# THE POWER AND THE PROMISE OF DNA METHYLATION MARKERS

*Peter W. Laird*

The past few years have seen an explosion of interest in the epigenetics of cancer. This has been a consequence of both the exciting coalescence of the chromatin and DNA methylation fields, and the realization that DNA methylation changes are involved in human malignancies. The ubiquity of DNA methylation changes has opened the way to a host of innovative diagnostic and therapeutic strategies. Recent advances attest to the great promise of DNA methylation markers as powerful future tools in the clinic.

EARLY DETECTION

GENOMIC IMPRINTING
Process by which genes are selectively expressed by the maternal or paternal homologue of a chromosome.

X-INACTIVATION
Mammalian females have two X chromosomes per genome, whereas males have only one. In female mammals, one X chromosome is functionally silenced during embryogenesis to ensure that the stoichiometry of X-chromosomal and autosomal gene products is the same in males and females.

*University of Southern California, Norris Comprehensive Cancer Center, Room 6418, 1441 Eastlake Avenue, Los Angeles, California 90089-9176, USA.*
*e-mail: plaird@usc.edu*
doi:10.1038/nrc1045

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[1–3]. In recent years, it has become apparent that the transcriptional silencing that is associated with 5-methylcytosine is important in mammalian development[4,5], protection against intragenomic parasites[6], GENOMIC IMPRINTING[3], X-INACTIVATION[3], mental health[1] and cancer[7].

Cytosine methylation occurs after DNA synthesis, by enzymatic transfer of a methyl group from the methyl donor *S*-adenosylmethionine to the carbon-5 position of 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[8]. 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[9,10]. 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 ISLANDS.

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[11–14]. Nevertheless, most methylated cytosine residues are found in CpG dinucleotides that are located outside of CpG islands, primarily in repetitive sequences[15]. Methylation of some CpG islands in non-malignant tissues increases with age[11,12], but the total genomic content of 5-methylcytosine declines[16]. These opposing events are also present, but are much more pronounced, in cancer cells[17–19]. Cancer-specific DNA methylation changes at individual gene loci have so far focused primarily on HYPERMETHYLATION of CpG islands[7,19,25,26]. Cancer-specific DNA HYPOMETHYLATION

## Summary

- Cytosine-5 DNA methylation occurs in mammals at CpG dinucleotides. About 70% of the CpG dinucleotides in the mammalian genome are methylated.
- The complexity of varying distributions of methylated cytosines across the approximately 50 million CpG dinucleotides of each mammalian genome in a DNA sample that is derived from a heterogenous tissue sample is a diagnostic dream and an analytical nightmare.
- A nomenclature for the three principal approaches to methylation analysis — methylation content, methylation levels and methylation patterns — is proposed. The latter two types of analysis can be performed at multiple sites in the genome to yield methylation profiles.
- In the past decade, DNA methylation analysis has been revolutionized by two technological advances — bisulphite modification of DNA and methylation-specific polymerase chain reaction (MSP).
- CpG islands are approximately 1-kb stretches of DNA-containing clusters of CpG dinucleotides that are usually unmethylated in normal cells and are often located near the 5′ ends of genes. Methylation of promoter CpG islands is associated with a closed chromatin structure and transcriptional silencing of the associated gene.
- Hypermethylation of CpG islands is a common event in carcinogenesis. The transcriptional silencing of tumour-suppressor genes by promoter CpG island hypermethylation can contribute to oncogenesis.
- DNA methylation profiles represent a more chemically and biologically stable source of molecular diagnostic information than RNA or most proteins. The diagnostic potential of DNA methylation profiles is still largely untapped.
- Cancer-specific DNA methylation patterns can be detected in tumour-derived free DNA in the bloodstream and in epithelial tumour cells shed into the lumen, offering a promising approach to the early detection of cancer. Clinical application will first require further validation and will ultimately be based on standardized MSP-based technologies, such as MethyLight, rather than on gel-based techniques.
- A distinction between the clinical and analytical sensitivities of DNA methylation biomarkers is proposed.


CpG ISLAND
First described by Adrian Bird as an unmethylated HpaII tiny fragment (HTF) and formerly commonly defined as a contiguous window of DNA of at least 200 base pairs in which the G:C content is at least 50% and the ratio of observed CpG frequency over the expected frequency exceeds 0.6. Recently, a more stringent definition of a 500-base-pair window with a G:C content of at least 55% and an observed over expected CpG frequency of at least 0.65 has been proposed to exclude most Alu repeat sequences.

DNA HYPERMETHYLATION
An increased level of DNA methylation in a DNA sample at either an individual CpG dinucleotide or at a group of CpG dinucleotides relative to a reference DNA sample, usually derived from a normal tissue.

DNA HYPOMETHYLATION
A decreased level of DNA methylation in a DNA sample at either an individual CpG dinucleotide or at a group of CpG dinucleotides (or even the entire genome) relative to a reference DNA sample, usually derived from a normal tissue.

POLYMORPHISM
A sequence variant that is present among human populations.

events at individual unique sequences remain largely unexplored. Global 5-methylcytosine content is influenced by the nutritional availability of folate[20–22] and by POLYMORPHISMS in folate metabolic enzymes[23,24].

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[27–29], 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[30]. However, the great strength of DNA methylation in the clinic promises to be in the area of molecular diagnostics and early detection.

## Detecting DNA methylation

*A rich tapestry.* The distribution of 5-methylcytosine residues in a sample of genomic DNA that is derived from a heterogeneous tissue source resembles a rich tapestry, and there are several main ways in which this distribution can be analysed[31] (FIG. 1). Although global genomic DNA methylation content (FIG. 1a) might have an important role in carcinogenesis, its measurement in cancer cells has little to offer as a molecular marker, either in sensitivity or in informational content. Conversely, methylation levels at individual CpG dinucleotides (FIG. 1b) are useful for quantifying differences at important regulatory sequences. Methylation patterns (FIG. 1c) have two important uses. Bisulphite genomic sequencing of cloned polymerase chain reaction (PCR) products can provide detailed information on the pattern of 5-methylcytosine distribution along a stretch of DNA, representing a single individual DNA strand[32]. In addition, methylation-specific PCR (MSP)-based techniques can detect the presence of specific DNA patterns with very high sensitivity and specificity[33]. Methylation profiles (FIG. 1d and e) yield information on the methylation status across many sites in the genome, providing a unique approach to genome-wide molecular diagnostics. Profiles can be constructed using measurements of individual CpG dinucleotides, as in most restriction-enzyme- and microarray-based methods[34–36] (FIG. 1d), or they can be compiled from multiple MSP-based analyses[37–39] (FIG. 1e).

*Enzymes to the rescue.* The delay in appreciation of the importance of epigenetic events in carcinogenesis, compared with the celebrated role that genetic mechanisms have enjoyed in the past two decades, is rooted in the difficulty of analysing DNA methylation. Standard molecular biology techniques to analyse individual gene loci erase DNA methylation information, leaving the investigator oblivious to the epigenetic information that was present in the original genomic DNA. Scientists were first able to analyse DNA methylation at specific gene loci with the advent of methylation-sensitive restriction enzymes, in conjunction with Southern blot analysis. This experimental approach is reliable, but cumbersome, and requires a substantial amount of high-quality DNA. The use of post-digestion PCR to circumvent these limitations is prone to false-positive results that occur because of incomplete enzyme cleavage. More recently, methylation-sensitive enzymes have been successfully used in genome-wide methylation analysis and marker discovery techniques.

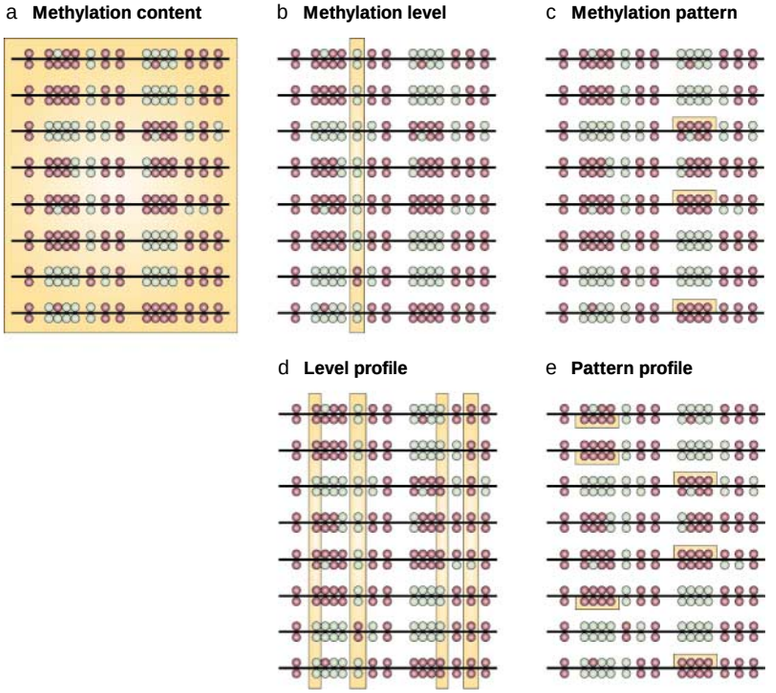


Figure 1 | **Principles of DNA methylation analysis.** A genomic DNA sample usually consists of a heterogeneous mix of DNA molecules that are derived from many different cells. In this figure, each horizontal bar is intended to represent an entire double-stranded haploid genome. Eight such haploid genomes are aligned above each other in each panel. Circles represent cytosine residues in the context of CpG dinucleotides on the top or bottom strand of the DNA double helix. Methylated cytosines are represented by maroon circles, unmethylated cytosines by green circles, and hemi-methylated CpG dinucleotides by opposing maroon and green circles. Shaded areas cover the CpGs participating in the relevant measurement principle. The total genomic methylation content shown in panel **a** represents the total amount of methylated cytosine in the genome, expressed either as a fraction of cytosines, or as a fraction of CpG dinucleotides. Methylation level (panel **b**) refers to the average methylation occupancy at a single CpG dinucleotide. If the methylation level is measured by digestion with a methylation-sensitive restriction enzyme, then methylation of either the top or bottom strand might contribute to the methylation determination, as indicated by the shaded area covering both DNA strands. Methylation pattern (panel **c**) refers to a series of CpG dinucleotides located *in cis* on a single DNA strand. DNA methylation patterns can be determined by bisulphite genomic sequencing of subclones (not shown). Methylation-specific polymerase chain reaction measures the presence of one particular methylation pattern in a pool of DNA, as indicated. A methylation profile refers to the measurement of DNA methylation at multiple sites throughout the genome. This can either be at individual CpG dinucleotides as a level profile (panel **d**), or by the measurement of multiple methylation patterns (panel **e**).

*The bisulphite revolution.* The analysis of DNA methylation was revolutionized by the introduction of sodium bisulphite conversion of genomic DNA (FIG. 2). 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.* showed the usefulness of this chemical reaction in conjunction with PCR amplification and sequencing[32] 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 (FIG. 2b).

*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[40,41] and of primer design[42]. 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 (FIG. 2b). Inclusion of such bases in the primer design helps to avoid amplification of any residual unconverted DNA[43]. Third, the sequence variation that reflects the methylation status at CpG dinucleotides in the original genomic DNA offers two approaches to PCR primer design (FIG. 2c). 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 (FIG. 2c). This method is used in bisulphite genomic sequencing[32], COBRA[44], Ms-SNuPE[45] and several other techniques (TABLE 1). 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 (FIG. 2c). This is the basis for MSP[33], 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.

