Cancer epigenomics: DNA methylomes and histone-modification maps

Manel Esteller

Abstract | An altered pattern of epigenetic modifications is central to many common human diseases, including cancer. Many studies have explored the mosaic patterns of DNA methylation and histone modification in cancer cells on a gene-by-gene basis; among their results has been the seminal finding of transcriptional silencing of tumour-suppressor genes by CpG-island-promoter hypermethylation. However, recent technological advances are now allowing cancer epigenetics to be studied genome-wide — an approach that has already begun to provide both biological insight and new avenues for translational research. It is time to 'upgrade' cancer epigenetics research and put together an ambitious plan to tackle the many unanswered questions in this field using epigenomics approaches.

Endoparasitic sequences
Repeated sequences, most of which are derived from transposable elements. These sequences are propagated by inserting new copies of themselves into random sites in the genome.

Cancer Epigenetics Laboratory, Spanish National Cancer Centre (CNIO), Melchor Fernandez Almagro 3, 28029 Madrid, Spain.
e-mail: mesteller@cnio.es
doi:10.1038/nrg2005
Published online
6 March 2007

The epigenetic landscape of cancer cells is profoundly distorted. Human tumours undergo a massive overall loss of DNA methylation, but also acquire specific patterns of hypermethylation at certain promoters[1–4]. In addition, these DNA-methylation changes are linked with the presence of an aberrant pattern of histone modification[5–9]. Small-scale studies of epigenetic marks have provided important insights into cancer biology; for example, the hypermethylation of tumour-suppressor genes, which is associated with their transcriptional silencing, is recognized as a key feature of cancer pathogenesis[1–4]. However, several important unanswered questions remain. How many genes undergo epigenetic disruption in a given tumour? Do these changes differ between distinct types of cancer cell? What are the molecular and genetic mechanisms that underlie these altered epigenetic profiles? And can a more global knowledge of the epigenetic characteristics of cancer cells be used for translational purposes?

Many of these questions can be answered by applying 'omics' approaches to cancer epigenetics. Here I provide a broad picture of how our current knowledge of epigenetic defects in cancer cells is beginning to be extended following the recent advent of genome-scale technologies to map DNA methylation and histone modifications. I begin by providing a brief overview of the current understanding of how epigenetic alterations contribute to tumorigenesis, and then discuss the powerful high-resolution epigenomics approaches that are extending these findings. This is followed by an exploration of the biological insight that is emerging from epigenomic studies of DNA methylation and histone modification patterns in cancer cells, and the prognostic, diagnostic and therapeutic impact of epigenomic profiling. I conclude by discussing the future of cancer epigenomics in the light of large-scale, coordinated efforts to catalogue the human epigenome.

Epigenetic marks in normal and cancer cells

DNA methylation and histone modifications have crucial roles in the control of gene activity and nuclear architecture. The most widely studied epigenetic modification in humans is the cytosine methylation of DNA within the dinucleotide CpG; 3–6% of all cytosines are methylated in normal human DNA[4]. Potentially 'methylable' CpG dinucleotides are not randomly distributed in the human genome; instead, CpG-rich regions known as CpG islands, which span the 5′ end region (promoter, untranslated region and exon 1) of many genes, are usually unmethylated in normal cells. This unmethylated status corresponds with the ability of CpG-island-containing genes to be transcribed in the presence of the necessary transcriptional activators. In cancer cells, the transcriptional silencing of tumour-suppressor genes by CpG-island-promoter hypermethylation is key to the tumorigenic process, contributing to all of the typical hallmarks of a cancer cell that result from tumour-suppressor inactivation[10] (FIG. 1).

By contrast, repetitive genomic sequences are heavily methylated. The maintenance of this DNA methylation could have a role in the protection of chromosomal integrity, by preventing chromosomal instability, translocations and gene disruption through the reactivation of endoparasitic sequences[11–13]. Although less well studied

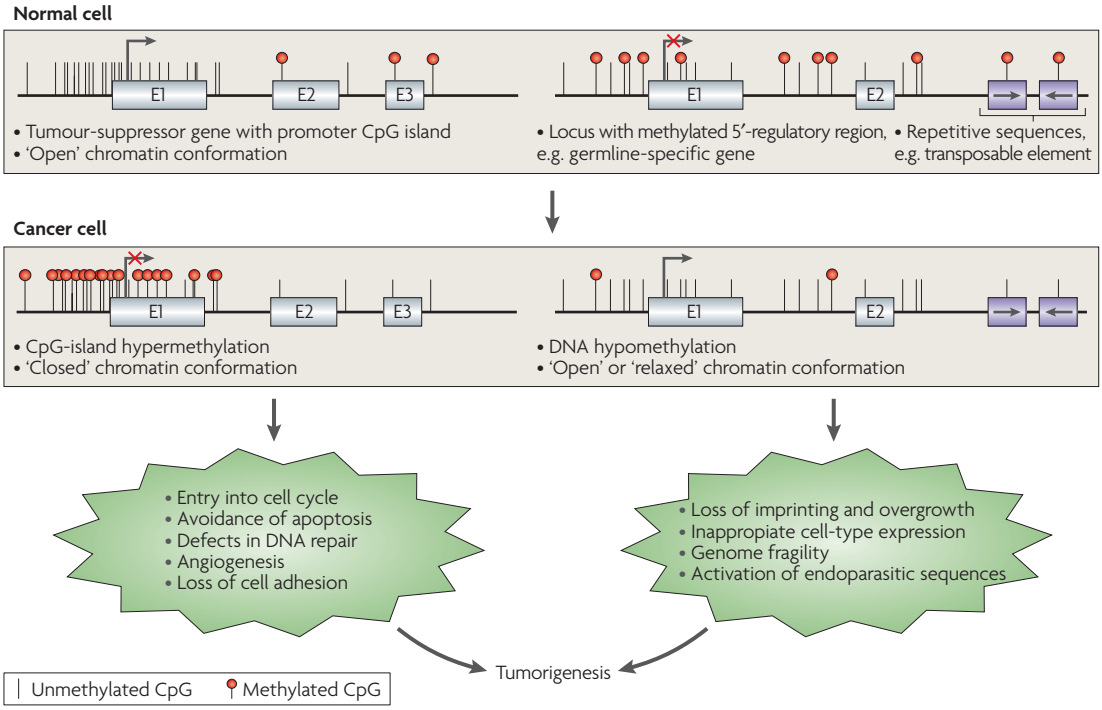

Figure 1 | **Altered DNA-methylation patterns in tumorigenesis.** The hypermethylation of CpG islands of tumour-suppressor genes is a common alteration in cancer cells, and leads to the transcriptional inactivation of these genes and the loss of their normal cellular functions. This contributes to many of the hallmarks of cancer cells. At the same time, the genome of the cancer cell undergoes global hypomethylation at repetitive sequences, and tissue-specific and imprinted genes can also show loss of DNA methylation. In some cases, this hypomethylation is known to contribute to cancer cell phenotypes, causing changes such as loss of imprinting, and might also contribute to the genomic instability that characterizes tumours. E, exon.

than DNA hypermethylation events, the global DNA hypomethylation that characterizes cancer cells is likely to contribute to the large-scale genetic changes that are a feature of tumorigenesis[2,4].

DNA methylation also provides one of the layers of epigenetic control of germline- and tissue-specific genes. Genomic imprinting requires DNA hypermethylation at one of the two parental alleles of a gene in order to establish monoallelic expression[14], and a similar gene-dosage reduction is involved in X-chromosome inactivation in females[14]. In addition, DNA methylation is required for the germline-specific expression of genes such as those of the MAGE family[15]. Finally, methylation has been postulated as a mechanism for silencing tissue-specific genes in cell types in which they should not be expressed, although the evidence for this is limited[16]. Among these specific regulatory roles of DNA methylation, loss of imprinting (LOI) is well established as a mechanism of gene activation in some types of cancer[17].

DNA methylation occurs in the context of other epigenetic modifications. It is associated with the formation of nuclease-resistant chromatin, and methyl-CpG-binding proteins and DNA methyltransferases (DNMTs) are associated with histone deacetylases and histone methyltransferases, two key regulators of histone modification[18–20]. Histones are no longer considered to be simple 'DNA-packaging' proteins; they are recognized as being dynamic regulators of gene activity that undergo many post-translational chemical modifications, including acetylation, methylation, phosphorylation, ubiquitylation and sumoylation[20]. The status of acetylation and methylation of specific lysine residues contained within the tails of nucleosomal core histones is known to have a crucial role in regulating chromatin structure and gene expression[20].

Overall, histone hypoacetylation and hypermethylation are characteristic of DNA sequences that are methylated and repressed in normal cells — such as the inactive X chromosome in females, and silenced imprinted and tissue-specific genes. However, the emerging model is that specific combinations of histone modifications confer the overall expression status of a region of chromatin, a theory known as the 'histone code' hypothesis, the details of which are just beginning to be understood[20]. In addition to their influence on gene expression, emerging evidence indicates that specific histone modifications interface with other nuclear processes, such as DNA repair pathways[21]. Histone modifications, together with DNA methylation, also have a vital role in organizing nuclear architecture[13–22], which, in turn, is involved in regulating transcription and other nuclear processes. Clearly, global alterations of histone modification patterns have the potential to affect the structure and integrity of the genome and to disrupt normal patterns of gene expression, which — like alterations in DNA methylation — might be causal factors in cancer[8,23].

Genomic imprinting
The epigenetic marking of a locus on the basis of parental origin, which results in monoallelic gene expression.

