

Review

Tackling the methylome: recent methodological advances in genome-wide methylation profiling

Marcos RH Estécio and Jean-Pierre J Issa

Address: Department of Leukemia, UT MD Anderson Cancer Center, Houston, TX 77030, USA.

Correspondence: Jean-Pierre J Issa. Email: jpissa@mdanderson.org

Abstract

DNA methylation of promoter CpG islands is strongly associated with gene silencing and is known as a frequent cause of loss of expression of tumor suppressor genes, as well as other genes involved in tumor formation. DNA methylation of driver genes is very likely outnumbered by the number of methylated passenger genes, though these can be useful as tumor markers. Much of what is known about the importance of DNA methylation in cancer was gained through small- and moderate-scale analysis of gene promoters and tumor samples. A much better understanding of the role of DNA methylation in cancer, either as a marker of disease or as an active driver of tumorigenesis, will likely be gained from genome-wide studies of this modification in normal and malignant cells. This goal has become more attainable with the recent introduction of large-scale genome analysis methodologies and these have been modified to allow for investigation of DNA methylation. Several research groups have been formed to coordinate efforts and apply these methodologies to decipher the methylome of healthy and diseased tissues. In this article we review technological advances in genome-wide methylation profiling.

Introduction

In mammals, DNA methylation is predominantly, if not exclusively, found in CpG dinucleotides, due to site specificity of the known DNA methyltransferases [1]. Although it was reported in the early 1960s that cytosines can be methylated, it was not until two decades later that DNA methylation was fully recognized as an important player in gene regulation [2-4]. By acting coordinately with histone tail modifications and recruitment of an array of proteins involved in chromatin condensation, DNA methylation participates in gene silencing, independently of changes in DNA sequence [5]. The large majority of CpG dinucleotides in the human genome are methylated, and this results in a depletion of CpG sites due to conversion to thymines by deamination [6,7]. Unmethylated CpG sites escape depletion and are clustered in relatively small areas

called CpG islands. A widely accepted definition of CpG islands was formulated by Gardiner-Garner and Frommer and takes into account local GC content, observed-to-expected frequency of CpGs and length of the region [8]. The exact meaning of these parameters has been disputed in recent publications and alternative definitions have been proposed in an attempt to better match definition of CpG islands to biological function [9-11]. Regardless of the definition, roughly one-third of CpG islands overlap with gene promoters, and as many as 70% of human promoters are associated with a CpG island. The vast majority of these promoter-associated CpG islands are unmethylated in normal tissues in both active and inactive genes, thus do not explain tissue-specific gene expression [12]. Exceptions to this general pattern are imprinted genes, X-inactivated genes in women, and germ-cell-restricted genes where promoter CpG island methylation is present [13]. Outside of CpG islands, the bulk of methylated cytosines in normal tissues is found in repetitive DNA elements, mostly retrotransposons of LINE and SINE classes [14].

DNA methylation is an extremely dynamic process during fertilization and embryogenesis. Almost complete loss of methylation occurs very early, and selective re-methylation occurs during implantation [15,16]. The pattern of methylation established after this stage is remarkably stable, although as discussed above, somewhat rare in *bona fide* promoter CpG islands in adult tissues. Remodeling of these patterns is found in human diseases, especially cancer, with global demethylation (mainly at repetitive DNA) and local hypermethylation (frequent in promoter CpG islands) being hallmarks of most neoplasias [17-19]. Since DNA methylation results in gene silencing, it has been recognized as a frequent cause of inactivation of tumor suppressor genes and other genes important for tumor development [20]. There is a vast literature on promoter CpG island methylation in cancer, with evidence supporting its

BSPP, bisulfite padlock probe; CHARM, comprehensive high-throughput arrays for relative methylation; CIMP, CpG island methylator phenotype; ESC, embryonic stem cell; FDA, Food and Drug Administration; HELP, *Hpa*II tiny fragment enrichment by ligation-mediated PCR; MCA, methylated CpG island amplification; MCAM, methylated CpG island amplification microarray; MeDIP, methylated DNA immunoprecipitation; MIRA, methylated-CpG island recovery assay; MSCC, methyl-sensitive cut counting; PCR, polymerase chain reaction; RRBS, reduced representation bisulfite sequencing.

role in disease progression [21]. Also of note is the existence of a subset of tumors with extensive, concomitant methylation of multiple genes, which has been termed CpG island methylator phenotype (CIMP) [22,23]. Additionally, DNA methylation has proven to be an important therapeutic target. Two drugs with demethylating activity (azacitine and decitabine) have been approved by the Food and Drug Administration (FDA) for treatment of myelodysplastic syndrome, and are being tested in clinical trials for treatment of other leukemias as well as solid tumors [24-26]. These broad implications support the in-depth study of DNA methylation in cancer and normal tissues.

