Published in final edited form as:

Reprod Toxicol. 2007 ; 23(3): 267–282.

Techniques used in studies of epigenome dysregulation due to aberrant DNA methylation: an emphasis on fetal-based adult diseases

Shuk-mei Ho^{a,b} and Wan-ye Tang^a

^a Department of Environmental Health, College of Medicine, University of Cincinnati, Cincinnati, Ohio, U.S.A.

^b Cancer Center, College of Medicine, University of Cincinnati, Cincinnati, Ohio, U.S.A.

Abstract

Epigenetic changes are heritable modifications that do not involve alterations in the primary DNA sequence. They regulate crucial cellular functions such as genome stability, X-chromosome inactivation, and gene imprinting. Epidemiological and experimental observations now suggest that such changes may also explain the fetal basis of adult diseases such as cancer, obesity, diabetes, cardiovascular disorders, neurological diseases, and behavioral modifications. The main molecular events known to initiate and sustain epigenetic modifications are histone modification and DNA methylation. This review specifically focuses on existing and emerging technologies used in studying DNA methylation, which occurs primarily at CpG dinucleotides in the genome. These include standard exploratory tools used for global profiling of DNA methylation and targeted gene investigation: methylation sensitive restriction fingerprinting (MSRF), restriction landmark genomic scanning (RLGS), methylation CpG island amplification-representational difference analysis (MCA-RDA), differential methylation hybridization (DMH), and cDNA microarrays combined with treatment with demethylating agents and inhibitors of histone deacetylase. The basic operating principals, resource requirements, applications, and benefits and limitations of each methodology are discussed. Validation methodologies and functional assays needed to establish the role of a CpG-rich sequence in regulating the expression of a target or candidate gene are outlined. These include *in silico* database searches, methylation status studies (bisulfite genomic sequencing, COBRA, MS-PCR, MS-SSCP), gene expression studies, and promoter activity analyses. Our intention is to give readers a starting point for choosing methodologies and to suggest a workflow to follow during their investigations. We believe studies of epigenetic changes such as DNA methylation hold great promise in understanding the early origins of adult diseases and in advancing their diagnosis, prevention, and treatment.

Keywords

Cytosine methylation; chromatin remodeling; fetal based adult disease; epigenetics; genome-wide methylation profiling; methylation sensitive restriction fingerprinting; restriction landmark genomic scanning (RLGS)

Corresponding author: Tel: 513-558-5701; fax: 513-558-0071. E-mail address: shuk-mei.ho@uc.edu (S-M Ho)

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. DNA methylation as an epigenetic mechanism of fetal-based adult diseases

1.1 Epigenetic modulation of fetal programming

Past research has identified mutations, deletions, gene fusion, tandem duplications, and gene amplifications as key mechanisms that dysregulate expression of disease-predisposing or disease-determining genes at the linear DNA level [1-7]. However, it has become clear that disruption of epigenetic regulation of gene expression plays an equally important role in the development of diseases [8-10].

The term *epigenetic* means *outside conventional genetics* [11]. Epigenetic changes are reversible, heritable modifications that do not involve alterations in the primary DNA sequence. Three distinct and intertwined mechanisms are now known to initiate and sustain epigenetic modifications: small-interfering RNAs, DNA methylation, and histone modification [12-14]. These processes affect transcript stability, DNA folding, nucleosome positioning, chromatin compaction, and ultimately nuclear organization. Singularly or conjointly, they determine whether a gene is silenced or activated. Dysregulation of these processes certainly is the possible mechanism underpinning the epigenetic basis of disease development [8,9]. Disease susceptibility, therefore, is a result of a complex interplay between one's genetic endowment and epigenetic modulations induced by endogenous or exogenous environmental cues.

Epigenetic modifications of disease risk could begin as early as during fetal development and be transmitted transgenerationally [15-20]. The paradigm of fetal basis of adult disease first emerged from large-cohort epidemiological studies linking poor growth *in utero* with adult diseases [16,21-22]. During pregnancy, maternal conditions such as nutritional deficits, infection, hypertension, diabetes, or hypoxia expose the fetus to hormonal and metabolic cues that induce “fetal programming.” It alters the courses of cellular and organ differentiation *in utero* and permanently affects the functional capacity of adult organs in later stages of life [15,22]. From an evolutionary perspective, fetal programming is an “adaptive” trait since it allows the fetus to make anticipatory responses to the external environment to gain advantages for later life challenges. However, contemporary human life is greatly influenced by lifestyle choices that are in conflict with the programmed adaptive changes made during fetal development. In addition, synthetic agents that mimic internal cues can alter the course of fetal programming adversely. Both could cause insufferable consequences in later life, leading to heightened disease susceptibility. Classical examples include the association of lower birth weight with a greater risk for adult onset of cardiovascular disease [23], Type 2 diabetes mellitus [24], osteoporosis [25], and depression [26]. The link between exposure to the synthetic estrogen diethylstilbestrol (DES) *in utero* and increased incidence of reproductive tract cancers in “DES daughters” has been a difficult lesson learned by the health care community [27]. Genetic factors, such as telomere attrition [28] and polymorphisms in mitochondrial DNA [29], may in part mediate fetal programming. However, epigenetic dysregulation of gene expression is currently a widely accepted mechanism of fetal-based adult disease [15-20].

The two main epigenetic mechanisms currently recognized as playing a role in the fetal basis of adult disease are histone modification and DNA methylation [15-20]. A comprehensive review of how these processes affect gene transcription is beyond the scope of this review. In simple terms, the histone modification refers to post-translational modifications of histone tails, while DNA modification involves methylation of cytosine at the carbon-5 position in CpG dinucleotides. The two processes work together to affect chromatin packaging of DNA, which, in turn, determines which gene or gene set is transcribable. Changes mediated by either process are heritable, not only transmittable to the daughter cells, but to subsequent generations [19, 30]. Thus, interest in the field of epigenetic control of fetal-based disease has increased dramatically within the last few years. Our intention with this review is to give readers a starting

point for choosing methodologies and a workflow to follow during their investigations. The past decade has witnessed an exponential increase in novel approaches to the conduct of epigenetic analysis [31]. Methodologies applicable for the investigation of histone modifications have recently been reviewed [13]. This review will focus primarily on contemporary methods designed to elucidate DNA methylation—regulated gene expression as a mechanism of early origins of adult diseases.

1.2 DNA methylation: chemistry, developmental dynamics, and proposed functions

DNA methylation refers to the covalent addition of a methyl group derived from S-adenosyl-L-methionine to the fifth carbon of the cytosine ring to form the fifth base, 5-methyl cytosine (5mC; Figure 1) [32,33]. The reaction is catalyzed by DNA methyltransferases and accessory proteins (Dnmt1, Dnmt3a, Dnmt 3b, Dnmt2, and Dnmt 3L). Across eukaryotic species, methylation occurs predominantly in cytosines located 5' of guanines, known as CpG dinucleotides (CpGs), although methylation also takes place in non-CpGs such as CpNpG and nonsymmetrical CpA and CpT at a lower frequency.

In the mammalian genome, the distribution of CpG sequences is nonrandom [34]. Because of the high susceptibility of 5mC to undergo spontaneous deamination to yield thymine (Figure 1), the mammalian genome has become progressively depleted of CpGs through the course of evolution. CpGs are normally under-represented, appearing at a low frequency of 1 per 80 dinucleotides in 98% of the mammalian genome. In contrast, CpGs are found as clusters known as CpG islands (CGIs) in 1–2% of the genome. Typically, a CGI ranges in length from 200 bp to 5 kb, has a high percentage of CG (>60%), and a ratio of CpG to GpC of at least 0.6. CGIs are normally unmethylated and presumably protected from spontaneous deamination. Important exceptions to the unmethylated status of CGIs include those that are associated with imprinted genes, genes subject to X-chromosome inactivation, and transposable elements [33,35]. About 70% of CGIs are associated with DNA sequences 1- to 2-kb long located in the promoter, the first and second exons, and the first intron regions of all genes (5' CGIs) [36]. Most 5' CGIs are found overlapping with the transcription start site, suggesting that they are important for gene regulation.

It has been proposed that over half of the 5' CGIs participate in the regulation of gene transcription [33,35], although, it is important to note, not all CGIs are involved in gene regulation. An inverse relationship usually exists between the extent of methylation of a regulatory CGI and gene transcription [35]. Two mechanisms have been proposed to explain how cytosine methylation leads to repression of gene transcription [33]. First, the methyl group of the 5mC extends into the major groove of DNA and inhibits binding of transcription factors (TFs) to their CpG-containing recognition sites. Second, a class of proteins known as methyl-binding proteins (MBDs) specifically bind methylated CGIs and create steric hindrance to access by TFs to their regulatory elements. Both scenarios will suppress gene transcription. Furthermore, upon binding to methylated CGIs, MBDs recruit histone deacetylases (HDACs) and histone methyltransferases (HMTs). These enzymes mediate complex histone modifications and result in the establishment of repressive chromatin structures that permanently silent gene transcription [35,37].

DNA methylation is viewed, at the global level, as a generalized repression system in more complex genomes [38]. It is thought to repress inappropriate expression of endogenous transposons that may disrupt the organization and integrity of the genome and to be the key mechanism responsible for X-chromosome inactivation and genomic imprinting [33,35]. Recent evidence suggests that DNA methylation plays a crucial role in the establishment and/or maintenance of cell- or tissue-specific gene expression in adult somatic tissues.

DNA methylation pattern is established through defined phases during the development of an organism. Gamete methylation patterns are erased by a genome-wide demethylation at around the eight-cell stage of blastocyst formation. During the implantation stage, methylation patterns are established via *de novo* methylation. In adulthood, the amount and pattern of methylation are tissue- and cell-type specific. Disruption of these preset patterns of DNA methylation in adult life has been linked to aging and disease development [39-41]. Furthermore, deregulation of fetal programming by suboptimal maternal factors or environmental mimics is now believed to involve abnormal DNA methylation of specific genes that permit them to undergo either inappropriate or untimely expression in adult life, leading to disease development. Recent and limited evidence supports this as a key mechanism of the fetal basis of adult disease (see below).

1.3 DNA methylation in fetal basis of adult disease

Several lines of direct and indirect evidence support the hypothesis that DNA methylation is a key mechanism mediating fetal-based adult disease development. Fetal exposure of agouti mice to methyl donors (folic acid, choline, and betaine), which is expected to induce global DNA methylation changes, causes shifts in the coat color from yellow to brown and reduces the risk of obesity, diabetes, and cancer [42]. Global gene profiling identified altered expression of 27 genes with regulatory 5' CGIs in mice born to mothers on a choline-deficient diet [43]. These mice also exhibited abnormal hippocampal development and loss of memory function in adult life.

McLachlan and associates [44] provided the first evidence in support of the effects of early-life exposure on DNA methylation of specific genes that may show dysregulated expression in disease target organs later in life. These investigators found a high incidence of uterine cancer and elevated expression of *c-fos* and *lactoferrin* genes in mice exposed neonatally to diethylstilbestrol (DES) [45]. In concordance with epigenetic regulation of gene expression as a underpinning the early origin of adult disease, these investigators reported demethylation of a single CpG site in the lactoferrin promoter and hypomethylation of the promoter and intron-1 regions of *c-fos* in the uterus of adult mice exposed neonatally to DES [46]. In a multi-generations study, Skinner and associates [19,47] demonstrated induction of 15 imprinted-like genes/DNA sequences with altered methylation patterns in sperm obtained from mice exposed developmentally to vinclozolin. Importantly, the altered methylation patterns in these sequences persist in sperm of subsequent non-exposed generations. Although these findings did not provide a direct link between changes in methylation status of the susceptible genes and disruption of their expression in later life, they are highly suggestive that such a relationship exists. More direct evidence linking epigenetic re-programming via DNA methylation and aberrant gene expression in adult tissues and altered disease susceptibility was provided by a recent study showing increased prostate cancer risk in rats exposed neonatally to bisphenol A or estradiol [20]. In this study, the investigators demonstrated concomitant, aberrant overexpression of *phosphodiesterase 4 variant 4 (PDE4D4)* and hypomethylation in a 5' CGI of *PDE4D4* in the prostate of rats exposed early in life to the estrogen and its mimic. Additional cell-culture experiments further substantiated the claim for this interrelationship.

