correspondence to: Bogomir Dimitrijević; Institute for Nuclear Sciences "Vinča", Laboratory for Molecular Genetics and Radiobiology, Mike Alasa 14, POB 522, 11000 Belgrade, Serbia & Montenegro E-mail: bogomir@beotel.yu, The manuscript was received: 15. 09. 2005, Accepted for publication: 11. 10. 2005

© 2005, Institute of Oncology Sremska Kamenica, Serbia & Montenegro

ABSTRACT

Cancer genomics that normally relies on mutational analysis of oncogenes and tumor suppressor genes has approached its inherent limits. This was not much of a surprise having in mind the genome dynamics and the resulting complexity of cancer phenotype and genotype. In response to this challenge, molecular genetics offered a new armory for the analysis of the genetic basis of cancer. This refers to the analysis of molecular features that regulate gene activity and the analysis of products of this activity. In the focus of tuning transcription is the methylation surveillance of the genome of the cell. Modification of proteins associated with chromatin and methylation of CpG sites in DNA was found to affect profoundly gene expression and is commonly termed epigenomics. Quantitative and qualitative characterization of the methylation profile of the cancer cell genome is formidable but necessary task with great potential for molecular pathology of cancer. There is little doubt that this line of research will add a great deal to the clinical practice and the basic science of oncology. The only question is how to make a large database large enough and how select the most reliable and sensitive technological approach with the highest throughput.

KEY WORDS: *Neoplasms; Genome, Human*

INTRODUCTION

We are accustomed to the idea that the coding potential of the genome lies within the arrangement of the four bases adenine, cytosine, guanine and thymine; however, additional information that affects phenotype is stored in the distribution of the modified base 5-methylcytosine. This form of information storage is flexible enough to be adapted for different somatic cell types, yet is stable enough to be retained during mitosis and/or meiosis. It is a modification of the genome, as opposed to being part of the genome, so is known as "epigenetics" (Greek for "upon" genetics). Dense methylation of promoter regions is associated with a compacted chromatin structure, and accompanying transcriptional silencing of the affiliated gene. In recent years, it has become apparent that the transcriptional silencing that is associated with 5-methylcytosine is important in mammalian development, protection against intragenomic parasites, genomic imprinting, x-inactivation, mental health, aging and cancer.

Cytosine methylation occurs after DNA synthesis, by enzymatic transfer of a methyl group from the methyl donor S-adenosylmethionine to the carbon-5 position on cytosine. The enzymatic reaction is performed by one of a family of dedicated enzymes called DNA methyltransferases (DNMTs). DNMT1 is the main enzyme in mammals, and is responsible for the post-replicative restoration of hemi-methylated sites to full methylation, referred to as maintenance methylation, whereas DNMT3A and DNMT3B are thought to be involved primarily in methylating new sites, a process called *de novo* methylation. The predominant sequence recognition motif for mammalian DNA methyltransferases is 5'-CpG-3', although non-CpG methylation in mammals has also been reported. CpG is the only dinucleotide to be severely under-represented in the human genome, and this is thought to be due to the high rate of methylcytosine-to-thymine transition mutations. The remaining CpG dinucleotides are unequally distributed across the human genome - vast stretches of sequence are deficient for CpGs, and these are interspersed by CpG clusters called CpG.

CpG islands were traditionally thought to be unmethylated in normal cells, with the exception of those that are associated with imprinted genes and genes on the inactive X chromosome. It now seems that some non-imprinted autosomal CpG islands are methylated in normal cells, and might even use this mechanism for the control of gene expression. Nevertheless, most methylated cytosine residues are found in CpG dinucleotides that are located outside of CpG islands, primarily in repetitive sequences. Methylation of some CpG islands in non-malignant tissues increases with age, but the total genomic content of 5-methylcytosine declines. These opposing events are also present, but are much more pronounced, in cancer cells. Cancer-specific DNA methylation changes at individual gene loci have so far focused primarily on hypermethylation of CpG islands. Cancer-specific DNA hypomethylation events at individual unique sequences remain largely unexplored. Global 5-methylcytosine content is influenced by the nutritional availability of folate and by polymorphisms in folate metabolic enzymes.

The past few years have seen a tremendous advance in our understanding of the functional consequences of DNA methylation and its interaction with chromatin structure and the transcriptional machinery. We have also obtained some first insights into what causes DNA methylation patterns to undergo changes in cancer cells, although this fundamental process remains, for the most part, an enigma. From a clinical perspective, DNA methylation

changes in cancer represent an attractive therapeutic target, as epigenetic alterations are, in principle, more readily reversible than genetic events. However, the great strength of DNA methylation in the clinic promises to be in the area of molecular diagnostics and early detection.

THE "-OMICS" REVOLUTION

Cancer epigenome and transcriptome emerge from frontline research areas referred to as epigenomics and transcriptomics. The syllable "omics" is not only a linguistic curiosity but deserves scientific clarification. Attaching the "omics" label to one's discipline increases visibility, opens avenues for funding, and a willingness to consider problems at level of the whole cell, or whole organism. This is a view of a cynic who believes that this terminology is nothing but a glamour arising from human genome euphoria. However, most scientists disagree or ignore this malice.