Array-based methodologies for large-scale analysis

One of the main obstacles to DNA methylation analysis is that methylated cytosines cannot be detected simply by sequencing. During polymerase chain reaction (PCR) amplification, methylated cytosines are not differentiated by the DNA polymerase and, similarly to unmethylated cytosines, they are paired with guanosine dinucleotides. Thus, reading of methylated cytosines depends on indirect methods. The most commonly used are (1) restriction enzyme-based approaches, which take advantage of methylation-sensitive enzymes, (2) affinity-based approaches, where antibodies against either 5-methylcytosine or methyl-binding domain proteins are used to collect the methylated fraction of the genome, and (3) bisulfite conversion of non-methylated cytosines to thymidine through a hydrolytic deamination reaction, which takes advantage of the non-reactivity of methylated cytosines to free hydroxyl groups. Each one of these methods has an important application in studying the epigenome and has been individually, or in combination, applied to individual genes and also to large-scale analyses (Table 1). Among these methods, bisulfite conversion is the gold standard, due to its potential high resolution when combined with sequencing methods. In this way, every single cytosine can be identified as methylated or unmethylated.

All the above-mentioned strategies to unveil methylated cytosines have been applied to microarray platforms to achieve moderate- and high-resolution coverage of the human genome. In the first generation of methylation microarrays, methylated genomic fragments were selectively amplified in a ligation-mediated PCR after DNA digestion with one or more methylation-sensitive enzymes and, after labeling with fluorescent dyes, hybridized against a normal control [27,28]. Soon thereafter, the gold-standard status of bisulfite modification to study DNA methylation prompted the generation of microarray platforms exploiting this chemical to study methylated cytosines. These arrays mostly targeted a few genes by tiling oligonucleotide probes representing the bisulfite-converted methylated and unmethylated versions of the promoter sequence [29,30]. These methods suffered from low

throughput and complicated probe design and were soon abandoned in favor of restriction-enzyme-based methods.

Since then, the microarray platforms have increased in gene density, and genome-wide coverage can be achieved with tiling arrays. Concomitantly, variations of the restriction-enzyme-based methods were developed to maximize the number of studied genomic targets and to increase the sensitivity and specificity of the method. Our group developed a strategy based on the well-established methylated CpG island amplification protocol (MCA). The advantage of the method is the use of two isoschizomer enzymes with differential sensitivity to methylated cytosines (*SmaI* and *XmaI*) which, due to their recognition site, preferentially target CpG islands [31]. Done this way, our method is a positive representation of methylated fragments (Figure 1), which results in higher sensitivity and specificity compared to other methods. Since then, this method has been applied to study the methylome of leukemias, liver cancer and normal peripheral blood lymphocytes [12,21,32]. Other enzymes tested by other groups include *HpaI/MspI* (HELP - *HpaII*-tiny fragment enrichment by ligation-mediated PCR [33]) and *McrBc*, which, contrary to methylation-sensitive enzymes, preferentially fragments the DNA between a pair of methylated CpGs at a critical distance.

The success of restriction-enzyme-based methods is largely dependent upon their capacity to simplify the genome prior to PCR amplification (thus allowing a more uniform, unbiased amplification), generating what has been called a reduced representation. However, since only selected sites can be studied at once, these methods are not truly genome-wide and can be biased to genome compartments (for example, CG-rich versus CG-poor areas). Two affinity-based strategies were developed to circumvent this limitation. In one method, termed methylated DNA immunoprecipitation (MeDIP), antibodies against 5-methyl-cytosine were used to pull-down the methylated fraction of the genome, and were co-hybridized against the unprocessed DNA from the same sample [34]. In another strategy, antibodies against the methyl-binding domain proteins MBD2 and MBD3L1 were used to capture methylated DNA fragments. This methylated-CpG island recovery assay (MIRA) was performed similarly to MeDIP, in the sense that the control sample is the unprocessed DNA. A recent comparison of the sensitivity and specificity of HELP, MeDIP and *McrBc* fragmentation methods showed that each was biased in a different way [35]. Among these, the authors found *McrBc* fragmentation to have the highest potential for improvement, and modified it to achieve more precise mapping of methylated CpG sites, a method they called comprehensive high-throughput arrays for relative methylation (CHARM).