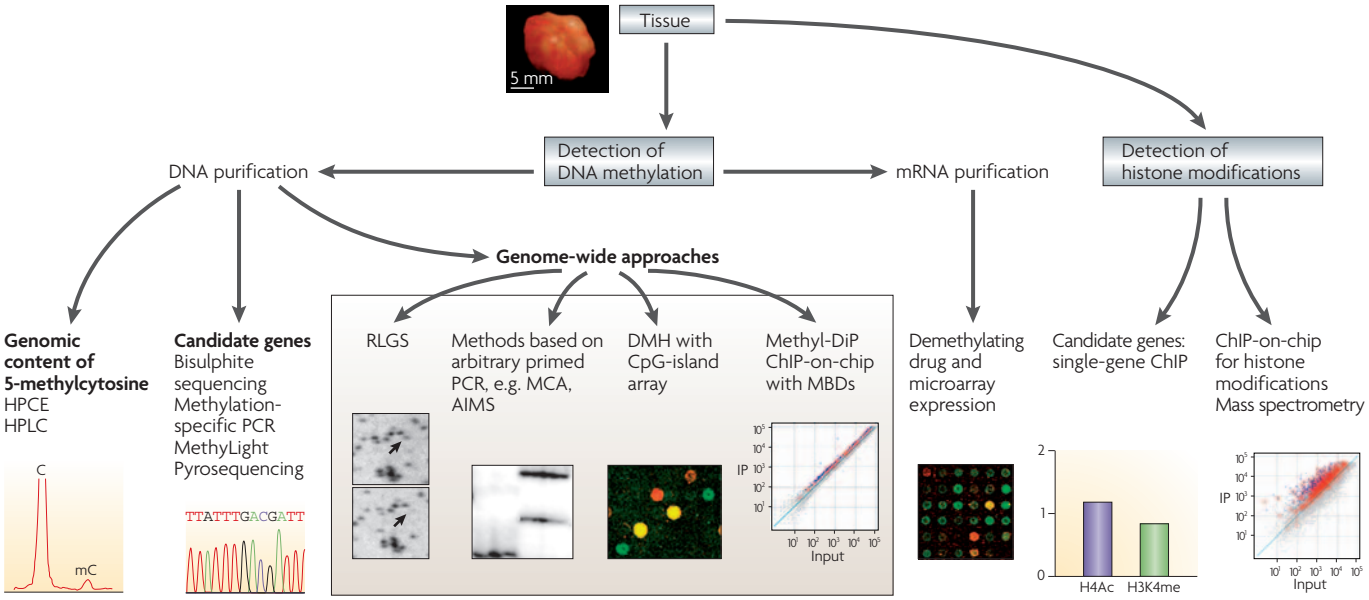

Figure 2 | **Techniques for studying epigenetic changes in cancer.** Most approaches to detecting DNA methylation start with the purification of DNA from cell samples. Subsequently, the overall DNA 5-methylcytosine content can be determined using high-performance capillary electrophoresis (HPCE) or high-performance liquid chromatography (HPLC), or the DNA methylation of specific candidate genes can be detected with methylation-sensitive methods. Recently, several genome-wide approaches to detecting DNA methylation have been developed (described in more detail in FIG. 3). An alternative approach for profiling DNA-methylation patterns is based on the extraction of mRNA, followed by microarray expression analysis. Analysis of the same sample in the presence or absence of a demethylating agent indicates genes that show increased expression owing to the removal of DNA methylation marks. For profiling histone modifications, marks at candidate genes can be detected using mass spectrometry or single-gene chromatin immunoprecipitation (ChIP) using antibodies against specific histone modifications. For global profiling, ChIP is combined with DNA arrays (ChIP-on-chip) to detect patterns across the genome. AIMS, amplification of intermethylated sites; ChIP, chromatin immunoprecipitation; DMH, differential methylation hybridization; MBDs, methyl-CpG-binding domain proteins; MCA, methylated CpG-island amplification; Methyl-DIP, methyl-DNA immunoprecipitation; RLGS, restriction landmark genomic scanning.

High-performance liquid chromatography
(HPLC). A technique for separating DNA or protein molecules by molecular weight and conformation. The molecules are resolved by differences in their distribution between a stationary phase and a mobile phase. The resolution is increased by increasing the pressure of the system.

High-performance capillary electrophoresis
(HPCE). A class of separation techniques that use narrow-bore fused-silica capillaries to separate a complex mixture of chemical compounds. Molecules are separated on the basis of differences in charge, size, structure and hydrophobic potential using strong electric fields.

Methylation-specific PCR
This DNA-methylation assay entails initial modification of DNA by sodium bisulphite, converting all unmethylated, but not methylated, cytosines to uracil, and subsequent amplification with primers that are specific for methylated versus unmethylated DNA.

Until recently, studies of cancer epigenetics have generally relied on investigating DNA methylation and histone modification at specific genes. However, although much insight has been gained into the role of epigenetic defects in cancer, answering questions about overall patterns of epigenetic modification in cancer cells will be possible only by using unbiased, genome-wide approaches.

Technological approaches to cancer epigenomics

A range of approaches are available for assessing patterns of epigenetic modification in normal and cancer cells, in a gene-specific or genome-wide manner (FIG. 2). Here I provide an overview of these methods, focusing on recent technological developments that have allowed a genome-wide approach to be taken, and discussing the merits and disadvantages of these approaches with respect to studies of cancer cells.

Approaches to detecting DNA methylation. Since the discovery of the first oncogene 25 years ago, genetics has taken the leading role in cancer research. Epigenetics initially lagged behind, despite the fact that the first seminal findings in the fields of cancer genetics and epigenetics occurred at almost the same time[24]. One of the main reasons for this delay was technical limitations. Although it has long been possible to accurately quantify the total amount of 5-methylcytosine using techniques such as high-performance liquid chromatography (HPLC) and high-performance capillary electrophoresis (HPCE)[25], the study of DNA methylation of particular sequences was initially almost entirely based on the use of enzymes that can distinguish between methylated and unmethylated recognition sites in genes of interest. This approach has many drawbacks, from incomplete restriction-enzyme cutting to limitation of the regions that can be studied. Furthermore, it usually involves Southern blotting, which requires substantial amounts of DNA of high molecular weight — a particular challenge when studying primary tumours because of the limited quantities of material available.

The popularization of the bisulphite treatment of DNA, which reproducibly changes unmethylated cytosines to uracil but leaves methylated cytosines unchanged, was an important advance in cancer epigenetics. Combining this approach with genomic sequencing[26] or amplification by methylation-specific PCR[27] has allowed any laboratory to study DNA methylation, using even limited amounts of material from old archives. More recently, quantitative PCR-based

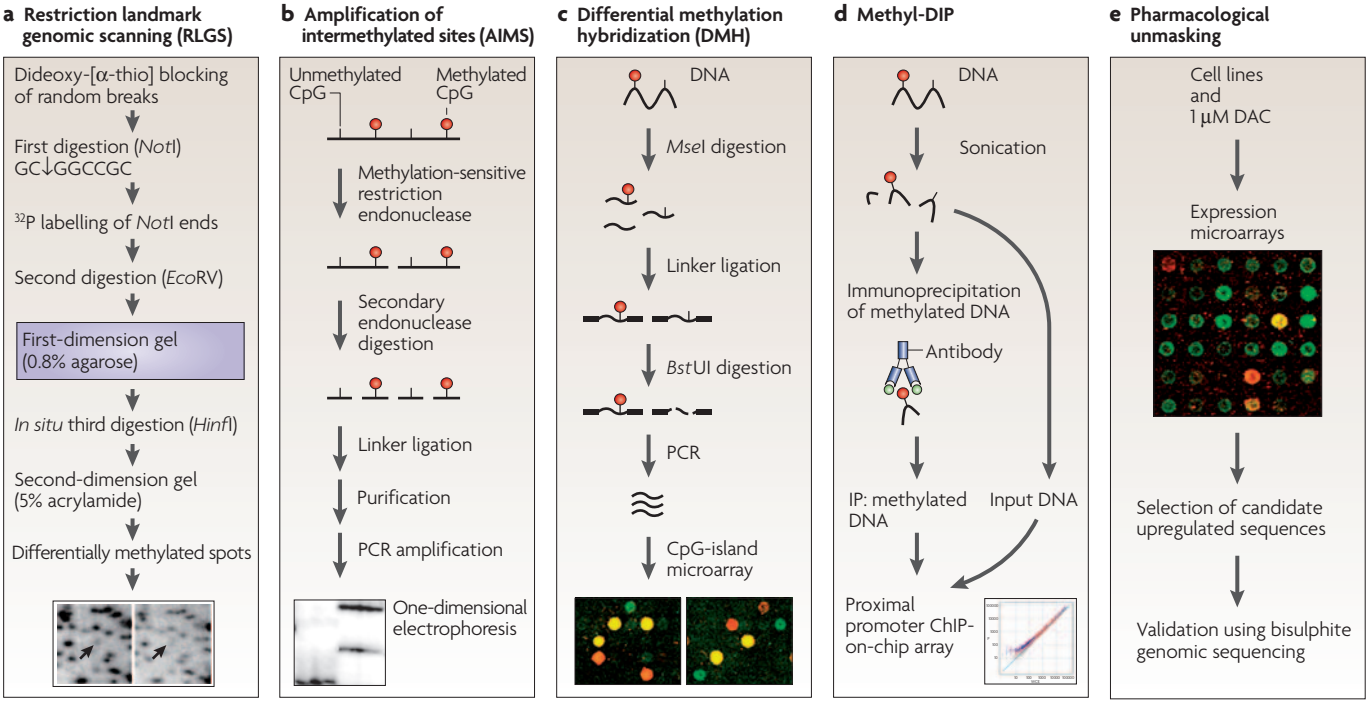

Figure 3 | **Methods for profiling genome-wide DNA-methylation patterns. a** | Restriction landmark genomic scanning (RLGS). After random breaks have been blocked, DNA is radioactively labelled at methylation-specific cleavage sites (in this example, the restriction enzyme *Not*I is used) and size-fractionated in one dimension. The digestion products are further digested with two more restriction endonucleases that are specific for high-frequency targets, and the fragments are separated by two-dimensional electrophoresis, with an *in situ* third digestion using *Hinf*I, yielding a number of scattered hot spots of DNA methylation. The images at the bottom illustrate RLGS results for two differentially methylated samples. In the left panel, the arrow points to unmethylated DNA at a particular location, as indicated by a spot on the gel. In the right panel, there is no corresponding spot on the gel, because the DNA at this location is methylated and is cut by the restriction enzyme. **b** | Amplification of intermethylated sites (AIMS) is an example of a method based on arbitrary primed PCR, which does not rely on prior knowledge of sequence information for amplification because it includes ligation to a linker oligonucleotide. In AIMS, the DNA templates for amplification are enriched in an initial step that involves digestion with a methylation-sensitive restriction enzyme. Specificity is provided by using linkers that bind specifically to the ends of fragments that are cut by this enzyme. After a secondary digestion and electrophoresis, cloning and sequencing are used to determine the methylated sequence. **c** | Differential methylation hybridization (DMH). CpG-island library DNA fragments are gridded on high-density arrays. Genomic DNA from the tissueof interest is digested with methylation-sensitive enzymes and the digestion products are used as templates for linker PCR. The resulting oligonucleotides are used as probes to screen for hypermethylated sequences within the CpG-island library. A comparison between two differentially methylated samples is shown. **d** | In methylated DNA immunoprecipitation (methyl–DIP), DNA is first fragmented by sonication and methylated fragments are then immunoprecipitated using a methylation-specific antibody. These fragments can then by hybridized to an array of proximal promoter regions to assess DNA-methylation patterns specifically in gene-regulatory regions. In this case, the example of a human proximal promoter array is shown. The graph represents the results obtained using immunoprecipitated (IP) DNA on the *y*-axis and input DNA (control) on the *x*-axis. **e** | In pharmacological unmasking, RNA is collected from a cancer cell line before and after a treatment with a DNA-demethylating agent — in this case, decitabine (DAC) — and is hybridized to an expression microarray. Genes that are methylated in the cell line can be identified on the basis of their upregulation after demethylation.