The proliferation of DNA methylation analysis technologies (TABLE 1) has prevented the development of uniform standards and has made cross-validation studies problematic. At the same time, it should be emphasized that no one technique or general approach is universally superior, as the competing goals of quantitative accuracy, sensitive detection, high local or global informational content, compatibility with formalin-fixed tissues and compatibility with automation are not all found in a single technique. Therefore, the method of choice will depend on the desired application.

*Marker discovery.* How do we identify the DNA methylation markers that will be of greatest clinical use among the tens of millions of CpG dinucleotides and tens of thousands of gene-associated CpG islands[46]? The techniques described above can be used to analyse methylation differences between two groups of clinical DNA samples. The nature of these two groups will determine, in part, the type of methylation markers that arise from such an analysis. Researchers are often initially interested in identifying differences between tumour tissue and equivalent histologically normal tissue from the same patient. This can be extended to include genomic DNA

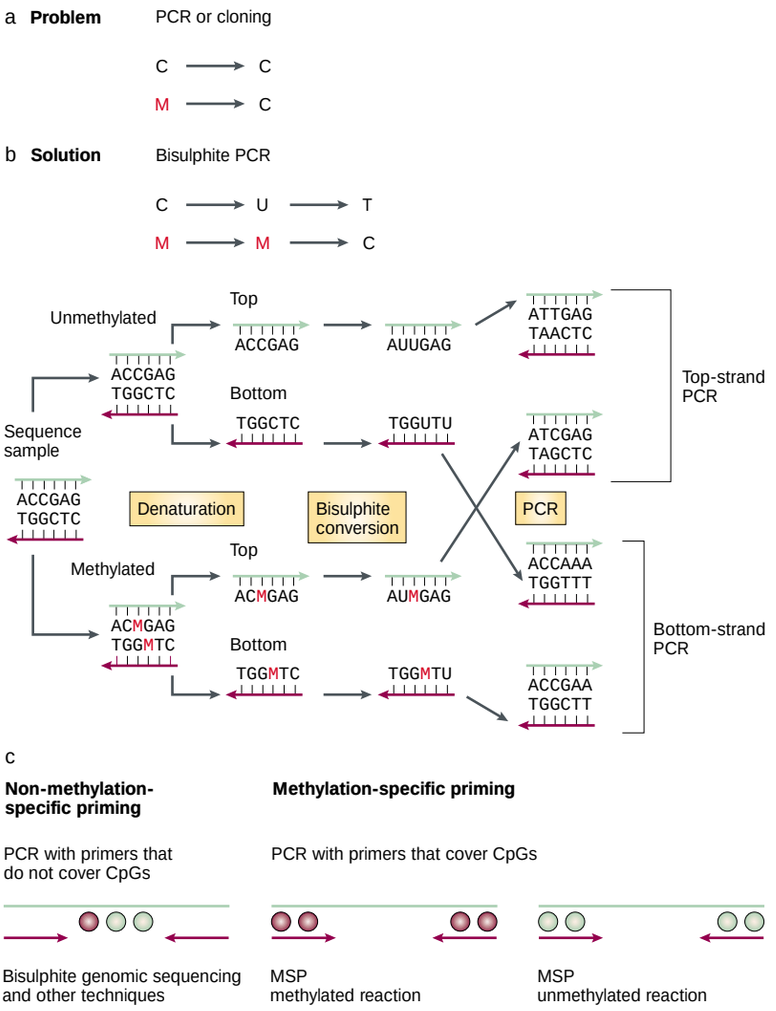


Figure 2 | **Principle of sodium bisulphite conversion.** Standard molecular biology techniques to analyse individual gene loci, such as polymerase chain reaction (PCR) and biological cloning, erase DNA methylation information, leaving the investigator oblivious to the epigenetic information that was present in the original genomic DNA (panel **a**). 5-methylcytosine residues are indicated as red Ms. 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 (not shown), or by treating the genomic DNA with sodium bisulphite (panel **b**), which converts unmethylated cytosines to uracil residues. As a consequence, the converted DNA is no longer self-complementary, and amplification of either the top or bottom DNA strand requires different primers. Priming can be either universal, or methylation specific (panel **c**). MSP, methylation-specific PCR.

samples from different stages of disease progression, such as metaplasia, dysplasia, carcinoma or metastatic tissue. This type of screening approach can lead to the identification of methylation markers that are useful for the sensitive detection of disease or markers associated with disease progression. It is important to keep in mind that these studies generally do not include samples of histologically normal tissue from control individuals, as such tissues are difficult to obtain. If methylation abnormalities arise early in normal tissues, leading to systemic or regional epigenetic defects, then a comparison between histologically normal tissues from cancer patients and healthy controls could lead to the identification of methylation markers that are useful in risk assessment. Indeed, it has been reported that loss of genomic imprinting (LOI) occurs at increased frequency in normal cells, including blood cells, of individuals with colorectal cancer, compared to controls[47]. This important finding opens the way to an exciting potential new approach to risk assessment.

An alternative to the more or less routine comparison between tumour and normal tissues is to analyse methylation differences among samples of identical tissue type or stage. Such stratification approaches can be used to reveal the existence of previously unrecognized subgroups, which is referred to as 'class discovery' or 'unsupervised learning'[31], or it can be used to distinguish between previously known subgroups, a strategy referred to as 'class discrimination' or 'supervised learning'[31]. Both of these approaches offer opportunities for the development of powerful new markers that have the ability to distinguish between samples or subjects that seem to be otherwise very similar. However, the screening of hundreds or thousands of markers on limited numbers of samples can lead to spurious associations, which do not hold up with subsequent testing in independent data sets. The judicious use of mathematical algorithms and the use of separate training and test data sets can help to reduce such false associations, although the fundamental problem remains that many such studies are underpowered. The stratification methods described here can be used to identify aetiological factors that are involved in the causation of methylation abnormalities. In this case, DNA methylation markers are treated as outcome variables. Alternatively, methylation markers can be assessed for their ability to predict clinical outcome. In this case, the methylation markers are themselves considered the source of phenotypic variation. DNA methylation markers are therefore uniquely positioned as intermediate phenotypes — reflective of genetic variation and environmental exposures on the one hand, but contributing actively to the phenotype on the other. DNA methylation markers are therefore useful both in epidemiological and in clinical studies.

DNA methylation markers have been developed using both targeted candidate gene approaches, as well as by systematic screening for markers using genome-scanning techniques (TABLE 1). Most of the genome-scanning techniques rely on methylation-sensitive restriction digestion. Although these techniques have proven to be powerful tools in the discovery of methylation markers, their reliance on methylation-sensitive restriction enzymes renders them poorly suited for the routine analysis of formalin-fixed clinical specimens. Therefore, most researchers in the field envision a boom period of marker discovery in the next few years, based on restriction enzyme technology, followed by a shift back to bisulphite-based methods of methylation analysis as assays are developed for use in the clinic. Ideally, a more robust method of methylation discrimination than either restriction digestion or bisulphite conversion would be developed for the ultimate clinical assays. The large-scale discovery of methylation markers for cancer has just begun, and it should be recognized that even the

Table 1a | DNA methylation detection and marker discovery techniques

| Technology | Methylation discrimination principle | Application | References |
|---|---|---|---|
| ***Methylation content*** | | | |
| Anti-mC immunological techniques | 5-Methylcytosine antibody | Global and regional content | 108,109 |
| Reversed phase HPLC — UV | HPLC separation | Global methylation analysis | 110 |
| Postlabelling/TLC | TLC separation | Global methylation analysis | 111 |
| LC — mass spectrometry | LC/MS separation | Global methylation analysis | 112 |
| *Sss*I methyl acceptance assay | DNA methyltransferase substrate | Global methylation analysis | 113,114 |
| Chloracetaldehyde reaction | Bisulphite conversion | Global methylation analysis | 115 |
| Nearest neighbour TLC | TLC separation | Global methylation analysis | 8 |
| HPLC-TLC | HPLC-TLC separation | Trace detection of methylcytosine | 116 |
| HPCE | Capillary electrophoretic separation | Global methylation analysis | 117 |
| ERMA | Bisulphite conversion | Regional methylation content | 118 |
| ***Methylation levels*** | | | |
| Southern blot | Restriction digestion | Quantitative analysis | 119,120 |
| Hydrazine/permanganate LM-PCR | Hydrazine/permanganate reactions | Quantitative analysis | 121 |
| *Hpa*II PCR | Restriction digestion | Qualitative analysis | 122 |
| Restriction LM-PCR | Restriction digestion | Quantitative analysis | 123 |
| Solid-phase primer extension | Restriction digestion | Quantitative analysis | 124 |
| Bisulphite sequencing direct | Bisulphite conversion | Quantitative analysis | 32 |
| MS-SNuPE | Bisulphite conversion | Quantitative analysis | 45 |
| COBRA | Bisulphite conversion | Quantitative analysis | 44 |
| Hemi-methylation assay | Restriction and bisulphite conversion | Quantitative analysis of hemi-methylation | 125 |
| Bisulphite primer extension IP-RP-HPLC | Bisulphite conversion | Quantitative analysis | 126 |
| McCOBRA | Bisulphite conversion | High-throughput analysis | 127 |
| SNaPmeth | Bisulphite conversion | Quantitative high-throughput analysis | 128 |
| PyroMeth | Bisulphite conversion | Quantitative high-throughput analysis | 128 |

COBRA, combined bisulphite restriction analysis; ERMA, enzymatic regional methylation assay; HPCE, high-performance capillary electrophoresis; HPLC, high-performance liquid chromatography; IP-RP-HPLC, ion pair-reverse-phase high-performance liquid chromatography; LC, liquid chromatography; LM-PCR, ligation-mediated PCR; mC, methylcytosine; McCOBRA, melting curve combined bisulphite restriction analysis; MS-SNuPE, methylation-sensitive single nucleotide primer extension; PCR, polymerase chain reaction; SNaPmeth, single nucleotide polymorphism methylation; TLC, thin-layer chromatography; UV, ultraviolet light.

genome-scanning techniques are confined by the distribution of the relevant restriction enzyme recognition motifs throughout the genome.

## 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 shown in FIG. 1, 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.

*Disease detection.* Early detection of disease results in an improved clinical outcome for most types of cancer. Therefore, much effort is being put into the development of early-detection strategies. DNA methylation changes have been reported to occur early in carcinogenesis and therefore are potentially good early indicators of existing disease[48], and even of risk assessment for the future development of disease. The first signs of cancer usually come from one or more of the following sources: presentation of symptoms, direct palpation or visual detection, histopathological analysis of a biopsy specimen, remote imaging or the detection of a cancer BIOMARKER in a tissue or bodily fluid specimen. For many types of solid malignancies, symptoms often do not arise until after the primary tumour has metastasized. Direct palpation, visual detection and biopsy analysis are generally limited to accessible sites of the body. For some types of cancer, such as ovarian, pancreatic and lung cancer, poor accessibility and the late presentation of symptoms thwart the timely detection of malignancy, contributing to high mortality rates. For such diseases, improved remote imaging, such as spiral computed tomography scanning, and the development of cancer biomarkers, offer the best hope for early detection.

BIOMARKER
In cancer research and detection, a biomarker refers to a substance or process that is indicative of the presence of cancer in the body. It might be either a molecule secreted by a malignancy itself, or it can be a specific response of the body to the presence of cancer.