"Omics" is a general term used to describe several rapidly growing fields of scientific endeavor, the best-known member of which is genomics (1). Genomics is the study of a genome, the complete genetic complement of an individual or species, rather than the study of single genes. The suffix-omics generally refers to the study of a complete set of biological molecules. Just as genomics is the study of an organism's genome, proteomics is the study of an organism's entire complement of proteins and metabolomics is the study of the complete set of low-molecular weight metabolites present in a cell or organism at any one time (2,3). The word "omics" to twentieth century pop culture brings to mind genomics, specifically the human genome project, is what comes to mind. The science of omics is much more diverse, however. Omics includes the sub classifications of genomics, transcriptomics, proteomics, phenomics, lipidomics, epigenomics, ligandomics, proteasomics, and other biochemically important sub classifications. These omic technologies allow for a large number of endpoints to be simultaneously measured on biological samples from human and animal subjects (4,5). The vast amount of generated data generates an unforeseen panoramic view of the biological complexity of malignant phenotype.

CANCER EPIGENOME

EPIGENETIC CHARACTERISTICS OF DNA

The double helical structure of DNA is remarkable in its simplicity, yet correct gene function requires not only that the base sequence is faithfully transcribed, but also that expression is both spatially and temporally regulated in a tightly controlled manner. The process of development in multi cellular organisms depends on the differential repression or activation of particular genes in a cell type-specific manner, and this programming information must be maintained throughout the life of the individual. This is referred to as "cellular memory" and is controlled by epigenetic mechanisms (6). Epigenetic regulation is therefore fundamentally important to the control of gene expression. Two major classes of epigenetic modification are instrumental in determining this complex level of gene regulation: histone acetylation and cytosine methylation.

Chromatin structure

The human genome consists of approximately three billion base pairs of DNA divided amongst 22 pairs of autosomes and two sex chromosomes. If left in its native form, the genome would be several meters in length. The problem of packaging and organizing the DNA within the nucleus is overcome by the coordinated compaction of the genome by specialized DNA-binding proteins, including the histone proteins. Chromatin refers to the combination of DNA with these proteins and is organized into two major types of sub chromosomal domains. Heterochromatin is tightly compacted and less transcriptionally active while euchromatin is less compacted and more likely to be transcribed. The chromatin structure also varies with the phases of the cell cycle. It is relaxed and transcriptionally active during interphase whereas it is condensed and inactive during mitosis.

Histones are a family of proteins that provide the scaffolding for chromatin assembly and, consistent with this function, are among the most highly conserved proteins throughout evolution. The histones share the same basic structure consisting of a globular head and a positively charged, nonglobular tail. Histones H2A, H2B, H3 and H4 together form a hetero-octamer core around which approximately 150 base pairs of DNA are wrapped to form discrete units called nucleosomes. Inter-nucleosomal segments are bound by histone H1 and link each nucleosome to the next. The nucleosomes are further bundled into higher order structures to form compacted and organized chromatin. Histone-DNA interactions are modulated in part by reversible acetylation of the ε-amino groups of the histones lysine side chains. The lysine residues carry positive charges thought to contribute to the histone's affinity for negatively charged DNA. Acetylation of these lysine side chains neutralizes the positive charge thereby decreasing the ability of histones to interact with DNA, resulting in a more open chromatin configuration and increased transcription. Acetylation of histones by histone acetyltransferase (HAT) is reversed by the action of histone deacetylase (HDAC). The latter enzyme reduces transcriptional activity by promoting chromatin condensation and inhibiting access of the transcription machinery to the DNA. Histone modifications at distinct lysine residues may also allow for the recruitment of proteins capable of regulating transcription in a gene-specific manner. The combinatorial nature of histone amino-terminal modifications reveals a "histone code" that considerably extends the information potential of the genetic code.

CpG methylation

Genome structure is influenced by cytosine methylation, the only known biological base modification of DNA. Indeed, methylated cytosine has been referred to as the "fifth base" because of its ability to convey heritable information. The extent of cytosine methylation plays a major role in the organization of the genomic DNA. Densely methylated DNA is located in condensed heterochromatin while sparsely methylated DNA is located in the more relaxed euchromatin. Methylation is the best-studied epigenetic modification that occurs in cancer, and will be the focus of this text. The genomic methylation patterns in gametic DNA are erased by a genome-wide demethylation shortly after fertilization. This is followed by de novo re-establishment of the methylation patterns after implantation. The mechanisms involved and the proteins guiding the erasure and resetting of the methylation patterns in the genome during embryogenesis are not yet fully understood. Several known DNA methyltransferases (DNMTs) in mammals, including DNMT1, DNMT3a, and DNMT3b, catalyze the transfer of a methyl group from S-adenosylmethionine (SAM) to the 5-carbon position of cytosine. This reaction occurs most commonly when cytosine is in a CpG dinucleotide sequence. All three enzymes are capable of performing both de novo and maintenance methylation, but DNMT3a and DNMT3b appear to act by transferring methyl groups to previously unmodified CpG sequences, a process that occurs predominantly during embryogenesis.

In contrast, DNMT1 is thought to be the major maintenance DNA methyltransferase enzyme. The palindromic nature of the CpG target of DNMT1 is a key feature in the heritability of the DNA methylation profile during cell replication. During DNA synthesis, the nascent daughter strand is methylated by the methyltransferase enzyme positioned at the replication fork. DNMT1 preferentially recognizes the hemimethylated state of the two strands, and copies the methylation pattern of the parent strand on to the daughter strand. Thus, somatic cell methylation profiles represent epigenetic information that is faithfully replicated from one generation to the next. The incidence of CpG dinucleotides in the genome is about 5 to 10-

fold less than the approximate 6% frequency expected from the random distribution of the 16 possible dinucleotide combinations. The mammalian genome is thought to have progressively lost many of the methylated cytosines within CpG dinucleotide pairs during the course of evolution. This most likely occurred by endogenous deamination of methylated cytosines to form thymine. Although 60-80% of the CpGs within the mammalian genome are methylated, cytosines that reside in CpG islands, ranging from 200 to 4000 bases in length, are protected from methylation. They are therefore resistant to mutation by methyl group-driven deamination. These CpG islands are located near the promoter regions of approximately 50-60% of the genes within the genome, including all housekeeping genes. It is not unusual to find CpG islands also located within the 5' coding region of genes or even in downstream introns.