Two-dimensional electrophoresis
A gel electrophoresis method in which the proteins in a sample are separated by their isoelectric points in one dimension, and by size in a second, perpendicular dimension.

MethyLight
A high-throughput quantitative methylation assay that uses fluorescence-based real-time PCR (TaqMan) technology and requires no further manipulations after the PCR step. This technique is carried out in combination with bisulphite treatment (in which unmethylated cytosine residues are converted to uracil), and sequence discrimination is achieved by designing the primers to overlap with potential sites of DNA methylation (CpG dinucleotides).

Pyrosequencing
A DNA-sequencing method in which light is emitted as a result of an enzymatic reaction, each time a nucleotide is incorporated into the growing DNA chain. As applied to methylation detection, methylation-dependent DNA sequence variation, which is achieved by sodium bisulphite treatment, is treated as a kind of SNP of the C–T type, and is subjected to conventional SNP typing.

methods, such as bisulphite treatment in combination with MethyLight[28] or pyrosequencing[29], have provided another expanding area in the DNA-methylation field, because of their ability to detect minimal amounts of aberrant DNA methylation. Because these assays rely on PCR of CG-rich DNA, the use of positive and negative control samples for methylated DNA (such as *in vitro* methylated DNA) and unmethylated DNA (normal tissue) is always a requirement when using this approach. It is also important to bear in mind that these PCR-based methods interrogate the methylation status only at CpG sites that are complementary to the primers that are designed for use in such experiments. Therefore, the predominant methylation pattern in a sample is not necessarily reflected in the results of such experiments, and bisulphite treatment followed by genomic sequencing might be needed to provide a complete picture of the heterogeneous methylation patterns that exist in cancer cells.

All of the techniques that were initially limited to studying DNA methylation at candidate genes —

bisulphite treatment, the use of methylation-sensitive restriction enzymes and quantitative-PCR-based methods — can also be coupled with genomic approaches for detecting DNA-methylation patterns (FIG. 3). Restriction landmark genomic scanning (RLGS) was one of the earliest methods to be adapted for genome-wide methylation analysis[30]. With this technique, DNA is first radioactively labelled at unmethylated sites within methylation-sensitive restriction enzyme targets, and is then size-fractionated in one dimension. The digestion products are digested with a second restriction endonuclease that is specific for high-frequency targets, and the fragments are separated in the second dimension, yielding a number of scattered hot spots of DNA methylation. When normal and tumoural tissues are compared, the position and strength of a spot reveals its location and the copy number of the corresponding restriction site, respectively. By allowing the simultaneous quantification of gene copy number and methylation status[31], this method provides the added advantage of relating genetic and epigenetic changes in cancer. However, disadvantages of RLGS include its reliance on specific digestion sites that are not present in all CpG islands and the fact that not all of the resulting fragments can be resolved in the electrophoresis steps. Bearing in mind these caveats, the global analysis of the methylation status of ~1,000 unselected CpG islands can be achieved using this method[30].

Other important tools for detecting altered patterns of DNA methylation across the genome involve variations of arbitrary primed PCR[32], in which no prior sequence information is required for amplification. Examples include methylation-sensitive arbitrary primed PCR[33], methylated CpG-island amplification (MCA)[34] and amplification of intermethylated sites (AIMS)[35]. These methods are particularly useful because arbitrary primed PCR is carried out using DNA templates that have been enriched for methyl sequences, resulting in preferential amplification of CpG islands and gene-rich regions[35]. However, all of these techniques require further validation by bisulphite genomic sequencing, and a background of PCR 'noise' from repetitive sequences must be taken into account.

Undoubtedly, one of the most efficient means of studying CpG-island methylation at a genome-wide scale involves novel technologies that make use of CpG-island and promoter microarrays. Among other features, such approaches avoid the further cloning and sequencing that are needed when a positive signal has been obtained using arbitrary primed PCR methods. A widely used example of such an approach is differential methylation hybridization (DMH), which allows the simultaneous determination of the methylation levels of a large number of CpG-island loci[36]. The CpG-island library DNA fragments are gridded on high-density arrays, genomic DNA from the tissues of interest is digested with methylation-sensitive enzymes, and digestion products are used as templates for PCR after ligation to linkers. The resulting oligonucleotides are used as probes to screen for hypermethylated sequences within the CpG-island library to identify sequences that are hypermethylated in cancer cells but not in normal control cells. A recently developed related method is the HELP assay (HpaII tiny fragment enrichment by ligation-mediated PCR), which involves co-hybridization of the DNA samples to a genomic DNA microarray after cutting with a methylation-sensitive restriction enzyme or its methylation-insensitive isoschizomer[37]. This assay has revealed a large number of tissue-specific, differentially methylated regions[37], and can be applied to cancer genomes by comparing normal and tumour cells.

Techniques based on chromatin immunoprecipitation using the ChIP-on-chip approach have provided another important recent advance in the epigenomic profiling of cancer cells. For example, DNA that was immunoprecipitated from breast cancer cells using antibodies against methyl-CpG-binding domain proteins (MBDs), which have a great affinity for binding to methylated cytosines[38], has been used to identify hypermethylated genes in mammary tumorigenesis[7]. Importantly, the key finding that DNA that is immunoprecipitated with an antibody against 5-methylcytosine (methyl-DIP) can be used as a probe for hybridization to genomic microarray platforms[39,40] promises to simplify and universalize the analysis of the DNA methylome, because it allows the rapid identification of multiple CpG sites. In *Arabidopsis thaliana*, the combination of the methyl-DIP strategy with the affinity purification of methylated DNA using the methylcytosine-binding domain of the human methy-CpG-binding protein (MeCP2) has provided the first comprehensive DNA-methylation map of an entire genome at 35-bp resolution[41].

However, it is important to bear in mind that, although several promoter, CpG-island and tiling microarrays are available from different companies, the entire human genome is not yet represented in any microarray. In addition, the need for whole-genome amplification after immunoprecipitation can introduce PCR biases, a consideration that should be kept in mind when using the methyl-DIP assay.

Finally, another means of assessing genome-wide DNA-methylation patterns is gene-expression profiling using microarrays, which is now becoming widely used and has the potential to be particularly useful in cancer epigenomics. This method involves comparing mRNA levels from cancer cell lines before and after treatment with a demethylating drug[42–44], and has proved successful in identifying hypermethylated genes, avoiding the cloning step that is required when other technologies are used. However, it is important to note that not all of the genes that became re-expressed after the use of the demethylating agent are necessarily methylated, and false positives are not uncommon.

Because the approaches described above do not provide direct proof of the presence of DNA methylation, bisulphite genomic sequencing is always required for confirmation. Additional evidence, such as that provided by luciferase reporter assays, is also necessary to confirm the link between gene silencing and DNA methylation.

Arbitrary primed PCR
Amplification of genomic DNA using arbitrary primers. The first amplification cycles are carried out at a low annealing temperature, such that the primer hybridizes to many non-specific sequences. The temperature is then increased, so that only the 'best' products of the initial annealing events are amplified further, generating a number of discrete bands that provide a fingerprint of the genome.

Amplification of intermethylated sites
(AIMS). A DNA-methylation fingerprinting technique that uses methyl-isoschizomers and arbitrary PCR amplification to obtain many anonymous bands, which represent DNA sequences flanked by two methylated sites.

Chromatin immunoprecipitation
(ChIP). The isolation, using specific antibodies, of chromatin fragments that are bound by a particular nuclear factor or associated with a particular histone-modification signature. The immunoprecipated DNA can subsequently be analysed with specific PCR primers.

ChIP-on-chip
A combination of chromatin immunoprecipitation with hybridization to genomic microarrays that is used to identify DNA sequences bound to a particular nuclear factor or with a specific histone-modification profile.

Methyl-DIP
(Methylated DNA immunoprecipitation). Immunoprecipitation with anti-5-methylcytosine antibodies followed by hybridization to genomic microarrays, allowing the identification of methyl-CpG-rich sequences.

Tiling microarrays
Microarrays that contain a set of overlapping oligonucleotides or other probes that span either the entire genome or, for a more specialized approach, a subregion of interest.

Approaches to detecting histone modifications. The implementation of epigenomic technologies for the study of histone modifications presents greater challenges than exist for the analysis of DNA methylomes. If for DNA-methylation studies the gold standard for accuracy is bisulphite treatment coupled with genomic sequencing, then the equivalent for post-translational histone modification is mass spectrometry, which is also the most accurate technique for identifying histone modifications. However, mass spectrometry requires a high degree of technical expertise and is difficult to apply genome-wide.