Since the fetal basis of adult disease is a rapidly growing field, it would be timely to provide a comprehensive review of methods and techniques available to interested investigators. It is our intention to provide readers with a starting point in their choice of methodologies and a workflow to follow during their investigations. The basic operating principals, resource requirements, applications, and benefits and limitations of each methodology are discussed below. Validation methodologies and functional assays necessary to establishing the role of a CpG-rich sequence in regulating the expression of a target or candidate gene are briefly outlined. We recognize that this is a rapidly growing field, with many innovative methodologies

and approaches emerging daily, and thus that it will be impossible for us to cover this topic exhaustively.

2. Techniques and methods

In this review we will outline methodologies that would allow investigators to globally discover CpGs or CGIs that are differentially methylated in different tissues/cells or under different experimental conditions. Many of the methods are also applicable to investigations in which researchers already have target genes in mind. In the course of describing methodologies, we will first introduce the operating principals, then give examples of applications, and finally discuss the benefits and limitations for each technique. Table 1 summarizes some of the key features of each methodology and their advantages and limitations for easy reading and for facilitating decision-making regarding the choice of a methodology for a particular study.

To identify candidates under methylation regulation and to determine whether a newly identified CGI is involved in gene regulation, one needs undertake a comprehensive and vigorous investigative approach (Figure 2). The typical experimental workflow for a discovery platform involves 1) identification of candidates through genome-wide screening methods, 2) *in silico* database analyses to shortlist candidates that contain 5' CGIs or CpGs, 3) confirmation of changes in methylation status in the candidate CpGs under the expected experimental conditions, 4) establishment of a causal or correlative relationship between increased cytosine methylation and gene silencing, and 5) other supportive functional assays to uncover the functionality of the cognate genes. However, if a target gene is to be studied, the entry point could begin at step 2). All the subsequent steps are applicable for target-gene studies.

2.1 Methods for identification of candidates

Currently, there is a wide range of approaches to obtaining quantitative and qualitative information on changes in genomic DNA methylation. Several standard exploratory tools, including methylation sensitive restriction fingerprinting (MSRF), restriction landmark genomic scanning (RLGS), methylation CpG island amplification-representational difference analysis (MCA-RDA), differential methylation hybridization (DMH), methylation-specific oligonucleotide microarrays (MSO), methylation target array (MTA), and cDNA microarrays combined with treatment of demethylating agents, are routinely used to identify putative regulatory CGIs at the genome-wide level, as well as to analyze methylation status of specific CGIs of known genes in large sample sets. Other advanced techniques, such as chromatin immunoprecipitation (ChIP)-on chips, polymerase extension assay by pyrosequencing, mass spectrometry, and HPLC, have emerged as high-throughput methods for obtaining information of CpG methylation outside/inside CGIs. These techniques also have the potential for being used in a whole-genome screen.

2.1.1 Methylation sensitive restriction fingerprinting (MSRF)—MSRF (Figure 3) is a sensitive PCR-based method that allows one to screen for novel CpG-rich sequences that exhibit differential changes in cytosine methylation under different physiological or pathological conditions [48]. Multiple samples can be compared simultaneously. Genomic DNA isolated from tissues is subjected to *MseI* restriction digestion. Since the restriction site of *MseI* is TTAA, which is rarely found in CG-rich regions, cellular DNA is digested into small fragments, leaving most of the CG-rich region intact. The *MseI*-digested DNA is then divided into two aliquots with one aliquot left unmodified and the other subjected to digestion with *BstUI*, a methylation-sensitive restriction enzyme. *BstUI* cuts at CGCG, a sequence that occurs in more than 80% of CGIs, and therefore will cut most CG-rich regions unless they are protected by methylation. Short DNA sequences with methylated *BstUI* sites are left uncut and can subsequently be amplified by PCRs, whereas sequences with unmethylated *BstUI* sites are cut

and will not yield any PCR products. Aliquots of the single (*MseI*)- or double (*MseI* and *BstUI*)-digested DNA are then subjected to PCR amplification using different pairs of short arbitrary primers (10mer with at least one CG site in the primer) [48] in the presence of radiolabeled [³²P]-dNTPs. Labeled PCR products are size-fractionated on 6% polyacrylamide gels. “Candidates” with differentially methylated patterns among different samples are visualized with autoradiography, which is used to guide excision of “candidates” from the dried sequencing gels. Amplification and sequencing of the candidates followed by sequence database analyses reveal the identities of the candidates and possibly their innate genes.

We have successfully used MSRF to globally profile the prostate epigenome for genes whose promoter/5' CGIs undergo aberrant cytosine methylation following neonatal exposure to estrogens or bisphenol A [20]. Among the initial 50 candidate sequences identified, three were confirmed to be part of the promoter regions of genes regulated by methylation. One of the genes, phosphodiesterase 4D4, was further characterized for its functional role in modulating prostate cancer risk in adulthood [20]. In an earlier study [49], using a similar approach, we identified hypermethylation of peroxisome membrane protein 24 as a marker/modifier for the transition of prostate cancer cells from an androgen-dependent to an androgen-independent state. Others have used this method to identify genes whose promoters are aberrantly methylated in breast cancers [48], nasopharyngeal carcinoma [50], and hepatocellular carcinoma [51]. This method has also been used to identify novel epigenetically modified candidates during stem cell differentiation [52].

With radiolabeling of PCR products, only a small amount of genomic DNA from each sample is required (100 ng–1 µg) (Table 1). Therefore, this method is highly applicable to studies using laser-capture microdissected samples or for developmental studies in which tissue availability is a key concern. Usually products from 4–8 sets of arbitrary primers are run on each gel and only a limited number of candidates (30–40) can be identified in one autoradiograph. The ease of setting up MSRF for methylation profiling of multiple samples is definitely an advantage of this methodology, since it only requires the use of a simple sequencing gel set-up, standard PCR protocols, and routine amplification and sequencing techniques. By increasing the number of arbitrary primers used in MSRF, one can increase the coverage of the genome and the number of candidate sequences identified. For example, in two of the experiments we have conducted [20,49], both using 4 primers and 6 permutations of primer pairs, we identified approximately 50 candidate sequences. However, if 9 primers were to be used, a total of $36 [n \times (n-1)/2; n \text{ is the number of primers}]$ permutations for arbitrary PCRs could be achieved, likely yielding over 1,800 candidates.

2.1.2 Restriction landmark genomic scanning (RLGS)—The RLGS approach was first developed to identify imprinting genes [53] and later adapted for genome-wide screening of methylation changes in CGIs (Figure 4). It can evaluate the methylation status of thousands of CG-rich sequences and simultaneously obtain information on copy number of their cognate genes [53–56]. Methylation detection in RLGS profiles depends on the methylation sensitivity of the restriction enzymes used to cut the genomic DNA. *NotI*, which recognizes GC-rich sequences, is most commonly used to generate thousands of landmark sites on the gel. Differences in digestion are assessed by radiolabeling the DNA at cleaved *NotI* sites. Following further endonuclease digestion, two-dimensional electrophoretic separation, and autoradiography, the intensity of a DNA fragment on the resulting RLGS profile quantitatively reflects the copy number and methylation status of the *NotI* fragment (Figure 4). Early applications of RLGS for identification of novel differentially methylated genes were hampered by the tedious task of having to obtain sequence information on the identified sequences. With the completion of several commonly used genome databases (human and mouse), development of interactive informatics tools, and tailored software programs, it is now possible to conduct automated RLGS fragment prediction and to download corresponding

sequences for mouse [57] and human studies [58]. Users can directly upload or query RLGS databases at <http://genome.gsc.riken.go.jp/RLGS/RLGShome.html> or <http://dot.ped.med.umich.edu:2000/VGS/index.html>).

RLGS has been used to analyze epigenetic changes due to aberrant DNA methylation for breast cancer [59], ovarian cancer [60], and hepatocellular carcinoma [61]. Recently, Sato and co-workers applied this method to illustrate the aberrant methylation of genomic DNA in the epididymis of mice neonatally exposed to DES [62]. Seven loci of the genomic DNA were found to be aberrantly demethylated, and one locus in the epididymis of DES-treated mice was found to be methylated. Among these putative candidates, four were confirmed to have promoter or 5' CGIs. However, additional experiments necessary to establish the regulatory role of these CGIs on their cognate gene expression were not included in this study. This omission of validation experiments has prevented the authors from drawing definitive conclusions on the significance of DNA methylation changes in these CG-rich sequences or their cognate genes in mediating DES-induced abnormalities in this organ.

An obvious advantage of RLGS is its ability to identify thousands of loci/landmark fragments in a single run (Table 1). With continued improvements on the various RLGS sequence databases, sequence identification of the loci should no longer be a limitation of this methodology. At least for human and mouse studies, the current automated RLGS fragment prediction programs and sequence databases are quite adequate, yet challenges still exist for studies using samples from other species. Furthermore, since the methodology requires at least a few micrograms of DNA, sample availability may be a limitation for some investigations. Other demands for successful utilization of RLGS are the requirements for a fairly elaborate gel electrophoresis set-up and a powerful image analysis system. Finally, limited library coverage is another constraint for this approach unless new advanced genome sequence-based tools are developed in the future.

2.1.3 Methylation CpG island amplification-representational difference analysis (MCA-RDA)

Methylated CGI amplification (MCA) coupled with representational difference analyses (RDA) [63,64] provides a powerful approach [65] for simultaneous identification and cloning of novel differentially methylated sequences between two samples (Figure 5). The fundamental principle of MCA involves amplification of DNA sequences with closely spaced (<1 kb) methylated *Sma*I (CCCGGG) sites, which are commonly found in CGIs. Since only short fragments (400–600 bp) flanked by two *Sma*I sites are amplified, MCA ensures enrichment of CGIs [65]. A pair of enzymes, *Sma*I (methylation-sensitive) and *Xma*I (methylation insensitive), that cut the same recognition sequence CCCGGG are used in the protocol (Figure 5). The two DNA samples under investigation are first subjected to *Sma*I digestion to remove unmethylated *Sma*I sites, generating only blunt end fragments in DNA regions that with no methylated *Sma*I sites; i.e., hypomethylated CGI regions. These blunt end sequences will not be amplified by subsequent steps of the protocol and thus are eliminated from the DNA pools. The samples are then digested with the *Sma*I isoschizomer *Xma*I, which cuts at methylated *Sma*I sites, generating fragments with a four-base overhang. Adaptors are ligated to this overhang before the fragments are amplified by PCR using primers complementary to the adaptors. Southern or dot blotting analysis is employed to determine if a candidate CGI is differentially methylated among samples.

To identify novel CGIs that are differentially methylated in two different samples, RDA, a subtraction technique developed to clone small differences between genomes [63], is used following MCA, a technique referred to as MCA-RDA. The RDA technique is based on subtraction of “tester” sequences from the “driver” pool of sequences, followed by PCR amplification of the tester sequences left un-hybridized. Differentially present clones, after the subtraction, are identified by subcloning and sequencing.

MCA-RDA was first used to identify novel hypermethylated CGIs in colorectal cancer [65]. More recently, several groups have used this method to identify genomic clones that are hypermethylated in microsatellite instability-positive sporadic colorectal cancers [66], hypermethylated DNA fragments as putative biomarkers of prostate cancer [67], and hypermethylated sequences uniquely expressed in pancreatic carcinoma but not in normal pancreatic cells [68].