The biological function of CpG methylation is not clearly understood. Methylation of genomic DNA is a modification employed by numerous species, including bacteria, plants, and mammals; however, methylation is not detectable in yeast, Drosophila or Caenorhabditis elegans. Methylation serves as a host defense mechanism in prokaryotes to protect against the introduction of foreign DNA. DNA methylation in eukaryotes is proposed to serve in host defense by protecting cells from transcription and transposition of endogenous retroviral sequences, and/or to reduce transcriptional "noise" from very large genomes (7,8). The CpG islands associated with many of the genes located on the inactive X chromosome in females are also extensively methylated whereas the same alleles on the active X chromosome are unmethylated. Furthermore, the silenced allele of imprinted genes usually exhibits parent of origin-dependent dense methylation of at least one associated CpG island.

Hypermethylation of promoter region CpG dinucleotides is strongly correlated with the transcriptional silencing of genes. The causal relationship between cytosine methylation and gene silencing in mammals is supported by studies both in vitro and in vivo. Transfection experiments using reporter constructs with a methylated promoter region show reduced transcription relative to that for constructs with unmethylated promoters. The DNMT inhibitor 5-azacytidine (5-azaC) causes transcriptional reactivation of endogenous genes with hypermethylated promoters. Furthermore, homozygous disruption of DNMT in mice results in a three-fold reduction in genomic 5-methylcytosine content, embryonic death (9), and biallelic expression of imprinted genes (10). These results emphasize the importance of cytosine methylation in gene regulation and embryogenesis. The deviations from normal methylation patterns frequently observed in cancer cells further suggest that epigenetic perturbations are mechanistically involved in carcinogenesis.

Epigenetic characteristics of tumor cells

There is widespread documentation of significant changes in the epigenome of cancer cells. These changes include an overall level of genomic hypomethylation coupled with gene-specific hypermethylation. Hypomethylation events are more generalized, and could lead to the activation of endogenous retroviral elements and dormant proto-oncogenes. However, gene-specific hypomethylation is unlikely to play a major role in oncogenesis since promoter CpG islands are normally unmethylated, with the notable exceptions of imprinted alleles and genes on the inactive X chromosome. In contrast, promoter-specific hypermethylation can lead to the silencing of tumor-suppressor genes.

Inherent mutability of methylated cytosine

The methylation of CpG dinucleotides creates mutagenic susceptibility targets that can subsequently undergo endogenous deamination to form TpG (CpA on the opposite strand) dinucleotide pairs. In contrast to cytosine deamination that results in DNA containing uracil, 5-methylcytosine deamination creates a C to T base substitution that is not as readily recognized by DNA repair proteins as being misplaced in the DNA strand. This contributes to inefficient repair of these lesions, and subsequent accumulation of this type of mutation in the genome. Deamination of 5-methylcytosine in p53 and HPRT is a frequent mutational event associated with human cancers (11). Although the spectrum of mutations within p53 varies between different forms of cancer, 50% of all point mutations in colon cancer arise from transition mutations of normally methylated CpG dinucleotides. In contrast, only 10% of liver and lung cancers contain these same mutations. In addition to the potential direct mutational inactivation of a gene by a CpG to TpG transition, these mutated sequences could hinder the interaction between DNA and specific proteins involved in transcriptional regulation.

Alterations in CpG island methylation

The genome regions subject to hypermethylation in cancer cells are the CpG islands associated with gene promoters. A study undertaken to identify all differentially methylated CpG islands in cancer estimated that of the 45 000 CpG islands in the human genome, 600 exhibit methylation patterns in tumors different from those in normal tissues (12). These methylation changes appear to occur early in the neoplastic process, and some are even cancer-type specific, suggesting that CpG island hypermethylation is mechanistically involved in carcinogenesis rather than being a consequence of neoplastic transformation. An increase in DNA methylation also occurs with ageing. Consequently, the increased cancer predisposition observed with ageing may be partially attributable to the age-dependent increases in genome methylation.

A tumor-suppressor gene inactivated by CpG island hypermethylation would have a number of important characteristics. These include:

1. Dense methylation of the normally unmethylated CpG island present in the promoter region
2. An absence of coding region mutations in the tumor
3. A deficiency of gene-specific transcripts in the tumor
4. Gene reactivation in the tumor with DNMT inhibitors (e.g. 5-azaC)
5. Loss of gene function from hypermethylation comparable to that seen for inactivating mutations (13)

It is important to note that alterations in CpG methylation are mitotically heritable, and can potentially result in the clonal expansion of neoplastic cells if the epigenetic changes provide a selective growth advantage.

The normally unmethylated promoter for many genes is increasingly methylated during neoplastic progression. Table 1 lists genes whose promoter regions are hypermethylated in cancer. This list is certainly not final. An extensive list of candidate tumor-suppressor and other genes are waiting to be epigenetically linked with cancer.