Acceptable data on global levels of histone modification can currently be obtained by combining other methods. For example, all histones (H3, H4, H2A, H2B and H1) can be isolated by HPLC, and the corresponding eluted fractions analysed by HPCE and liquid chromatography–electrospray mass spectrometry (LC–ES/MS)[8]. Specific modifications at each amino-acid residue can also be characterized using antibodies in Western blots, immunostaining[23] or tandem mass spectrometry (MS/MS)[8].

However, in addition to determining the types and relative amounts of histone modification that are present in a particular cell or tumour type, information that couples histone modification with particular DNA sequences is also needed. At present, the most powerful technique with which to accomplish this goal is the use of ChIP with antibodies against specific histone modifications. The immunoprecipitated DNA is typically analysed by PCR with specific primers to investigate the presence of a candidate DNA sequence. Recently, a modification of the ChIP technology has been reported that might allow the analysis of as few as 100 cells[45]. This would be of particular use for studying small amounts of cancer material.

The use of ChIP-on-chip with genomic platforms has started to provide extensive maps of histone modifications in model organisms such as *A. thaliana*[46], yeast[47,48], *Drosophila melanogaster*[49] and mice[50,51], including mouse embryonic stem cells[52]. Importantly, the first preliminary histone modification maps for normal human cells[51], including stem cells[53], are also available.

As described above for the application of ChIP-on-chip to the detection of DNA-methylation patterns, this approach has shortcomings that need to be overcome. In addition to the current lack of representation of the whole human genome, it must be borne in mind that ChIP-on-chip relies on the quality of the antibodies that are available. Ideally, antibodies that are highly specific for each particular histone modification epitope should be used in order to produce results that can be compared across experiments. However, most ChIP-on-chip experiments still use polyclonal antibodies, which can differ in specificity between batches. The generation and use of monoclonal antibodies for ChIP-on-chip is therefore an important goal.

The epigenomes of cancer cells

DNA methylation in the cancer epigenome. An important way in which epigenomic methods have already contributed to our understanding of cancer biology is in gaining a more global picture of aberrant DNA methylation at promoters. A long list of hypermethylated genes has now been obtained from various human neoplasias (see TABLE 1 for a selective list), with the result that this type of epigenetic alteration is now considered to be a common hallmark of all types of human cancer, affecting genes involved in all cellular pathways[1–4]. During the past few years, genes with key roles in cancer biology, such as the gene that encodes the cell-cycle inhibitor $p16^{INK4a}$ and the DNA-repair genes *MLH1* and *BRCA1*, have been shown to undergo methylation-associated silencing in cancer cells[1–4].

Many of these findings fit well with the existing genetic knowledge of tumour-suppressor genes. For example, several genes that are now known to be inactivated by CpG-island hypermethylation in transformed cells have antiproliferative roles, and in many instances there are familial cancer cases with associated germline mutations (for example, *MLH1*, *BRCA1*, *VHL*, $p16^{INK4a}$ and *WRN*). Interestingly, the use of unbiased epigenomic technologies has also identified genes showing methylation-associated silencing that belong to non-classical pathways that lead to tumorigenesis, such as the genes that encode the ID4 transcription factor[54] and insulin-like growth factor binding protein 3 (REF. 44). Importantly, profiles of CpG-island hypermethylation have also been shown to vary with tumour type[30,55,56]. Each tumour subtype can now be assigned a DNA hypermethylome that almost completely defines that particular malignancy, in a similar way to genetic and cytogenetic markers. As described later, this has important diagnostic and prognostic implications.

It has recently been proposed that about 200 genes are mutated in human breast and colon cancers, with an average of 11 mutations for each tumour[57]. Similarly, the average number of hypermethylated CpG islands in a particular cancer cell is a question of great interest. The answer could shed light on the relative contribution of genetic and epigenetic events to cancer development, and the synergy between them. Results obtained using various approaches[7,39,40,42–44,58] indicates a range of 100–400 promoter hypermethylated CpG islands in a given tumour, although these numbers are likely to change as epigenomic studies are carried out across a wider range of tumour types. Furthermore, the recent demonstration that microRNAs that have tumour-suppressor function can also undergo DNA-methylation-associated silencing in tumour cells[59,60] might indicate an additional contribution of DNA hypermethylation events to cancer development.

As well as questions about the number of hypermethylated genes, it is currently unclear why some genes become hypermethylated in certain tumours, whereas others with similar properties — a typical CpG island, a history of loss of expression in certain tumours and the absence of mutations — remain methylation-free. Putting this question into an epigenomics context might be helpful. For example, genome-wide analysis of DNA methylation using the methyl-DIP approach in colon and prostate cancer cells indicates that there might be common sequence motifs in promoters that

Mass spectrometry
An analytical technique determining molecular mass. This involves an ion source in which gas-phase molecular ions are produced from the analyte molecules, a mass analyser in which electrical and/or magnetic fields are used to separate the analyte ions by their different mass-to-charge ratios, and a detector for recording the separated ions.

Liquid chromatography–electrospray mass spectrometry
A mass spectrometry technique in which ionization of molecules is carried out within aerosols of small droplets. Molecules are then identified using electric and magnetic fields.

Tandem mass spectrometry
An analytical system in which two linked mass spectrometers are used to measure small amounts of metabolites. The analytes are separated according to their mass and charge. By programming the instrument to respond to only certain masses, a high degree of specificity and sensitivity can be achieved.

Table 1 | **A catalogue of genes silenced by CpG island promoter hypermethylation in human cancer**

Gene	Function	Location	Tumour type	Consequences
MLH1	DNA mismatch repair	3p21.3	Colon, endometrium, stomach	Frameshift mutations
BRCA1	DNA repair, transcription	17q21	Breast, ovary	Double-strand breaks?
p16^{INK4a}	Cyclin-dependent kinase inhibitor	9p21	Multiple types	Entrance in cell cycle
p14^{ARF}	MDM2 inhibitor	9p21	Colon, stomach, kidney	Degradation of p53
p15^{INK4b}	Cyclin-dependent kinase inhibitor	9p21	Leukaemia	Entrance into cell cycle
MGMT	DNA repair of 06–alkyl-guanine	10q26	Multiple types	Mutations, chemosentivity
GSTP1	Conjugation to glutathione	11q13	Prostate, breast, kidney	Adduct accumulation?
p73	p53 homologue	1p36	Lymphoma	Unknown
LKB1/STK11	Serine–threonine kinase	19p13.3	Colon, breast, lung	Unknown
ER	Oestrogen receptor	6q25.1	Breast	Hormone insensitivity
PR	Progesterone receptor	11q22	Breast	Hormone insensitivity
AR	Androgen receptor	Xq11	Prostate	Hormone insensitivity
PRLR	Prolactin receptor	5p13–p12	Breast	Hormone insensitivity
TSHR	Thyroid-stimulating hormone receptor	14q31	Thyroid	Hormone insensitivity
RARβ2	Retinoic acid receptor-β2	3p24	Colon, lung, head and neck	Vitamin insensitivity?
CRBP1	Retinol-binding protein	3q21–q22	Colon, stomach, lymphoma	Vitamin insensitivity?
RASSF1A	Ras effector homologue	3p21.3	Multiple types	Unknown
NORE1A	Ras effector homologue	1q32	Lung	Unknown
VHL	Ubiquitin ligase component	3p25	Kidney, haemangioblastoma	Loss of hypoxic response?
Rb	Cell-cycle inhibitor	13q14	Retinoblastoma	Entrance into cell cycle
THBS1	Thrombospondin-1, Anti-angiogenic	15q15	Glioma	Neovascularization
CDH1	E cadherin, cell adhesion	16q22.1	Breast, stomach, Leukaemia	Dissemination
CDH13	H cadherin, cell adhesion	16q24	Breast, lung	Dissemination?
FAT	Cadherin, tumour suppressor	4q34–35	Colon	Dissemination?
HIC1	Transcription factor	17p13.3	Multiple types	Unknown
APC	Inhibitor of β-catenin	5q21	Aerodigestive tract	Activation β-catenin route
SFRP1	Secreted frizzled-related protein 1	8p12–p11	Colon	Activation Wnt signalling
DKK1	Extracellular Wnt inhibitor	10q11.2	Colon	Activation Wnt signalling
WIF1	Wnt inhibitory factor	12q14.3	Colon, lung	Activation Wnt signalling
COX2	Cyclooxygenase-2	1q25	Colon, stomach	Anti-inflammatory resistance?
SOCS1	Inhibitor of JAK–STAT pathway	16p13.13	Liver, mieloma	JAK2 activation
SOCS3	Inhibitor of JAK–STAT pathway	17q25	Lung	JAK2 activation
GATA4	Transcription factor	8p23–p22	Colon, stomach	Silencing of target genes
GATA5	Transcription factor	20q13	Colon, stomach	Silencing of target genes
ID4	Transcription factor	6p22–p21.3	Leukaemia, stomach	Unknown
SRBC	BRCA1-binding protein	1p15	Breast, lung	Unknown
SYK	Tyrosine kinase	9q22	Breast	Unknown
RIZ1	Histone/protein methyltransferase	1p36	Breast, liver	Aberrant gene expression?
DAPK	Pro-apoptotic	9q34.1	Lymphoma, lung, colon	Resistance to apoptosis
TMS1	Pro-apoptotic	16p11	Breast	Resistance to apoptosis
IGFBP3	Growth-factor-binding protein	7p14–p12	Lung, skin	Resistance to apoptosis
TPEF/HPP1	Transmembrane protein	2q33	Colon, bladder	Unknown
SLC5A8	Sodium transporter	12q23	Glioma, colon	Unknown
HOXA9	Homeobox protein	7p15–p14	Neuroblastoma	Unknown
EXT1	Heparan sulphate synthesis	8q24	Leukaemia, skin	Cellular detachment
Lamin A/C	Nuclear intermediate filament	1q21.2	Lymphoma, leukaemia	Unknown
WRN	DNA repair	8p12–p11.2	Colon, stomach, sarcoma	DNA breakage, chemosensitivity

