Use of DNA Methylation for Cancer Detection and Molecular Classification

Jingde Zhu1,2,3,* and Xuebiao Yao4,5

1Cancer Genetics and Gene Therapy program, The State-key Laboratory for Oncogenes and Related Genes, Shanghai Cancer Institute, Shanghai Jiaotong University and 2Epigenomics program, Shanghai Center for Systems Biomedicine, Shanghai Jiao Tong University, LN 252200, Xie-tu Road, Shanghai 200032, China
3School of Life Sciences, Zhejiang Science and Technology University, Xiasha University District, Hangzhou, Zhejiang 310018, China
4Division of Cellular Dynamics, Hefei National Laboratory, Hefei 230027, China
5Cancer Biology Program, Morehouse School of Medicine, 720 Westview Drive, Atlanta, GA 30310, USA

Received 20 February 2007

Conjugation of the methyl group at the fifth carbon of cytosines within the palindromic dinucleotide 5'-CpG-3' sequence (DNA methylation) is the best studied epigenetic mechanism, which acts together with other epigenetic entities: histone modification, chromatin remodeling and microRNAs to shape the chromatin structure of DNA according to its functional state. The cancer genome is frequently characterized by hypermethylation of specific genes concurrently with an overall decrease in the level of 5-methyl cytosine, the pathological implication of which to the cancerous state has been well established. While the latest genome-wide technologies have been applied to classify and interpret the epigenetic layer of gene regulation in the physiological and disease states, the epigenetic testing has also been seriously explored in clinical practice for early detection, refining tumor staging and predicting disease recurrence. This critique reviews the latest research findings on the use of DNA methylation in cancer diagnosis, prognosis and staging/classification.

Keywords: Cancer DNA Methylation, Epigenetics, Methylation Specific PCR

Introduction

Having the genetic blueprints for human and the increasing number of model organisms available has ushered in a new era for the functional elucidation of genetic makeup in both development and diseases. It chiefly concerns the epigenetic layer of genetic information; that is, inheritable component without changes in DNA sequence which consists of four definable entities, i.e., DNA methylation (cytosine [CpG] methylation), non-coding RNA, histone modification, and chromatin remodeling. The epigenetic interface sits between the genetic blueprints stored in DNA sequences and phenotypes dictated by the pattern of gene expression and more readily responds to the changing environment than its genetic counterparts with the profound biological consequences. Addition of the methyl group at cytosine ring within 5'-CpG-3' sequence (Fig. 1) is carried out by one of three DNA methyl transferase genes (DNMT1, DNMT3a, and DNMT3b) with S-adenosyl methionine as the methyl donor (Bird, 1992). The DNA methylation pattern in the parental cells can be faithfully duplicated and distributed into daughter cells in a fashion similar to the semi-conservative mechanism for precise replication of the genetic information encoded in DNA sequence. For the lasting transcription memory, DNA methylation is the preferred mechanism, regardless of the nature of the first epigenetic hit responsible for the changed transcription state. Between the major epigenetic reprogramming taken place in the early embryonic development and germ cell maturation, the epigenetic make-up is also moderately altered throughout the life of living organisms (Jaenisch and Bird, 2003). An integrated epigenetic homeostatic mechanism is required to monitor and repair any epigenetic abnormalities, failure of which leads to an accumulation of the epigenetic lesions beyond repair and ultimately the various disease states, including cancer (Baylin and Ohm, 2006; Ting et al., 2006).

Cancer is a disease of aberrant DNA methylation

Cancers are extremely complex diseases with distinct combinations of genetic and epigenetic defects that vary with both types of diseases and individuals (Hanahan and Weinberg, 2000), thus challenging the current biomedical and...
molecular knowledge and technology. Epigenetic mechanisms comprise of DNA methylation, histone modification, microRNAs and chromatin remodeling that work together to control the chromatin structure of DNA to meet the functional requirements. DNA methylation, an enzymatic process to add the methyl group at the fifth carbon of cytosines within the palindromic dinucleotide 5’-CpG-3’ sequence (DNA methylation) (Fig. 1) is the best studied epigenetic mechanism and the focal point of the present cancer epigenetic study for better management of these devastating diseases.