Table 1b | DNA methylation detection and marker discovery techniques

| Technology | Methylation discrimination principle | Application | References |
|---|---|---|---|
| ***Methylation patterns*** | | | |
| Hydrazine/permanganate | Hydrazine/permanganate reactions | Informational content sequencing | 129,130 |
| Methylation-dependent DGGE | DNA melting characteristics | Complex informational analysis | 131,132 |
| Bisulphite sequencing clones | Bisulphite conversion | Informational content | 32 |
| MSP | Bisulphite conversion | Sensitive detection | 33 |
| Bisulphite-DGGE | Bisulphite conversion | Complex informational analysis | 132,133 |
| MSP-ISH | Bisulphite conversion | *In situ* detection | 134 |
| MethyLight | Bisulphite conversion | Quantitative high-throughput analysis | 54–56 |
| Bisulphite-SSCP | Bisulphite conversion | Complex informational analysis | 135–137 |
| MSP/DHPLC (not MSP-based!) | Bisulphite conversion | Complex informational analysis | 138 |
| In-tube fluorescence melting curve | Bisulphite conversion | Complex informational analysis | 139 |
| McMSP | Bisulphite conversion | High-throughput MSP | 127 |
| ConLight-MSP | Bisulphite conversion | Quantitative high-throughput analysis | 43 |
| HeavyMethyl | Bisulphite conversion | Quantitative high-throughput analysis | 140 |
| ***Methylation profiles*** | | | |
| RLGS | Restriction digestion | Marker discovery | 141 |
| Methylated DNA-binding column | MeCP2 affinity | Marker discovery | 142 |
| Methylation-sensitive AP-PCR | Restriction digestion | Marker discovery | 143 |
| MCA | Restriction digestion | Marker discovery | 144 |
| DMH | Restriction digestion | Marker discovery | 99 |
| ICEAMP | Methyl-CpG binding column | Marker discovery | 145 |
| *Not*I CODE | Restriction digestion | Genome analysis | 103 |
| SMR/DMR cloning | Restriction digestion | Marker discovery | 14 |
| MSO | Bisulphite conversion | Disease stratification | 146 |
| Epigenomics microarray | Bisulphite conversion | Disease stratification | 35 |
| AIMS | Restriction digestion | Marker discovery | 147 |
| MAD | Restriction digestion | Marker discovery | 104 |
| Two-colour oligo array | Bisulphite conversion | Localized information | 105 |
| Demethylation/expression arrays | 5-aza-CdR activation | Marker discovery | 148,149 |
| Size fractionation profiling | Restriction digestion | Marker discovery | 150 |

AIMS, amplification of inter-methylated sites; AP-PCR, arbitrarily primed PCR; CODE, cloning of deleted sequences; ConLight-MSP, conversion-specific MethyLight MSP; DGGE, denaturing gradient gel electrophoresis; DHPLC, denaturing high-performance liquid chromatography; DMH, differential methylation hybridization; ICEAMP, identification of CpG islands exhibiting altered methylation patterns; MAD, methylation amplification DNA chip; MCA, methylated CpG island amplification; McMSP, melting curve methylation-specific PCR; MSO, methylation-specific oligonucleotide array; MSP, methylation-specific PCR; MSP-ISH, methylation-specific PCR *in situ* hybridization; PCR, polymerase chain reaction; RLGS, restriction landmark genome scanning; SMR/DMR, similarly methylated regions/differentially methylated regions; SSCP, single-strand conformational polymorphism.

The efficacy of a biomarker assay is determined by its sensitivity and specificity. Unfortunately, these terms are not always used consistently. Experimental researchers might use the term 'sensitivity' to describe the minimal amount of analyte that an assay can detect in a laboratory test. This analytical sensitivity can be measured as an absolute sensitivity or as a relative sensitivity. For example, the absolute sensitivity of a DNA hypermethylation assay might refer to the minimal quantity of pure, methylated target DNA that the assay is able to detect, whereas the relative sensitivity would refer to the smallest fraction of methylated DNA that the assay could detect in the presence of an excess of unmethylated DNA. This latter measurement is sometimes also referred to as 'specificity'.

The terms 'sensitivity' and 'specificity' take on precise meanings in the development of biomarker tests for population-based screening, or for clinic-based surveillance of high-risk populations (TABLE 2). Here, the clinical sensitivity of a biomarker refers to the proportion of case subjects (individuals with confirmed disease) who test positive for the biomarker assay, whereas the specificity refers to the proportion of control subjects (individuals without the disease)

Table 2 | Sensitivity and specificity of methylation markers

| Methylation marker detection in tumour | Methylation marker detection in remote sample | Classification |
|---|---|---|
| + (Case) | + | True positive (TP) |
| N/A (Control) | – | True negative (TN) |
| N/A (Control) | + | False positive (FP) |
| + (Case) | – | False negative (FN) |
| – (Case) | + | Methylation false positive (MFP) (true positive for disease detection) |
| – (Case) | – | Methylation true negative (MTN) (false negative for disease detection) |

Subjects with clinically confirmed cancer are described as cases. The first column represents the detection of a methylation biomarker in the primary tumour in cases of confirmed cancer. However, this measurement is not applicable (N/A) to control individuals with no evidence of malignant disease. The second column represents the detection of the same methylation biomarker in a remote sample, which might be serum, plasma or another bodily fluid, or a tissue sample not taken from the histologically defined primary tumour itself. Clinical sensitivity (sensitivity of disease detection) is defined as the ratio of the number of individuals in the classifications (TP + MFP)/(TP + FN + MFP + MTN). Analytical sensitivity (sensitivity of methylation detection) is defined here as the ratio of the number of individuals in the classifications TP/(TP + FN). It should be emphasized that this definition differs from the use of the term 'analytical sensitivity' in an *in vitro* laboratory simulation of assay sensitivity. The definition described in this table is a composite of the *in vitro* analytical sensitivity efficiency of the test and the efficiency of transfer of the methylation marker from the primary tumour to the remote sample. Specificity is defined as the ratio of the number of individuals in the classifications TN/(TN + FP).

who test negative for the biomarker assay[49]. Clinical sensitivity is a composite of the marker prevalence in the tumour, the efficiency of transfer of the marker to the remote media being tested, and the analytical sensitivity of the assay. The tumour prevalence of many methylation markers is considerably higher than that of genetic markers, such as tumour-suppressor gene and proto-oncogene mutations and deletions. For example, methylation of *GSTP1* occurs in more than 90% of cases of prostate cancer[50,51]. Methylation of the *APC* tumour-suppressor gene 5′ CpG island was found in more than 90% of cases of oesophageal adenocarcinoma[52], but was also observed in normal gastric mucosa[52,53]. However, as gastric epithelial cells do not normally release DNA into the bloodstream, this marker still achieved 100% specificity in plasma[52].

Virtually all strategies for the sensitive detection of cancer-specific DNA methylation patterns rely on the principle of MSP[33], or fluorescence-based variants, such as MethyLight[54–56]. Cancer-specific DNA methylation patterns can be found in detached tumour cells in bodily fluids and biopsies, and they can be detected in free-floating DNA that is released from dead cancer cells[57]. It was shown more than 25 years ago that cancer patients have increased levels of free DNA in their serum[58], which is thought to be released from apoptotic or necrotic tumour cells[59,60]. This principle is the basis for a rapidly expanding number of studies of DNA methylation markers in blood serum and plasma (TABLE 3). Blood-borne tumour-derived DNA is often detected more frequently in patients with advanced-stage disease[52,61–65], although not all studies have reported such a correlation[66–72]. The presence of detectable tumour DNA in the plasma or serum is generally associated with a poor prognosis[52,61,64,73]. An alternative approach is to screen for tumour-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 (TABLES 3 and 4).

The methylation biomarker studies performed so far (TABLES 3 and 4) vary in methylation targets, source of DNA and type of tumour. Therefore, it is too early to give accurate estimates of analytical and clinical sensitivity of these approaches. Nevertheless, three preliminary insights can be gleaned from these studies. First, targeted luminal sources of DNA tend to give higher clinical sensitivities than serum or plasma analysis. Second, the specificity of plasma or serum detection of tumour-specific markers seems to be remarkably high (TABLE 3). Of the 599 CpG island hypermethylation analyses of plasma or serum samples from controls (325 independent subjects), all were negative for the methylation marker, resulting in an overall specificity of 100% (TABLE 3). Third, the analytical sensitivity of serum or plasma detection across all sites and all markers with confirmed presence in the tumour was 57% (502 out of 881) (TABLE 3). Therefore, if near-universal marker prevalence could be achieved, it is anticipated that the clinical sensitivity of serum or plasma methylation markers, using current technology, would be approximately 50%. Presumably, the clinical sensitivity of analysis of luminal fluids or cells will surpass that of systemic blood detection for most types of tumour. It is expected that markers with high prevalence will emerge for each type of cancer, as large numbers of markers are screened in marker discovery assays, as described above. Clinical sensitivity is also enhanced by the application of panels of multiple markers (TABLES 3 and 4). It is anticipated that such panels will provide more diagnostic information and cancer specificity than single-marker assays. Additional gains in analytical sensitivity can be expected to come from optimization of MSP primers and MethyLight probes, and from other advances, such as the use of nested PCR primers[74].

All the studies shown in TABLES 3 and 4 used MSP-based detection of CpG island hypermethylation — a method that can detect heavily methylated molecules against a background of excess unmethylated or partially methylated molecules (FIGS 1,2). DNA methylation technologies that measure methylation levels at individual CpG dinucleotides tend to be much less specific, and have not been successfully applied as sensitive biomarker assays. Conventional MSP produces data as a categorical variable (a positive or negative result for each sample for a particular reaction). For this type of measurement, the performance of a biomarker assay can be described simply by its clinical sensitivity and specificity. However, MethyLight technology combines the specificity of MSP with the ability to generate quantitatively accurate data as a continuous variable[54]. This data can then be dichotomized at a threshold to determine the sensitivity and specificity. However, the choice of threshold will affect these two variables differentially. Setting a higher threshold for a cancer-specific marker will increase the specificity of the assay, but reduce its sensitivity. Therefore, methods have been developed to describe the performance of a biomarker assay that take

*APC*
(Adenomatous polyposis coli). A tumour-suppressor gene that encodes a component of the WNT–β-catenin signalling pathway that is inactivated in most colorectal adenocarcinomas.

both measurements into account simultaneously. A commonly used statistical tool is to calculate the area under the Receiver Operating Characteristic (ROC) curve[75]. This curve can be plotted as the true positive rate against the false positive rate, in which each point in the curve represents the fraction of cancer cases with a biomarker measurement above a threshold (true-positive rate for that threshold) versus the corresponding fraction of control subjects above the same threshold (false-positive rate for that threshold)[49]. The area under the ROC curve then represents the probability that a randomly chosen cancer case is ranked as more likely to have the disease than a randomly chosen non-diseased control subject. This value is a useful way to describe the performance of a biomarker with a continuous output variable, regardless of the threshold level, and is identical to the non-parametric Wilcoxon statistic[75].

Although specificity does not seem to be an issue in serum and plasma detection (as long as the marker is not present in haematopoietic cells), it is more of a problem in other remote media, particularly in the analysis of samples from individuals at high risk or with premalignant stages of disease (TABLE 4). The development of new markers using the screening techniques described above is expected to yield assays that are capable of distinguishing between premalignant and malignant stages of disease. However, an inherent problem arises in the development of biomarkers with greater sensitivity than existing diagnostic methods. If the new biomarker is able to detect a cancer case earlier than with any other existing technology, then this case will be classified as a false positive, as the individual is judged to be disease-free — a classic Catch-22 situation. This quandary can be resolved by longitudinal prospective trials, in which such individuals are followed to see whether they subsequently develop overt disease. Such trials can take many years, and even then, it is difficult to discriminate between high-risk individuals with pre-malignant disease who have progressed to cancer, versus those that already had occult malignancy. If the risk of progression is very high for individuals with positive biomarkers, regardless of whether the disease is already present, then being able to distinguish between occult disease and a high risk for progression to malignancy might not matter for the clinical management of such patients.