Imprinted genes as cancer susceptibility loci

Genomically imprinted genes provide strong evidence that transcriptional silencing is a result of DNA methylation. This subset of genes is normally mono-allelically expressed in a parent of origin-dependent manner. Imprinted gene expression in somatic cells depends upon the sex of the parent from which the allele originated, but not the sex of the individual. Every imprinted gene thus far examined has been associated with at least one differentially methylated CpG island. The epigenetic changes that confer the heritable imprint mark have not yet been unambiguously defined, but cytosine methylation is the strongest candidate. Histone acetylation has also been proposed to he mechanistically involved in imprinted gene regulation since parental-specific differences in acetylation are associated with imprinted genes (14). Approximately 40 imprinted genes have been identified in

humans to date, and estimates predict the presence of 100-200 imprinted genes in the entire genome (15). The normal silencing of the imprinted allele is equivalent to a first 'hit' in the Knudson's "two-hit hypothesis" for carcinogenesis. A single genetic or epigenetic alteration in the expressed allele could therefore completely abrogate function of an imprinted gene. Since most imprinted genes are involved in cell growth and all are functionally haploid, they represent unique susceptibility loci for cancer development.

Tabela 1. Tumor suppressor genes subject to hypermethylation cancer

Gene	Locus	Function	Tumor type(s)
APC	5q21	Regulation of 6-catenin; cell adhesion	Colorectal, gastrointestinal
BRCA1	17q21	DNA damage repair	Breast, ovarian
CDH1 (E-cadherin)	16q22.1	Homotypic epithelial cell-cell adhesion	Bladder, breast, colon, liver
LKB1	19p 13.3	Serine, threonine kinase	Hamartomatous colon, papillary breast
MLHI	3p21.3	DNA mismatch repair	MSI positive colorectal and endometrial
p15(CDKN2B)	9p2 1	Cyclin-dependent kinase inhibitor	Acute leukaemias, Burkitt lymphoma, multiple myelomas
p16 (CDKN2A)	9p2 1	Cyclin-dependent kinase inhibitor	Lung, gliomas, breast, colon, bladder, nasopharyngeal, melanomas, prostate, thymomas, multiple myelomas, lymphomas
PTEN	10q23.3	Regulation of cell growth and apoptosis	Prostate
RB	13q 14.2	Sequesters E2F transcription factor	Retinoblastoma
VHL	3p25	Inhibits angiogenesis, regulates transcription	Renal cell carcinoma

There are several imprinted genes associated with inherited malignant or benign tumor syndromes (http://www.genecards.org/cgi-bin/cardsearch.pl?search=imprinting+AND+cancer+AND+inherited&search_type=kwd&speed=fast&mini=yes). These include:

1. *Human IGF2 and H19* genes are located in a chromosomal region (11p15. 5) harboring a cluster of imprinted genes. IGF2 encodes for a potent mitogenic factor involved in cell growth and embryonic development whereas H19 transcripts are non-coding. The reciprocally imprinted IGF2 and H19 genes are expressed from the paternally and maternally inherited alleles, respectively. The epigenetic regulation of this locus has been intensively studied in normal and malignant tissues. IGF2 overexpression occurs commonly in cancer (reviewed in 16), and loss of imprinting is one mechanism responsible for the dysregulation of this influential growth factor (17).

2. *RET* proto-oncogene (**re**arranged during **t**ransfection) is involved in the following heritable malignant diseases:

Disease: defects in ret are the cause of multiple neoplasia type iia (men2a) [mim:171400]; also called multiple neoplasia type II (MEN2). MEN2A, the most frequent form of MTC, is an inherited cancer syndrome characterized by MTC, phaeochromocytoma and/or hyperparathyroidism.

Disease: defects in ret are the cause of multiple neoplasia type iib (men2b) [mim:162300]. MEN2B is an uncommon inherited cancer syndrome characterized by predisposition to MTC and phaeochromocytoma, which is associated with marfanoid habitus, mucosal neuromas, skeletal and ophthalmic abnormalities, and ganglioneuromas of the intestine tract. Then the disease progresses rapidly with the development of metastatic MTC and a pheochromocytoma in 50% of cases.

3. *GNAS* - guanine nucleotide binding protein (G protein), is involved in the following heritable malignant diseases:

Disease: defects in gnas are the cause of McCune-Albright syndrome (mas) [mim:174800]. MAS is characterized by polyostotic fibrous dysplasia, cafe-au-lait lesions, and a variety of endocrine disorders, including precocious puberty, hyperthyroidism, hypercortisolism, growth hormone excess, and hyperprolactinemia. The mutations producing MAS lead to constitutive activation of GS alpha.

Disease: defects in gnas are the cause of a subset of growth hormone secreting pituitary tumors (somatotrophinoma) [mim:102200].

4. *MIP* - Lens fiber major intrinsic protein. Defects in mip are a cause of autosomal recessive congenital cataract [mim:154050] but also associated with brain neoplasms, glioma and glioblastoma.

5. *SDHC* - succinate dehydrogenase complex, subunit C, integral membrane protein Is associated with

Disease: defects in sdhc are the cause of autosomal dominant non- chromaffin paragangliomas type 3 (pgl3) [mim:605373]. Non-chromaffin paragangliomas are usually benign, neural crest derived tumors of parasympathetic ganglia.