The MCA assay can be adapted for medium-throughput determination of the methylation status of a large number of target genes through the utilization of “printed” membranes or highly reproducible dot-blots [69]. The costs and efforts for these set-ups will depend largely on the numbers of samples and genes under investigation. Similarly, the combined MCA-RDA approach has proven to be highly effective in identifying hundreds of CGIs containing genomic fragments between two samples. However, its limitation resides in the inability to compare more than two samples in a single subtraction experiment. Furthermore, the method is not conducive to the identification of aberrantly hypomethylated sequences.

2.1.4 Differential methylation hybridization (DMH) using CGI arrays—An affinity column that contained the methyl-CpG binding domain of the rat MeCP2 protein was constructed to isolate CGIs from human genomic DNA [70]. Using this CGI library (close to 8,000 CGIs), Huang et al. developed a novel array-based method called DMH that allows for a genome-wide screening for differentially methylated CGIs between two different samples [71]. The method has been widely used in the identification of aberrantly methylated gene promoters that are differentially expressed in various cancers [71-75]. In addition to its use in cancer research, DMH has been used to uncover specific CGIs that are altered in the prostates of mice fed genistein-rich diets [76]. Details of the methodology have been reviewed [77] and here we have illustrated only key steps and general principles underlying the procedures (Figure 6). Genomic DNA isolated from two different samples is first digested with *MseI* to generate small DNA fragments (100–200 bp), leaving most CGIs intact. The cut ends of the DNA fragments are ligated to linkers. Repetitive sequences are removed from the digests using Cot-1 subtractive hybridization [78]. The remaining DNA fragments are then subjected to methylation-sensitive endonuclease *BstUI* digestion. Following *BstUI* digestion, DNA fragments in the two samples are amplified and labeled with cyanine (Cy) 3 or Cy5. The fluorescent-labeled amplicons, representing differentially methylated DNA fragments between the two samples, are co-hybridized to a high-density CGI microarray with approximately 8,000 thousand probes. Cy3 and Cy5 fluorescence on each probe is detected in a two-channel scanner. Differences in methylation of a particular CGI between two samples are reflected in fluorescence intensities detected from the Cy3- and Cy5-labeled targets hybridized on the corresponding CGI probe in the microarray.

Although the “first-generation” CGI arrays contain approximately 8,000 probes, most published reports [71-75] found that only 0.5–1.5% (40–120 candidates) of the probes on the arrays were differentially methylated in the two samples under comparison. Following additional *post hoc* conformational analyses, most studies identified only a handful (8–15) of genes with promoter CGIs that play a role in gene regulation. The original DMH utilized CGIs dotted on nylon membranes, which have high background signal, thus creating problems with a high noise-to-signal ratio. The fabrication of current arrays on glass slides has minimized this problem. Recently, the availability of new high-density human and mouse CGI arrays have greatly expanded the power of using DMH for a whole-genome screen of differential methylation [79].

In summary, DMH is a useful tool for the discovery of differentially methylated CGIs in two samples. When combined with bioinformatic tools for analyzing expression microarray data, DMH can readily be used to classify samples into biologically or clinically relevant groups on

the basis of their methylation profiles (see above references). However, the method does present some limitations for individual investigators if their institutional microarray core facility is not set up to run the arrays. On the technical side, since most genes are single-copy genes, the differences in Cy3 and Cy5 signal intensities for most loci (spots) may not be as strong and readily discernable as those in transcriptional profiling arrays in which fold changes for many probes could be more significant. Furthermore, problems arising from dye bias (Cy5 versus Cy3) would affect DMH more than transcriptional profiling. Furthermore, since the method requires approximately 2–5 µg of high-quality genomic DNA as starting materials, it may put constraints on experiments in which sample availability is an issue (e.g., microdissected samples). Finally, only two samples can be compared at a time. If multiple samples or treatment groups (several experimental groups over multiple time points) were to be compared, high-order bioinformatic analysis tools would be needed to facilitate these comparisons.

2.1.5 Bisulfite sequencing—Bisulfite sequencing is a “gold-standard” method used to determine the methylation status of each cytosine over an amplified region of a given gene, a method now used routinely for studies of CGIs [80]. The underlying principle is based on the ability of sodium bisulfite to deaminate cytosine (C) residues into uracil (U) in genomic DNA, whereas the methylation cytosine residues are resistant to this modification. After PCR amplification, the Us are amplified as thymines (Ts). Cloning and subsequent sequencing of the DNA fragments containing the CGIs then provide information on the methylation status of each C within the island (Figure 2, 11).

The method is used routinely in analyzing the methylation status of any target or candidate DNA sequence containing CpGs. It has the advantage of revealing the methylation status of each CpG dinucleotide within the sequence and also the interrelationship between the methylation status of multiple CpG sites. However, it is more labor intensive than other shotgun approaches, such as methylation-sensitive PCR (see below). Its successful application also depends on whether nested PCR primers could be designed to amplify the fragment of interest. Furthermore, DNA integrity that is less than optimal, as is the case of DNA isolated from microdissected samples or paraffin-embedded tissues, presents significant challenges for this application.

2.1.6 Methylation-specific oligonucleotide (MSO)/methylation target microarrays—As an alternative to bisulfite sequencing, methylation-sensitive oligonucleotide (MOS) array was initially designed to provide a high-throughput method for fine mapping of CpG sites in a known CGI using a hybridization-based microarray protocol [81–83] (Figure 7). Subsequently, the methodology was adapted to interrogate simultaneously the methylation status of multiple CGIs [84]. Test DNA is restricted with *KpnI* and *NdeI*, bisulfite modified and amplified, and labeled with Cy5, resulting in a pool of labeled targets with altered nucleotide sequences due to their differential methylation status. Sets of short oligonucleotides (~21–25-mers), corresponding to the methylated and unmethylated versions of the CGI, are designed to provide coverage of the entire region. These pair of oligonucleotides are synthesized and immobilized in triplicates as probes on glass slides. After hybridization of the targets to the probes, hybridization signals are captured, quantified, and analyzed. The percentage of methylation for each short CG-rich fragment (2–4 CpG sites) is determined by comparison of signal intensities between the paired “methylated” and “unmethylated” oligonucleotide probes.

Gitan et al. were successful in using this method to compare the methylation status of the estrogen receptor α CGI in breast cancer cell lines, normal fibroblasts, breast cancers, and control tissues [81]. Yang et al. used this method to examine the androgen receptor gene promoter in 76 cases of non-Hodgkin lymphoma [83]. Hou et al. also used MSO arrays to study

the methylation status of *p16* in gastric carcinomas [84]. Yu and associates identified a subset of aberrantly methylated genes/ESTs (25/105) in prostate cancer using MSO arrays printed with probes for 105 CGIs [85]. Similarly, MSO arrays were used to examine 156 loci involving 38 genes to identify those with DNA methylation patterns that differentiate between three types of small B-cell lymphomas [86].

To increase the throughput for screening of methylation changes in the promoter of a given gene among multiple samples simultaneously, Zhou and co-workers have described two versions of methylation target arrays (MTAs) [87]. In their investigation of the promoter region of *IGFBP7* gene in ten cases of breast cancer tissues and six cases of responding normal breast tissues, the investigators have fabricated two arrays, a bisulfite PCR product array and a bisulfite genomic DNA array, and compared the sensitivity and reliability of these two arrays to detect methylation changes in this promoter. They found that the bisulfite genomic DNA array is less suitable than the bisulfite PCR product array for analyzing methylation changes using tissue samples. Hence, the latter methods warrant future development for high-throughput analysis of target genes

Undoubtedly, MSO array hybridization is a versatile and high-throughput method for comparing the methylation status of known CGIs among multiple samples. Approximately 2–3 µg of genomic DNA is needed as starting material [85]. However, reproducibility of results could be drastically affected by the performance of the bisulfite treatment, which would impact the quality of the targets. Furthermore, since only a few samples could be processed with a single hybridization, the method is insufficient for effective analyses of hundreds of clinical samples. The efficiency of the various hybridizations may vary significantly. The success of the methodology also depends on the specificity of the methylated/unmethylated oligonucleotides.

2.1.7 Transcriptional profiling of genes reactivated by demethylating agents or inhibitors of histone deacetylation—DNA methylation and histone deacetylation are two major mechanisms resulting in epigenetic silencing of genes through direct and indirect methylation of CGIs. Hence, several investigators have taken the approach of using DNA demethylating agents or inhibitors of histone deacetylation to reactivate epigenetically silenced genes and then using transcription profiling [88] to identify the reactivated genes. This approach allows a pool of candidates to be identified, i.e., those genes that are reactivated after the treatment. *Post hoc* validation methods (see below) are then used to identify among these candidates which promoters play a regulatory role in transcription through cytosine methylation [89-91].

This approach was proven successfully in identifying dysregulated genes caused by aberrant cytosine methylation in colorectal cancer and gastric cancer [89]. A recent study [91] compared the gene sets reactivated by 5-aza-2' deoxycytosine (5-AZA), a DNA demethylating agent, and trichostatin (TSA), an inhibitor of histone deacetylation, in HepG2 cells. They found distinct and common gene sets that are reactivated by the individual agents and their combination, respectively, suggesting that the two epigenetic mechanisms may regulate different sets of genes. In contrast, Li and associates used an integrated “triple” microarray system to elucidate the “epigenetic hierarchies” on gene regulation. Analyses of data with established statistical models demonstrate that genes reactivated by DNA demethylation and histone hyperacetylation are highly correlative [92]. A comparable approach was recently applied to study the effects of maternal care, an early-life experience, in stable epigenetic programming of gene expression in the hippocampus that mediates stress responsivity. Adult offspring exhibiting anxiety behavior due to maternal care were centrally infused with TSA or methionine. Global transcriptional profiling identified over 900 genes that were reactivated by

these two reagents, and these genes are likely as candidates mediating adult anxiety disorder caused by early life experience.

5-AZA is a commonly used DNA demethylating agent [93]; it causes cytosine demethylation by covalently binding to the maintenance DNA methyltransferase (DNMT1) and inhibiting its “proofreading” ability during DNA replication. In this manner, the hemimethylated DNA can escape methylation and thereby become fully unmethylated after one further round of DNA replication. Treatment of cells with 5-AZA has repeatedly been demonstrated to reactivate genes silenced by DNA methylation. However, a few precautionary notes are noteworthy regarding its use in the aforementioned approaches: 1) 5-AZA could cause a low frequency of mutations that affect patterns of gene expression; 2) the agent has a short half-life and must be freshly prepared and added to the cell cultures repeatedly; 3) 5-AZA works best when cells are replicating rapidly; and 4) a period of recovery after treatment is necessary for the new DNA methylation patterns to be established.

Trichostatin A (TSA) belongs to the hydroxamic acid type of histone deacetylase (HDAC) inhibitors with very high potency [94]. TSA is widely used in many experimental settings due to its potency and wide coverage, i.e., it inhibits both class I and II HDAC isoforms. Treatment of cells with TSA has been shown to reactivate about 10% of genes but also to cause silencing of close to an equal number of expressed genes in melanoma cell lines [95]. Most HDAC inhibitors are able to induce expression of genes related to the cell cycle and apoptosis, and therefore genes reactivated by HDAC inhibitors are not necessarily related to epigenetic regulation. Hence, similar to the use of DNA demethylating agents in expression array studies, *post hoc* experiments are needed to verify that the candidates identified are directly regulated by cytosine methylation of their promoter CGIs.

Finally, since the promoters of many genes and ESTs remain unknown, the use of transcriptional profiling to identify DNA methylation-regulated genes still faces significant challenges since confirmation of a regulatory CGI for a given gene is possible only if its promoter region is published or in public databases. The use of genome walking techniques to find the 5' upstream sequence of a gene of interest could be tedious.