Another confounding category is the group of cancer cases with positivity for a tumour-specific methylation marker in the remote media, but in which the marker is ostensibly lacking in the primary tumour. Palmisano *et al.* report a case in which *MGMT* methylation was undetectable in the primary lung squamous-cell carcinoma, yet was positive in sputum of the same patient[74]. Similarly, Jeronimo *et al.* describe a prostate cancer patient with *GSTP1* methylation that was detectable in the urine, but not in the primary tumour[76]. Such cases are referred to as methylation false positives (MFPs), but they are actually true positives for disease detection, which is what counts in the calculation of clinical sensitivity (TABLE 2). Such findings are likely to be more common for luminal cells or fluids in high-risk individuals, in whom abnormal methylation might occur in premalignant epithelium that did not give rise to the malignancy, as opposed to serum assays of low-risk individuals, in whom methylation markers seem to have very high specificity (TABLE 3). If serum positivity for a highly specific methylation marker is observed in cancer cases that are negative for the marker in the primary tumour, then this could still be consistent with a true-positive classification if the serum marker is present because of tumour heterogeneity, a second primary tumour or occult metastases in which the marker has arisen. Alternatively, it could indeed be a false-positive result that reflects the presence of this marker in the blood of some subjects, independent of disease, or that is caused by an experimental error in the assay method, such as in the case of PCR contamination.

Recently, Sullivan Pepe *et al.* have proposed that biomarker development for early detection of cancer should be conceptualized as occurring in five consecutive phases[49] (see review by Hartwell *et al.* on page 243 in this issue). Most studies using DNA methylation markers for cancer detection are at Phase 2 at present — the development and validation stage of clinical assays. Some reports have shown preliminary data from Phase 3 retrospective, longitudinal studies[74].

Methylation markers offer many exciting opportunities for the clinical application of sensitive cancer detection. The analysis of readily obtainable bodily fluids, such as blood or urine, could form the basis of a population-based screening programme if the markers can be developed to have sufficient sensitivity and specificity. Although current technologies for the detection of DNA methylation markers in bodily fluids have fairly low sensitivity (TABLES 3 and 4), they excel in specificity. This is an asset in population-based screening approaches, in which the clinical follow-up of false positives can be costly and invasive. Therefore, methylation markers could be very useful as an ancillary tool in risk assessment or disease detection, by enhancing the specificity of existing screening methods with low specificity, such as prostate-specific antigen screening for prostate cancer. Such applications are not necessarily confined to the analysis of bodily fluids. The analysis of biopsy tissue specimens obtained from high-risk individuals participating in surveillance programmes, such as for Barrett's oesophagus, can be useful in risk assessment or the detection of disease progression. DNA methylation markers might be particularly useful as harbingers of impending malignancy in histologically normal tissue. Methylation of the *MLH1* promoter CpG island in normal colonic mucosa was found to be associated with the presence of microsatellite-unstable (MSI+) colorectal adenocarcinomas[77]. Loss of normal genomic imprinting patterns has been reported not only in cancer cells, but also at an increased frequency in normal cells of individuals with cancer[47]. The clinical applications of sensitive DNA methylation detection technology are not limited to early detection or risk assessment. Serum analysis could also be used to monitor patients for disease recurrence after surgical resection or to assess the efficacy of chemotherapy, particularly if marker positivity can be established *a priori* in the resected primary tumour.

Table 3 | Sensitive detection of cancer in plasma and serum using DNA methylation markers

| Disease | DNA source | Markers | Analytical sensitivity | Clinical sensitivity | Specificity | References |
|---|---|---|---|---|---|---|
| Bladder cancer | Plasma | *CDKN2A* (*ARF*) | 13/15 (87%) | 13/27 (48%) | N/A | 64 |
| | Plasma | *CDKN2A* (*INK4A*) | 2/5 (40%) | 2/27 (7%) | N/A | 64 |
| | Serum | *CDKN2A* (*INK4A*) | 12/14 (86%) | 19/86 (22%) | 31/31 (100%) | 151 |
| Breast cancer | Plasma | *CDKN2A* (*INK4A*) | 5/8 (63%) | 5/35 (14%)* | N/A | 152 |
| | Plasma | *CDKN2A* (*INK4A*) | 6/10 (60%) | 6/43 (14%)* | N/A | 61 |
| Colorectal cancer | Serum | *MLH1* | 3/9 (33%) | 3/18 (17%) | N/A | 63 |
| | Serum | *CDKN2A* (*INK4A*) | 14/20 (70%) | 14/52 (27%) | 44/44 (100%)‡ | 65 |
| | Serum | *CDKN2A* (*INK4A*) | 13/44 (30%) | 13/94 (11%) | N/A | 69 |
| | Plasma | *CDKN2A* (*INK4A*) | 21/31 (68%) | 21/58 (36%) | N/A | 73 |
| Oesophageal cancer | Plasma (AC) | *APC* | 13/48 (27%) | 13/52 (25%) | 54/54 (100%)‡ | 52 |
| | Plasma (SCC) | *APC* | 2/16 (13%) | 2/32 (6%) | 54/54 (100%)‡ | 52 |
| | Serum (SCC) | *CDKN2A* (*INK4A*) | 7/31 (23%) | 7/38 (18%) | N/A | 153 |
| Gastric cancer | Serum | *CDH1* | 31/41 (76%) | 31/54 (57%) | 30/30 (100%) | 70 |
| | Serum | *CDKN2A* (*INK4A*) | 28/36 (78%) | 28/54 (52%) | 30/30 (100%) | 70 |
| | Serum | *CDKN2B* (*INK4B*) | 30/37 (81%) | 30/54 (56%) | 30/30 (100%) | 70 |
| | Serum | *DAPK1* | 26/38 (68%) | 26/54 (48%) | 30/30 (100%) | 70 |
| | Serum | *GSTP1* | 8/10 (80%) | 8/54 (15%) | 30/30 (100%) | 70 |
| | Serum | Panel of five | 45/54 (83%) | 45/54 (83%) | 30/30 (100%) | 70 |
| Head and neck cancer | Serum | *CDKN2A* (*INK4A*) | 8/26 (31%) | 8/95 (8%) | N/A | 68 |
| | Serum | *DAPK1* | 3/17 (18%) | 3/95 (3%) | N/A | 68 |
| | Serum | *MGMT* | 14/31 (45%) | 14/95 (15%) | N/A | 68 |
| | Serum | Panel of three | 21/52 (40%) | 21/95 (22%) | N/A | 68 |
| | Plasma (nasopharyngeal) | *DAPK1* | 6/12 (50%) | N/A | N/A | 71 |
| Liver cancer | Plasma/serum | *CDKN2A* (*INK4A*) | 13/16 (81%) | 13/22 (45%) | 48/48 (100%) | 66 |
| | Plasma/serum | *CDKN2B* (*INK4B*) | 4/16 (25%) | 4/25 (16%) | 35/35 (100%) | 154 |
| | Plasma/serum | Panel of two | 17/23 (74%) | 17/25 (68%) | 35/35 (100%) | 154 |
| Lung cancer | Serum (NSCLC) | *CDKN2A* (*INK4A*) | 3/9 (33%) | 3/22 (14%) | N/A | 67 |
| | Serum (NSCLC) | *DAPK1* | 4/5 (80%) | 4/22 (18%) | N/A | 67 |
| | Serum (NSCLC) | *GSTP1* | 1/2 (50%) | 1/22 (5%) | N/A | 67 |
| | Serum (NSCLC) | *MGMT* | 4/6 (67%) | 4/22 (18%) | N/A | 67 |
| | Serum (NSCLC) | Panel of four | 11/15 (73%) | 11/22 (50%) | N/A | 67 |
| | Plasma | *CDKN2A* (*INK4A*) | 1/10 (10%) | N/A | N/A | 155 |
| | Plasma/serum | *APC* | N/A | 42/89 (47%) | 50/50 (100%) | 89 |
| | Plasma | *CDKN2A* (*INK4A*) | 64/73 (88%) | 77/105 (73%)§ | N/A | 72 |
| | Plasma (NSCLC) | *CDKN2A* (*INK4A*) | 12/22 (55%) | 12/35 (34%) | 15/15 (100%) | 156 |
| Prostate cancer | Plasma/serum | *GSTP1* | 12/16 (75%) | 23/33 (70%) | 22/22 (100%)‡ | 62 |
| | Plasma | *GSTP1* | 25/63 (40%) | 25/69 (36%) | 31/31 (100%) ‡ | 76 |

Analytical sensitivity is defined as the fraction of cases in which methylation of a marker is found in serum or plasma for cases with confirmed methylation of the same marker in the associated tumour (see TABLE 2). Clinical sensitivity is defined as the fraction of confirmed cases of disease, in which methylation of a marker is found in serum or plasma, regardless of whether methylation of that marker is present in the associated tumour, or regardless of whether the associated tumour has been analysed for the presence of the marker. Cases in which serum or plasma are not analysed are excluded from both sensitivity calculations. Cases of failed analyses of serum or plasma are included in the calculation of sensitivities, as these failures are largely due to limiting amounts of retrievable DNA in the serum or plasma, which should be factored into the calculation of the sensitivity of the technology. Specificity is defined as the fraction of controls without the disease that show a lack of detectable methylation in serum or plasma. In some cases, this includes control individuals that are known to be at high risk for the disease, or with premalignant conditions, as indicated‡. Sensitivities and specificities are listed as ratios of the actual number of cases, followed by the percentages in parentheses. These percentages have been newly calculated for this review, as some of the original publications did not perform these calculations, or did not use the same criteria for sensitivities and specificity. * Possible overlap in samples. ‡Included high-risk individuals or cases with non-malignant disease. §Nested methylation-specific polymerase chain reaction. AC, adenocarcinoma; N/A, not available; NSCLC, non-small-cell lung cancer; SCC, squamous-cell carcinoma.

*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 tumour[78]. 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 tumour response to 5-fluorouracil, but might also suffer more side effects from the drug[79]. 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 behaviour of the malignancy[80]. Both of these types of stratification markers are of clinical value and can assist physicians in their choice of treatment.