AGEING

DNA hypomethylation was originally suspected to be responsible for the gene expression changes often observed with the ageing process. Interestingly, age-related decreases in DNA methylation occur primarily in the coding and intronic regions of genes, and they correlate poorly with observed reductions in gene expression (18). This disparity was clarified by studies showing that reduced gene expression that occurs with age results from a progressive increase in gene specific promoter methylation rather than generalized genomic hypomethylation. These age-related increases in promoter CpG island methylation occur in a number of genes involved in cancer, including IGF2, Versican, FAX6, and N33 in colon cancer and HIC in prostate cancer (18). It is likely that many other genes will also fall into this category, because several studies designed to isolate differentially methylated CpG islands in cancer have identified a number of CpG islands that exhibit increased methylation with both ageing and neoplastic transformation (18). Not all age-dependent hypermethylation events result in cancer. The ER gene is hypermethylated in nearly all primary colon cancers, yet the normal colon of patients both with and without colon cancer has about the same yearly increase in ER promoter CpG island methylation (18). Since age-related hypermethylation varies among individuals of the same age, it is likely that genetic predisposition to epimutations, as well as exposure to environmental factors are involved in cancer formation. Thus, there is now compelling evidence of a mechanistic link between the ageing process and tumorigenesis in that age-related promoter hypermethylation frequently occurs in genes known to be involved in cancer formation (18).

THE METHYLOME: DIAGNOSTIC POTENTIAL OF EPIGENETIC ABNORMALITIES

An important future direction for these studies is the definition of the components of the methylome, i.e., the total methylation content of the cell and patterns of CpG methylation. The methylome cannot be found on the GenBank web site, because ordinary sequencing does not reveal it. Sequencing data obtained from the human genome project are currently undergoing analysis to construct a human epigenetic map based on CpG content. This knowledge coupled with cross-species comparisons of the epigenome will be invaluable in deciphering the epigenetic elements involved in gene regulation (19-21). Epigenetic alterations in genes are early oncogenic events in some cancers, and detection of these early abnormalities may aid in protecting people from cancer through dietary alterations or pharmacological intervention (22). With increasing awareness of the importance of epigenetics in cancer formation, and the advent of laboratory techniques such as bisulfite DNA sequencing, methylation-sensitive PCR and gene expression profiling by DNA microarrays, it is likely that methylation profiles will ultimately be used to predict an individual's predis-

position to cancer, assist in cancer diagnosis and determine optimal therapeutic approaches (23).

THE BISULPHITE REVOLUTION

The analysis of DNA methylation was revolutionized by the introduction of sodium bisulphite conversion of genomic DNA. The differential rates at which cytosine and 5-methylcytosine are deaminated by sodium bisulphite to yield uracil and thymine, respectively, had been known for some time, but it was not until Frommer et al. (24) showed the usefulness of this chemical reaction in conjunction with PCR amplification and sequencing that the method became widely adopted. Now, dozens of different techniques rely on the ability of sodium bisulphite to efficiently convert unmethylated cytosine to uracil, without affecting 5-methylcytosine. Among many different methods, the most frequently are illustrated in the Figure 1.

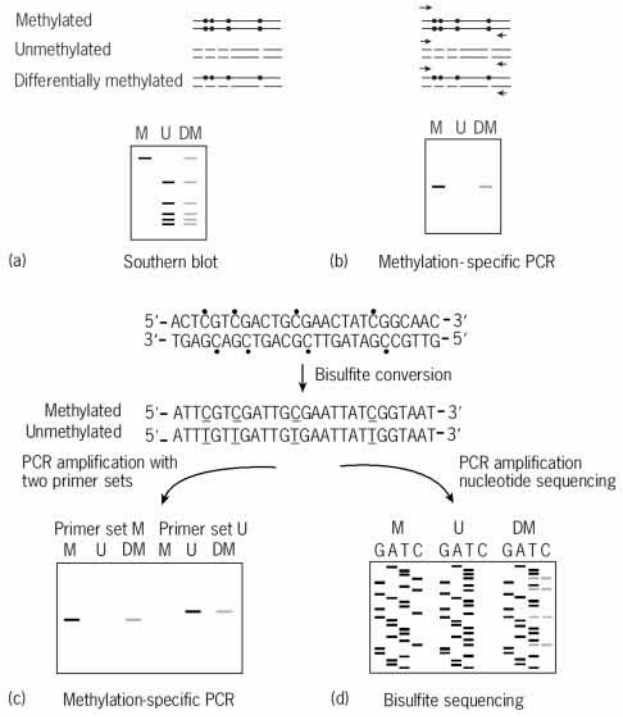

Figure 1. Methods used to analyze CpG methylation. Cytosine methylation is indicated by the circles. (a) Southern blotting depends on methylation-sensitive restriction enzymes to discriminate between methylated and unmethylated alleles. DNA is digested and fractionated on an agarose gel followed by blotting with a probe specific to the region of interest. Methylated recognition sites are resistant to digestion and will yield larger DNA fragments on the blot than unmethylated DNA. (b) Methylation-specific PCR requires digestion with methylation sensitive restriction enzymes, followed by PCR amplification. Methylated (uncut) DNA will yield an amplification product whereas unmethylated DNA will not be amplified. Bisulfite conversion of unmethylated cytosines precedes analysis by either a modification of methylation-specific PCR (c) or bisulfite sequencing (d). Sodium bisulfite treatment of DNA leads to the conversion of unmethylated cytosines to uracils while methylated cytosines are protected from conversion. Subsequent PCR, using two independent primer sets designed to bisulfite protected and bisulfite-converted sequence, amplifies methylated (M) and unmethylated (U) alleles, respectively. Bisulfite sequencing is the most direct means of analyzing the methylation status of individual cytosines. Fully methylated cytosines are evident in the C lane of a sequencing gel using this method, while unmethylated cytosines are converted to thymines in the PCR amplification step before sequencing. Alleles having differential cytosine methylation (DM) are evident by the presence of bands in both the C and T lanes.