Next-generation sequencing

Microarray-based methods, despite their high resolution, are generally far from being truly genome-wide analyses.

Table 1

Recent methodologies applied to whole human genome DNA methylation analysis

Technique	Platform	Reference	Description
Enzyme-based			
CHARM	Microarray	[35]	Digestion of methylated DNA is done using the McrBc enzyme, which cuts between two methylated CpG sites. Unprocessed DNA is used as control. Increased sensitivity and specificity of the method is achieved by smoothing the data of neighboring genomic locations.
HELP	Microarray	[33]	<i>HpaII</i> restriction enzyme is used to eliminate the methylated fraction of the genome, and the enrichment for unmethylated fragments is compared in an array platform with DNA digested with <i>MspI</i> .
MCAM	Microarray	[31]	The methylated fraction of the genome is selectively enriched by PCR after sequential digestion of the DNA with <i>SmaI</i> and <i>XmaI</i> restriction enzymes. CpG islands are preferentially represented in this method.
HELP-Seq	NextGen	[47]	The general procedure is done as for standard HELP, and the original adapters are removed by digestion with <i>MspI</i> before sequencing. DNA methylation is measured, and enrichment of <i>HpaII</i> compared to <i>MspI</i> sequences.
Methyl-Seq	NextGen	[43]	Massively parallel sequencing of <i>HpaII</i> -digested DNA is performed and methylation frequency is inferred from the frequency of tags per regions (fewer tags equals more methylation). The sequencing of <i>MspI</i> -digested DNA is used to identify regions refractory to sequencing, but unlike HELP-Seq, it is not used to calculate the enrichment of <i>HpaII</i> fragments.
MSCC	NextGen	[41]	The method is similar to Methyl-Seq; however, sequencing of <i>MspI</i> libraries was reported to have little effect on the measurement of methylation and was abolished to reduce costs.
Affinity-based			
MeDIP	Microarray	[34]	Methylated DNA is captured in using anti-5-methylcytosine antibodies and hybridized in an array platform. In this way, the method is unbiased towards recognition sites like enzyme-based methods, but it has been shown that dense CpG islands are preferentially captured.
MIRA	Microarray	[48]	Antibodies against methyl-binding domain proteins are used to capture methylated DNA.
MeDIP-Seq	NextGen	[49]	The procedure is the same as MeDIP, followed by massively parallel sequencing after DNA capture instead of microarray hybridization.
Bisulfite-based			
MethylC-Seq	NextGen	[44]	The genome is fragmented by sonication, and modified adapters are ligated to the DNA prior to bisulfite conversion. It is the only truly genome-wide method applied to the human genome at the moment, but the high cost of the method limits its application to large groups of samples.
Padlock, BSPPs	NextGen	[41,42]	Selected targets in the bisulfite-converted genome, typically thousands, are collected using molecular inversion probes. The method is extremely useful when there is interest in highly quantitative analysis of selected loci.

Close to genome-wide coverage can be achieved by the combination of one of the affinity-based methods and high-density tiling arrays, and this has been done to study the methylome of B lymphoid blood cells at 100-bp resolution [36]. Such an approach is quite expensive and time consuming, explaining why few research groups have used it to study whole-genome methylation. The introduction of what has been called next-generation sequencing brought a fresh excitement to genome and epigenome analysis. By making possible the reading of millions of sequences at once, next-generation sequencing equilibrated the usefulness of the methods to reveal genome-wide DNA methylation in favor of the gold-standard bisulfite-based

detection. Currently, there are four main competing next-generation sequencing technologies available: Illumina Genome Analyzer, generally referred to as Solexa sequencing, from Illumina, Inc.; SOLiD™ System, from Applied Biosystems; HeliScope Single Molecule Sequencer, from Helicos BioSciences; and 454 Sequencing, from Roche. Despite variations, all platforms take advantage of parallel processing of thousands to millions of DNA sequences at a time (massively parallel sequencing), and the base detection is either based on classical Sanger sequencing (using fluorescently labeled nucleotides) or the innovative pyrosequencing method. This is a rapidly advancing field and companies are strongly competing to