The presence of tumour-specific methylation markers in the serum or plasma of patients has been reported to be of prognostic significance[52,61,64,73,81]. However, the presence of a serum or plasma methylation marker is merely indicative of the release of sufficient amounts of

Table 4 | Sensitive detection of cancer in remote media using DNA methylation markers

| Disease | DNA source | Markers | Analytical sensitivity | Clinical sensitivity | Specificity | References |
|---|---|---|---|---|---|---|
| Bladder cancer | Urine sediment | *CDH1* (*E-cadherin*) | 13/17 (76%) | 13/22 (59%) | 17/17 (100%) | 157 |
| | Urine sediment | *CDKN2A* (*INK4A*) | 3/4 (75%) | 3/22 (14%) | 17/17 (100%) | 157 |
| | Urine sediment | *DAPK1* | 10/17 (58%) | 10/22 (45%) | 17/17 (100%) | 157 |
| | Urine sediment | *RARB* (*RAR-β*) | 15/20 (75%) | 15/22 (68%) | 13/17 (76%) | 157 |
| | Urine sediment | Panel of four | 20/22 (91%) | 20/22 (91%) | 13/17 (76%) | 157 |
| Breast cancer | Ductal lavage cells* | *CCND2* (*Cyclin D2*) | N/A | 8/19 (42%) | 42/45 (93%)‡ | 158 |
| | Ductal lavage cells* | *RARB* (*RAR-β*) | N/A | 12/20 (60%) | 43/45 (96%)‡ | 158 |
| | Ductal lavage cells* | *TWIST* | N/A | 13/18 (72%) | 35/35 (100%)‡ | 158 |
| | Ductal lavage cells* | Panel of three | 17/19 (89%) | 17/20 (85%) | 40/45 (89%)‡ | 158 |
| Head and neck cancer | Saliva | *CDKN2A* (*INK4A*) | 11/14 (79%) | 11/30 (37%) | 29/30 (97%)‡ | 159 |
| | Saliva | *DAPK1* | 6/10 (60%) | 6/30 (20%) | 30/30 (100%)‡ | 159 |
| | Saliva | *MGMT* | 4/7 (57%) | 4/30 (13%) | 29/30 (97%) | 159 |
| | Saliva | Panel of three | 11/17 (65%) | 11/30 (37%) | 30/30 (100%)‡ | 159 |
| | Buffy coat (nasopharyngeal) | *DAPK1* | 3/12 (25%) | N/A | N/A | 71 |
| Lung cancer | Sputum | *CDKN2A* (*INK4A*) | N/A | 3/7 (43%) | 21/26 (81%)‡ | 160 |
| | Broncheoalveolar lavage | *CDKN2A* (*INK4A*) | 12/19 (63%) | 12/50 (24%) | N/A | 161 |
| | Sputum | *CDKN2A* (*INK4A*) | 18/26 (69%) | 18/51 (35%) | 21/25 (84%)‡ | 162 |
| | Broncheoalveolar lavage | *CDKN2A* (*INK4A*) | 11/26 (42%) | 11/51 (22%) | 22/25 (88%)‡ | 162 |
| | Brushings | *CDKN2A* (*INK4A*) | 8/26 (31%) | 8/51 (16%) | 23/25 (92%)‡ | 162 |
| | Sputum | *CDKN2A* (*INK4A*) | 8/8 (100%)§ | 8/10 (80%)§ | N/A¶ | 74 |
| | Sputum | *MGMT* | 5/6 (N/A)§,‖ | 5/10 (50%)§,‖ | N/A¶ | 74 |
| | Sputum | Panel of two | 10/10 (100%)§ | 10/10 (100%)§ | N/A¶ | 74 |
| | Bronchial washings | *CDKN2A* (*INK4A*) | 4/17 (24%) | N/A | 10/10 (100%) | 155 |
| | Bronchial epithelium | *CDKN2A* (*INK4A*) | N/A | 23/52 (44%)§ | 23/41 (56%)‡,§ | 163 |
| | Bronchial epithelium | *DAPK1* | N/A | 5/52 (10%)§ | 37/41 (90%)‡,§ | 163 |
| | Sputum | *CDKN2A* (*INK4A*) | N/A | N/A | 43/66 (65%)‡,§ | 163 |
| | Sputum | *DAPK1* | N/A | N/A | 50/66 (78%)‡,§ | 163 |
| Prostate cancer | Ejaculate | *GSTP1* | N/A | 4/9 (44%) | N/A | 164 |
| | Buffy coat | *GSTP1* | 5/16 (31%) | 10/33 (30%) | 26/26 (100%)‡ | 62 |
| | Urine sediment (PPM) | *GSTP1* | 2/11 (18%) | 4/16 (25%) | 10/10 (100%)‡ | 62 |
| | Ejaculate | *GSTP1* | 1/5 (20%) | 4/8 (50%) | 6/6 (100%)‡ | 62 |
| | Urine sediment | *GSTP1* | 6/22 (27%) | 6/28 (21%) | N/A | 165 |
| | Urine sediment (PPM) | *GSTP1* | N/A | 29/40 (73%) | 44/45 (98%)‡ | 166 |
| | Biopsy washing | *GSTP1* | N/A | 10/10 (100%) | 10/10 (100%) | 167 |
| | Urine | *GSTP1* | 21/63 (33%) | 21/69 (30%) | 30/31 (97%)‡ | 76 |

Remote media are defined in this table as DNA sources not obtained directly from the tumour. These can include bodily fluids, lavages, detached cells, biopsies of non-malignant tissues and so on. Serum and plasma detection are shown in TABLE 3 and are therefore excluded from this table. See the footnotes of TABLE 3 for the definitions of sensitivity. Specificity is defined as the fraction of controls without the disease that show a lack of detectable methylation in the remote media. In some cases, this includes control individuals that are known to be at high risk for the disease, or with premalignant conditions, as indicated‡. Sensitivities and specificities are listed as ratios of the actual number of cases, followed by the percentages in parentheses. These percentages have been newly calculated for this review, as some of the original publications did not perform these calculations, or did not use the same criteria for sensitivities and specificity. *Obtained by ROBE (routine operative breast endoscopy[158]). ‡Included high-risk individuals or cases with non-malignant disease. §Nested methylation-specific polymerase chain reaction. ‖Detection of methylation in sputum/urine, but not in tumour (false positive). ¶This study reported the detection of methylation of either *CDKN2A* or *MGMT* in 100% of sputum samples from at-risk individuals up to 3 years before the diagnosis of cancer. N/A, not available; PPM, post-prostatic massage.

tumour DNA into the bloodstream, which is likely to be correlated with invasiveness. For this purpose, any tumour-specific methylation marker would suffice, and its presence in the tumour 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[82–92]. 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[84,86,88–90,92,93].

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[82,94,95]. The most extensive work has been done with CpG island hypermethylation of the $O^6$-methylguanine methyltransferase (*MGMT*) gene. Esteller *et al.* reported that *MGMT* methylation was associated with prolonged survival in glioma patients who were treated with carmustine[94], and in patients with diffuse large B-cell lymphoma who were treated with cyclophosphamide[95], as part of multidrug regimens. Others have expressed

**Box 1 | The CIMP phenotype**

As studies started to incorporate the analysis of multiple DNA methylation markers, Jean-Pierre Issa coined the term *CIMP* — an acronym for 'CpG island methylator phenotype' — to describe colorectal tumours with concurrent methylation of a number of cancer-specific markers. The CIMP phenomenon was subsequently reported for other types of tumour as well, and was found to be associated with distinct genetic and histopathological characteristics. However, some studies did not find evidence for such a separate cluster of tumours with excessive CpG island hypermethylation. At issue is whether CIMP represents a distinct grouping or merely the far end of a continuous spectrum. This is of great interest, because a distinct CIMP phenotype would indicate the existence of a systemic methylation defect in the cell, which could provide insight into the molecular mechanism of aberrant CpG island hypermethylation. This gets to the heart of how CpG island hypermethylation arises. On the one hand, the methylation changes could represent stochastic fluctuations in methylation states that provide a selective advantage when they result in the silencing of a growth-controlling gene. Alternatively, a defect in the transacting factors that are required for the protection of CpG islands from encroaching methylation could lead to the concurrent methylation of multiple CpG islands throughout the genome, including many for which the hypermethylation event offers no selective advantage to the cell, which is analogous to the mutator phenotype in mismatch-repair-deficient tumours. If such systemic defects exist, then they might come in various flavours. Different defects might lead to the hypermethylation of distinct groups of structurally similar CpG islands. Unsupervised clustering of hypermethylation events will reveal whether concordant methylation occurs in reproducible clusters, and, if so, whether the identity of the clustered events indicates structural similarities of the CpG islands, or functional cooperativity in particular oncogenic pathways.

reservations regarding the conclusions that can be drawn from such retrospective multidrug studies[96,97]. The identification of true predictive markers requires that care is 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 tumour 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[34,35,37,38,82,98,99]. DNA methylation patterns are a rich source of information, rivalling 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 seem to outnumber informative genetic alterations. The problem is not the lack of information in the epigenome, but the difficulty in accessing that information. Many of the DNA methylation profiling techniques that are used in global approaches rely on restriction enzyme digestion, which limits their clinical use, and all current sodium-bisulphite-based PCR approaches require the subsequent independent amplification of each individual marker (with some modest multiplexing possible), which limits their throughput capability. Nevertheless, the promise of DNA methylation profiles is so great that much effort is being put into the development of high-throughput multi-marker approaches to DNA methylation analysis[34,35,37,38,82,98–106]. As multi-gene approaches are developed, DNA methylation researchers are encountering many of the same bioinformatics issues that confront gene-expression microarray studies[31,107]. Several examples of hierarchical clustering using DNA methylation markers have been published[35,39,82]. However, most multi-marker methylation studies reported so far have not properly adjusted for multiple hypothesis testing, nor have multi-marker models been tested on independent data sets[31]. The nature of the CpG island methylator phenotype (CIMP) is a particularly interesting statistical issue in multi-marker DNA methylation studies (BOX 1).

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 tumour. 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.

## Future directions

In recent years, it has become apparent that cancer is as much a disease of misdirected epigenetics as it is a disease of genetic mutations. The next few years will bring a better understanding of the mechanisms that underlie the DNA methylation changes in cancer cells. We will also witness an explosion in the number of cancer-specific methylation markers, assisted by *in silico* tools, followed by a weeding process, ultimately resulting in small panels of markers that are optimized for particular clinical applications. As we gain insight into the functional significance of methylation changes, there will be a push to manipulate methylation patterns, initially using blunt, global tools, such as DNA methyltransferase inhibitors, but ultimately using targeted therapeutic agents.

However, in the short term, the greatest promise of DNA methylation markers lies in the power of diagnostics, in which panels of markers will provide a tool for risk assessment, early detection, molecular diagnostics of resected specimens, chemoprediction and monitoring for disease recurrence. Consider the following scene from a clinic in the future. As part of an annual routine physical examination of a patient, 10 ml of blood are drawn for a DNA methylation screen with a panel of 20 of the most common cancer-specific methylation markers. Three of the markers come up positive, and the sample is then selected for further analysis with more-specific, diagnostic methylation panels to determine the most likely tissue of origin. Based on this information, the physician uses sophisticated imaging technology and other non-invasive, or minimally invasive techniques to confirm the diagnosis. An incipient pancreatic carcinoma is caught before it has spread.