Standard molecular biology techniques to analyze individual gene loci, such as polymerase chain reaction (PCR) and biological cloning, erase DNA methylation information, leaving the investigator oblivious to the epigenetic information that was present in the original genomic DNA. The solution to this problem is to modify the DNA in a methylation-dependent way before amplification. This can be achieved either by digestion with a methylation-sensitive restriction enzyme or by treating the genomic DNA with sodium bisulphite which converts unmethylated cytosines to uracil residues. Consequently, the converted DNA is no longer self-complementary, and amplification of either the top or the bottom DNA strand requires different primers. Priming can be either universal, or methylation specific as in the MSP, methylation-specific PCR.

Amplifying bisulphite-converted DNA

Three factors need to be considered in the design of PCR primers for the amplification of bisulphite-converted DNA. First, a unique feature of bisulphite-converted DNA is that it is not self-complementary. Therefore, primers that are designed to amplify the top strand of a particular stretch of DNA will be different from those that are designed to amplify the bottom strand. Software tools are available to simplify the task of *in silico* bisulphite conversion of a DNA sequence and of primer design. Second, primers should cover several cytosines that are not part of CpG dinucleotides in the original sequence, and are therefore converted to uracils by bisulphite. Inclusion of such bases in the primer design helps to avoid amplification of any residual unconverted DNA. Third, the sequence variation that reflects the methylation status at CpG dinucleotides in the original genomic DNA offers two approaches to PCR primer design. One method uses primers that themselves do not cover any potential sites of DNA methylation. This yields a pool of PCR products with sequence variations at sites of differential methylation located between the two primers. This method is used in bisulphite genomic sequencing (24), COBRA (25), Ms-SNuPE (26), and several other techniques. The alternative method to amplify bisulphite-converted genomic DNA uses primers that are designed to anneal specifically with either the methylated or unmethylated version of the bisulphite-converted sequence. This is the basis for MSP (27), which is the most widely used method of DNA methylation analysis. MSP has had a significant impact on the burgeoning field of cancer epigenetics by making DNA methylation analysis accessible to a wide number of laboratories. Non-MSP amplification is particularly useful for the quantitative or detailed analysis of 5-methylcytosine distribution, whereas MSP excels at the sensitive detection of particular methylation patterns.

Clinical applications

DNA-methylation-based technologies have a promising future in both clinical diagnostics and therapeutics. DNA methylation markers have obvious applications in diagnostics, but can also contribute indirectly to therapeutics as predictors of response to therapy. Of the detection strategies, DNA methylation patterns have proven to be most useful in the sensitive detection of disease, whereas profiling methods are useful for the stratification approaches described above.

Virtually all strategies for the sensitive detection of cancer-specific DNA methylation patterns rely on the principle of MSP, or fluorescence-based variants, such as MethyLight. Cancer-specific DNA methylation patterns can be found in detached tumor cells in bodily fluids and biopsies, and they can be detected in free-floating DNA that is released from dead cancer cells. It was shown more than 25 years ago that cancer patients have increased levels of free DNA in their serum, which is thought to be released from apoptotic or necrotic tumor cells. This principle is the basis for a rapidly expanding number of studies of DNA methylation markers in blood serum and plasma. Blood-borne tumor-derived DNA is often detected more frequently in patients with advanced-stage disease, although not all studies have reported such a correlation. The presence of detectable tumor DNA in the plasma or serum is generally associated with a poor prognosis. An alternative approach is to screen for tumor-specific DNA methylation patterns in bodily fluids or detached cells that are derived from luminal content. This targeted and localized approach seems to give a higher sensitivity than serum or plasma detection.

DISEASE STRATIFICATION

Many types of cancer display significant variability in clinical outcome among patients with similar pathologies and disease stage. Recent advances indicate that it might be possible to more accurately predict clinical outcome from the molecular characteristics of a patient's tumor based on transcription profiling. Clinical outcome is affected by many factors, some of which are a function of the genetic composition and health status of the patient, whereas others are inherent to the malignancy itself. In addition, each of these factors might affect the response to clinical treatment. For example, a patient who is homozygous for a polymorphism in the thymidylate synthase gene, resulting in decreased enzyme activity, might be more likely to show a good tumor response to 5-fluorouracil, but might also suffer more side effects from the drug. It is important to distinguish between predictive markers, which are associated with the relative sensitivity to specific therapeutic strategies, and prognostic markers, which are associated with treatment-independent factors such as the growth rate and metastatic behavior of the malignancy. Both of these types of stratification markers are of clinical value and can assist physicians in their choice of treatment.

The presence of tumor-specific methylation markers in the serum or plasma of patients has been reported to be of prognostic significance (28). However, the presence of a serum or plasma methylation marker is merely indicative of the release of sufficient amounts of tumor DNA into the bloodstream, which is likely to be correlated with invasiveness. For this purpose, any tumor-specific methylation marker would suffice, and its presence in the tumor itself would not necessarily be correlated with clinical outcome. This should be distinguished from reports of associations between the presence of DNA methylation markers in malignancies themselves and clinical outcome (29). The presence of methylation markers was often found to be correlated with other known prognostic criteria. However, several studies have carefully documented independent prognostic values for DNA methylation markers (30).