Over 85% CpG dinucleotides are spread out in genome and located in the repetitive sequences having the transcription-dependent transposition potential. They are heavily hypermethylated/transcription-silenced in the normal cells, a state crucial to the integrity of the chromatin structure of genome. The global genomic hypomethylation found in all types of human cancer studied is instrumental to the unwanted transcription of the otherwise methylated/silenced repetitive sequences and the proto-oncogenes (Bird, 1992; Baylin and Ohm, 2006), predisposition to the chromosome instability (Eden et al., 2003; Gaudet et al., 2003). The remaining circa 15% CpG are clustered within the short DNA regions (approximately, 0.2 to 1kb in length, called “CpG island”) which accounts for 1% of the genome. Approximately 40-50% of the genes have CpG island within or around the promoter and are largely unmethylated in the normal somatic cells, except for the silenced allele of the imprinted genes and the cell lineage-specific genes in the lineages where their expression is not required. However, the hypermethylated-dependent silencing in cancer cells have been reported for the classic tumor suppressor genes, DNA repair genes, cell cycle control genes, anti-apoptotic genes, and genes that prevent abnormal activity of developmental pathways. The epigenetic regulation does not work independently from its genetic counterpart as vividly demonstrated by the fact that the hypermethylated promoter CpG island/transcription silencing affects exclusively the wild-type allele (Esteller et al., 2001) in colon cancer. To cope with the genotoxic effects of both external and internal origins, living organisms are equipped with a large array of mechanisms, including over 200 genes that involve in sensing and repairing the damages to protect the genome integrity. However, the expression of over 75% of the DNA repair genes are under the control by the DNA methylation mediated mechanisms (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?CMD=search&DB=gene). The hypermethylated/transcription silencing state of several DNA repairing genes has been suggested seminal to cancer formation, by leaving the mutated DNA sequences un-repaired. Indeed, the G-to-A type mutation of the Ras proto-oncogene and the p53 tumor suppressor gene were more frequent in the colorectal cancer tissues where the hypermethylated/transcriptionally silenced MGMT gene resides, the protein of which prevents the G-to-A mutation (Esteller, 2000; Esteller et al., 2001). Furthermore, methylation sensitizes the CG-to-TG transition in DNA, which is thought to be responsible for the so-called CpG deletion in mammals through evolution (Bird, 1986) and the frequent TG mutation in tumor-associated genes (p53) in cancer (http://www-p53.iarc.fr/). The UV radiation mediated CC→TT transversion takes place more frequently when C is methylated (Tommasi et al., 1997). Preference of the methylated to the unmethylated DNA has been reported for a number of the nucleic acid binding carcinogens (Chen et al., 1998). Finally, the epigenetic makeup, than its genetic counterpart, is more susceptible to the changing surroundings, placing it as a cause for the genetic changes in cancer, which are a direct consequence of the environmental insults by the traditional wisdom. The fact that the majority of carcinogens are non-genotoxic has further strengthened this view (Trosko and Upham, 2005).