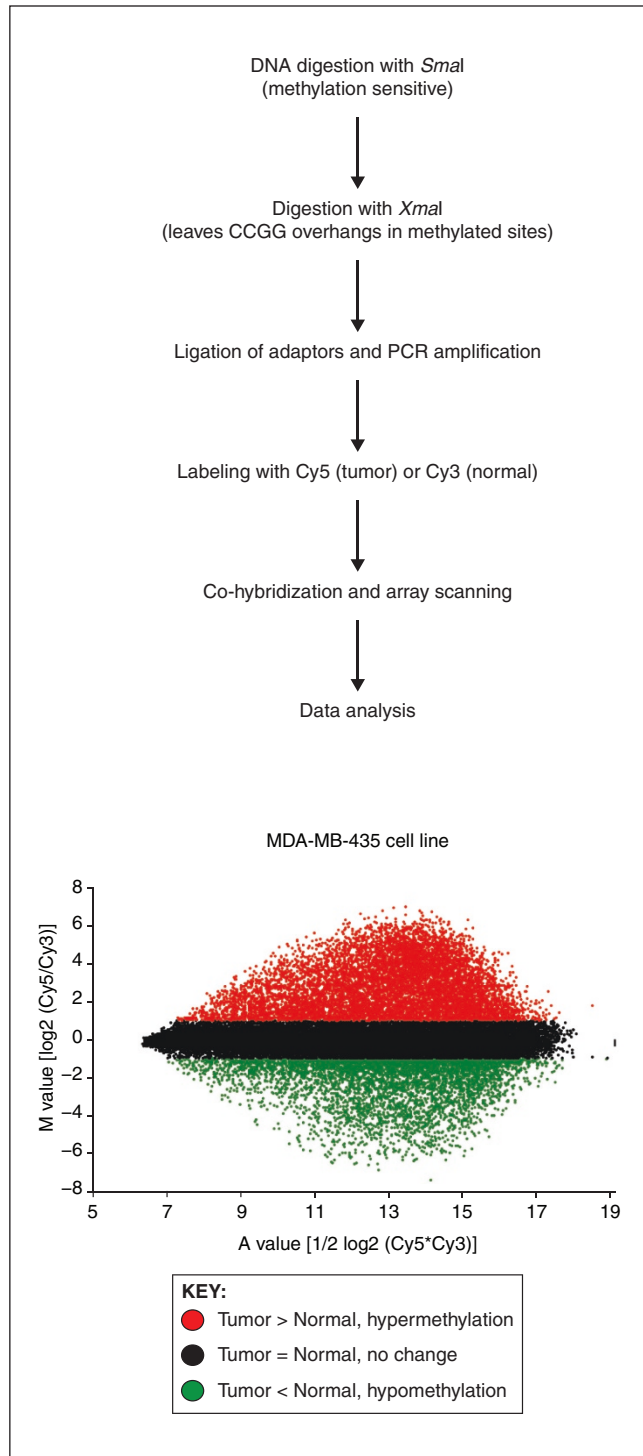


Figure 1

Schematic diagram of the methylated CpG island amplification microarray (MCAM) method. Enrichment for methylated DNA and reduction of genome complexity is achieved by serial digestion with *Sma*I (methylation sensitive) and *Xma*I (methylation insensitive) restriction enzymes, followed by ligation of adaptors and PCR amplification. The resulting amplicons, representative of the methylated fraction of tumor and normal cells, are labeled and co-hybridized in a microarray platform. Image acquisition and data analysis allow identification of methylated and non-methylated genes by comparing intensity values of Cy5 and Cy3 dyes for each pair of tumor and control samples. In this example, the M-A plot of normalized data from the cancer cell line MDA-MB-435 compared to normal peripheral blood is presented, from which amplicons were co-hybridized to a custom Agilent microarray containing 44,000 oligonucleotide probes targeting human promoter CpG islands.

increase genome coverage per run and to reduce the cost of their method.

As for whole-genome tiling microarrays [37], the first organism to have its methylome sequenced at single-base resolution was the plant *Arabidopsis thaliana* [38,39]. To do this, two groups fragmented the genomic DNA by sonication prior to ligation of PCR primer adaptors and bisulfite conversion, and performed shotgun sequencing using the Illumina Solexa platform. Compared to the human methylome (and the methylome of all mammals), the methylome of *Arabidopsis* is quite complex: in addition to methylation in CpG dinucleotides, there are also CHG and CHH methylation (H = A, C or T). From an analytical point of view, the possible combinations of methylated/unmethylated cytosines are less complex in humans than in *Arabidopsis*, making sequence matching and assembling less laborious. However, the *Arabidopsis* genome is just a fraction of the size of the human genome (119 Mb in *Arabidopsis* versus 3.1 Gb in human). Thus, the size of the human genome has been the main obstacle to whole-genome sequencing.