2.1.8. Genome-wide methylation array/promoter arrays—The initial CGI library was first produced by Cross and co-workers by isolating CG-rich DNA fragments with methyl-CpG binding protein affinity columns (see above). Subsets of these CG-rich clones (9-k set or 12-k set) were used in the construction of several arrays for the detection of differentially methylated CGIs [70,71]. However, not all the CG-rich clones correspond to CGIs and not all CGIs are in the promoter regions. Promoter arrays have recently become readily available through the University Health Network Microarray Center (UHNMC) in Toronto (www.microarray.ca). More than 20K CG-rich sequence clones can be assessed and used to distinguish promoter methylation status of human and mouse genome. The UHNMC's human CpG island array contains 12,192 spotted clones, while the mouse array contains 7,296 probes. These probe sets cover a large percentage of CGIs found in the human and mouse genomes. Moreover, the UHNMC provides a web interface for users to download the corresponding sequences of the clones. These arrays clearly provide a high-throughput, genome-wide screen of promoters with CGIs undergoing methylation changes due to experimental or pathological conditions. Since these arrays have become commercially available, an increasing number of studies using this approach have been published and yield insightful data on epigenetic gene regulation [75,79,96-98]. A new approach that combines methylation-sensitive enzyme digestion with the comparative genomic hybridization (CGH) technique was developed to screen the entire genome for changes in methylation pattern [99].

The above technology remains most suitable for running by a core facility, where inter-array variations could be minimized and high-throughput could be achieved. The expenditure for setting-up routine hybridization is relatively high since it requires hybridization stations, equipment for quality controls of the targets (e.g., Nanodrop), and a scanner for capturing signals. In addition, because of the large volume of data generated in one experiment, pre- and post-experimental consultations on experimental design and data mining are critical elements for this kind of study. Specific software packages provided by the vendors and others developed in-house are required to generate meaningful results.

Custom designed or small-scale promoter methylation arrays have recently become commercial available that simultaneously profile the methylation status of the gene of interests from one sample (NimbleGen System; Panomics). This is a high-throughput analysis of promoter methylation that costs less than the genome-wide CGI array.

2.1.9 Chromatin immunoprecipitation (ChIP) on DNA microarray (ChIP-chip)—

Chromatin immunoprecipitation (ChIP) was developed to identify and characterize the interactions of specific genomic DNA sequences associated with a target protein such as a transcriptional factor in the context of intact cells [100]. An antibody specific to the target protein is used to immunoprecipitate the protein-DNA complexes. After the crosslink between the two is reversed, the identities of the DNA sequences are then uncovered by amplification and sequencing. Recently, with the advent of bioinformatics on promoter/transcription start site [101] and DNA microarrays, it has become possible to use the DNA fragments isolated from a ChIP assay as targets to probe a microarray in a protocol known as the ChIP-chip [102-106] (Figure 8). The logical approach to applying the ChIP-chip methodology to study the effects of DNA methylation/chromatin structure on gene expression is to use groups of antibodies specific for histone deacetylase/methylases to pull down DNA, followed by DNA array analysis. A human promoter array has recently been developed for ChIP experiments (Affymetrix). It is a single array comprising more than 4.6 million probes titled through more than 25,500 human promoter regions proximal to transcriptional start sites and probes for approximately 59% CpG islands annotated by UCSC in the NCBI human genome assembly (Build 34).

An example of the use of the ChIP-chip approach for methylation studies is the identification of the “methylome,” which refers to the complete set of DNA sequences susceptible to methylation in a cell [107]. Genomic DNA is sonicated into smaller fragments. Anti-5'-methylcytosine antibody is used to enrich methyl cytosine-rich genomic DNA fragments, a procedure known as methylated DNA immunoprecipitation (MeDIP). The MeDIP-enriched fragments are then labeled with Cy3, while the non-MeDIP DNA (or the input DNA) is labeled with Cy5. The two samples are co-hybridized to a sub-megabase resolution tiling (SMRT) array or a CGI array. Relative signal intensity at each locus (spot) represented on the array indicates methylation status. DNA methylation profiles could be obtained simultaneously at both genome-wide and locus-specific levels.

This approach provides an unbiased, whole-genome view of changes in methylation status that are mapped to chromosomal regions. It is one of the most exciting of the recently developed applications of the ChIP-chip technology for DNA methylation studies. However, the limitations in this technique remain the high cost of setting up the array technology platform and the dependence on the efficiency of the antibody to pull down DNA fragments. The number of probes on the CGI arrays will continue to grow, hence increasing the coverage.

2.1.10 Luminometric methylation assay (LUMA) using Pyrosequencing—The underlying principal of LUMA is based on DNA cleavage by a pair of isoschizomer endonucleases followed by bioluminometric polymerase extension to quantify the extent of

restriction cleavage using Pyrosequencing [108]. The method is a modern version of one of the earliest method used to detect methylated cytosine in genomic DNA [109] in which *HpaII* (methylation-sensitive) and *MspI* (methylation-insensitive) were used to cut at the recognition site CCGG. *HpaII* is not able to cut if the internal internal cytosine is methylated (C^m CGG). After the DNA cleavage, the methylation status can be determined by Southern blotting or by PCR analysis [109]. With the advent of Pyrosequencing [110], a new sequencing methodology that dispenses labeled primers, labeled nucleotides, and gel-electrophoresis, the extent of methylation at each CpG sites could be determined accurately and rapidly. The LUMA workflow (Figure 9) started with DNA digestion in parallel reactions using either *MspI* or *HpaII* in combination with *EcoRI*. *MspI* or *HpaII* both leave 5'CG overhangs after DNA cleavage, but in the *HpaII* reaction, C^m CGG sites are protected. As an internal reference, *EcoRI* generates 5'-AATT overhangs. These overhangs are then filled in a polymerase extension assay with stepwise addition of dNTPs (four steps: dATP α S, dGTP+dCTP, dTTP, and dGTP+dCTP). As each dNTP is extended, inorganic pyrophosphate (PPi) is released and converted to ATP by ATP-sulfurylase and adenosine-5-phosphosulfate. This reaction is coupled to the conversion of luciferin to oxyluciferin by luciferase and ATP to generate a proportional amount of visible light, which is quantified by a charge-coupled-device camera in the Pyrosequencer. The signals corresponding to dATP α S and dTTP additions both represent *EcoRI* cleavage and are equal in the two DNA samples (*MspI*+*EcoRI*-treated versus *HpaII* +*EcoRI*-treated). The signal intensities generated by the additions of dCTP and dGTP are added together, and the sum represents an *HpaII* or *MspI* cleavage. The degree of methylation at a CpG site can be derived from the (*HpaII*/*EcoRI*)/(*MspI*/*EcoRI*) ratio.

LUMA can be used to access cytosine methylation level of the whole genome [108]. However, the technology has more commonly applied to studies of targeted CGIs [111]. LUMA has been used to determine changes in methylation status associated with disease development [112-114] or as diagnostic/prognostic markers of cancers [115-118]. It is limited by the length of the sequence read and thereby the number of CpGs that can be analyzed in one sequencing reaction. In addition, the relatively high cost of the pyrosequencing machine has limited the use of the technology without the support of a core facility.

2.1.11 Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS)/high-performance liquid chromatography (HPLC)—Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS) [119,120] is a promising and powerful technique for DNA methylation analysis. Various MS approaches have been developed for the measurement of levels of DNA methylation, including rapid screening of single nucleotide polymorphisms (SNPs), and quantitative allele studies. It has been used to monitor nucleotide digestion and DNA sequencing [121,122]. The commonly used approach involves bisulfite conversion of genomic DNA [80] followed by DNA sequencing, resulting in accurate determination of methylation status in genomic DNA.

As an alternative, Schatz et al. introduced RNA-mediated analysis of methylation status of individual CpGs using MALDI-TOF-MS based on *in vitro* transcription of bisulfite PCR product and base-specific cleavage [123]. After treatment with sodium bisulfite, DNA samples are subjected to PCR in which the amplicons are tagged with T7 RNA polymerase promoter as well as with a complementary sequence with low guanosine content. After T7-mediated transcription, these guanosine residues in the newly transcribed RNA are subjected to cleavage by RNase T1 and the methylation fingerprint (RNA fragmentation pattern) is visualized by MALDI-MS (Figure 10). This approach was applied to the analysis of artificially methylated and unmethylated DNA, mixtures thereof, and colon DNA samples.

Recently, Tos et al. combined the bisulfite conversion genomic DNA method with the GOODs [124,125] assay for accurate quantification of methylation status of CpG dinucleotides using MALDI-MS. They analyzed several paraffin-embedded tissue biopsies and found the MS results comparable to those of chip hybridization. Furthermore, Ehrich et al. applied base-specific cleavage of DNA following MALDI-TOF-MS analysis to measure the degree of methylation in normal and neoplastic lung cancer tissue samples, allowing accurate classification of samples according to their histopathology [126]. This method is applicable to methylation studies of target genes.

Another related technology for global profiling involves the use of high-performance liquid chromatograph (HPLC). It is the routine and most widely used separation technique for the measurement of 5-methylcytosine and/or DNA methylation [127-129] involving the digestion of DNA to nucleotides, nucleosides, or bases, followed by high-resolution separation and quantification with UV detection. HPLC has been implicated in the study of p16 methylation status of gastric dysplasia/carcinomas [130,131] and a study of tamoxifen-induced hepatocarcinogenesis in rat [132]. Recently, Song et al. reported a novel method utilizing liquid chromatography-electrospray ionization tandem mass spectrometry to measure 5-methyl-2'-deoxycytidine levels [133]. Tissue-specific differences in DNA methylation in various mammals has been reported by using HPLC analysis [134]. Several other chromatographic and electrophoretic techniques, including gas chromatography, thin layer chromatography, and capillary electrophoresis, have also been used to determine DNA methylation [135-137].

Both MALDI-TOF-MS and HPLC are tools suitable for high-throughput, multi-channel detection with the benefits of high speed, accuracy, and automation. They either can measure the content of methyl-cytosine in the whole genome or can detect methylation patterns of specific target genes. Relatively high cost or the need for standardization of the protocols would be the major pitfalls of applying these techniques to methylation studies.

2.2 Characterization of a target or candidate sequence and validation of its involvement in gene regulation

The various global discovery platforms mentioned above would undoubtedly identify many novel genes that are epigenetically regulated and that might be important in the development of diseases. However, the major challenge remains in designing an efficient strategy for *post hoc* validation of the involvement of the CGI in gene regulation. It has to be reemphasized that multiple validation approaches need to be conducted before one can lay claim to its role in gene regulation. These include, but are not limited to, database searches and *in silico* analyses, studies of methylation status using bisulfite sequencing (2.1.5), combined bisulfite restriction analysis (COBRA), methylation-sensitive PCR (MSPCR), methylation-sensitive single strand conformational polymorphism (MS-SSCP), MethyLight analysis, and gene expression studies (Figures 2 and 11). Some of these techniques are summarized briefly below, and most of them can be used both for target gene studies and for studies of novel candidates identified by discovery platforms.

2.2.1 In silico database analysis—The first step in *post hoc* validation starts at *in silico* analysis of the candidate genes of interest. Detailed searches of public databases such as the National Centre for Biotechnology Information (NCBI) and the University of California Santa Cruz (UCSC) Genome Bioinformatics group are the logical first steps. If the CGI is located in the promoter or 5' region of a gene, it usually has a higher chance of being involved in transcriptional regulation. After promoter/5' sequences of a candidate gene are identified, the next step is to locate the CGI and discern its location relative to the transcriptional start site and to the 5' exons and introns. Information generated from these analyses provides clues as to 1) whether a CGI exists in the sequence and 2) if it has a high probability of playing a

regulatory role in gene expression. These *in silico* analyses could be laborious, but several open-source tools are now available through the Internet: MetPrimer (<http://www.urogene.org/methprimer/index.html>), Promoter Inspector (<http://www.genomatix.de>), BIMAS (<http://thr.cit.nih.gov/molbio/index.shtml>), and Database of Transcription Start Sites (<http://dbtss.hgc.jp>).