1. Robertson, K. D. & Wolffe, A. P. DNA methylation in health and disease. *Nature Rev. Genet.* **1**, 11–19 (2000).
2. Ordway, J. M. & Curran, T. Methylation matters: modeling a manageable genome. *Cell Growth Differ.* **13**, 149–162 (2002).
3. Li, E. Chromatin modification and epigenetic reprogramming in mammalian development. *Nature Rev. Genet.* **3**, 662–673 (2002).
4. Li, E., Bestor, T. H. & Jaenisch, R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69**, 915–926 (1992).
5. Bird, A. DNA methylation patterns and epigenetic memory. *Genes Dev.* **16**, 6–21 (2002).
6. Walsh, C. P., Chaillet, J. R. & Bestor, T. H. Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. *Nature Genet.* **20**, 116–117 (1998).
**Provides experimental support for the importance of DNA methylation in the protection against intragenomic parasites in vertebrate genomes.**
7. Jones, P. A. & Baylin, S. B. The fundamental role of epigenetic events in cancer. *Nature Rev. Genet.* **3**, 415–428 (2002).
**An excellent overview of the field of gene silencing by aberrant promoter CpG island hypermethylation in cancer.**
8. Ramsahoye, B. H. *et al.* Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proc. Natl Acad. Sci. USA* **97**, 5237–5242 (2000).
9. Chan, M. F. *et al.* Reduced rates of gene loss, gene silencing, and gene mutation in DNMT1-deficient embryonic stem cells. *Mol. Cell. Biol.* **21**, 7587–7600 (2001).
10. Sved, J. & Bird, A. The expected equilibrium of the CpG dinucleotide in vertebrate genomes under a mutation model. *Proc. Natl Acad. Sci. USA* **87**, 4692–4696 (1990).
11. Issa, J. P. *et al.* Methylation of the oestrogen receptor CpG island links ageing and neoplasia in human colon. *Nature Genet.* **7**, 536–540 (1994).
12. Ahuja, N., Li, Q., Mohan, A. L., Baylin, S. B. & Issa, J. P. Aging and DNA methylation in colorectal mucosa and cancer. *Cancer Res.* **58**, 5489–5494 (1998).
**References 11 and 12 reported the age-related hypermethylation of CpG islands in histologically normal tissues.**
13. Futscher, B. W. *et al.* Role for DNA methylation in the control of cell type specific maspin expression. *Nature Genet.* **31**, 175–179 (2002).
**Shows that methylation of CpG islands can be important in the establishment and maintenance of cell-type-restricted gene expression.**
14. Strichman-Almashanu, L. Z. *et al.* A genome-wide screen for normally methylated human CpG islands that can identify novel imprinted genes. *Genome Res.* **12**, 543–554 (2002).
15. Yoder, J. A., Walsh, C. P. & Bestor, T. H. Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet.* **13**, 335–340 (1997).
16. Hoal-van Helden, E. G. & van Helden, P. D. Age-related methylation changes in DNA may reflect the proliferative potential of organs. *Mutat. Res.* **219**, 263–266 (1989).
17. Feinberg, A. P. & Vogelstein, B. Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature* **301**, 89–92 (1983).
**The first to describe DNA hypomethylation of single genes in human tumour DNA.**
18. Gama-Sosa, M. A. *et al.* The 5-methylcytosine content of DNA from human tumors. *Nucleic Acids Res.* **11**, 6883–6894 (1983).
**The first to describe genome-wide DNA hypomethylation in human tumour DNA.**
19. Jones, P. A. & Laird, P. W. Cancer epigenetics comes of age. *Nature Genet.* **21**, 163–167 (1999).
20. Rampersaud, G. C., Kauwell, G. P., Hutson, A. D., Cerda, J. J. & Bailey, L. B. Genomic DNA methylation decreases in response to moderate folate depletion in elderly women. *Am. J. Clin. Nutr.* **72**, 998–1003 (2000).
21. Kim, Y. I. *et al.* Effects of folate supplementation on two provisional molecular markers of colon cancer: a prospective, randomized trial. *Am. J. Gastroenterol.* **96**, 184–195 (2001).
22. Cravo, M. L. *et al.* Effect of folate supplementation on DNA methylation of rectal mucosa in patients with colonic adenomas: correlation with nutrient intake. *Clin. Nutr.* **17**, 45–49 (1998).
23. Friso, S. *et al.* A common mutation in the 5,10-methylenetetrahydrofolate reductase gene affects genomic DNA methylation through an interaction with folate status. *Proc. Natl Acad. Sci. USA* **99**, 5606–5611 (2002).
24. Paz, M. F. *et al.* Germ-line variants in methyl-group metabolism genes and susceptibility to DNA methylation in normal tissues and human primary tumors. *Cancer Res.* **62**, 4519–4524 (2002).
25. Esteller, M. CpG island hypermethylation and tumor suppressor genes: a booming present, a brighter future. *Oncogene* **21**, 5427–5440 (2002).
26. Plass, C. Cancer epigenomics. *Hum. Mol. Genet.* **11**, 2479–2488 (2002).
27. Di Croce, L. *et al.* Methyltransferase recruitment and DNA hypermethylation of target promoters by an oncogenic transcription factor. *Science* **295**, 1079–1082 (2002).
**The first paper to report the targeted recruitment of a DNA methyltransferase in the hypermethylation of a CpG island.**
28. Velicescu, M. *et al.* Cell division is required for *de novo* methylation of CpG islands in bladder cancer cells. *Cancer Res.* **62**, 2378–2384 (2002).
29. Song, J. Z., Stirzaker, C., Harrison, J., Melki, J. R. & Clark, S. J. Hypermethylation trigger of the glutathione-*S*-transferase gene (*GSTP1*) in prostate cancer cells. *Oncogene* **21**, 1048–1061 (2002).
30. Karpf, A. R. & Jones, D. A. Reactivating the expression of methylation silenced genes in human cancer. *Oncogene* **21**, 5496–5503 (2002).
31. Siegmund, K. D. & Laird, P. W. Analysis of complex methylation data. *Methods* **27**, 170–178 (2002).
32. Frommer, M. *et al.* A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc. Natl Acad. Sci. USA* **89**, 1827–1831 (1992).
**Revolutionized the field of DNA methylation analysis by introducing the sodium bisulphite conversion technique.**
33. Herman, J. G., Graff, J. R., Myohanen, S., Nelkin, B. D. & Baylin, S. B. Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. *Proc. Natl Acad. Sci. USA* **93**, 9821–9826 (1996).
**Introduced a technique that made DNA methylation analysis in human tumours widely accessible.**
34. Costello, J. F. *et al.* Aberrant CpG-island methylation has non-random and tumour-type-specific patterns. *Nature Genet.* **24**, 132–138 (2000).
**This paper made the first convincing case for the wide extent of DNA methylation changes in human tumours.**
35. Adorjan, P. *et al.* Tumour class prediction and discovery by microarray-based DNA methylation analysis. *Nucleic Acids Res.* **30**, E21 (2002).
36. Yan, P. S. *et al.* Applications of CpG island microarrays for high-throughput analysis of DNA methylation. *J. Nutr.* **132**, 2430S–2434S (2002).
37. Eads, C. A. *et al.* Epigenetic patterns in the progression of esophageal adenocarcinoma. *Cancer Res.* **61**, 3410–3418 (2001).
38. Esteller, M., Corn, P. G., Baylin, S. B. & Herman, J. G. A gene hypermethylation profile of human cancer. *Cancer Res.* **61**, 3225–3229 (2001).
39. Virmani, A. K. *et al.* Hierarchical clustering of lung cancer cell lines using DNA methylation markers. *Cancer Epidemiol. Biomarkers Prev.* **11**, 291–297 (2002).
40. Singal, R. & Grimes, S. R. Microsoft Word macro for analysis of cytosine methylation by the bisulfite deamination reaction. *Biotechniques* **30**, 116–120 (2001).
41. Anbazhagan, R., Herman, J. G., Enika, K. & Gabrielson, E. Spreadsheet-based program for the analysis of DNA methylation. *Biotechniques* **30**, 110–114 (2001).
42. Li, L. C. & Dahiya, R. MethPrimer: designing primers for methylation PCRs. *Bioinformatics* **18**, 1427–1431 (2002).
43. Rand, K., Qu, W., Ho, T., Clark, S. J. & Molloy, P. Conversion-specific detection of DNA methylation using real-time polymerase chain reaction (ConLight-MSP) to avoid false positives. *Methods* **27**, 114–120 (2002).
44. Xiong, Z. & Laird, P. W. COBRA: a sensitive and quantitative DNA methylation assay. *Nucleic Acids Res.* **25**, 2532–2534 (1997).
45. Gonzalgo, M. L. & Jones, P. A. Rapid quantitation of methylation differences at specific sites using methylation-sensitive single nucleotide primer extension (Ms-SNuPE). *Nucleic Acids Res.* **25**, 2529–2531 (1997).
46. Takai, D. & Jones, P. A. Comprehensive analysis of CpG islands in human chromosomes 21 and 22. *Proc. Natl Acad. Sci. USA* **99**, 3740–3745 (2002).
47. Cui, H., Horon, I. L., Ohlsson, R., Hamilton, S. R. & Feinberg, A. P. Loss of imprinting in normal tissue of colorectal cancer patients with microsatellite instability. *Nature Med.* **4**, 1276–1280 (1998).
**Provides the first indication that loss of imprinting might represent a systemic defect that is present even in haematopoietic cells of individuals with colorectal cancer.**
48. Laird, P. W. Oncogenic mechanisms mediated by DNA methylation. *Mol. Med. Today* **3**, 223–229 (1997).
49. Sullivan Pepe, M. *et al.* Phases of biomarker development for early detection of cancer. *J. Natl Cancer Inst.* **93**, 1054–1061 (2001).
**The first paper to provide a systematic approach to biomarker development for the early detection of cancer.**
50. Lee, W. H. *et al.* Cytidine methylation of regulatory sequences near the pi-class glutathione *S*-transferase gene accompanies human prostatic carcinogenesis. *Proc. Natl Acad. Sci. USA* **91**, 11733–11737 (1994).
51. Jeronimo, C. *et al.* Quantitation of GSTP1 methylation in non-neoplastic prostatic tissue and organ-confined prostate adenocarcinoma. *J. Natl Cancer Inst.* **93**, 1747–1752 (2001).
52. Kawakami, K. *et al.* Hypermethylated APC DNA in plasma and prognosis of patients with esophageal adenocarcinoma. *J. Natl Cancer Inst.* **92**, 1805–1811 (2000).
53. Eads, C. A. *et al.* Fields of aberrant CpG island hypermethylation in Barrett's esophagus and associated adenocarcinoma. *Cancer Res.* **60**, 5021–5026 (2000).
54. Eads, C. A. *et al.* MethyLight: a high–throughput assay to measure DNA methylation. *Nucleic Acids Res.* **28**, E32 (2000).
55. Eads, C. A. *et al.* CpG island hypermethylation in human colorectal tumors is not associated with DNA methyltransferase overexpression. *Cancer Res.* **59**, 2302–2306 (1999).
56. Lo, Y. M. *et al.* Quantitative analysis of aberrant p16 methylation using real-time quantitative methylation-specific polymerase chain reaction. *Cancer Res.* **59**, 3899–3903 (1999).