Not long after the *Arabidopsis* methylome was fully sequenced, the mouse methylome of pluripotent and differentiated cells from various tissues was sequenced with moderate coverage. To circumvent the genome size obstacle (the mouse genome is 2.7 Gb in size), the authors took advantage of the reduced representation generated from DNA digestion with the *Msp*I restriction enzyme, which has a recognition site (CCGG) abundant in CpG islands [40]. In this technique (reduced representation bisulfite sequencing, RRBS), bisulfite treatment is done for size-selected DNA fragments, targeting the most CpG island-enriched fraction, followed by bisulfite-treatment and Illumina Solexa sequencing. While analysis of the human methylome by RRBS has not yet been reported, this ingenious technique is very promising for such investigation. Meanwhile, the human methylome has been studied using other reduced representation strategies. A target-specific approach using 'padlock' probes was recently introduced by two different groups [41,42]. By presenting a unique sequence in each end, designed to match the bisulfite-converted genome, these probes capture targeted regions and create a circular molecule. The internal part of these probes is a universal sequence that allows for simultaneous amplification of all circularized, captured

sequences prior to massively parallel sequencing. Coincidentally, in their initial articles, both groups demonstrated the feasibility of their method by sequencing 10,000 targets, but the method can be extended to more or fewer targets according to the research goal. Interestingly, there seems to be an inherent bias in the process, with some circularized DNA being preferentially amplified or sequenced. Thus, some additional optimization of the method will be necessary prior to increasing the number of targets per analysis. It is also important to note that, since target selection is part of the procedure, these methods do not represent a genome-wide method. However, they are of extreme practical use when there is a strong interest in genome regions or promoter CpG islands alone. In one of these reports, the authors go one step further and introduce a less biased approach, termed MSCC (for methyl-sensitive cut counting) [41]. In this method, the authors use the methylation-sensitive restriction enzyme *HpaII*, which, similarly to its methylation-insensitive isoschizomer *MspI*, cuts the genome at CCGG sites and thus covers 90% or more of the human CpG islands. The ligation of adaptors to the generated fragments, followed by PCR and massively parallel sequencing, results in mapping of unmethylated cytosines in the CCGG context. The authors present an inverse correlation between the abundance of MSCC tags and measured cytosine methylation per regions, but recognize that a much larger sequencing effort is necessary to increase accuracy at low methylation densities. In another independent publication, Brunner *et al.* [43] published a similar approach to MSCC, but they introduced the *MspI*-digested DNA as a control in the procedure, to discriminate CpG sites that can be assayed and mapped uniquely in the genome from those that cannot, to reduce the rate of false-positive methylation.

The first human methylome at single-base resolution was published earlier this year [44] and the authors employed the MethylC-Seq method, previously used to sequence the *Arabidopsis* methylome, to investigate the human methylome at single-base resolution. This landmark report is industrious both in methodology and in its findings. One embryonic stem cell (ESC) and one fetal lung fibroblast were sequenced and, to achieve a 14-fold coverage of the genome, more than 1 billion Solexa reads were generated for each. The results support that the methylome is very different between undifferentiated and differentiated cells, and the authors' unexpected findings of significant non-CpG methylation in ESCs (up to 25% of the methylated cytosines were in CHG and CHH contexts, similar to *Arabidopsis* cytosine methylation) strongly support that the physiological impact of DNA methylation will be better captured in whole-genome, deep, unbiased analyses. However, until sequencing costs are significantly reduced, the human methylome analysis at single-base resolution will be restricted to a few samples at a time. Studies in

cancer, however, will need more extensive analysis. At the minimum, cancer studies require the sequencing of dozens, if not hundreds, of samples due to their inherent genetic and epigenetic heterogeneity, and the various disease grades and prognostic groups. Additionally, genome-wide mapping of methylated cytosines must be quantitative rather than just qualitative; thus, massively parallel sequencing requires several-fold coverage of each individual CpG dinucleotide, which makes the task prohibitively expensive. As a compromise, strategies based on reduced representation of the genome are currently more practical for whole-methylome analysis.