The critical role of the epigenetic abnormality at the initiation stage of cancer formation was first suggested by the association of the loss of genetic imprinting (LOI) for the maternal source in teratocarcinoma and the paternal source in hydatidiform moles/choriocarcinomas (Jelinec and Shaw, 2007). Indeed, the LOI of IGF2 gene, over-expression of which promotes cell proliferation, was found in normal-appearing colonic epithelium of patients with colorectal cancer and the increased risk of colon cancer has been reported for the individuals with LOI of this gene in circulating white cells (Cui et al., 2003). The hypermethylated/transcription silencing of the tumor suppressor and DNA repairing genes prevails at the pre-neoplastic stage of cancer formation (Belinsky et al., 1998; Belinsky, 2004). For instance, the hypermethylated pl6ink4A (tumor suppressor gene) and MGMT (the DNA repairing gene) were found in the sputum DNA as early as 35 months before lung cancer have been diagnosed (Palmituso et al., 2000). The genetic studies with the animal models of the germline mutations have demonstrated the role at the initiation stage in cancer formation of pl6ink4A in lung cancer, which correlates with a prolonged expansion life of the stem cell population (Janzen et al., 2006) and the subsequent epigenetic silencing of other crucial genes (Reynolds et al., 2006). This is consistent with the notion that would increase the stem cell expansion with an increased chance of abnormalities (Hamlin and Weinberg, 2000). Indeed, the Wnt signaling circuit has been found overactive in colon cancer via both mutations in downstream effector genes such as APC and β-catenin (Kinzler and Vogelstein, 1996; Radtke and Clevers, 2005) and epigenetic silencing of the SFRPs, which encode for proteins that antagonize Wnt activation at the cell membrane (Suzuki et al., 2002; Suzuki et al., 2004). In favor of this notion, embryonic fibroblasts from the mice with substantial loss of gene imprinting by the regulated disruption of the maintenance DNA methyltransferase, Dnmt1 (Holm et al., 2005) formed tumors in immuno-compromised mice. The same genetically modified cells behave as the spontaneously immortalized cells in vitro and create multiple tumors in chimerical animals originated (Holm et al., 2005). Therefore, a switch in heritable gene expression patterns, in the absence of mutations, may lead to abnormal expansion of
stem/progenitor cells and, thus increase the chance that full transformation and evolution of cancer occur. Recent studies in gastric cancer has strongly suggested for the association of the risk factors as such Helicobacter pylori infection and a unique aberrant DNA methylation pattern (Ushijima et al., 2006). Therefore, utility of the DNA methylation where serum, stool, sputum and urine sediments are used as the tumor DNA resources in population screening scheme for occult neoplasia sufferers or the high risk groups have been also seriously considered (Jubb et al., 2003).

As genetic defects, the epigenetic abnormalities can occur at all the stages of cancer progression, some of which are responsible for the altered behavior specific to given stages, such as the dysplasia, local benign and malignant tumor and finally metastatic lesions. Therefore, the information of the altered DNA methylation in the context of the cancer progression should have a great value in refining the staging and classification of cancer and provide the guidance for the better clinical management of cancers.

DNA methylation for cancer detection

The cancer associated altered CpG methylation can readily be detected against a background of normal DNA where the targeted CpG is at an opposite state of DNA methylation to the former, by the methylation-specific PCR (MSP) (Herman et al., 1996) of the bisulphite treated DNA where the unmethylated by converting to uridine but not the methylated CpG MSP is a sensitive assay, detecting one hypermethylated allele in the presence of 1,000-10,000 the unmethylated alleles (Herman et al., 1996). Using the quantitative version of MSP assay (MethyLight), the sensitivity reaches 1/100,000 without a loss of specificity and the need for gel electrophoresis (Eads et al., 2003). The nested PCR version of MSP could further increases the sensitivity, but at the expense of specificity (Palmisano et al., 2000). The methylation status of DNA is both chemically and biologically stable. Therefore, tumor-associated altered methylated DNA has been identified in various body fluids and/or the circulation of patients with esophagus, lung, breast, head and neck, prostate, and colorectal cancer by MSP or quantitative MSP (Table 1).

DNA methylation offers other distinct advantages over genetic markers (Table 1). It is a positive signal, opposed to negative signals (e.g., LOH), which are masked by contamination with normal cells (Kolbke et al., 1999). Secondly, an oncogene or tumor suppressor gene may be altered/inactivated by any one of many nucleotide sequences of the coding and noncoding regions of the genes. Alteration of an amino acid can be changes across three nucleotides of codon, each of which can be substituted by any one of a few nucleotides, thus necessitating several PCR reactions to identify all possible alterations responsible for a single amino acid alteration. There are over 23,544 somatic mutations, 376 germline mutations have been reported for the p53 coding region in cancer (http://www-p53.iarc.fr/). It is not possible to have all the possible variations to cover all the known mutation in p53 gene analyzed even with the most powerful platform technology at the present time. Furthermore, the genetic effects within it’s non-coding region has not or rarely been surveyed in cancer. On the contrary, epigenetic “defects” are almost always associated with hypermethylation within a defined promoter-region CpG island, and a single MSP PCR reaction (usually for the methylated allele) is needed.