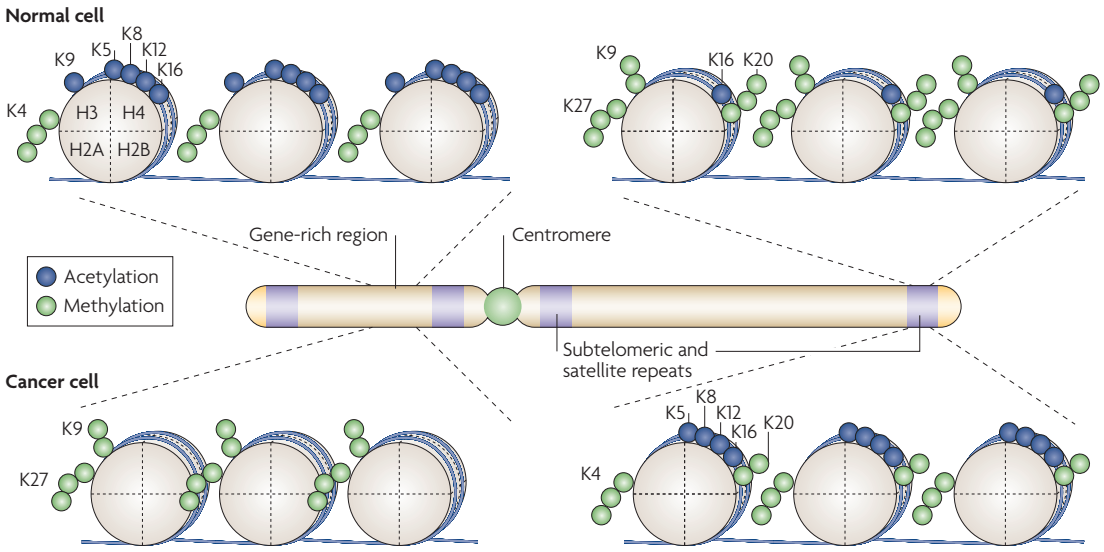

Figure 4 | **Histone-modification maps for a typical chromosome in normal and cancer cells.** Nucleosomal arrays are shown in the context of chromosomal location and transcriptional activity. Octamers consisting of histones H2A, H2B, H3 and H4 are represented as grey cylinders. Histone acetylation and methylation (di- and tri-) are shown. In 'normal' cells, genomic regions that include the promoters of tumour-suppressor genes are enriched in histone-modification marks associated with active transcription, such as acetylation of H3 and H4 lysine residues (for instance K5, K8, K9, K12 and K16) and trimethylation of K4 of H3. In the same cells, DNA repeats and other heterochromatic regions are characterized by trimethylation of K27 and dimethylation of K9 of H3, and trimethylation of K20 of H4, which function as repressive marks. In transformed cells, this scenario is disrupted by the loss of the 'active' histone-marks on tumour-suppressor gene promoters, and by the loss of repressive marks such as the trimethylation of K20 of H4 or trimethylation of K27 of histone H3 at subtelomeric DNA and other DNA repeats. This leads to a more 'relaxed' chromatin conformation in these regions.

undergo CpG-island hypermethylation[40]. In this regard, DNA sequences that are targeted by transcriptional repressors such as Polycomb proteins early in development might mark those genes that are more prone to be hypermethylated later in cancer cells[61,62]. It is also possible — as has been proposed for genetic mutations — that, in certain tumour types, a particular gene is preferentially methylated with respect to others because its inactivation confers a selective clonal advantage.

Alternatively, aberrant DNA methylation might be directly targeted as a result of preceding genetic alterations. It has been proposed that fusion proteins, such as the promyelocytic leukaemia–retinoic acid receptor-α (PML–RARA) fusion protein, which is expressed in some leukaemias, contribute to aberrant CpG-island methylation by recruiting DNMTs and histone deacetylases (HDACs)[63] to specific target genes, although this does not seem to be a general mechanism in this type of cancer[64]. In another example, the oncogene *MYC* has been shown to associate with a DNMT *in vivo* in osteosarcoma cells[65]. It seems likely that selection for hypermethylation and targeting as a result of genetic defects are not exclusive events, and both probably occur during the generation and maintenance of hypermethylated CpG islands of tumour-suppressor genes. Epigenomic analyses of genetically engineered cells in which some of these putative oncogenic 'drivers' have been knocked out could provide a way of investigating the relative contributions of the two mechanisms.

At the same time that CpG islands become hypermethylated, cancer cell genomes undergo global hypomethylation[1–4], with malignant cells having 20–60% less genomic 5-methylcytosine than their normal counterparts. This loss is accomplished mainly by hypomethylation of the 'body' of the genes involved (the coding region and introns) and through demethylation of repetitive DNA sequences, which account for 20–30% of the human genome. Epigenomic technologies have confirmed these concepts, but have also provided new insight. For example, the study of DNA methylation in a colorectal cancer cell line using the methyl-DIP approach has detected large hypomethylated genomic regions in gene-poor areas[39]. This raises the important question of whether and how hypomethylation in such regions might contribute to the cancer cell phenotype. A combination of genomic and epigenomic approaches will be needed to determine whether these hypomethylated regions coincide with sites of genomic instability.

Histone modifications in the cancer epigenome. So far, little is known about the patterns of histone modification disruption in human tumours. Promoter CpG-island hypermethylation in cancer cells is known to be associated with a particular combination of histone marks: deacetylation of histones H3 and H4, loss of histone H3 lysine K4 (H3K4) trimethylation, and gain of H3K9 methylation and H3K27 trimethylation[6,7,66] (FIG. 4). It is also recognized that certain genes with tumour-suppressor-like properties, such as $p21^{WAF1}$, are silent

Table 2 | **Disrupted DNA-methylation and histone-modification genes in cancer**

Gene	Alteration	Tumour type
DNA methyltransferases		
DNMT1	Overexpression	Multiple types
DNMT3b	Overexpression	Multiple types
Methyl-CpG-binding proteins		
MeCP2	Overexpression, rare mutations	Multiple types
MBD1	Overexpression, rare mutations	Multiple types
MBD2	Overexpression, rare mutations	Multiple types
MBD3	Overexpression, rare mutations	Multiple types
MBD4	Inactivating mutations in MSI+	Colon, stomach, endometrium
Histone acetyltransferases		
p300	Mutations in MSI+	Colon, stomach, endometrium
CBP	Mutations, translocations, deletions	Colon, stomach, endometrium, lung, leukaemia
pCAF	Rare mutations	Colon
MOZ	Translocations	Haematological malignancies
MORF	Translocations	Haematological malignancies, leiomyomata
Histone deacetylases		
HDAC1	Aberrant expression	Multiple types
HDAC2	Aberrant expression, mutations in MSI+	Multiple types
Histone methyltransferases		
MLL1	Translocation	Haematological malignancies
MLL2	Gene amplification	Glioma, pancreas
MLL3	Deletion	Leukaemia
NSD1	Translocation	Leukaemia
EZH2	Gene amplification, overexpression	Multiple types
RIZ1	CpG-island hypermethylation	Multiple types
Histone demethylase		
GASC1	Gene amplification	Squamous cell carcinoma

MSI+, microsatellite instable tumours.

at the transcriptional level in the absence of CpG-island hypermethylation when hypoacetylated and hypermethylated histones H3 and H4 are present[67].

Until recently, a genome-wide profile of histone modifications and their locations was not available for any transformed cell type. Post-translational modifications of histone H4 have now been profiled at a global level in a comprehensive panel of normal human tissues, cancer cell lines and primary tumours[8]. In this study, cancer cells were shown to exhibit a loss of monoacetylated and trimethylated forms of histone H4. Interestingly, these changes appear early and accumulate during the tumorigenic process, as shown in a mouse model of multistage skin carcinogenesis[8]. The finding that these changes in histone modification patterns occur so soon in the course of tumorigenesis, similar to the manner in which CpG-island hypermethylation precedes *KRAS* mutations in small colorectal adenomas[68], indicates that they might be relevant steps in the transformation process. By mass spectrometry, these losses were found to occur predominantly at the acetylated K16 and trimethylated K20 residues of histone H4, and were associated with the well characterized hypomethylation of DNA repetitive sequences (FIG. 4). Similar data have been obtained in breast and liver tumorigenesis[69,70], indicating that the global loss of monoacetylation and trimethylation of histone H4 might be a common hallmark of human tumour cells, as has now been accepted for global DNA hypomethylation and CpG-island hypermethylation.

Disruption of the epigenetic machinery. An emerging question, touched on briefly above, is whether mutations in the DNA methylation or histone modification machineries themselves cause the global epigenomic changes seen in cancer cells, or whether other mechanisms are responsible. A preliminary list of genes involved in epigenetic modifications that are disrupted in human cancer is provided in TABLE 2. Mutations that affect the DNA-methylation machinery have not been identified in cancer cells, although moderately elevated levels of DNMTs and MBD-containing proteins are commonly observed in human tumours[1–4,38]. How do

increased levels of DNMTs fit in with the observation of global DNA hypomethylation? In plants, a defect in the gene *SUPERMAN*, which is a transcription factor involved in flower development, results in specific areas of hypermethylation in the context of global DNA hypomethylation[71]. However, an equivalent mechanism has not yet been described for transformed human cells. In mouse models, crossing cancer-prone lines with strains carrying genetic defects in DNMTs and MBD-containing proteins changes the risk and time of onset of tumorigenesis[72], indicating that the potential role of these proteins as driving forces of oncogenesis needs further investigation. Epigenomic studies using ChIP-on-chip approaches with antibodies against DNMTs and MBD-containing proteins in normal versus cancer cells might provide one way to explore this further.

In terms of histone modifications, an extensive analysis of expression patterns of histone-modifying enzymes was able to discriminate tumour samples from their normal counterparts and cluster the tumour samples according to cell type[73]. This indicates that changes in the expression of these proteins have important and tumour-specific roles in cancer development. For the two most common histone-modification changes in human neoplasia — the reduction of monoacetylated H4K16 and trimethylated H4K20 — some clues as to how these cancer-specific alterations arise are already available. The delicate acetylation and deacetylation balance of H4K16 is mediated by specific histone acetyltransferases (HATs), such as MORF, MOZ, MOF, TIP60, and HBO1, and a dedicated HDAC, sirtuin 1 (SIRT1). Leukaemias[74] and uterine myomas[75] carry translocations that generate fusion proteins such as CREB-binding protein (CBP)–MOZ and CBP–MORF that disrupt the acetylation of H4K16 (REF. 8). In epithelial tumours, it has been shown by ChIP-on-chip that there is a specific loss of recruitment of these H4K16 HATs to repeat sequences[8]. It is also possible that there is increased recruitment of SIRT1 to the H4K16 position on repeat DNA sequences in transformed cells, but this remains to be investigated.