2.2.2 Validation assays of gene of interest—Once a “potential” CGI has been identified in the 5′ region of a gene, physical and functional characterization of the CGI is in order. The first criterion of a regulatory CGI is that an inverse relationship can be established between its degree of methylation and the expression of its cognate gene. Determination of the methylation status of the CGI and its adjacent sequences could be achieved by one or more of the following methods: MSPCR, COBRA, MSSSCP, MethyLight, or bisulfite DNA sequencing. For cell-culture experiments, DNA could be extracted from cultures under different conditions or exposed to DNA-demethylating agents or inhibitors of histone deacetylase. By extracting both DNA for methylation studies and RNA for transcript quantification, one can determine if an inverse relationship exists between the extent of methylation in the CGI and gene transcription [20]. Bisulfite genomic sequencing [80] is the preferred method (see section 2.1.5) of initial characterization because it can reveal the methylation status of individual CpG in the CGI. Detailed mapping of a CGI could uncover important “methylation hotspots” that could be used to design methylation oligonucleotides for use in gene silencing experiments by targeting these hotspot sequences [138].

After initial characterization of the putative CGI, other methods are available for more-rapid or higher-throughput analyses. MS-PCR [139,140] can be used to assess the methylation status of a cluster of CpGs within a CGI by conducting PCRs with primers designed for the methylated or unmethylated version of the sequence of interest. MS-PCR has the advantage of requiring small amounts of DNA, with a sensitivity of detecting 0.1% methylated alleles in a given sample. It can be applied to DNA extracted from paraffin-embedded samples which most of the DNA is fragmented.

One of the limitations of MS-PCR is that it is not highly quantitative. Hence, other more quantitative methods have recently been developed. COBRA is similar to MS-PCR but provides more quantitative information on the degree of methylation of the targeted sequence. Restriction-enzyme digestion is used to reveal methylation-dependent sequence differences in PCR products of sodium bisulfite-treated DNA as described previously [141]. The method has the advantage of accessing methylation levels of a target sequence in a small amount of DNA samples and provides linear quantification across a wide spectrum of DNA methylation levels. The technique can be reliably applied to DNA obtained from microdissected paraffin-embedded tissue samples. COBRA thus combines the powerful features of ease of use, quantitative accuracy, and compatibility with archival samples. MS-SSCP also provides a fairly quantitative method to access the methylated and unmethylated allele populations. It uses high-resolution gel electrophoresis to generate a specific methylation pattern for determining the percentages of methylation in a targeted sequence. In SSCP, methylated alleles, due to the difference in conformation from the unmethylated alleles, can be separated by MDE-gel electrophoresis and visualized by SYBR-Gold staining or radiolabeling [142,143]. A high-throughput technology, known as MethyLight, has recently been developed for cytosine analysis. It utilizes fluorescence-based real-time PCR Taq-Man technology [144,145] and thus is highly quantitative. MethyLight assays have good precision and linearity. It can be used effectively in a high-throughput manner for analysis of DNA methylation of small amounts of DNA.

After the methylation status of the CGI has been established by the aforementioned methods, it is imperative to determine the expression of its cognate gene. Modern laboratory procedures

such as quantitative reverse transcriptase-PCR or real-time PCR are most appropriate for such studies because they provide quantitative data and are adaptable to high-throughput application. If levels of transcripts can not be accessed directly (e.g., if only archival paraffin-embedded samples are available), then immunostaining of sections [138] or *in-situ* hybridization could be attempted to ascertain gene expression. In addition, cell cultures treated with DNA demethylating agents and inhibitors of histone deacetylase can further elucidate the regulatory role of the CGI. More recently, we have used a class of methylation oligonucleotides to induce sequence-specific methylation *in cellulo* that subsequently leads to gene silencing. In this manner, we have provided additional evidence that DNA methylation is involved in the regulation of estrogen receptor β [138].

3. Discussion

In this section, we would like to point out some important issues that should be noted when methylation techniques are employed in epigenetic studies. The above validation assays could be used for whole-tissue analyses [20]. However, a tissue contains multiple cell types and any change in the methylation status and gene expression of that CGI might occur in only one cell type. In this scenario, it would be beneficial to have laser-capture microscope-assisted microdissection to enrich a specific cell population [138]. Alternatively, enzymatic digestion of the tissue followed by isolation of the specific cell population would also enhance the experimental outcome.

Methods including BSPCR, MSPCR, MSSSCP, COBRA, or MethyLight depend on whether nested PCR primers could be designed to amplify the fragment of interest. Furthermore, if the integrity of the DNA is less than optimal, as is in the case of DNA isolated from microdissected samples or paraffin-embedded tissues, the amplified fragment would be limited in size. Also, efficiency of bisulfite modification on DNA samples would be another major pitfall affecting the sensitivity and accuracy of the PCR results.

On the other hand, if 5AZA and/or TSA treatments are performed to determine the relationship between the transcription level and methylation status of target genes, dosage of the modifiers and incubation period for the treatment are critical. Optimization of the protocol is necessary. Additional studies such as site-directed mutagenesis and luciferase reporter assay *in vitro* and *in vivo* can be involved to establish the gene regulatory function of the CGIs. Briefly, deletion constructs at specific CG sites are generated by using a commercial site-directed mutagenesis kit according to the manufacturer's recommendations. Modified promoters are resubcloned into reporter vector followed by a promoter reporter assay. ChIP or transcription-factors binding assay may also be included to determine which regulatory elements on promoters are associated with the methylation changes.

In summary, multiple methods and technologies are available to determine changes in global and regional methylation of cytosine in genomic DNA; each has advantages, disadvantages, and areas of applicability. Because of variations in sample size, the nature of the samples, the number of samples in the studies, the experience of the investigators, and the resources of the laboratory or the institution, there is no "gold" standard or "standard operation procedure" for conducting a DNA methylation analysis. This review simply attempts to provide an overview of the currently available techniques and to discuss some of the advantages and limitations of each technology. With the rapid growth in interest in understanding the epigenetic regulation of disease development, a variety of new and improved methodologies are certain to emerge in the coming years. These technologies will undoubtedly change the landscape of epigenetic studies and provide valuable new insights into areas such as the developmental basis of disease and reproductive toxicology.

Acknowledgments

Grant support: NIH grants ES12281 (S-M Ho) and ES13071 (S-M Ho) and the Department of Defense awards DAMD W81XWH-06-1-0373 (W-Y Tang).

We thank Dawn Ho for technical contribution to the generation of the artwork. We also thank Suresh Babu, Yuet-kin Leung, and Neville NC Tam for discussions in preparation of the manuscript.

4. References

1. Liu Y, Freedman BI. Genetics of progressive renal failure in diabetic kidney disease. *Kidney Int Suppl* 2005;S94–S97. [PubMed: 16336585]
2. Kroll TG. Molecular events in follicular thyroid tumors. *Cancer Treat Res* 2004;122:85–105. [PubMed: 16209039]
3. Moore MA. Converging pathways in leukemogenesis and stem cell self-renewal. *Exp Hematol* 2005;33:719–37. [PubMed: 15963848]
4. Scher HI, Sawyers CL. Biology of progressive, castration-resistant prostate cancer: directed therapies targeting the androgen-receptor signaling axis. *J Clin Oncol* 2005;23:8253–61. [PubMed: 16278481]
5. Tusie Luna MT. Genes and type 2 diabetes mellitus. *Arch Med Res* 2005;36:210–22. [PubMed: 15925011]
6. Soussi T, Ishioka C, Claustres M, Beroud C. Locus-specific mutation databases: pitfalls and good practice based on the p53 experience. *Nat Rev Cancer* 2006;6:83–90. [PubMed: 16397528]
7. Garg V. Insights into the genetic basis of congenital heart disease. *Cell Mol Life Sci* 2006;63:1141–8. [PubMed: 16568242]
8. Verschure PJ, Visser AE, Rots MG. Step out of the groove: epigenetic gene control systems and engineered transcription factors. *Adv Genet* 2006;56:163–204. [PubMed: 16735158]
9. Jiang YH, Bressler J, Beaudet AL. Epigenetics and human disease. *Annu Rev Genomics Hum Genet* 2004;5:479–510. [PubMed: 15485357]
10. Rodenhiser D, Mann M. Epigenetics and human disease: translating basic biology into clinical applications. *CMAJ* 2006;174:341–8. [PubMed: 16446478]
11. Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet* 2003;33(Suppl):245–54. [PubMed: 12610534]
12. Morris KV. siRNA-mediated transcriptional gene silencing: the potential mechanism and a possible role in the histone code. *Cell Mol Life Sci* 2005;62:3057–66. [PubMed: 16314933]
13. Cheung P, Lau P. Epigenetic regulation by histone methylation and histone variants. *Mol Endocrinol* 2005;19:563–73. [PubMed: 15677708]
14. Esteller M. Aberrant DNA methylation as a cancer-inducing mechanism. *Annu Rev Pharmacol Toxicol* 2005;45:629–56. [PubMed: 15822191]
15. Gluckman PD, Hanson MA. The developmental origins of the metabolic syndrome. *Trends Endocrinol Metab* 2004;15:183–7. [PubMed: 15109618]
16. Dolinoy DC, Weidman JR, Jirtle RL. Epigenetic gene regulation: Linking early developmental environment to adult disease. *Reprod Toxicol*. 2006
17. Santos F, Dean W. Epigenetic reprogramming during early development in mammals. *Reproduction* 2004;127:643–51. [PubMed: 15175501]
18. Weidman JR, Maloney KA, Jirtle RL. Comparative phylogenetic analysis reveals multiple non-imprinted isoforms of opossum Dlk1. *Mamm Genome* 2006;17:157–67. [PubMed: 16465595]
19. Anway MD, Cupp AS, Uzumcu M, Skinner MK. Epigenetic transgenerational actions of endocrine disruptors and male fertility. *Science* 2005;308:1466–9. [PubMed: 15933200]
20. Ho SM, Tang WY, Belmonte dF, Prins GS. Developmental exposure to estradiol and bisphenol A increases susceptibility to prostate carcinogenesis and epigenetically regulates phosphodiesterase type 4 variant 4. *Cancer Res* 2006;66:5624–32. [PubMed: 16740699]
21. Dolinoy DC, Weidman JR, Waterland RA, Jirtle RL. Maternal genistein alters coat color and protects Avy mouse offspring from obesity by modifying the fetal epigenome. *Environ Health Perspect* 2006;114:567–72. [PubMed: 16581547]