The development of DNA methylation markers that are predictive of a response to chemotherapy is still in its infancy. Several studies have reported associations between DNA methylation markers and response to chemotherapy (31). The most extensive work has been done with CpG island hypermethylation of the O^6-methylguanine methyltransferase (*MGMT*) gene. Esteller et al. (32) reported that *MGMT* methylation was associated with prolonged survival in glioma patients who were treated with carmustine, and in patients with diffuse large B-cell lymphoma who were treated with cyclophosphamide (31), as part of multidrug regimens. Others have expressed reservations regarding the conclusions that can be drawn from such retrospective multidrug studies. The identification of true predictive markers requires that care be taken to separate out any prognostic associations of the marker. This is best achieved in a prospective randomized clinical trial, in which the candidate marker(s) is used to predict response in a directed therapy arm, and in which prognostic associations can be separately evaluated in the control arm. Such prospective clinical trials using DNA methylation markers have yet to be conducted. Nevertheless, a flood of reports on predictive DNA methylation markers is predicted in the near future.

Going global

Molecular profiling of cancer has, so far, focused primarily on the use of gene-expression (cDNA) microarrays. However, this technique is poorly compatible with formalin fixation and paraffin embedding of tumor tissues, both of which are used in routine histopathology. DNA methylation markers offer an alternative approach to molecular profiling that is just starting to be explored. DNA methylation patterns are a rich source of information, rivaling that of gene-expression profiles. The haploid human genome contains approximately 50 million CpG dinucleotides, which are capable of encoding $2^{50,000,000}$ different permutations per haploid genome, and the number of methylation changes in cancer cells seems to outnumber informative genetic alterations. The problem is not the lack of information in the epigenome, but the difficulty in accessing that information.

DNA methylation profiles offer several advantages over gene-expression microarrays and proteomic approaches. First, the DNA molecule is very stable, surviving routine processing for histopathology. Second, measurements of DNA methylation can be compared with absolute reference points (completely methylated or completely unmethylated DNA). This greatly simplifies the design of internal references for methylation assays. Third, abnormal methylation patterns in cancer cells differ qualitatively from normal cells, not just quantitatively. This allows for the development of assays with high specificity and sensitivity. This can be an advantage in analyses of samples that contain substantial amounts of stromal tissue or non-malignant epithelium. In such cases, it can be very difficult to detect decreased gene expression or loss of heterozygosity in the cancer cells. A fourth advantage is that methylation assays for individual markers tend to be universal, just like gene-expression markers. Genetic mutation assays usually have to be tailored to the individual tumor. Finally, a fifth advantage is that DNA methylation patterns are fairly stable over time. They do not fluctuate in response to short-term stimuli, as gene-expression profiles do.

EPIGENOME BASED CANCER TREATMENT

In the context of cancer treatment, it is important to bear in mind that epigenetic alterations are reversible, and possibly easily affected by the environment, unlike conventional genetic mutations. Epigenetic cancer therapy also has major potential advantages over conventional therapeutic approaches. First, intact copies of tumour-suppressor genes do not need to be transfected into cells because they are already present in the cancer cell genome; they only need to be reactivated. Second, if gene-specific approaches are used to reactivate epigenetically silenced tumour-suppressor genes there should be little normal tissue toxicity, enabling them to be safely combined with therapies that are more conventional. While it is difficult to conceive a therapeutic strategy to replace a mutated gene in patients with a normal copy, it is much less fantastic to imagine restoring a normal pattern of imprinting and methylation to cells. For example, 5aza-2-deoxycytidine was shown to restore a normal pattern of imprinting to tumor cells with LOI (Loss of Imprinting), without disrupting imprinting on the normally marked allele. The studies of Cui et al. (33) are particularly exciting in this regard, in that one may be able to identify an epigenetic alteration affecting the entire organ system, *prior* to the development of cancer and attempt timely prophylaxis and treatment.

The two major pharmacological targets associated with these epigenetic changes are DNMT and HDAC. The DNMT inhibitor 5-azaC is structurally similar to cytosine but when incorporated into DNA it forms a stable covalent bond with DNMT that inhibits further-methylation by the sequestered enzyme. Consequently, overall genomic hypomethylation develops with subsequent rounds of DNA replication.

5-AzaC is efficacious in treating patients with acute leukemia. It has also undergone clinical testing for the treatment of solid tumors; however, 5-azaC produces high level of normal tissue toxicity and mutagenicity. These untoward side effects are not due to the resulting hypomethylation, but are attributed to the presence of the incorporated DNMT-5-azaC complexes in the genomic DNA (22). More specific strategies to inhibit the action of DNMT are being developed, including the use of antisense molecules. In this approach, antisense DNAs complementary to the DNMT mRNA inhibit methyltransferase activity by preventing DNMT translation. HDAC inhibitors, such as trichostatin A and sodium butyrate, have been shown to increase the level of histone acetylation in cultured cells, and to cause growth arrest, differentiation, and apoptosis. Consequently, they are currently being tested in clinical trials as therapeutic agents for cancer.