For the trimethyl-H4K20 mark, the observed loss in cancer cells[8] and the demonstration that knockout mice for the histone methyltransferase (HMT) SUV39H are prone to developing cancer[76] indicates that HMTs for H4K20 could function as tumour-suppressor genes. These matters can now be addressed owing to the development of specific antibodies against SUV4-20H[8], which can be used to determine whether these HMTs are lost in human tumours — a typical feature of tumour-suppressor genes.

A recent study has identified HDAC2 as another component of the epigenetic machinery that is targeted for mutational inactivation in human cancer[77]. This enzyme deacetylates various histone-tail lysines, including those that show an altered acetylation profile in cancer[23]. Cancers associated with microsatellite instability were chosen for study, with the aim of screening those tumours with the highest probability of carrying mutations in any gene. When the exonic repeats of numerous genes involved in DNA methylation and histone modification were analysed, *HDAC2* alone presented an inactivating frameshift mutation[77]. In light of the increasing interest surrounding a human cancer genome project[78], it would be informative to include in the sequencing effort all of the described genes that encode components of the epigenetic machinery.

Translational epigenomics

Epigenomic profiles as cancer cell markers. As described above, recent years have seen the mapping of increasing numbers of genes in which promoter CpG islands are hypermethylated in cancer[1–4]. Such DNA-methylation mapping has revealed unique profiles of hypermethylated CpG islands that define each neoplasia[30,55,56]. The specificity of the assay is increased only if those DNA-methylation markers that are always unmethylated in normal 'healthy' cells are included in this panel. In some cases, such as prostate cancer, a single hypermethylated marker, glutathione S-transferase-π (GSTP1), is informative in 80–90% of cases[79,80]. So far, however, the finding of a highly informative hypermethylation marker such as GSTP1 in other tumour types has been uncommon, and a larger panel is usually needed. For the various cancers for which DNA-methylation profiles are available, CpG-island hypermethylation has been used as a tool to detect cancer cells in all types of biological fluid and biopsy[81]. One of its main advantages over other classical markers is the extreme sensitivity of some of the methods used for the detection of aberrant methylation, such as methylation-specific PCR and MethyLight[81].

Another important finding has been that the CpG-island hypermethylation of tumour-suppressor genes occurs early in tumorigenesis. For example, CpG-island hypermethylation is seen in $p16^{INK4a}$, $p14^{ARF}$ and *MGMT* (*O*-6-methylguanine DNA methyltransferase) in colorectal adenomas and *MLH1* aberrant methylation in atypical endometrial hyperplasia[4]. In this regard, it has been proposed that these epigenetic changes might commit cancer cells to specific altered signal-transduction pathways during the early stages of tumour development[82].

The number of genes undergoing CpG-island-promoter hypermethylation increases during tumorigenic development[44,83]. This finding might be useful in early-detection screenings, especially in individuals with a high familial risk of developing cancer who have similar patterns of CpG-island hypermethylation as sporadic cases[84]. The loss of monoacetylated and trimethylated histone H4 is also an early alteration; for example, both types of alteration first appear in the small skin papillomas that precede the development of non-melanoma skin cancer[8,44].

Epigenomic profiles as markers of tumour prognosis. From the DNA-methylation standpoint, there are instances in which a tumour suppressor that undergoes methylation-associated silencing is a potential candidate for testing as a predictor of tumour prognosis. For example, death-associated protein kinase (DAPK), $p16^{INK4a}$ and epithelial membrane protein 3 (EMP3) hypermethylation have been linked to tumour aggressivity in lung, colorectal and brain cancer patients[4]. Further candidates

Adenoma
A cellular growth of glandular origin, which can arise from organs including the colon and the adrenal, pituitary and thyroid glands. These growths are benign, but some are known to have the potential, over time, to transform to malignancy (at which point they become known as adenocarcinoma.)

awaiting analysis include genes related to increased metastatic potential (such as members of the cadherin and ADAMs families) and angiogenesis (for example, the thrombospondin family). Prognostic dendograms similar to those used in expression microarray analyses involving a combination of hypermethylated markers[85,86] or CpG-island microarrays[87] have also been developed. These epigenomic profiles are complementary to profiles of gene-expression patterns and genetic alterations, but have the advantage that they can be assayed using DNA that has been extracted from archived material[4].

Changes in global levels of individual histone modifications are predictive of the clinical outcome of prostate cancer[23]. Through immunohistochemical staining of primary prostatectomy tissue samples, Seligson and colleagues were able to distinguish two disease subtypes with distinct risks of tumour recurrence in patients with low-grade prostate cancer on the basis of differential staining for the histone acetylation and dimethylation of five residues in histones H3 and H4. These histone-modification patterns were considered to be predictors of outcome independently of other features such as tumour stage, pre-operative prostate-specific antigen levels and capsule invasion[23]. These patterns will be analysed in other tumour types in the near future to assess the potential prognostic use of histone-modification profiles in a more widespread manner.

Pharmacoepigenetics: epigenomics as a predictor of response to chemotherapy. The most compelling evidence that epigenomic profiles can predict responses of cancer to therapy is provided by the methylation-associated silencing of the DNA-repair protein MGMT in human brain tumours. MGMT is directly responsible for reversing the addition of alkyl groups to the guanine base of DNA[68], and this base is the preferred point of attack in the DNA of several alkylating chemotherapeutic drugs, including BCNU (carmustine), ACNU (nimustine), procarbazine, streptozotocin and temozolamide. MGMT hypermethylation is the best independent predictor of response to BCNU[88] and temozolomide[89] in gliomas. The potential of MGMT to predict the chemoresponse of human tumours to alkylating agents can also be extended to other drugs with similar modes of action, such as cyclophosphamide[90]. Additional examples for which hypermethylation predicts drug-responsiveness have been highlighted for other DNA-repair genes identified using CpG-island microarrays[91]. Interestingly, mutations in the genes that encode the epigenetic modification machinery might also predict responses to certain drugs. This is the case for mutations of *HDAC2*, which render those tumours more resistant to the action of certain subtypes of HDAC inhibitor[77].

Epigenomic marks as therapeutic targets. Unlike genetic changes in cancer, epigenetic changes are potentially reversible. In cultured cancer cell lines, it has been possible for years to re-express genes that had been silenced by methylation by using DNA-demethylating agents[92]. When given to patients at low doses, these drugs have shown a significant antitumoral activity, and the US Food and Drug Administration (FDA) has approved the use of two such agents, 5-azacytidine and 5-aza-2′-deoxycytidine, as elective treatments for a pre-leukaemic disease, myelodisplastic syndrome[92].

HDAC inhibitors constitute another promising group of agents for the epigenetic therapy of cancer. One of the main therapeutic mechanisms of action of HDAC inhibitors is their transcriptional reactivation of 'dormant' tumour-suppressor genes, such as *p21*^{WAF1}. However, the pleiotropic nature of these inhibitors raises the possibility that their well-known abilities to induce differentiation, cell-cycle arrest and apoptosis are accompanied by other less desirable outcomes. Despite these concerns, many phase I clinical trials indicate that HDAC inhibitors are well tolerated and, recently, the first drug of this type, suberoylanilide hydroxamic acid (SAHA), has been approved by the FDA for the treatment of cutaneous T-cell lymphoma[92,93]. Our increasing knowledge of the epigenomic profiles of cancer cells, which show an imbalance of histone acetylation[8,23], provides an experimental basis for these treatments and, when characterized in more detail, might provide insights into the specificity of these drugs. This is particularly interesting for the loss of monoacetylated H4K16 (REF. 8), which can be reversed by a new class of drugs that inhibit sirtuins[9], the specific subclass of HDACs that deacetylate H4K16. The ability of sirtuin inhibitors to restore the expression of epigenetically silenced tumour-suppressor genes remains to be comprehensively tested, but would make this class of drugs of high interest for potential clinical use[9].

The future of cancer epigenomics

Although interesting biological insights and promising translational tools are beginning to emerge, cancer epigenomics is still in its infancy. There is a basic need to carry out comprehensive epigenomic profiling in a greater variety of both normal and cancerous cell types, and across a wider range of stages of the tumorigenic process. There is also a need to look at other layers of epigenetic complexity, which are only briefly mentioned here, such as the genomic occupancy profiles of chromatin-remodelling proteins such as the Polycomb and SWI/SNF families[94].

Calls have been made for a comprehensive human epigenome project, with the aim of cataloguing genome-wide profiles for an extensive range of epigenetic marks in various human cell types, as well as in some of the key animal models for human development and disease[95–98]. The first DNA methylomes of *A. thaliana* have recently been published, paving the way for other epigenomes[41,99]. Many researchers from different countries are starting to combine their expertise to work towards characterizing the human epigenome. Given the implications for human health, it would be highly beneficial if this project allowed some of the key questions in cancer epigenomics to be addressed. This ambitious enterprise will involve many decisions, such as the selection of the epigenetic marks to be profiled and the samples to be studied, and the degree of resolution at which profiles are to be obtained.

Prognostic dendogram
A tree diagram that represents the relative similarities among different samples corresponding to human patients in terms of outcome prediction. Samples clustering in the same branch of the dendrogram have the same prognostic markers (for example, age, stage, chromosomal deletions or gains, or specific gene expression) and are likely to have the same outcome.