Emerging technologies: single-molecule sequencing

Much of the excitement about advances in DNA sequencing technologies has emerged from the race to achieve genome-wide analysis of the human genome for \$1,000 or less. At the same time as improvements to the performance of next-generation sequencing are being carried out to reduce costs, totally new technologies are emerging. One of the most promising new technologies uses nanopores to achieve fast and reliable DNA sequencing. An electric current is generated by passing the DNA molecule through these nanopores and, although very weak, this current can be accurately measured and is dependent on the nucleotide base passing through the pore [45]. Importantly, done this way, DNA sequencing is possible without prior DNA amplification or use of labeled nucleotides. In terms of methylome analysis, this is very exciting: it has been reported that the electric current-based nanopore detection can differentiate methylated from unmethylated cytosines directly, bypassing the need for bisulfite treatment [46]. There is still much improvement to be made before this technology is ready to be commercialized, and one of the main technical difficulties is to pass the DNA molecule through the nanopores at the right speed, enabling correct base detection without gaps.

Conclusions

Genome-wide methods for methylome analysis have evolved at a pace. The methodological advances achieved in the last five years have moved the field from single-gene detection to the possibility of whole-genome studies at the single-base level, or at least high resolution. A better understanding of the function of DNA methylation in healthy and diseased tissues is likely to arise from these more detailed investigations and their correlation with both genetic and other epigenetic studies. Specifically in cancer, the study of the methylome of various disease stages and response to therapies will improve patient care by providing markers of progression and response to treatment.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

MRHE and J-PJI jointly wrote the review.

Acknowledgements

The authors are supported by the Leukemia Specialized Program of Research Excellence grant P50 CA100632. J-PJI is an American Cancer Society Clinical Research Professor.