Furthermore, genetic lesions are not specific for particular tumor tissue types. Consequently, if the circulating DNA in serum or plasma is used for screening, it is not possible to pinpoint the location of the occult tumor. On the contrary, the set of hypermethylated CpG islands can be specific to a given tumor type, allowing the investigator to identify the origin of the altered DNA (Table 2). In last few years, we have methylation-profiled of six types of common cancer in China to identify the informative set of genes for differentially detection of one from other cancer types (Yu et al., 2002; Yu et al., 2003; Ding et al., 2004; Gu et al., 2006; Li et al., 2004; Xu et al., 2004; Xu et al., 2004; Yang et al., 2004; Yu et al., 2004; Yu et al., 2004; Zhu, 2005; Zhu, 2006; Huang et al., 2006; unpublished observations). From no less 31 genes profiled in the cancer tissues of no less than 70 cases, no less 19 genes are informative to a given type of cancer. As listed in

| Table 1. Molecular biomarkers for cancer detection |

<table>
<thead>
<tr>
<th>Genetic</th>
<th>Epigenetic</th>
<th>Expressional</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mutation, SNP, LOH</td>
<td>DNA methylation</td>
<td>mRNA</td>
</tr>
<tr>
<td>Stability</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>PCRable</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Target/gene</td>
<td>Multiple</td>
<td>Single</td>
</tr>
<tr>
<td>Nature</td>
<td>Quantitative</td>
<td>Qualitative</td>
</tr>
<tr>
<td>Sample purity</td>
<td>Essential</td>
<td>Moderate</td>
</tr>
<tr>
<td>Fluctuation</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Tumor type specificity</td>
<td>No</td>
<td>High</td>
</tr>
</tbody>
</table>

NA, not applicable; multiple, more than one target (single) needs to be analyzed; fluctuation, whether the amount of the biomarker changes according to the fluctuation of nontumor factors (emotional, physiological, or pathological factors)
Both RNA and protein based biomarkers, underlining the phenotypic uniqueness of cancer cells have been most intensively evaluated for their utility in cancer detection for more than a decade. However, the following features prevent their use in clinical practice. Both markers are biologically and biochemically unstable, making recovery of them in bodily fluid difficult and their readout suffers from sample contamination and illegitimate expression. (reviewed in Macdonald, 1999; Tsavellas et al., 2001).

Despite of all the advantages, DNA methylation based method for cancer detection remains at its infant stage. Lack of the informative set of the promoter CpG islands for particular cancer types and the practical and regulatory concerns in the clinical evaluation are two major hurdles. Therefore, our efforts to establish a DNA methylated method to detect bladder cancer in urine sediments started with screening for the informative set of 58 tumor associated genes in three established bladder cancer cell lines, bladder cancer tissues and urine sediments. Then 21 genes were methylated profiled in the urine sediments of the bladder cancer (132 cases), the non-cancerous urinary lesion (23 cases), the non-urinogenitary disease (6) cases and the healthy volunteers (7) cases. Among the controls, only 3 of 23 non-cancerous urinary lesion patients were found associated with three hypermethylated genes. 121 among 132 (91.7%) bladder cancer cases have at least one hypermethylated of 11 genes analyzed, with a decent specificity: 87% (manuscript in submission). We are currently evaluating this set of the methylated genes for bladder cancer detection in a large cohort of patients with the urinary symptoms. At the same time, searching for more informative targets by a genome-wide approach involving probing the oligonucleotide CpG island array with the affinity purified heavy methylated DNA fragments is underway (unpublished observation).

Table 2, the occurrence of the hypermethylated state of a five gene set varies significantly with the cancer type (Table 2), thus distinguishing one from the other types of cancer.