22. Fernandez-Twinn DS, Ozanne SE. Mechanisms by which poor early growth programs type-2 diabetes, obesity and the metabolic syndrome. *Physiol Behav* 2006;88:234–43. [PubMed: 16782139]
23. Barker DJ, Osmond C, Simmonds SJ, Wield GA. The relation of small head circumference and thinness at birth to death from cardiovascular disease in adult life. *BMJ* 1993;306:422–6. [PubMed: 8461722]
24. Ravelli AC, van der Meulen JH, Michels RP, et al. Glucose tolerance in adults after prenatal exposure to famine. *Lancet* 1998;351:173–7. [PubMed: 9449872]
25. Dennison EM, Arden NK, Keen RW, et al. Birthweight, vitamin D receptor genotype and the programming of osteoporosis. *Paediatr Perinat Epidemiol* 2001;15:211–9. [PubMed: 11489147]
26. Thompson C, Syddall H, Rodin I, et al. Birth weight and the risk of depressive disorder in late life. *Br J Psychiatry* 2001;179:450–5. [PubMed: 11689404]
27. Veurink M, Koster M, Berg LT. The history of DES, lessons to be learned. *Pharm World Sci* 2005;27:139–43. [PubMed: 16096877]
28. Demerath EW, Cameron N, Gillman MW, et al. Telomeres and telomerase in the fetal origins of cardiovascular disease: a review. *Hum Biol* 2004;76:127–46. [PubMed: 15222684]
29. Lee HK, Park KS, Cho YM, et al. Mitochondria-based model for fetal origin of adult disease and insulin resistance. *Ann N Y Acad Sci* 2005;1042:1–18. [PubMed: 15965040]
30. Rakyan VK, Blewitt ME, Druker R, et al. Metastable epialleles in mammals. *Trends Genet* 2002;18:348–51. [PubMed: 12127774]
31. Tollefsbol TO. Methods of epigenetic analysis. *Methods Mol Biol* 2004;287:1–8. [PubMed: 15273399]
32. Chiang PK, Gordon RK, Tal J, et al. S-Adenosylmethionine and methylation. *FASEB J* 1996;10:471–80. [PubMed: 8647346]
33. Singal R, Ginder GD. DNA methylation. *Blood* 1999;93:4059–70. [PubMed: 10361102]
34. Murphy SK, Jirtle RL. Imprinted genes as potential genetic and epigenetic toxicologic targets. *Environ Health Perspect* 2000;108(Suppl 1):5–11. [PubMed: 10698719]
35. Costello JF, Plass C. Methylation matters. *J Med Genet* 2001;38:285–303. [PubMed: 11333864]
36. Antequera F, Bird A. Number of CpG islands and genes in human and mouse. *Proc Natl Acad Sci U S A* 1993;90:11995–9. [PubMed: 7505451]
37. Lyko F, Brown R. DNA methyltransferase inhibitors and the development of epigenetic cancer therapies. *J Natl Cancer Inst* 2005;97:1498–506. [PubMed: 16234563]
38. Bird A, Tate P, Nan X, et al. Studies of DNA methylation in animals. *J Cell Sci Suppl* 1995;19:37–9. [PubMed: 8655645]
39. Monk M, Boubelik M, Lehnert S. Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. *Development* 1987;99:371–82. [PubMed: 3653008]
40. Kafri T, Ariel M, Brandeis M, et al. Developmental pattern of gene-specific DNA methylation in the mouse embryo and germ line. *Genes Dev* 1992;6:705–14. [PubMed: 1577268]
41. Issa JP. CpG-island methylation in aging and cancer. *Curr Top Microbiol Immunol* 2000;249:101–18. [PubMed: 10802941]
42. Waterland RA, Jirtle RL. Transposable elements: targets for early nutritional effects on epigenetic gene regulation. *Mol Cell Biol* 2003;23:5293–300. [PubMed: 12861015]
43. Niculescu MD, Craciunescu CN, Zeisel SH. Gene expression profiling of choline-deprived neural precursor cells isolated from mouse brain. *Brain Res Mol Brain Res* 2005;134:309–22. [PubMed: 15836926]
44. Li S, Washburn KA, Moore R, et al. Developmental exposure to diethylstilbestrol elicits demethylation of estrogen-responsive lactoferrin gene in mouse uterus. *Cancer Res* 1997;57:4356–9. [PubMed: 9331098]
45. Newbold RR, Padilla-Banks E, Jefferson WN. Adverse effects of the model environmental estrogen diethylstilbestrol are transmitted to subsequent generations. *Endocrinology* 2006;147:S11–S17. [PubMed: 16690809]

46. Li S, Hansman R, Newbold R, et al. Neonatal diethylstilbestrol exposure induces persistent elevation of c-fos expression and hypomethylation in its exon-4 in mouse uterus. *Mol Carcinog* 2003;38:78–84. [PubMed: 14502647]
47. Anway MD, Leathers C, Skinner MK. Endocrine disruptor vinclozolin induced epigenetic transgenerational adult onset disease. *Endocrinology*. 2006Epub ahead of print
48. Huang TH, Laux DE, Hamlin BC, et al. Identification of DNA methylation markers for human breast carcinomas using the methylation-sensitive restriction fingerprinting technique. *Cancer Res* 1997;57:1030–4. [PubMed: 9067264]
49. Wu M, Ho SM. PMP24, a gene identified by MSRF, undergoes DNA hypermethylation-associated gene silencing during cancer progression in an LNCaP model. *Oncogene* 2004;23:250–9. [PubMed: 14712230]
50. Lo KW, Tsang YS, Kwong J, et al. Promoter hypermethylation of the EDNRB gene in nasopharyngeal carcinoma. *Int J Cancer* 2002;98:651–5. [PubMed: 11920632]
51. Lv Z, Zhang M, Bi J, et al. Promoter hypermethylation of a novel gene, ZHX2, in hepatocellular carcinoma. *Am J Clin Pathol* 2006;125:740–6. [PubMed: 16707376]
52. Rodic N, Oka M, Hamazaki T, et al. DNA methylation is required for silencing of ant4, an adenine nucleotide translocase selectively expressed in mouse embryonic stem cells and germ cells. *Stem Cells* 2005;23:1314–23. [PubMed: 16051982]
53. Hatada I, Hayashizaki Y, Hirotune S, et al. A genomic scanning method for higher organisms using restriction sites as landmarks. *Proc Natl Acad Sci U S A* 1991;88:9523–7. [PubMed: 1946366]
54. Akama TO, Okazaki Y, Ito M, et al. Restriction landmark genomic scanning (RLGS-M)-based genome-wide scanning of mouse liver tumors for alterations in DNA methylation status. *Cancer Res* 1997;57:3294–9. [PubMed: 9242463]
55. Yoshikawa H, de la MS, Nagai H, et al. Chromosomal assignment of human genomic NotI restriction fragments in a two-dimensional electrophoresis profile. *Genomics* 1996;31:28–35. [PubMed: 8808276]
56. Costello JF, Plass C, Cavenee WK. Aberrant methylation of genes in low-grade astrocytomas. *Brain Tumor Pathol* 2000;17:49–56. [PubMed: 11210171]
57. Matsuyama T, Kimura MT, Koike K, et al. Global methylation screening in the Arabidopsis thaliana and Mus musculus genome: applications of virtual image restriction landmark genomic scanning (Vi-RLGS). *Nucleic Acids Res* 2003;31:4490–6. [PubMed: 12888509]
58. Rouillard JM, Erson AE, Kuick R, et al. Virtual genome scan: a tool for restriction landmark-based scanning of the human genome. *Genome Res* 2001;11:1453–9. [PubMed: 11483587]
59. Asaga S, Ueda M, Jinno H, et al. Identification of a new breast cancer-related gene by restriction landmark genomic scanning. *Anticancer Res* 2006;26:35–42. [PubMed: 16475676]
60. Wu R, Lin L, Beer DG, et al. Amplification and overexpression of the L-MYC protooncogene in ovarian carcinomas. *Am J Pathol* 2003;162:1603–10. [PubMed: 12707044]
61. Motiwala T, Ghoshal K, Das A, et al. Suppression of the protein tyrosine phosphatase receptor type O gene (PTPRO) by methylation in hepatocellular carcinomas. *Oncogene* 2003;22:6319–31. [PubMed: 14508512]
62. Sato K, Fukata H, Kogo Y, et al. Neonatal exposure to diethylstilbestrol alters the expression of DNA methyltransferases and methylation of genomic DNA in the epididymis of mice. *Endocr J* 2006;53:331–7. [PubMed: 16714842]
63. Lisitsyn N, Lisitsyn N, Wigler M. Cloning the differences between two complex genomes. *Science* 1993;259:946–51. [PubMed: 8438152]
64. Schutte M, da Costa LT, Moskaluk CA, et al. Isolation of YAC insert sequences by representational difference analysis. *Nucleic Acids Res* 1995;23:4127–33. [PubMed: 7479076]
65. Toyota M, Ho C, Ahuja N, et al. Identification of differentially methylated sequences in colorectal cancer by methylated CpG island amplification. *Cancer Res* 1999;59:2307–12. [PubMed: 10344734]
66. Koinuma K, Kaneda R, Toyota M, et al. Screening for genomic fragments that are methylated specifically in colorectal carcinoma with a methylated MLH1 promoter. *Carcinogenesis* 2005;26:2078–85. [PubMed: 16033773]
67. Yamada Y, Toyota M, Hirokawa Y, et al. Identification of differentially methylated CpG islands in prostate cancer. *Int J Cancer* 2004;112:840–5. [PubMed: 15386391]

68. Ueki T, Toyota M, Skinner H, et al. Identification and characterization of differentially methylated CpG islands in pancreatic carcinoma. *Cancer Res* 2001;61:8540–6. [PubMed: 11731440]
69. Asada K, Asada R, Yoshiji H, et al. DNA cytosine methylation profile in various cancer-related genes is altered in cultured rat hepatocyte cell lines as compared with primary hepatocytes. *Oncol Rep* 2006;15:1241–8. [PubMed: 16596193]
70. Cross SH, Charlton JA, Nan X, Bird AP. Purification of CpG islands using a methylated DNA binding column. *Nat Genet* 1994;6:236–44. [PubMed: 8012384]
71. Huang TH, Perry MR, Laux DE. Methylation profiling of CpG islands in human breast cancer cells. *Hum Mol Genet* 1999;8:459–70. [PubMed: 9949205]
72. Yan PS, Perry MR, Laux DE, et al. CpG island arrays: an application toward deciphering epigenetic signatures of breast cancer. *Clin Cancer Res* 2000;6:1432–8. [PubMed: 10778974]
73. Ahluwalia A, Yan P, Hurteau JA, et al. DNA methylation and ovarian cancer. I. Analysis of CpG island hypermethylation in human ovarian cancer using differential methylation hybridization. *Gynecol Oncol* 2001;82:261–8. [PubMed: 11531277]
74. Yan PS, Shi H, Rahmatpanah F, et al. Differential distribution of DNA methylation within the RASSF1A CpG island in breast cancer. *Cancer Res* 2003;63:6178–86. [PubMed: 14559801]
75. van Doorn R, Zoutman WH, Dijkman R, et al. Epigenetic profiling of cutaneous T-cell lymphoma: promoter hypermethylation of multiple tumor suppressor genes including BCL7a, PTPRG, and p73. *J Clin Oncol* 2005;23:3886–96. [PubMed: 15897551]
76. Day JK, Bauer AM, DesBordes C, et al. Genistein alters methylation patterns in mice. *J Nutr* 2002;132:2419S–23S. [PubMed: 12163704]
77. Yan PS, Wei SH, Huang TH. Differential methylation hybridization using CpG island arrays. *Methods Mol Biol* 2002;200:87–100. [PubMed: 11951657]
78. Craig JM, Kraus J, Cremer T. Removal of repetitive sequences from FISH probes using PCR-assisted affinity chromatography. *Hum Genet* 1997;100:472–6. [PubMed: 9272175]
79. Heisler LE, Torti D, Boutros PC, et al. CpG Island microarray probe sequences derived from a physical library are representative of CpG Islands annotated on the human genome. *Nucleic Acids Res* 2005;33:2952–61. [PubMed: 15911630]
80. Frommer M, McDonald LE, Millar DS, et al. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc Natl Acad Sci U S A* 1992;89:1827–31. [PubMed: 1542678]
81. Gitan RS, Shi H, Chen CM, et al. Methylation-specific oligonucleotide microarray: a new potential for high-throughput methylation analysis. *Genome Res* 2002;12:158–64. [PubMed: 11779841]
82. Shi H, Wei SH, Leu YW, et al. Triple analysis of the cancer epigenome: an integrated microarray system for assessing gene expression, DNA methylation, and histone acetylation. *Cancer Res* 2003;63:2164–71. [PubMed: 12727835]
83. Yang H, Chen CM, Yan P, et al. The androgen receptor gene is preferentially hypermethylated in follicular non-Hodgkin's lymphomas. *Clin Cancer Res* 2003;9:4034–42. [PubMed: 14519624]
84. Hou P, Shen JY, Ji MJ, et al. Microarray-based method for detecting methylation changes of p16 (Ink4a) gene 5'-CpG islands in gastric carcinomas. *World J Gastroenterol* 2004;10:3553–8. [PubMed: 15534905]
85. Yu YP, Paranjpe S, Nelson J, et al. High throughput screening of methylation status of genes in prostate cancer using an oligonucleotide methylation array. *Carcinogenesis* 2005;26:471–9. [PubMed: 15485992]
86. Guo J, Burger M, Nimrich I, et al. Differential DNA methylation of gene promoters in small B-cell lymphomas. *Am J Clin Pathol* 2005;124:430–9. [PubMed: 16191512]
87. Zhou D, Qiao W, Yang L, Lu Z. Bisulfite-modified target DNA array for aberrant methylation analysis. *Anal Biochem* 2006;351:26–35. [PubMed: 16516134]
88. Kramer JA, Pettit SD, Amin RP, et al. Overview on the application of transcription profiling using selected nephrotoxics for toxicology assessment. *Environ Health Perspect* 2004;112:460–4. [PubMed: 15033596]
89. Suzuki H, Gabrielson E, Chen W, et al. A genomic screen for genes upregulated by demethylation and histone deacetylase inhibition in human colorectal cancer. *Nat Genet* 2002;31:141–9. [PubMed: 11992124]