References

- Jones PA: **DNA methylation and cancer.** *Cancer Res* 1986, **46**:461-466.
- Baylin SB, Hoppener JW, de Bustros A, Steenbergh PH, Lips CJ, Nelkin BD: **DNA methylation patterns of the calcitonin gene in human lung cancers and lymphomas.** *Cancer Res* 1986, **46**:2917-2922.
- Bird AP: **DNA methylation - how important in gene control?** *Nature* 1984, **307**:503-504.
- Jones PA, Taylor SM: **Cellular differentiation, cytidine analogs and DNA methylation.** *Cell* 1980, **20**:85-93.
- Baylin SB: **DNA methylation and gene silencing in cancer.** *Nat Clin Pract Oncol* 2005, **2**(Suppl 1):S4-11.
- Bird AP: **DNA methylation and the frequency of CpG in animal DNA.** *Nucleic Acids Res* 1980, **8**:1499-1504.
- Cooper DN, Krawczak M: **Cytosine methylation and the fate of CpG dinucleotides in vertebrate genomes.** *Hum Genet* 1989, **83**:181-188.
- Gardiner-Garden M, Frommer M: **CpG islands in vertebrate genomes.** *J Mol Biol* 1987, **196**:261-282.
- Irizarry RA, Wu H, Feinberg AP: **A species-generalized probabilistic model-based definition of CpG islands.** *Mamm Genome* 2009, 24 September epub ahead of print.
- Takai D, Jones PA: **Comprehensive analysis of CpG islands in human chromosomes 21 and 22.** *Proc Natl Acad Sci U S A* 2002, **99**:3740-3745.
- Weber M, Hellmann I, Stadler MB, Ramos L, Paabo S, Rebhan M, Schubeler D: **Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome.** *Nat Genet* 2007, **39**:457-466.
- Shen L, Kondo Y, Guo Y, Zhang J, Zhang L, Ahmed S, Shu J, Chen X, Waterland RA, Issa JP: **Genome-wide profiling of DNA methylation reveals a class of normally methylated CpG island promoters.** *PLoS Genet* 2007, **3**:2023-2036.
- Scarano MI, Strazzullo M, Matarazzo MR, D'Esposito M: **DNA methylation 40 years later: Its role in human health and disease.** *J Cell Physiol* 2005, **204**:21-35.
- Yoder JA, Walsh CP, Bestor TH: **Cytosine methylation and the ecology of intragenomic parasites.** *Trends Genet* 1997, **13**:335-340.
- Reik W, Dean W, Walter J: **Epigenetic reprogramming in mammalian development.** *Science* 2001, **293**:1089-1093.
- Santos F, Dean W: **Epigenetic reprogramming during early development in mammals.** *Reproduction* 2004, **127**:643-651.
- Ehrlich M: **Cancer-linked DNA hypomethylation and its relationship to hypermethylation.** *Curr Top Microbiol Immunol* 2006, **310**:251-274.
- Issa JP: **CpG island methylator phenotype in cancer.** *Nat Rev Cancer* 2004, **4**:988-993.
- Laird PW: **Cancer epigenetics.** *Hum Mol Genet* 2005, **14**:R65-76.
- Jones PA, Baylin SB: **The fundamental role of epigenetic events in cancer.** *Nat Rev Genet* 2002, **3**:415-428.
- Kroeger H, Jelinek J, Estecio MR, He R, Kondo K, Chung W, Zhang L, Shen L, Kantarjian HM, Bueso-Ramos CE, Issa JP: **Aberrant CpG island methylation in acute myeloid leukemia is accentuated at relapse.** *Blood* 2008, **112**:1366-1373.
- Toyota M, Ahuja N, Ohe-Toyota M, Herman JG, Baylin SB, Issa JP: **CpG island methylator phenotype in colorectal cancer.** *Proc Natl Acad Sci U S A* 1999, **96**:8681-8686.
- Weisenberger DJ, Siegmund KD, Campan M, Young J, Long TI, Faasse MA, Kang GH, Widschwendter M, Weener D, Buchanan D, Koh H, Simms L, Barker M, Leggett B, Levine J, Kim M, French AJ, Thibodeau SN, Jass J, Haile R, Laird PW: **CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer.** *Nat Genet* 2006, **38**:787-793.
- Gore SD, Jones C, Kirkpatrick P: **Decitabine.** *Nat Rev Drug Discov* 2006, **5**:891-892.
- Issa JP, Byrd JC: **Decitabine in chronic leukemias.** *Semin Hematol* 2005, **42**:S43-49.
- Issa JP, Kantarjian HM: **Targeting DNA methylation.** *Clin Cancer Res* 2009, **15**:3938-3946.
- Shi H, Yan PS, Chen CM, Rahmatpanah F, Lofton-Day C, Caldwell CW, Huang TH: **Expressed CpG island sequence tag microarray for dual screening of DNA hypermethylation and gene silencing in cancer cells.** *Cancer Res* 2002, **62**:3214-3220.
- Yan PS, Chen CM, Shi H, Rahmatpanah F, Wei SH, Caldwell CW, Huang TH: **Dissecting complex epigenetic alterations in breast cancer using CpG island microarrays.** *Cancer Res* 2001, **61**:8375-8380.
- Hou P, Ji M, Liu Z, Shen J, Cheng L, He N, Lu Z: **A microarray to analyze methylation patterns of p16(Ink4a) gene 5'-CpG islands.** *Clin Biochem* 2003, **36**:197-202.
- Yan PS, Shi H, Rahmatpanah F, Hsiao TH, Hsiao AH, Leu YW, Liu JC, Huang TH: **Differential distribution of DNA methylation within the RASSF1A CpG island in breast cancer.** *Cancer Res* 2003, **63**:6178-6186.
- Estecio MR, Yan PS, Ibrahim AE, Tellez CS, Shen L, Huang TH, Issa JP: **High-throughput methylation profiling by MCA coupled to CpG island microarray.** *Genome Res* 2007, **17**:1529-1536.
- Gao W, Kondo Y, Shen L, Shimizu Y, Sano T, Yamao K, Natsume A, Goto Y, Ito M, Murakami H, Osada H, Zhang J, Issa JP, Sekido Y: **Variable DNA methylation patterns associated with progression of disease in hepatocellular carcinomas.** *Carcinogenesis* 2008, **29**:1901-1910.
- Khulan B, Thompson RF, Ye K, Fazzari MJ, Suzuki M, Stasiak E, Figueroa ME, Glass JL, Chen Q, Montagna C, Hatchwell E, Selzer RR, Richmond TA, Green RD, Melnick A, Grealley JM: **Comparative isoschizomer profiling of cytosine methylation: the HELP assay.** *Genome Res* 2006, **16**:1046-1055.
- Weber M, Davies JJ, Wittig D, Oakeley EJ, Haase M, Lam WL, Schubeler D: **Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells.** *Nat Genet* 2005, **37**:853-862.
- Irizarry RA, Ladd-Acosta C, Carvalho B, Wu H, Brandenburg SA, Jeddelloh JA, Wen B, Feinberg AP: **Comprehensive high-throughput arrays for relative methylation (CHARM).** *Genome Res* 2008, **18**:780-790.
- Rauch TA, Wu X, Zhong X, Riggs AD, Pfeifer GP: **A human B cell methylome at 100-base pair resolution.** *Proc Natl Acad Sci U S A* 2009, **106**:671-678.
- Zhang X, Yazaki J, Sundaresan A, Cokus S, Chan SW, Chen H, Henderson IR, Shinn P, Pellegrini M, Jacobsen SE, Ecker JR: **Genome-wide high-resolution mapping and functional analysis of DNA methylation in arabidopsis.** *Cell* 2006, **126**:1189-1201.
- Cokus SJ, Feng S, Zhang X, Chen Z, Merriman B, Haudenschild CD, Pradhan S, Nelson SF, Pellegrini M, Jacobsen SE: **Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning.** *Nature* 2008, **452**:215-219.
- Lister R, O'Malley RC, Tonti-Filippini J, Gregory BD, Berry CC, Millar AH, Ecker JR: **Highly integrated single-base resolution maps of the epigenome in Arabidopsis.** *Cell* 2008, **133**:523-536.
- Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, Zhang X, Bernstein BE, Nusbaum C, Jaffe DB, Gnirke A, Jaenisch R, Lander ES: **Genome-scale DNA**