One important aspect would be for a human epigenome project to provide a well characterized set of profiles from normal human cells, with which cancer cell epigenomes could be compared. Several individual efforts have been meritorious in allowing such comparisons, including the DNA-methylation profiling of the human major histocompatibility complex in various normal tissues from several individuals[100], the characterization of DNA-methylation differences among different tissues in monozygotic twins[101], the profiling of a 4-Mb band of chromosome 2q14.2 in normal and transformed colon cancer cells[102], the coupling of transcriptome and histone-methylation maps in bladder tumours[103], and a recent DNA-methylation profile of human chromosomes 6, 20 and 22 (REF. 104). However, a far wider range of comparisons is required. Sample selection is likely to be a highly discussed issue in the context of epigenome projects. From the cancer perspective, having homogeneous cell populations would be highly advantageous, but achieving this is often challenging owing to the heterogeneity of cell types that are present in many tumour samples. Therefore, although ultimately we would want to sample a wide range of cancer genomes, the best samples to start off with might be tissues in which the progenitor cell, the differentiated cell and the altered disease state have been well characterized, such as in the case of tumours that arise from the haematopoietic system.

Another important decision will relate to whether prioritizing profiling of particular regions of the genome can be used to maximize the information that is likely to be gained. For example, a first draft of the DNA-methylation patterns at the 5′-regulatory regions of all human genes would be a coveted goal. The same can be said for the bisulphite genomic sequencing of all the non-repeat regions of several 'small' chromosomes (such as chromosomes 21, 18, 13, 22 and the Y chromosome). With regard to histone modifications, focusing on those marks that yield more reliable results with the current antibodies, such as the trimethylation of H3K4, would be beneficial, and ChIP-on-chip assays should be developed for these marks.

Ultimately, epigenomic technologies promise to provide us with detailed maps of the DNA methylomes and histone-modification patterns of healthy and cancerous cells. However, increased funding is needed to accomplish these ambitious goals, and continuing international collaboration will clearly play an important part in carrying out the large amount of work that will be required.

1. Herman, J. G. & Baylin, S. B. Gene silencing in cancer in association with promoter hypermethylation. *N. Engl. J. Med.* **349**, 2042–2054 (2003).
2. Feinberg, A. P. & Tycko, B. The history of cancer epigenetics. *Nature Rev. Cancer* **4**, 143–153 (2004).
3. Egger, G., Liang, G., Aparicio, A. & Jones, P. A. Epigenetics in human disease and prospects for epigenetic therapy. *Nature* **429**, 457–463 (2004).
4. Esteller, M. Aberrant DNA methylation as a cancer-inducing mechanism. *Annu. Rev. Pharmacol. Toxicol.* **45**, 629–656 (2005).
5. Nguyen, C. T., Gonzales, F. A. & Jones, P. A. Altered chromatin structure associated with methylation-induced gene silencing in cancer cells: correlation of accessibility, methylation, MeCP2 binding and acetylation. *Nucleic Acids Res.* **29**, 4598–4606 (2001).
6. Fahrner, J. A., Eguchi, S., Herman, J. G. & Baylin, S. B. Dependence of histone modifications and gene expression on DNA hypermethylation in cancer. *Cancer Res.* **62**, 7213–7218 (2002).
7. Ballestar, E. *et al.* Methyl-CpG binding proteins identify novel sites of epigenetic inactivation in human cancer. *EMBO J.* **22**, 6335–6345 (2003).
8. Fraga, M. F. *et al.* Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nature Genet.* **37**, 391–400 (2005).
The first report of a common disruption of histone-modification patterns in cancer cells.
9. Pruitt, K. *et al.* Inhibition of SIRT1 reactivates silenced cancer genes without loss of promoter DNA hypermethylation. *PLoS Genet.* **2**, e40 (2006).
10. Hanahan, D. & Weinberg, R. A. The hallmarks of cancer. *Cell* **100**, 57–70 (2000).
11. 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).
12. Gaudet, F. *et al.* Induction of tumors in mice by genomic hypomethylation. *Science* **300**, 489–492 (2003).
13. Esteller, M. & Almouzni, G. How epigenetics integrates nuclear functions. Workshop on epigenetics and chromatin: transcriptional regulation and beyond. *EMBO Rep.* **6**, 624–628 (2005).
14. Reik, W. & Lewis, A. Co-evolution of X-chromosome inactivation and imprinting in mammals. *Nature Rev. Genet.* **6**, 403–410 (2005).
15. Bodey, B. Cancer-testis antigens: promising targets for antigen directed antineoplastic immunotherapy. *Expert Opin. Biol. Ther.* **2**, 577–584 (2002).
16. Futscher, B. W. *et al.* Role for DNA methylation in the control of cell type specific maspin expression. *Nature Genet.* **31**, 175–179 (2002).
17. Kaneda, A. & Feinberg, A. P. Loss of imprinting of IGF2: a common epigenetic modifier of intestinal tumor risk. *Cancer Res.* **65**, 11236–11240 (2005).
18. Wade, P. A. Methyl CpG-binding proteins and transcriptional repression. *BioEssays* **23**, 1131–1137 (2001).
19. Dobosy, J. R. & Selker, E. U. Emerging connections between DNA methylation and histone acetylation. *Cell. Mol. Life Sci.* **58**, 721–727 (2001).
20. Wang, Y. *et al.* Beyond the double helix: writing and reading the histone code. *Novartis Found. Symp.* **259**, 3–17 (2004).
21. Sanders, S. L. *et al.* Methylation of histone H4 lysine 20 controls recruitment of Crb2 to sites of DNA damage. *Cell* **119**, 603–614 (2004).
22. Espada, J. *et al.* Human DNA methyltransferase 1 is required for maintenance of the histone H3 modification pattern. *J. Biol. Chem.* **279**, 37175–37184 (2004).
23. Seligson, D. B. *et al.* Global histone modification patterns predict risk of prostate cancer recurrence. *Nature* **435**, 1262–1266 (2005).
The first profile of histone modifications in cancer with prognostic relevance.
24. Barbacid, M. *ras* genes. *Annu. Rev. Biochem.* **56**, 779–827 (1987).
25. Fraga, M. F. & Esteller, M. DNA methylation: a profile of methods and applications. *Biotechniques* **33**, 632–649 (2002).
26. Clark, S. J., Harrison, J., Paul, C. L. & Frommer, M. High sensitivity mapping of methylated cytosines. *Nucleic Acids Res.* **22**, 2990–2997 (1994).
27. Herman, J. G. *et al.* Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. *Proc. Natl Acad. Sci. USA* **93**, 9821–9826 (1996).
28. Eads, C. A. *et al.* MethyLight: a high-throughput assay to measure DNA methylation. *Nucleic Acids Res.* **28**, e32 (2000).
29. Uhlmann, K. *et al.* Evaluation of a potential epigenetic biomarker by quantitative methyl-single nucleotide polymorphism analysis. *Electrophoresis* **23**, 4072–4079 (2002).
30. Costello, J. F. *et al.* Aberrant CpG-island methylation has non-random and tumour-type-specific patterns. *Nature Genet.* **24**, 132–138 (2000).
One of the first studies to define the number of hypermethylated CpG islands in cancer cells and their distribution according to tumour type.
31. Zardo, G. *et al.* Integrated genomic and epigenomic analyses pinpoint biallelic gene inactivation in tumors. *Nature Genet.* **32**, 453–458 (2002).
32. Welsh, J. & McClelland, M. Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Res.* **18**, 7213–7218 (1990).
33. 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).
34. Toyota, M. *et al.* Identification of differentially methylated sequences in colorectal cancer by methylated CpG island amplification. *Cancer Res.* **59**, 2307–2312 (1999).
35. 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).
36. 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).
37. Khulan, B. *et al.* Comparative isoschizomer profiling of cytosine methylation: the HELP assay. *Genome Res.* **16**, 1046–1055 (2006).
38. Lopez-Serra, L. *et al.* A profile of MBD protein occupancy of hypermethylated promoter CpG islands of tumor suppressor genes in human cancer. *Cancer Res.* **66**, 8342–8346 (2006).
39. Weber, M. *et al.* Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nature Genet.* **37**, 853–862 (2005).
The first report of the use of methyl-DIP, a powerful epigenomic technique for identifying DNA-methylation changes in transformed cells.
40. Keshet, I. *et al.* Evidence for an instructive mechanism of *de novo* methylation in cancer cells. *Nature Genet.* **38**, 149–153 (2006).
This paper reports extensive mapping of DNA-methylation changes in cancer cells using methyl-DIP.
41. Zhang, X. *et al.* Genome-wide high-resolution mapping and functional analysis of DNA methylation in *Arabidopsis*. *Cell* **126**, 1189–1201 (2006).
A ground-breaking study that combines methyl-DIP and tiling arrays, providing the first example of an almost complete DNA methylome in one organism.
42. 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).