57. Sidransky, D. Emerging molecular markers of cancer. *Nature Rev. Cancer* **2**, 210–219 (2002).
58. Leon, S. A., Shapiro, B., Sklaroff, D. M. & Yaros, M. J. Free DNA in the serum of cancer patients and the effect of therapy. *Cancer Res.* **37**, 646–650 (1977).
**The first to describe the presence of increased levels of free DNA in the serum of cancer patients.**
59. Vasioukhin, V. *et al.* Point mutations of the N-ras gene in the blood plasma DNA of patients with myelodysplastic syndrome or acute myelogenous leukaemia. *Br. J. Haematol.* **86**, 774–779 (1994).
**Provided the first proof that the free DNA in serum was of tumour origin.**
60. Jahr, S. *et al.* DNA fragments in the blood plasma of cancer patients: quantitations and evidence for their origin from apoptotic and necrotic cells. *Cancer Res.* **61**, 1659–1665 (2001).
61. Silva, J. M. *et al.* Presence of tumor DNA in plasma of breast cancer patients: clinicopathological correlations. *Cancer Res.* **59**, 3251–3256 (1999).
62. Goessl, C. *et al.* Fluorescent methylation-specific polymerase chain reaction for DNA-based detection of prostate cancer in bodily fluids. *Cancer Res.* **60**, 5941–5945 (2000).
63. Grady, W. M., Rajput, A., Lutterbaugh, J. D. & Markowitz, S. D. Detection of aberrantly methylated hMLH1 promoter DNA in the serum of patients with microsatellite unstable colon cancer. *Cancer Res.* **61**, 900–902 (2001).
64. Dominguez, G. *et al.* p14ARF promoter hypermethylation in plasma DNA as an indicator of disease recurrence in bladder cancer patients. *Clin. Cancer Res.* **8**, 980–985 (2002).
65. Zou, H. Z. *et al.* Detection of aberrant p16 methylation in the serum of colorectal cancer patients. *Clin. Cancer Res.* **8**, 188–191 (2002).
66. Wong, I. H. *et al.* Detection of aberrant p16 methylation in the plasma and serum of liver cancer patients. *Cancer Res.* **59**, 71–73 (1999).
67. Esteller, M. *et al.* Detection of aberrant promoter hypermethylation of tumor suppressor genes in serum DNA from non-small cell lung cancer patients. *Cancer Res.* **59**, 67–70 (1999).
68. Sanchez-Cespedes, M. *et al.* Gene promoter hypermethylation in tumors and serum of head and neck cancer patients. *Cancer Res.* **60**, 892–895 (2000).
69. Nakayama, H. *et al.* Molecular detection of p16 promoter methylation in the serum of colorectal cancer patients. *Cancer Lett.* **188**, 115–119 (2002).
70. Lee, T. L. *et al.* Detection of gene promoter hypermethylation in the tumor and serum of patients with gastric carcinoma. *Clin. Cancer Res.* **8**, 1761–1766 (2002).
71. Wong, T. S. *et al.* High frequency of promoter hypermethylation of the death-associated protein-kinase gene in nasopharyngeal carcinoma and its detection in the peripheral blood of patients. *Clin. Cancer Res.* **8**, 433–437 (2002).
72. An, Q. *et al.* Detection of p16 hypermethylation in circulating plasma DNA of non-small cell lung cancer patients. *Cancer Lett.* **188**, 109–114 (2002).
73. Lecomte, T. *et al.* Detection of free-circulating tumor-associated DNA in plasma of colorectal cancer patients and its association with prognosis. *Int. J. Cancer* **100**, 542–548 (2002).
74. Palmisano, W. A. *et al.* Predicting lung cancer by detecting aberrant promoter methylation in sputum. *Cancer Res.* **60**, 5954–5958 (2000).
75. Hanley, J. A. & McNeil, B. J. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* **143**, 29–36 (1982).
76. Jeronimo, C. *et al.* Quantitative GSTP1 hypermethylation in bodily fluids of patients with prostate cancer. *Urology* **60**, 1131–1135 (2002).
77. Nakagawa, H. *et al.* Age-related hypermethylation of the 5′ region of MLH1 in normal colonic mucosa is associated with microsatellite-unstable colorectal cancer development. *Cancer Res.* **61**, 6991–6995 (2001).
78. van de Vijver, M. J. *et al.* A gene-expression signature as a predictor of survival in breast cancer. *N. Engl. J. Med.* **347**, 1999–2009 (2002).
79. Pullarkat, S. T. *et al.* Thymidylate synthase gene polymorphism determines response and toxicity of 5-FU chemotherapy. *Pharmacogenomics J.* **1**, 65–70 (2001).
80. Hayes, D. F., Trock, B. & Harris, A. L. Assessing the clinical impact of prognostic factors: when is 'statistically significant' clinically useful? *Breast Cancer Res. Treat.* **52**, 305–319 (1998).
81. Silva, J. M. *et al.* Persistence of tumor DNA in plasma of breast cancer patients after mastectomy. *Ann. Surg. Oncol.* **9**, 71–76 (2002).
82. Wei, S. H. *et al.* Methylation microarray analysis of late-stage ovarian carcinomas distinguishes progression-free survival in patients and identifies candidate epigenetic markers. *Clin. Cancer Res.* **8**, 2246–2252 (2002).
83. Tada, Y. *et al.* The association of death-associated protein kinase hypermethylation with early recurrence in superficial bladder cancers. *Cancer Res.* **62**, 4048–4053 (2002).
84. Powell, M. A. *et al.* Ribosomal DNA methylation in patients with endometrial carcinoma: an independent prognostic marker. *Cancer* **94**, 2941–2952 (2002).
85. Maruyama, R. *et al.* Aberrant promoter methylation profile of prostate cancers and its relationship to clinicopathological features. *Clin. Cancer Res.* **8**, 514–519 (2002).
86. Maruyama, R. *et al.* Aberrant promoter methylation profile of bladder cancer and its relationship to clinicopathological features. *Cancer Res.* **61**, 8659–8663 (2001).
87. Widschwendter, M. & Jones, P. A. The potential prognostic, predictive, and therapeutic values of DNA methylation in cancer. *Clin. Cancer Res.* **8**, 17–21 (2002). Commentary re: Kwong, J. *et al.* Promoter hypermethylation of multiple genes in nasopharyngeal carcinoma. *Clin. Cancer Res.* **8**, 131–137 (2002), and Zou, H.-Z. *et al.* Detection of aberrant p16 methylation in the serum of colorectal cancer patients. *Clin. Cancer Res.* **8**, 188–191 (2002).
88. Brabender, J. *et al.* Adenomatous polyposis coli gene promoter hypermethylation in non-small cell lung cancer is associated with survival. *Oncogene* **20**, 3528–3532 (2001).
89. Usadel, H. *et al.* Quantitative adenomatous polyposis coli promoter methylation analysis in tumor tissue, serum, and plasma DNA of patients with lung cancer. *Cancer Res.* **62**, 371–375 (2002).
90. Ogi, K. *et al.* Aberrant methylation of multiple genes and clinicopathological features in oral squamous cell carcinoma. *Clin. Cancer Res.* **8**, 3164–3171 (2002).
91. Garcia-Manero, G. *et al.* DNA methylation of multiple promoter-associated CpG islands in adult acute lymphocytic leukemia. *Clin. Cancer Res.* **8**, 2217–2224 (2002).
92. Tang, X. *et al.* Hypermethylation of the death-associated protein (DAP) kinase promoter and aggressiveness in stage I non-small-cell lung cancer. *J. Natl Cancer Inst.* **92**, 1511–1516 (2000).
93. Brabender, J. *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.* **9**, 223–227 (2003).
94. Esteller, M. *et al.* Inactivation of the DNA-repair gene MGMT and the clinical response of gliomas to alkylating agents. *N. Engl. J. Med.* **343**, 1350–1354 (2000).
95. Esteller, M. *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.* **94**, 26–32 (2002).
96. Buckner, J. C. & Moynihan, T. J. The DNA-repair gene MGMT and the clinical response of gliomas to alkylating agents. *N. Engl. J. Med.* **344**, 686–688 (2001).
97. Quinn, D. I. The DNA-repair gene MGMT and the clinical response of gliomas to alkylating agents. *N. Engl. J. Med.* **344**, 686–688 (2001).
98. Novik, K. L. *et al.* Epigenomics: genome-wide study of methylation phenomena. *Curr. Issues Mol. Biol.* **4**, 111–128 (2002).
99. Huang, T. H., Perry, M. R. & Laux, D. E. Methylation profiling of CpG islands in human breast cancer cells. *Hum. Mol. Genet.* **8**, 459–470 (1999).
100. Toyota, M. *et al.* Methylation profiling in acute myeloid leukemia. *Blood* **97**, 2823–2829 (2001).
101. Toyota, M. *et al.* CpG island methylator phenotype in colorectal cancer. *Proc. Natl Acad. Sci. USA* **96**, 8681–8686 (1999).
102. Liang, G. *et al.* DNA methylation differences associated with tumor tissues identified by genome scanning analysis. *Genomics* **53**, 260–268 (1998).
103. Li, J. *et al.* NotI subtraction and NotI-specific microarrays to detect copy number and methylation changes in whole genomes. *Proc. Natl Acad. Sci. USA* **99**, 10724–10729 (2002).
104. Hatada, I. *et al.* A microarray-based method for detecting methylated loci. *J. Hum. Genet.* **47**, 448–451 (2002).
105. Balog, R. P. *et al.* Parallel assessment of CpG methylation by two-color hybridization with oligonucleotide arrays. *Anal. Biochem.* **309**, 301–310 (2002).
106. Shi, H. *et al.* Oligonucleotide-based microarray for DNA methylation analysis: principles and applications. *J. Cell Biochem.* **88**, 138–143 (2003).
107. Model, F., Adorjan, P., Olek, A. & Piepenbrock, C. Feature selection for DNA methylation based cancer classification. *Bioinformatics* **17**, S157–S164 (2001).
108. Miller, O. J., Schnedl, W., Allen, J. & Erlanger, B. F. 5-Methylcytosine localised in mammalian constitutive heterochromatin. *Nature* **251**, 636–637 (1974).
109. Sano, H., Royer, H. D. & Sager, R. Identification of 5-methylcytosine in DNA fragments immobilized on nitrocellulose paper. *Proc. Natl Acad. Sci. USA* **77**, 3581–3585 (1980).
110. Kuo, K. C., McCune, R. A., Gehrke, C. W., Midgett, R. & Ehrlich, M. Quantitative reversed-phase high performance liquid chromatographic determination of major and modified deoxyribonucleosides in DNA. *Nucleic Acids Res.* **8**, 4763–4776 (1980).
111. Wilson, V. L. *et al.* Genomic 5-methylcytosine determination by 32P-postlabeling analysis. *Anal. Biochem.* **152**, 275–284 (1986).
112. Annan, R. S., Kresbach, G. M., Giese, R. W. & Vouros, P. Trace detection of modified DNA bases via moving-belt liquid chromatography-mass spectrometry using electrophoric derivatization and negative chemical ionization. *J. Chromatogr.* **465**, 285–296 (1989).
113. Balaghi, M. & Wagner, C. DNA methylation in folate deficiency: use of CpG methylase. *Biochem. Biophys. Res. Commun.* **193**, 1184–1190 (1993).
114. Wu, J. *et al.* Expression of an exogenous eukaryotic DNA methyltransferase gene induces transformation of NIH 3T3 cells. *Proc. Natl Acad. Sci. USA* **90**, 8891–8895 (1993).
115. Oakeley, E. J., Schmitt, F. & Jost, J. P. Quantification of 5-methylcytosine in DNA by the chloroacetaldehyde reaction. *Biotechniques* **27**, 744–746 (1999).
116. Gowher, H., Leismann, O. & Jeltsch, A. DNA of *Drosophila melanogaster* contains 5-methylcytosine. *EMBO J.* **19**, 6918–6923 (2000).
117. Fraga, M. F. *et al.* High-performance capillary electrophoretic method for the quantification of 5-methyl 2′-deoxycytidine in genomic DNA: application to plant, animal and human cancer tissues. *Electrophoresis* **23**, 1677–1681 (2002).
118. Galm, O. *et al.* Enzymatic regional methylation assay: a novel method to quantify regional CpG methylation density. *Genome Res.* **12**, 153–157 (2002).
119. Bird, A. P. & Southern, E. M. Use of restriction enzymes to study eukaryotic DNA methylation: I. The methylation pattern in ribosomal DNA from *Xenopus laevis*. *J. Mol. Biol.* **118**, 27–47 (1978).
120. Waalwijk, C. & Flavell, R. A. DNA methylation at a CCGG sequence in the large intron of the rabbit beta-globin gene: tissue-specific variations. *Nucleic Acids Res.* **5**, 4631–4634 (1978).
121. Pfeifer, G. P., Steigerwald, S. D., Mueller, P. R., Wold, B. & Riggs, A. D. Genomic sequencing and methylation analysis by ligation mediated PCR. *Science* **246**, 810–813 (1989).
122. Singer-Sam, J., LeBon, J. M., Tanguay, R. L. & Riggs, A. D. A quantitative HpaII–PCR assay to measure methylation of DNA from a small number of cells. *Nucleic Acids Res.* **18**, 687 (1990).
123. Steigerwald, S. D., Pfeifer, G. P. & Riggs, A. D. Ligation-mediated PCR improves the sensitivity of methylation analysis by restriction enzymes and detection of specific DNA strand breaks. *Nucleic Acids Res.* **18**, 1435–1439 (1990).
124. Heiskanen, M., Syvanen, A. C., Siitari, H., Laine, S. & Palotie, A. A novel method to quantitate methylation of specific genomic regions. *PCR Methods Appl.* **4**, 26–30 (1994).
125. Liang, G. *et al.* Cooperativity between DNA methyltransferases in the maintenance methylation of repetitive elements. *Mol. Cell. Biol.* **22**, 480–491 (2002).
126. Matin, M. M., Baumer, A. & Hornby, D. P. An analytical method for the detection of methylation differences at specific chromosomal loci using primer extension and ion pair reverse phase HPLC. *Hum. Mutat.* **20**, 305–311 (2002).
127. Akey, D., Akey, J., Zhang, K. & Jin, L. Assaying DNA methylation based on high-throughput melting curve approaches. *Genomics* **80**, 376 (2002).
128. Uhlmann, K., Brinckmann, A., Toliat, M. R., Ritter, H. & Nurnberg, P. Evaluation of a potential epigenetic biomarker by quantitative methyl- single nucleotide polymorphism analysis. *Electrophoresis* **23**, 4072–4079 (2002).
129. Ohmori, H., Tomizawa, J. I. & Maxam, A. M. Detection of 5-methylcytosine in DNA sequences. *Nucleic Acids Res.* **5**, 1479–1485 (1978).
130. Fritzsche, E., Hayatsu, H., Igloi, G. L., Iida, S. & Kossel, H. The use of permanganate as a sequencing reagent for identification of 5-methylcytosine residues in DNA. *Nucleic Acids Res.* **15**, 5517–5528 (1987).
131. Collins, M. & Myers, R. M. Alterations in DNA helix stability due to base modifications can be evaluated using denaturing gradient gel electrophoresis. *J. Mol. Biol.* **198**, 737–744 (1987).
132. Uhlmann, K., Marczinek, K., Hampe, J., Thiel, G. & Nurnberg, P. Changes in methylation patterns identified by two-dimensional DNA fingerprinting. *Electrophoresis* **20**, 1748–1755 (1999).