<table>
<thead>
<tr>
<th>Type (No. of cases)</th>
<th>No. of genes tested</th>
<th>No. of genes informative</th>
<th>CDH13</th>
<th>MYOD1</th>
<th>MGMT</th>
<th>p16INK4a</th>
<th>RASSF1A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bladder Cancer (209)</td>
<td>66</td>
<td>28</td>
<td>3.8</td>
<td>14.3</td>
<td>0</td>
<td>6.3</td>
<td>11.9</td>
</tr>
<tr>
<td>Colon Cancer (64)</td>
<td>31</td>
<td>19</td>
<td>63.5</td>
<td>68.3</td>
<td>19.0</td>
<td>15.9</td>
<td>3.2</td>
</tr>
<tr>
<td>Glioma (57)</td>
<td>95</td>
<td>28</td>
<td>5.7</td>
<td>1.9</td>
<td>35.8</td>
<td>1.9</td>
<td>69.8</td>
</tr>
<tr>
<td>Liver Cancer (184)</td>
<td>136</td>
<td>47</td>
<td>28.9</td>
<td>27.1</td>
<td>3.6</td>
<td>34.3</td>
<td>97.0</td>
</tr>
<tr>
<td>Lung Cancer (89)</td>
<td>100</td>
<td>42</td>
<td>48.1</td>
<td>35.1</td>
<td>22.1</td>
<td>26.0</td>
<td>51.9</td>
</tr>
<tr>
<td>Ovarian Cancer (235)</td>
<td>47</td>
<td>39</td>
<td>28.7</td>
<td>10.4</td>
<td>2.3</td>
<td>7.5</td>
<td>46.6</td>
</tr>
</tbody>
</table>

Hepatocellular carcinoma (HCC) is one of the most aggressive malignancies (http://www-depdb.iarc.fr/globocan/GLOBOframe.htm). Difficulties in the early diagnosis and clinical management, such as both inherent and adaptive resistance to the common chemotherapeutic modalities, and its rapidly advancing nature contribute to the poor clinical outcome of HCC sufferers, which is largely attributed to the inadequacy of the staging and classification system based upon clinical observations, imaging, and biochemical data (Gospodarowicz et al., 2004). In last few years, we have methylation profiled by MSP of methylation 136 promoter CpG island in 184 liver cancer tissues (with the pair non-cancerous tissues) and 6 healthy liver tissue, found that the 47 genes, while unmethylated in the normal control, were hypermethylated to various extents in the liver cancer tissues (unpublished observations). It has been generally accepted...
that the pathologically defined neighboring non-cancerous tissues represent the cells residing at the early stage of carcinogenesis (Fig. 2A). Assuming that the genes significantly more frequent methylated in tumor (C) in comparison with the neighboring tissues (N) as the late phase targets and the otherwise as the early phase one, 27 genes were assigned in the late phase and 20 in the early phase (Fig. 2B, unpublished observation). The current efforts at the larger scale involving over 500 patients and 500 tumor associated are under way for a DNA methylation based classification and staging system of liver cancer.

**Future perspectives**

The recent surges in interest in DNA methylation and other epigenetic phenomena have best been demonstrated by the ongoing international efforts to decipher both the components and the mechanisms of the epigenetic layer of heredity on a gigantic scale: the epigenome project (initiated in 1998, http://www.epigenome.org), the Epigenetic Plasticity of the Genome project (initiated in 2004, http://www.epigenome-nce.net), and the Encyclopedia of DNA Elements project (ENCODE, initiated in 2003, http://www.genome.gov/1005107). It is fully anticipated that there will be a rapid improvement in technological capability and in the gaining of new insights into the role and mechanisms of epigenetic phenomena in both development and disease (including cancer) in the near future. The concordant epigenetic behavior of multiple genes up to the whole-genome level will be dealt with in a highly precise manner prediction, with considerable ease, and at ease in the near future.

**Acknowledgments**

This work was supported by grants to Jingde Zhu from the Shanghai Science Foundation (04DJZ14006 and 05DJZ19318), the National Science Foundation (30570850, 10574134), the National Research Program for Basic Research of China (2004CB518804), the National Research Program for High Technology (2006AA02Z197, 2006AA02Z520) and the European 6th program (LSHB-CT-2005-019067). Thanks also go to Professor Jianren Gu for his constructive comments on this manuscript.

**References**


methylguanin-DNA methyltransferase is associated with the presence of G:C to A:T transition mutations in p53 in human colorectal tumorigenesis. Cancer Res. 61, 4689-4692.