43. Yamashita, K. *et al.* Pharmacologic unmasking of epigenetically silenced tumor suppressor genes in esophageal squamous cell carcinoma. *Cancer Cell* **2**, 485–495 (2002).
References 42 and 43 describe pioneering studies that combine DNA-demethylating agents and expression microarrays to reveal methylation-associated silencing in cancer cells.
44. Fraga, M. F. *et al.* A mouse skin multistage carcinogenesis model reflects the aberrant DNA methylation patterns of human tumors. *Cancer Res.* **64**, 5527–5534 (2004).
45. O'Neill, L. P., Vermilyea, M. D. & Turner, B. M. Epigenetic characterization of the early embryo with a chromatin immunoprecipitation protocol applicable to small cell populations. *Nature Genet.* **38**, 835–841 (2006).
46. Lippman, Z. *et al.* Role of transposable elements in heterochromatin and epigenetic control. *Nature* **430**, 471–476 (2004).
47. Kurdistani, S. K., Tavazoie, S. & Grunstein, M. Mapping global histone acetylation patterns to gene expression. *Cell* **117**, 721–733 (2004).
A pioneering study in which histone-modification changes in yeast were analysed by ChIP-on-chip.
48. Pokholok, D. K. *et al.* Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell* **122**, 517–527 (2005).
49. Schubeler, D. *et al.* The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. *Genes Dev.* **18**, 1263–1271 (2004).
50. Martens, J. H. *et al.* The profile of repeat-associated histone lysine methylation states in the mouse epigenome. *EMBO J.* **24**, 800–812 (2005).
51. Bernstein, B. E. *et al.* Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* **120**, 169–181 (2005).
52. Bernstein, B. E. *et al.* A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**, 315–326 (2006).
53. Azuara, V. *et al.* Chromatin signatures of pluripotent cell lines. *Nature Cell Biol.* **8**, 532–538 (2006).
54. Yu, L. *et al.* Global assessment of promoter methylation in a mouse model of cancer identifies ID4 as a putative tumor-suppressor gene in human leukemia. *Nature Genet.* **37**, 265–274 (2005).
55. Esteller, M., Corn, P. G., Baylin, S. B. & Herman, J. G. A gene hypermethylation profile of human cancer. *Cancer Res.* **8**, 3225–3229 (2001).
The most comprehensive analysis of promoter CpG-island hypermethylation in human cancer using a candidate-gene approach.
56. Paz, M. F. *et al.* A systematic profile of DNA methylation in human cancer cell lines. *Cancer Res.* **63**, 1114–1121 (2003).
57. Sjoblom, T. *et al.* The consensus coding sequences of human breast and colorectal cancers. *Science* **314**, 268–274 (2006).
58. Paz, M. F. *et al.* Genetic unmasking of epigenetically silenced tumor suppressor genes in colon cancer cells deficient in DNA methyltransferases. *Hum. Mol. Genet.* **12**, 2209–2219 (2003).
59. Saito, Y. *et al.* Specific activation of microRNA-127 with downregulation of the proto-oncogene *BCL6* by chromatin-modifying drugs in human cancer cells. *Cancer Cell* **9**, 435–443 (2006).
60. Lujambio, A. *et al.* Genetic unmasking of an epigenetically silenced microRNA in human cancer cells. *Cancer Res.* **67**, 1424–1429 (2007).
References 59 and 60 describe the first examples of CpG-island methylation-associated silencing of microRNAs in human cancer.
61. Schlesinger, Y. *et al.* Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for *de novo* methylation in cancer. *Nature Genet.* **39**, 232–236 (2006).
62. Widschwendter, M. *et al.* Epigenetic stem cell signature in cancer. *Nature Genet.* **39**, 157–158 (2006).
63. Di Croce, L. *et al.* Methyltransferase recruitment and DNA hypermethylation of target promoters by an oncogenic transcription factor. *Science* **295**, 1079–1082 (2002).
64. Esteller, M. *et al.* Cancer epigenetics and methylation. *Science* **297**, 1807–1808 (2002).
65. Brenner, C. *et al.* Myc represses transcription through recruitment of DNA methyltransferase corepressor. *EMBO J.* **24**, 336–346 (2005).
66. Vire, E. *et al.* The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* **439**, 871–874 (2006).
67. Richon, V. M., Sandhoff, T. W., Rifkind R. A. & Marks, P. A. Histone deacetylase inhibitor selectively induces p21WAF1 expression and gene-associated histone acetylation. *Proc. Natl Acad. Sci. USA* **97**, 10014–10019 (2000).
68. Esteller, M. & Herman, J. G. Generating mutations but providing chemosensitivity: the role of O6-methylguanine DNA methyltransferase in human cancer. *Oncogene* **23**, 1–8 (2004).
69. Tryndyak, V. P., Kovalchuk, O. & Pogribny, I. P. Loss of DNA methylation and histone H4 lysine 20 trimethylation in human breast cancer cells is associated with aberrant expression of DNA methyltransferase 1, Suv4–20h2 histone methyltransferase and methyl-binding proteins. *Cancer Biol. Ther.* **5**, 65–70 (2006).
70. Pogribny, I. P. *et al.* Histone H3 lysine 9 and H4 lysine 20 trimethylation and the expression of Suv4–20h2 and Suv-39h1 histone methyltransferases in hepatocarcinogenesis induced by methyl deficiency in rats. *Carcinogenesis* **27**, 1180–1186 (2006).
71. Jacobsen, S. E. & Meyerowitz, E. M. Hypermethylated *SUPERMAN* epigenetic alleles in *Arabidopsis*. *Science* **277**, 1100–1103 (1997).
72. Mager, J. & Bartolomei, M. S. Strategies for dissecting epigenetic mechanisms in the mouse. *Nature Genet.* **37**, 1194–1200 (2005).
73. Ozdag, H. *et al.* Differential expression of selected histone modifier genes in human solid cancers. *BMC Genomics* **7**, 90 (2006).
74. Yang, X. J. The diverse superfamily of lysine acetyltransferases and their roles in leukemia and other diseases. *Nucleic Acids Res.* **32**, 959–976 (2004).
75. Moore, S. D. *et al.* Uterine leiomyomata with t(10;17) disrupt the histone acetyltransferase *MORF*. *Cancer Res.* **64**, 5570–5577 (2004).
76. Peters, A. H. *et al.* Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell* **107**, 323–337 (2001).
77. Ropero, S. *et al.* A truncating mutation of HDAC2 in human cancers confers resistance to histone deacetylase inhibition. *Nature Genet.* **38**, 566–569 (2006).
The first-described disruption of a histone deacetylase in human cancer provides a basis for further pharmacogenomic studies.
78. Varmus, H. & Stillman, B. Support for the Human Cancer Genome Project. *Science* **310**, 1615 (2005).
79. 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).
80. Cairns, P. *et al.* Molecular detection of prostate cancer in urine by GSTP1 hypermethylation. *Clin. Cancer Res.* **7**, 2727–2730 (2001).
81. Laird, P. W. The power and the promise of DNA methylation markers. *Nature Rev. Cancer* **3**, 253–266 (2003).
82. Baylin, S. B. & Ohm, J. E. Epigenetic gene silencing in cancer — a mechanism for early oncogenic pathway addiction? *Nature Rev. Cancer* **6**, 107–116 (2006).
83. Gallagher, W. M. *et al.* Multiple markers for melanoma progression regulated by DNA methylation: insights from transcriptomic studies. *Carcinogenesis* **26**, 1856–1867 (2005).
84. Esteller, M. *et al.* DNA methylation patterns in hereditary human cancers mimic sporadic tumorigenesis. *Hum. Mol. Genet.* **10**, 3001–3007 (2001).
85. Alaminos, M. *et al.* Clustering of gene hypermethylation associated with clinical risk groups in neuroblastoma. *J. Natl Cancer Inst.* **96**, 1208–1219 (2004).
86. Tsou, J. A. *et al.* Distinct DNA methylation profiles in malignant mesothelioma, lung adenocarcinoma, and non-tumor lung. *Lung Cancer* **47**, 193–204 (2005).
87. Wei, S. H. *et al.* Prognostic DNA methylation biomarkers in ovarian cancer. *Clin. Cancer Res.* **12**, 2788–2794 (2006).
88. 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).
The first translational use of a hypermethylated CpG island in the management of cancer patients.
89. Hegi, M. E. *et al. MGMT* gene silencing and benefit from temozolomide in glioblastoma. *N. Engl. J. Med.* **352**, 997–1003 (2005).
90. 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).
91. Glasspool, R. M., Teodoridis, J. M. & Brown, R. Epigenetics as a mechanism driving polygenic clinical drug resistance. *Br. J. Cancer* **94**, 1087–1092 (2006).
92. Mack, G. S. Epigenetic cancer therapy makes headway. *J. Natl Cancer Inst.* **98**, 1443–1444 (2006).
93. Thompson, C. A. Vorinostat approved for rare lymphoma. *Am. J. Health Syst. Pharm.* **63**, 2168 (2006).
94. Seo, S. & Kroll, K. L. Geminin's double life: chromatin connections that regulate transcription at the transition from proliferation to differentiation. *Cell Cycle* **5**, 374–379 (2006).
95. Jones, P. A. & Martienssen, R. A blueprint for a Human Epigenome Project: the AACR Human Epigenome Workshop. *Cancer Res.* **65**, 11241–11246 (2005).
96. Rauscher, F. J. It is time for a Human Epigenome Project. *Cancer Res.* **65**, 11229 (2005).
97. Garber, K. Momentum building for Human Epigenome Project. *J. Natl Cancer Inst.* **98**, 84–86 (2006).
98. Esteller, M. The necessity of a Human Epigenome Project. *Carcinogenesis* **27**, 1121–1125 (2006).
99. Zilberman, D. *et al.* Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. *Nature Genet.* **39**, 61–69 (2006).
A high-resolution map of the *A. thaliana* DNA methylome that addresses its relationship to gene-expression patterns.
100. Rakian, V. K. *et al.* DNA methylation profiling of the human major histocompatibility complex: a pilot study for the Human Epigenome Project. *PLoS Biol.* **2**, e405 (2004).
101. Fraga, M. F. *et al.* Epigenetic differences arise during the lifetime of monozygotic twins. *Proc. Natl Acad. Sci. USA* **102**, 10604–10609 (2005).
A comprehensive epigenomic study that demonstrates the presence of DNA methylation and histone modifications in individuals who have the same genetic background.
102. Frigola, J. *et al.* Epigenetic remodeling in colorectal cancer results in coordinate gene suppression across an entire chromosome band. *Nature Genet.* **38**, 540–549 (2006).
103. Stransky, N. *et al.* Regional copy number-independent deregulation of transcription in cancer. *Nature Genet.* **38**, 1386–1396 (2006).
104. Eckhardt, F. *et al.* DNA methylation profiling of human chromosomes 6, 20 and 22. *Nature Genet.* **38**, 1378–1385 (2006).

Acknowledgements
M.E. research is supported by the Science Department of the Spanish Government.

Competing interests statement
The authors declare no competing financial interests.

DATABASES
The following terms in this article are linked online to:
Entrez Gene: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene
BRCA1 | CBP | DAPK | EMP3 | GSTP1 | HBO1 | HDAC2 | ID4 | MeCP2 | *MGMT* | *MLH1* | MORF | MOZ | $p16^{INK4a}$ | SIRT1 | *SUPERMAN* | SUV39H | *VHL* | *WRN*

FURTHER INFORMATION
Cancer Epigenetics homepage: http://www.cnio.es/ing/grupos/plantillas/presentacion.asp?grupo=50004270
Epigenetics Society: http://www.dnamethsoc.com/
Epigenome Network of Excellence: http://www.epigenome-noe.net/
Human Epigenome Project: http://www.epigenome.org/
Manel Esteller's Laboratory: http://www.cancerepigenetics.com/
Access to this links box is available online.