Support

Grant from Serbian Ministry for science and technology, No 102019

REFERENCES

1. Waters MD, Fostel JM. Toxicogenomics and systems toxicology: aims and prospects. Nat Rev Genet 2004;5:936-48.
2. de Hoog CL and Mann M. Proteomics. Annu Rev Genom Hum Genet 2004;5:267-93.
3. Whitfield PD, German AJ, Noble PJ. Metabolomics: an emerging post-genomic tool for nutrition. Br J Nutr 2004;92:549-55.
4. Hanash S. Integrated global profiling of cancer. Nat Rev Cancer 2004;4:638-44.
5. Waters MD, Selkirk JK, Olden K. The impact of new technologies on human population studies. Mutat Res 2003;544:349-60.
6. Riggs AD, Porter TN. Overview of epigenetic mechanisms. In: Russo VEA et al, editors. Epigenetic Mechanisms of Gene Regulation. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 1996. p. 29-45.
7. Baylin SB, Herman JG, Graff JR, Vertino PM, Issa JP. Alterations in DNA methylation: a fundamental aspect of neoplasia. Cancer Res 1998;72:141-96.
8. Robertson KD, Wolffe AP. DNA methylation in health and disease. Nature Rev Genet 2000; 1:11-9.
9. Li E, Bestor TH, Jaenisch R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell 1992;69:915-26.
10. Li E, Beard C, Jaenisch R. Role for DNA methylation in genomic imprinting. Nature 1993; 366:362-65.
11. Pfeifer GP, Tang M, Denissenko MF. Mutation hotspots and DNA methylation. Current Topics in Microbiology and Immunology 2000;249:1-20.
12. Costello JF, Fruhwald MC, Smiraglia DJ, Rush LJ, Robertson GP, Gao X, et al. Aberrant CpG-island methylation has non-random and tumor-type-specific patterns. Nature Genetics 2000; 25:132-8.
13. Baylin SB, Herman JG, Graff JR, Vertino PM, Issa JP. Alterations in DNA methylation: a fundamental aspect of neoplasia. Cancer Res 1998;72:141-96.
14. Hu JF, Oruganti H, Vu TH, Hoffman AR. The role of histone acetylation in the allelic expression of the imprinted human insulin-like growth factor II gene. Biochem Biophys Res Com 1998; 251:403-8.
15. Barlow D P. Gametic imprinting in mammals. Science 1995;270:1610-3.
16. Reik W, Constancia M, Dean W, Davies K, Bowden L, Murrell A, Feil R, et al. Igf2 imprinting in development and disease. Int J Develop Biol 2000;44:145-50.
17. Rainier S, Johnson LA, Dobry CJ, Ping AJ, Grundy PE, Feinberg AP. Relaxation of imprinted genes in human cancer. Nature 1993;362:747-9.
18. Issa JP. CpG-island methylation in aging and cancer. Current Topics in Microbiology and Immunology 2000;249:101-18.
19. Hardison RC. Conserved noncoding sequences are reliable guides to regulatory motifs. Trends in Genetics 2000;16:369-72.
20. Vu TH, Hoffman AR. Comparative genomics sheds light on mechanisms of genomic imprinting. Genome Res 2000;10:1660-3.
21. Wylie AA, Murphy SK, Orton TC, Jirtle RL. Regulatory motifs of the novel imprinted domain, DLKJ/GTL2, mimic those of IGF2/H19. Genome Res 2000;10:1711-8.
22. Laird PW, Jaenisch R. The role of DNA methylation in cancer genetics and epigenetics. Ann Rev Genet 1996;30:441-64.
23. Esteller M, Corn PG, Baylin SB, Herman JG. A gene hypemiethylation profile of human cancer. Cancer Res 2001;61:3225-9.
24. Frommer M, McDonald LE, Millar DS, Collis CM, Watt F, Grigg GW, et al. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. Proc Natl Acad Sci USA 1992;89:1827-31.
25. Xiong Z, Laird P W. COBRA: a sensitive and quantitative DNA methylation assay. Nucleic Acids Res 1997;25:2532-4.
26. Gonzalgo M, Jones PA. Rapid quantitation of methylation differences at specific sites using methylation-sensitive single nucleotide primer extension (Ms-SNuPE). Nucleic Acids Res 1997; 25:2529-31.
27. Herman JG, Graff J R, Myohanen S, Nelkin BD, Baylin SB. Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. Proc Natl Acad Sci USA 1996;93:9821-6.
28. Silva JM, Garcia JM, Dominguez G, Silva J, Miralles C, Cantos B, et al. Persistence of tumor DNA in plasma of breast cancer patients after mastectomy. Ann Surg Oncol 2002;9:71-6.
29. Tang X, Khuri FR, Lee JJ, Kemp BL, Liu D, Hong WK, Mao L. Hypermethylation of the death-associated protein (DAP) kinase promoter and aggressiveness in stage I non-small-cell lung cancer. J Natl Cancer Inst 2000;92:1511-6.
30. Brabender J, Usadel H, Metzger R, Schneider PM, Park J, Salonga D, et al. Quantitative O(6)-methylguanine DNA methyltransferase methylation analysis in curatively resected non-small cell lung cancer: associations with clinical outcome. Clin Cancer Res 2003;9:223-7.
31. Esteller M, Gaidano G, Goodman SN, Zagonel V, Capello D, Botto B, et al. Hypermethylation of the DNA repair gene O(6)-methylguanine DNA methyltransferase and survival of patients with diffuse large B-cell lymphoma. J Natl Cancer Inst 2002; 94:26-32.
32. Esteller M, Garcia-Foncillas J, Andion E, Goodman SN, Hidalgo OF, Vanaclocha V, et al. Inactivation of the DNA-repair gene MGMT and the clinical response of gliomas to alkylating agents. N Engl J Med 2000;343:1350-4.
33. Cui H, Horon IL, Ohlsson R, Hamilton SR, Feinberg AP. Loss of imprinting in normal tissue of colorectal cancer patients with microsatellite instability. Nat Med 1998;4:1276-80.