90. Chiba T, Yokosuka O, Arai M, et al. Identification of genes up-regulated by histone deacetylase inhibition with cDNA microarray and exploration of epigenetic alterations on hepatoma cells. *J Hepatol* 2004;41:436–45. [PubMed: 15336447]
91. Dannenberg LO, Edenberg HJ. Epigenetics of gene expression in human hepatoma cells: Expression profiling the response to inhibition of DNA methylation and histone deacetylation. *BMC Genomics* 2006;7:181. [PubMed: 16854234]
92. Li L, Shi H, Yiannoutsos C, et al. Epigenetic hypothesis tests for methylation and acetylation in a triple microarray system. *J Comput Biol* 2005;12:370–90. [PubMed: 15857248]
93. Holliday R, Ho T. DNA methylation and epigenetic inheritance. *Methods* 2002;27:179–83. [PubMed: 12095278]
94. Boyle GM, Martyn AC, Parsons PG. Histone deacetylase inhibitors and malignant melanoma. *Pigment Cell Res* 2005;18:160–6. [PubMed: 15892712]
95. Glaser KB, Staver MJ, Waring JF, et al. Gene expression profiling of multiple histone deacetylase (HDAC) inhibitors: defining a common gene set produced by HDAC inhibition in T24 and MDA carcinoma cell lines. *Mol Cancer Ther* 2003;2:151–63. [PubMed: 12589032]
96. Weber M, Davies JJ, Wittig D, et al. Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nat Genet* 2005;37:853–62. [PubMed: 16007088]
97. Gebhard C, Schwarzfischer L, Pham TH, et al. Genome-wide profiling of CpG methylation identifies novel targets of aberrant hypermethylation in myeloid leukemia. *Cancer Res* 2006;66:6118–28. [PubMed: 16778185]
98. Shi H, Guo J, Duff DJ, et al. Discovery of novel epigenetic markers in non-Hodgkin's lymphoma. *Carcinogenesis*. 2006
99. Wojdacz TK, Hansen LL. Techniques used in studies of age-related DNA methylation changes. *Ann N Y Acad Sci* 2006;1067:479–87. [PubMed: 16804030]
100. Kuras L. Characterization of protein-DNA association in vivo by chromatin immunoprecipitation. *Methods Mol Biol* 2004;284:147–62. [PubMed: 15173614]
101. Davuluri RV, Grosse I, Zhang MQ. Computational identification of promoters and first exons in the human genome. *Nat Genet* 2001;29:412–7. [PubMed: 11726928]
102. Weinmann AS, Yan PS, Oberley MJ, et al. Isolating human transcription factor targets by coupling chromatin immunoprecipitation and CpG island microarray analysis. *Genes Dev* 2002;16:235–44. [PubMed: 11799066]
103. Ren B, Robert F, Wyrick JJ, et al. Genome-wide location and function of DNA binding proteins. *Science* 2000;290:2306–9. [PubMed: 11125145]
104. Ren B, Cam H, Takahashi Y, et al. E2F integrates cell cycle progression with DNA repair, replication, and G(2)/M checkpoints. *Genes Dev* 2002;16:245–56. [PubMed: 11799067]
105. Mao DY, Watson JD, Yan PS, et al. Analysis of Myc bound loci identified by CpG island arrays shows that Max is essential for Myc-dependent repression. *Curr Biol* 2003;13:882–6. [PubMed: 12747840]
106. Oberley MJ, Tsao J, Yau P, Farnham PJ. High-throughput screening of chromatin immunoprecipitates using CpG-island microarrays. *Methods Enzymol* 2004;376:315–3. [PubMed: 14975315]
107. Wilson IM, Davies JJ, Weber M, et al. Epigenomics: mapping the methylome. *Cell Cycle* 2006;5:155–8. [PubMed: 16397413]
108. Karimi M, Johansson S, Stach D, et al. LUMA (LUMinometric Methylation Assay)-A high throughput method to the analysis of genomic DNA methylation. *Exp Cell Res* 2006;312:1989–95. [PubMed: 16624287]
109. Cedar H, Solage A, Glaser G, Razin A. Direct detection of methylated cytosine in DNA by use of the restriction enzyme MspI. *Nucleic Acids Res* 1979;6:2125–32. [PubMed: 223125]
110. Ronaghi M. Pyrosequencing sheds light on DNA sequencing. *Genome Res* 2001;11:3–11. [PubMed: 11156611]
111. Tost J, El Abdalaoui H, Gut IG. Serial pyrosequencing for quantitative DNA methylation analysis. *Biotechniques* 2006;40:721–2. [PubMed: 16774114]724, 726

112. Shiao YH, Crawford EB, Anderson LM, et al. Allele-specific germ cell epimutation in the spacer promoter of the 45S ribosomal RNA gene after Cr(III) exposure. *Toxicol Appl Pharmacol* 2005;205:290–6. [PubMed: 15922014]
113. Mill J, Dempster E, Caspi A, et al. Evidence for monozygotic twin (MZ) discordance in methylation level at two CpG sites in the promoter region of the catechol-O-methyltransferase (COMT) gene. *Am J Med Genet B Neuropsychiatr Genet* 2006;141:421–5. [PubMed: 16583437]
114. White HE, Durston VJ, Harvey JF, Cross NC. Quantitative analysis of SNRPN (correction of SRNPN) gene methylation by pyrosequencing as a diagnostic test for Prader-Willi syndrome and Angelman syndrome. *Clin Chem* 2006;52:1005–13. [PubMed: 16574761]
115. Shaw RJ, Liloglou T, Rogers SN, et al. Promoter methylation of P16, RARbeta, E-cadherin, cyclin A1 and cytoglobin in oral cancer: quantitative evaluation using pyrosequencing. *Br J Cancer* 2006;94:561–8. [PubMed: 16449996]
116. Xinarianos G, McDonald FE, Risk JM, et al. Frequent genetic and epigenetic abnormalities contribute to the deregulation of cytoglobin in non-small cell lung cancer. *Hum Mol Genet* 2006;15:2038–44. [PubMed: 16698880]
117. Kang S, Kim J, Kim HB, et al. Methylation of p16INK4a is a non-rare event in cervical intraepithelial neoplasia. *Diagn Mol Pathol* 2006;15:74–82. [PubMed: 16778587]
118. Mirmohammadsadeh A, Marini A, Nambiar S, et al. Epigenetic silencing of the PTEN gene in melanoma. *Cancer Res* 2006;66:6546–52. [PubMed: 16818626]
119. Smith LM. The future of DNA sequencing. *Science* 1993;262:530–2. [PubMed: 8211178]
120. Hillenkamp F, Karas M, Beavis RC, Chait BT. Matrix-assisted laser desorption/ionization mass spectrometry of biopolymers. *Anal Chem* 1991;63:1193A–203A. [PubMed: 1897719]
121. Puapaboon U, Jai-Nhuknan J, Cowan JA. Characterization of a multi-functional metal-mediated nuclease by MALDI-TOF mass spectrometry. *Nucleic Acids Res* 2001;29:3652–6. [PubMed: 11522836]
122. Fu DJ, Tang K, Braun A, et al. Sequencing exons 5 to 8 of the p53 gene by MALDI-TOF mass spectrometry. *Nat Biotechnol* 1998;16:381–4. [PubMed: 9555731]
123. Schatz P, Dietrich D, Schuster M. Rapid analysis of CpG methylation patterns using RNase T1 cleavage and MALDI-TOF. *Nucleic Acids Res* 2004;32:e167. [PubMed: 15576674]
124. Tost J, Schatz P, Schuster M, et al. Analysis and accurate quantification of CpG methylation by MALDI mass spectrometry. *Nucleic Acids Res* 2003;31:e50. [PubMed: 12711695]
125. Sauer S, Lechner D, Berlin K, et al. A novel procedure for efficient genotyping of single nucleotide polymorphisms. *Nucleic Acids Res* 2000;28:E13. [PubMed: 10666474]
126. Ehrich M, Nelson MR, Stanssens P, et al. Quantitative high-throughput analysis of DNA methylation patterns by base-specific cleavage and mass spectrometry. *Proc Natl Acad Sci U S A* 2005;102:15785–90. [PubMed: 16243968]
127. Fraga MF, Esteller M. DNA methylation: a profile of methods and applications. *Biotechniques* 2002;33:632, 634, 636–2, 634–649. [PubMed: 12238773]
128. Ramsahoye BH. Measurement of genome wide DNA methylation by reversed-phase high-performance liquid chromatography. *Methods* 2002;27:156–61. [PubMed: 12095275]
129. Chakrabarty D, Yu KW, Paek KY. Detection of DNA methylation changes during somatic embryogenesis of Siberian ginseng (*Eleutherococcus senticosus*). *Plant Science* 2003;165:61–8.
130. Sun Y, Deng D, You WC, et al. Methylation of p16 CpG islands associated with malignant transformation of gastric dysplasia in a population-based study. *Clin Cancer Res* 2004;10:5087–93. [PubMed: 15297411]
131. Luo D, Zhang B, Lv L, et al. Methylation of CpG islands of p16 associated with progression of primary gastric carcinomas. *Lab Invest* 2006;86:591–8. [PubMed: 16534497]
132. Tryndyak VP, Muskhelishvili L, Kovalchuk O, et al. Effect of long-term tamoxifen exposure on genotoxic and epigenetic changes in rat liver: implications for tamoxifen-induced hepatocarcinogenesis. *Carcinogenesis*. 2006
133. Song L, James SR, Kazim L, Karpf AR. Specific method for the determination of genomic DNA methylation by liquid chromatography-electrospray ionization tandem mass spectrometry. *Anal Chem* 2005;77:504–10. [PubMed: 15649046]