- methylation maps of pluripotent and differentiated cells.** *Nature* 2008, **454**:766-770.
41. Ball MP, Li JB, Gao Y, Lee JH, LeProust EM, Park IH, Xie B, Daley GQ, Church GM: **Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells.** *Nat Biotechnol* 2009, **27**:361-368.
 42. Deng J, Shoemaker R, Xie B, Gore A, LeProust EM, Antosiewicz-Bourget J, Egli D, Maherali N, Park IH, Yu J, Daley GQ, Eggan K, Hochedlinger K, Thomson J, Wang W, Gao Y, Zhang K: **Targeted bisulfite sequencing reveals changes in DNA methylation associated with nuclear reprogramming.** *Nat Biotechnol* 2009, **27**:353-360.
 43. Brunner AL, Johnson DS, Kim SW, Valouev A, Reddy TE, Neff NF, Anton E, Medina C, Nguyen L, Chiao E, Oyolu CB, Schroth GP, Absher DM, Baker JC, Myers RM: **Distinct DNA methylation patterns characterize differentiated human embryonic stem cells and developing human fetal liver.** *Genome Res* 2009, **19**:1044-1056.
 44. Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo QM, Edsall L, Antosiewicz-Bourget J, Stewart R, Ruotti V, Millar AH, Thomson JA, Ren B, Ecker JR: **Human DNA methylomes at base resolution show widespread epigenomic differences.** *Nature* 2009, [Epub ahead of print].
 45. Branton D, Deamer DW, Marziali A, Bayley H, Benner SA, Butler T, Di Ventra M, Garaj S, Hibbs A, Huang X, Jovanovich SB, Krstic PS, Lindsay S, Ling XS, Mastrangelo CH, Meller A, Oliver JS, Pershin YV, Ramsey JM, Riehn R, Soni GV, Tabard-Cossa V, Wanunu M, Wiggins M, Schloss JA: **The potential and challenges of nanopore sequencing.** *Nat Biotechnol* 2008, **26**:1146-1153.
 46. Clarke J, Wu HC, Jayasinghe L, Patel A, Reid S, Bayley H: **Continuous base identification for single-molecule nanopore DNA sequencing.** *Nat Nanotechnol* 2009, **4**:265-270.
 47. Oda M, Glass JL, Thompson RF, Mo Y, Olivier EN, Figueroa ME, Selzer RR, Richmond TA, Zhang X, Dannenberg L, Green RD, Melnick A, Hatchwell E, Bouhassira EE, Verma A, Suzuki M, Grealley JM: **High-resolution genome-wide cytosine methylation profiling with simultaneous copy number analysis and optimization for limited cell numbers.** *Nucleic Acids Res* 2009, **37**:3829-3839.
 48. Rauch T, Pfeifer GP: **Methylated-CpG island recovery assay: a new technique for the rapid detection of methylated-CpG islands in cancer.** *Lab Invest* 2005, **85**:1172-1180.
 49. Down TA, Rakyen VK, Turner DJ, Flicek P, Li H, Kulesha E, Graf S, Johnson N, Herrero J, Tomazou EM, Thorne NP, Backdahl L, Herberth M, Howe KL, Jackson DK, Miretti MM, Marioni JC, Birney E, Hubbard TJ, Durbin R, Tavare S, Beck S: **A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis.** *Nat Biotechnol* 2008, **26**:779-785.
-
- Published: 16 November 2009
doi:10.1186/gm106
© 2009 BioMed Central Ltd