133. Aggerholm, A., Guldberg, P., Hokland, M. & Hokland, P. Extensive intra- and interindividual heterogeneity of p15INK4B methylation in acute myeloid leukemia. *Cancer Res.* **59**, 436–441 (1999).
134. Nuovo, G. J., Plaia, T. W., Belinsky, S. A., Baylin, S. B. & Herman, J. G. *In situ* detection of the hypermethylation-induced inactivation of the p16 gene as an early event in oncogenesis. *Proc. Natl Acad. Sci. USA* **96**, 12754–12759 (1999).
135. Burri, N. & Chaubert, P. Complex methylation patterns analyzed by single-strand conformation polymorphism. *Biotechniques* **26**, 232–234 (1999).
136. Bianco, T., Hussey, D. & Dobrovic, A. Methylation-sensitive, single-strand conformation analysis (MS-SSCA): a rapid method to screen for and analyze methylation. *Hum. Mutat.* **14**, 289–293 (1999).
137. Maekawa, M. *et al.* DNA methylation analysis using bisulfite treatment and PCR-single-strand conformation polymorphism in colorectal cancer showing microsatellite instability. *Biochem. Biophys. Res. Commun.* **262**, 671–676 (1999).
138. Baumer, A., Wiedemann, U., Hergersberg, M. & Schinzel, A. A novel MSP/DHPLC method for the investigation of the methylation status of imprinted genes enables the molecular detection of low cell mosaicisms. *Hum. Mutat.* **17**, 423–430 (2001).
139. Worm, J., Aggerholm, A. & Guldberg, P. In-tube DNA methylation profiling by fluorescence melting curve analysis. *Clin. Chem.* **47**, 1183–1189 (2001).
140. Cottrell, S. E. & Laird, P. W. Sensitive detection of DNA methylation. *Ann. NY Acad. Sci.* (in the press).
141. Kawai, J. *et al.* Methylation profiles of genomic DNA of mouse developmental brain detected by restriction landmark genomic scanning (RLGS) method. *Nucleic Acids Res.* **21**, 5604–5608 (1993).
142. Cross, S. H., Charlton, J. A., Nan, X. & Bird, A. P. Purification of CpG islands using a methylated DNA binding column. *Nature Genet.* **6**, 236–244 (1994).
143. Gonzalgo, M. L. *et al.* Identification and characterization of differentially methylated regions of genomic DNA by methylation-sensitive arbitrarily primed PCR. *Cancer Res.* **57**, 594–599 (1997).
144. Toyota, M. *et al.* Identification of differentially methylated sequences in colorectal cancer by methylated CpG island amplification. *Cancer Res.* **59**, 2307–2312 (1999).
145. Brock, G. J., Huang, T. H., Chen, C. M. & Johnson, K. J. A novel technique for the identification of CpG islands exhibiting altered methylation patterns (ICEAMP). *Nucleic Acids Res.* **29**, E123 (2001).
146. Gitan, R. S., Shi, H., Chen, C. M., Yan, P. S. & Huang, T. H. Methylation-specific oligonucleotide microarray: a new potential for high-throughput methylation analysis. *Genome Res.* **12**, 158–164 (2002).
147. Frigola, J., Ribas, M., Risques, R. A. & Peinado, M. A. Methylome profiling of cancer cells by amplification of inter-methylated sites (AIMS). *Nucleic Acids Res.* **30**, E28 (2002).
148. Suzuki, H. *et al.* A genomic screen for genes upregulated by demethylation and histone deacetylase inhibition in human colorectal cancer. *Nature Genet.* **31**, 141–149 (2002).
149. Liang, G., Gonzales, F. A., Jones, P. A., Orntoft, T. F. & Thykjaer, T. Analysis of gene induction in human fibroblasts and bladder cancer cells exposed to the methylation inhibitor 5-aza-2′-deoxycytidine. *Cancer Res.* **62**, 961–966 (2002).
150. Tompa, R. *et al.* Genome-wide profiling of DNA methylation reveals transposon targets of CHROMOMETHYLASE3. *Curr. Biol.* **12**, 65–68 (2002).
151. Valenzuela, M. T. *et al.* Assessing the use of p16(INK4a) promoter gene methylation in serum for detection of bladder cancer. *Eur. Urol.* **42**, 622–630 (2002).
152. Silva, J. M. *et al.* Aberrant DNA methylation of the p16INK4a gene in plasma DNA of breast cancer patients. *Br. J. Cancer* **80**, 1262–1264 (1999).
153. Hibi, K. *et al.* Molecular detection of p16 promoter methylation in the serum of patients with esophageal squamous cell carcinoma. *Clin. Cancer Res.* **7**, 3135–3138 (2001).
154. Wong, I. H., Lo, Y. M., Yeo, W., Lau, W. Y. & Johnson, P. J. Frequent p15 promoter methylation in tumor and peripheral blood from hepatocellular carcinoma patients. *Clin. Cancer Res.* **6**, 3516–3521 (2000).
155. Kurakawa, E. *et al.* Hypermethylation of p16(INK4a) and p15(INK4b) genes in non-small cell lung cancer. *Int. J. Oncol.* **19**, 277–281 (2001).
156. Bearzatto, A. *et al.* p16(INK4A) hypermethylation detected by fluorescent methylation- specific PCR in plasmas from non-small cell lung cancer. *Clin. Cancer Res.* **8**, 3782–3787 (2002).
157. Chan, M. W. *et al.* Hypermethylation of multiple genes in tumor tissues and voided urine in urinary bladder cancer patients. *Clin. Cancer Res.* **8**, 464–470 (2002).
158. Evron, E. *et al.* Detection of breast cancer cells in ductal lavage fluid by methylation-specific PCR. *Lancet* **357**, 1335–1336 (2001).
159. Rosas, S. L. *et al.* Promoter hypermethylation patterns of p16, O6-methylguanine-DNA-methyltransferase, and death-associated protein kinase in tumors and saliva of head and neck cancer patients. *Cancer Res.* **61**, 939–942 (2001).
160. Belinsky, S. A. *et al.* Aberrant methylation of p16(INK4a) is an early event in lung cancer and a potential biomarker for early diagnosis. *Proc. Natl Acad. Sci. USA* **95**, 11891–11896 (1998).
161. Ahrendt, S. A. *et al.* Molecular detection of tumor cells in bronchoalveolar lavage fluid from patients with early stage lung cancer. *J. Natl Cancer Inst.* **91**, 332–339 (1999).
162. Kersting, M. *et al.* Differential frequencies of p16(INK4a) promoter hypermethylation, p53 mutation, and K-ras mutation in exfoliative material mark the development of lung cancer in symptomatic chronic smokers. *J. Clin. Oncol.* **18**, 3221–3229 (2000).
163. Belinsky, S. A. *et al.* Aberrant promoter methylation in bronchial epithelium and sputum from current and former smokers. *Cancer Res.* **62**, 2370–2377 (2002).
164. Suh, C. I. *et al.* Comparison of telomerase activity and GSTP1 promoter methylation in ejaculate as potential screening tests for prostate cancer. *Mol. Cell Probes* **14**, 211–217 (2000).
165. Cairns, P. *et al.* Molecular detection of prostate cancer in urine by GSTP1 hypermethylation. *Clin. Cancer Res.* **7**, 2727–2730 (2001).
166. Goessl, C. *et al.* DNA-based detection of prostate cancer in urine after prostatic massage. *Urology* **58**, 335–338 (2001).
167. Goessl, C. *et al.* Methylation-specific PCR for detection of neoplastic DNA in biopsy washings. *J. Pathol.* **196**, 331–334 (2002).



## Acknowledgements

I apologize to colleagues whose contributions could not be cited due to space constraints. Supported by National Institutes of Health grants. Disclosure: Peter W. Laird is shareholder and Scientific Advisory Board Member of Epigenomics, AG, which has a commercial interest in the development of DNA methylation markers for disease detection and diagnosis. None of the work performed in the laboratory of Peter W. Laird is or has been supported by Epigenomics.


## Online links

DATABASES

The following terms in this article are linked online to:

Cancer.gov: http://www.cancer.gov/cancer_information/
colorectal cancer | lung cancer | ovarian cancer | pancreatic cancer | prostate cancer

LocusLink: http://www.ncbi.nih.gov/LocusLink/
DNMT1 | DNMT3A | DNMT3B | *GSTP1* | *MGMT* | *MLH1*

FURTHER INFORMATION

CpG Island Searcher: http://www.uscnorris.com/cpgislands/
DNA Methylation Society: http://www.dnamethylation.com
Early Detection Network: http://edrn.nci.nih.gov
MethDB – Public DNA Methylation Database: http://www.methdb.net

Access to this interactive links box is free online.