134. Gama-Sosa MA, Midgett RM, Slagel VA, et al. Tissue-specific differences in DNA methylation in various mammals. *Biochim Biophys Acta* 1983;740:212–9. [PubMed: 6860672]
135. Fisher DH, Giese RW. Determination of 5-methylcytosine in DNA by gas chromatography-electron-capture detection. *J Chromatogr* 1988;452:51–60. [PubMed: 2854138]
136. Leonard SA, Wong SC, Nyce JW. Quantitation of 5-methylcytosine by one-dimensional high-performance thin-layer chromatography. *J Chromatogr* 1993;645:189–92. [PubMed: 8408414]
137. Cortacero-Ramirez S, Segura-Carretero A, Cruces-Blanco C, et al. Simultaneous determination of multiple constituents in real beer samples of different origins by capillary zone electrophoresis. *Anal Bioanal Chem* 2004;380:831–7. [PubMed: 15490130]
138. Zhu X, Leav I, Leung YK, et al. Dynamic regulation of estrogen receptor-beta expression by DNA methylation during prostate cancer development and metastasis. *Am J Pathol* 2004;164:2003–12. [PubMed: 15161636]
139. Herman JG, Graff JR, Myohanen S, et al. Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. *Proc Natl Acad Sci U S A* 1996;93:9821–6. [PubMed: 8790415]
140. Derks S, Lentjes MH, Hellebrekers DM, et al. Methylation-specific PCR unraveled. *Cell Oncol* 2004;26:291–9. [PubMed: 15623939]
141. Xiong Z, Laird PW. COBRA: a sensitive and quantitative DNA methylation assay. *Nucleic Acids Res* 1997;25:2532–4. [PubMed: 9171110]
142. Bian YS, Yan P, Osterheld MC, et al. Promoter methylation analysis on microdissected paraffin-embedded tissues using bisulfite treatment and PCR-SSCP. *Biotechniques* 2001;30:66–72. [PubMed: 11196322]
143. Clement G, Bosman FT, Fontollet C, Benhattar J. Monoallelic methylation of the APC promoter is altered in normal gastric mucosa associated with neoplastic lesions. *Cancer Res* 2004;64:6867–73. [PubMed: 15466175]
144. Eads CA, Danenberg KD, Kawakami K, et al. MethyLight: a high-throughput assay to measure DNA methylation. *Nucleic Acids Res* 2000;28:E32. [PubMed: 10734209]
145. Ogino S, Kawasaki T, Brahmandam M, et al. Precision and performance characteristics of bisulfite conversion and real-time PCR (MethyLight) for quantitative DNA methylation analysis. *J Mol Diagn* 2006;8:209–17. [PubMed: 16645207]

Figure 1. DNA methylation machinery and transcriptional repression. A) Diagrams showing the biochemical pathways for cytosine methylation, demethylation, and mutagenesis of cytosine and 5-metC [33]. DNA methylation by addition of a methyl group to carbon 5 position of the cytosine ring is catalyzed by DNA methyltransferases (DNMTs), and demethylation is catalyzed by demethylase. 5-Methyl cytosine (5meC) undergoes hydrolytic deamination to thymidine. Mutation at CpG occurs because 5meC is more susceptible than cytosine to deamination and because some of the T-G mismatches produced by deamination are poorly repaired. B) Methylation at cytosine of DNA blocks transcription. Singal and Ginder have proposed three mechanisms of how methylation inhibits gene transcription [33]. Details are in the text.

Figure 2.

The work-flow chart outlines how to discover the methylated genes in a step-by-step manner. First, one of the methylation profiling techniques is used to pool out several candidate clones showing differential methylation patterns. After subcloning and sequencing, candidate clones are identified by aligning sequences into BLAST (from NCBI) or BLAT (from UCSC) database. Next, *in silico* database analysis, such as promoter and CpG island search on the candidate clones, is done to characterize the gene structure and design primers for subsequent data validation assays. Several methods, such as RT-PCR, Western blotting analysis, bisulfite sequencing (BS), methylation-specific PCR (MSPCR), combined bisulfite restriction analysis (COBRA), methylation-specific single strand conformation polymorphism (MSSSCP), and MethylLight, are used to validate the methylation level and gene expression of the target genes. If target genes are ready the methylation studies, it is not necessary to start from the first step. Methylation studies start from any step, depending on the researcher's needs.

Figure 3.

Principle of methylation sensitive restriction fingerprinting (MSRF). A) Methylated cellular DNA that will be digested by enzymes *MseI* and *BstUI* and DNA with an intact methyl group can be amplified in PCR. B) A schematic diagram depicting expected results. In this comparison, the methylation status was defined by comparing samples A and B. If there is no difference in methylation status, no difference in methylation status will be observed for both cases. If that candidate is being methylated in sample A, none of the *BstUI* sites in the CpG-rich sequence will be digested and they can be amplified in PCR. If vice versa, DNA fragments will be absent in the lane of sample A (*MseI/BstUI*) if this candidate is unmethylated in sample A.

Figure 4.

Diagram showing the procedures of restriction landmark genomic scanning (RLGS). Methylation detection in the RLGS profile depends on the methylation sensitivity of *NotI*. *NotI*, which recognizes CG-rich regions and cannot cleave DNA sequences when 5-cytosine is methylated, acts as the landmark in the profiles. Three types of methylation status are expected, depending on which allele is methylated. m*NotI* represents the methylated site of the enzymes. Following further endonuclease digestion of *EcoRV* (1D gel electrophoresis) and *HinfI* (2D gel electrophoresis), the RLGS profile will result in a change in spot intensity that quantitatively reflects the copy number and methylation status of the *NotI* fragment

Figure 5.

Schematic diagram for methylation CpG island amplification (MCA) coupled with representational difference analysis (RDA). CpG sites are labeled as 1 to 4. In the comparison of the methylation status of sample A and sample B, CpG site 2 is methylated in both cases whereas CpG site 4 is methylated only in sample A. For MCA, unmethylated CpGs are digested by methylation-sensitive restriction enzyme (*Sma*I), resulting in the formation of the blunt end fragments. Those methylated CpGs are then digested with *Xma*I and generate sticky end fragments. Following ligation into linker and PCR amplification, amplicons of short sequences can then be directly hybridized to enable study of the methylation status of the gene of interest for which a probe is available. As shown in the figure, differential methylation of CpG sites 2 and 4 between samples A and B can be found in dot-blot analysis. Besides, MCA amplicons (i.e., fragments containing CpG site 4) showing differentially methylation status in sample A can be cloned by RDA to identify any novel candidates.

Figure 6.

Schematic diagram for differential methylation hybridization (DMH). In brief, genomic DNA is first digested with *MseI* and ligated to linkers and then digested with methylation-sensitive *BstUI*. Both *MseI* or *Mse/BstUI* digestion products are amplified to generate probes for hybridization to a microarray that is prescreened by their CGI library array. The hybridization output is the measured intensities of the two fluorescence reporters with *red* (sample B) and *green* (sample A). *Yellow* spots indicate equal amounts of bound DNA from each amplicon, signifying no methylation differences between sample A and B genomes. Spots hybridized predominantly with sample B amplicon but not with sample A amplicon would appear *red*, which is indicative of the presence of hypermethylated CpG island loci in the tumor genome. Methylation level of each CG clone can be analyzed.

Figure 7.

Diagram illustrating how to analyze DNA methylation by methylation-specific oligonucleotide (MSO). Genomic DNA is first modified with sodium bisulfite before the assay. The probes on the MSO array are a set of short oligonucleotides (~20-mers) designed for specific methylated CG or unmethylated C/TG sites to test all the CpG sites within the CGI of a known gene. Both methylated and unmethylated DNA can be amplified, labeled with Cy5 dye, and then hybridized to oligonucleotide probes attached to the glass plate/membrane. Signals from MSO arrays can be recorded with a fluorescence scanner, and signal intensities between the pair of probes are compared to arrive at the percentage of methylation of the CpG within the short region represented by the oligonucleotides, usually 2–3 CpG sites [81].

Figure 8.

Workflow of chromatin immunoprecipitation (ChIP) on DNA microarray (ChIP-chip) analysis. At first, living cells are fixed by formaldehyde crosslinking. Intact genomic DNA with transcription factors (TF) is then isolated, and the chromatin-protein complex of interest is immunoprecipitated (IP) with a specific antibody to that intact TF. After reverse crosslinking, DNA is extracted and purified before PCR to generate chromatin amplicons. Amplicons from experimental immunoprecipitation are labeled with Cy5, and Cy3 is used to label input reference amplicons. Labeled probes are applied on the CGI array for hybridization. Data can be utilized to study the interaction between particular TFs and specific CpG sites of genes.

Figure 9.

Procedures for analysis global DNA methylation with luminometric methylation assay (LUMA). Genomic DNA is digested with combinations of restriction enzymes, *EcoRI*/*MspI* or *EcoRI*/*HpaII*, to leave the TTA (*EcoRI*) and CG (*HpaII* [methylation-sensitive] or *MspI* [methylation-insensitive]) overhangs. Next, the extent of the cleavage is determined by a polymerase extension assay based on a four-step pyrosequencing reaction. Inorganic pyrophosphate (PPI) is generated at each nucleotide addition in the polymerase extension assay. Signals are generated based on the utilization of PPI in a luciferase-based reaction. The amount of light generated is directly proportional to the number of overhangs produced by respective restriction enzymes. The A and T peaks represent signals from Steps 1 and 3. The C+G peaks resulting from Step 2 illustrate *HpaII* or *MspI* cleavage. The second C+G peak originating from Step 4 is an internal control for the completion of Step 2. Unmethylated CG is cleaved by *HpaII*, and a CG overhang is left after the cleavage and amplified in polymerase extension. Methylation status can be determined by analyzing the *HpaII*/*MspI* ratio at CG peaks from Step 2. The *HpaII*/*MspI* ratio is close to 1 in the unmethylated sample but close to zero in the methylated sample.

Figure 10.

Analysis of CpG methylation patterns using RNase T1 cleavage and Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS). Genomic DNA is modified by sodium bisulfite to convert unmethylated cytosines to uracils followed by PCR amplification with forward primer containing a control tag and reverse primer carrying T7 promoter. RNA transcription generates G-sites at originally methylated C sites and G-specific cleavage with RNase T1 is done. Control tag is used to monitor the successful full-length transcription and followed RNase T1 cleavage. RNA fragments are subjected to MALDI-TOF analysis. By comparison of the profile of m/z values of all fragments in the samples, methylation status of genes can be found.

Figure 11.

Different approaches can be chosen to yield the information on the overall characterization of genes showing differential methylation status. Unmethylated DNA is distinguished from methylated DNA with a sodium bisulfite modification of DNA used as a standard procedure prior to validation assays. The underlying principle is based on the ability of sodium bisulfite to deaminate cytosine (C) residues into uracil (U) in genomic DNA, whereas the methylation cytosine residues are resistant to this modification. After PCR amplification, the U residues are amplified as thymines (Ts). Cloning and subsequent sequencing of the DNA fragments containing the CpGs then provide information on the methylation status of each C within the CpGs. Details of methylation-specific PCR (MSPCR), combined bisulfite restriction analysis (COBRA), methylation-specific single-strand conformation polymorphism (MSSSCP), and MethyLight are discussed in the text.

Table 1

Advantages and limitations of different approaches to methylation profiling.

Methods	Global /Gene Specific	Sample amount	Advantages
MSRF	Global	100ng-1 μ g	Simple set-up Screen for novel genes
RLGS	Global	1–5 μ g	High-throughput
MCA-RDA	Gene Specific Global (RDA)	5 μ g	High-throughput CpG islands specific
DMH/MSO	Gene Specific	2 μ g	High-throughput CpG islands specific
CGI/ChIP-chip Promoter array	Global	1–10 μ g	High-throughput
LUMA	Global/Gene Specific	200–500ng	High-throughput
MALDI-TOF/HPLC	Gene Specific	10ng-1 μ g	High-throughput

Limitations	Examples
Small scale	[20], [49]
Landmark site only	[59-61]
Special gel electrophoresis set-up	
Software is not available for every species	
Need of prior knowledge of sequences (except RDA)	[63-65]
Need of prior knowledge of sequences	[71], [81]
High cost	[79]
Platform specific	[106-7]
Not popular	
Relative high cost	[108], [112-113]
Platform specific	
Limited length of sequence	
Relative high cost	[123], [130]
Platform specific	[131-132]
Still in stage of optimization of the protocol	
Mostly on specific target gene	