Techniques of DNA Methylation Analysis with Nutritional Applications

Maria L. Mansego  Fermín I. Milagro  Javier Campión  J. Alfredo Martínez

CIBERobn, Physiopathology of Obesity and Nutrition, Institute of Health Carlos III, Madrid, and Department of Nutrition, Food Science and Physiology, University of Navarra, Pamplona, Spain

Key Words
DNA methylation · Nutrient · Epigenetic · Methylation-specific PCR · Pyrosequencing · MassARRAY · Microarray · Next-generation sequencing

Abstract
Epigenetic mechanisms are likely to play an important role in the regulation of metabolism and body weight through gene-nutrient interactions. This review focuses on methods for analyzing one of the most important epigenetic mechanisms, DNA methylation, from single nucleotide to global measurement depending on the study goal and scope. In addition, this study highlights the major principles and methods for DNA methylation analysis with emphasis on nutritional applications. Recent developments concerning epigenetic technologies are showing promising results of DNA methylation levels at a single-base resolution and provide the ability to differentiate between 5-methylcytosine and other nucleotide modifications such as 5-hydroxymethylcytosine. A large number of methods can be used for the analysis of DNA methylation such as pyrosequencing® primer extension or real-time PCR methods, and genome-wide DNA methylation profile from microarray or sequencing-based methods. Researchers should conduct a preliminary analysis focused on the type of validation and information provided by each technique in order to select the best method fitting for their nutritional research interests.

Introduction

Nutrients promote changes in gene expression profile by affecting transcriptional and posttranscriptional mechanisms [1]. They may also induce transient or permanent alterations in the epigenetic marks that regulate the expression of genes involved in metabolic
processes and other biological networks, which could be one of the factors leading to the developmental origin of pathological conditions like obesity, insulin resistance, type 2 diabetes, cancer, neurological and cardiovascular diseases, neurodegenerative diseases, and immune diseases [2–7].

Epigenetics is defined as the heritable states of gene expression resulting from a set of reversible changes in chromatin structure without alterations in the DNA sequence [8]. These changes may be induced in response to environmental or internal factors, and include different components of epigenetic control, such as nucleic acid modification, chromatin remodeling and histone modifications or noncoding RNAs (fig. 1). Regarding nucleic acid modification, DNA methylation, mainly of cytosine nucleotide at the carbon 5 position (5-mC), is a common epigenetic mark involved in epigenetic regulation of gene expression in many eukaryotes and is often found in the sequence context CpG (cytosine followed by a guanine) or CpHpG (H: A, T, C) [9]. Promoter methylation is typically associated with repression, whereas genomic methylation correlates with transcriptional activity [10]. Recently, the Encyclopedia of DNA elements (ENCODE) Project Consortium reported that levels of DNA methylation correlated with chromatin accessibility. They used reduced representation bisulfite sequencing (RRBS) to quantitatively profile DNA methylation for an average of 1.2 million CpGs in each of 82 cell lines and tissues (8.6% of nonrepetitive genomic CpGs), including CpGs in intergenic regions, proximal promoters, and in intragenic regions (gene bodies) [11]. They found that 96% of CpGs exhibited differential methylation in at least one cell type or tissue assayed. The most variably methylated CpGs are found more often in gene bodies and intergenic regions, rather than in promoters and upstream regulatory regions although the underlying mechanisms are unclear [12].

In recent years, we are experiencing a revolution in the field of epigenetics. The development of new technologies like next-generation sequencing (NGS) has allowed to detect deep (high coverage, where coverage is number of reads representing a given nucleotide in the sequence) and new epigenetic modifications such as 5-hydroxymethylcytosine (5-hmC), 5-formylcytosine (5-fC) and 5-carboxycytosine (5-caC) in genomic DNA [13] or N(6)-methyladenosine (N6mA) in RNA [14]. In addition, new sensitive methods have been developed to analyze different conditions and types of samples [biopsies, i.e. from muscle or adipose tissue, low amounts of input sample DNA, formalin-fixed paraffin-embedded samples (FFPE), whole peripheral blood, urine, saliva, plasma or serum] [15].
In this review, we discuss the technologies that can be suitable, in a nutrigenomic context, for the quantitative global analysis of DNA methylation or the methylation levels at specific loci in DNA. These techniques include pyrosequencing™, primer extension or real-time PCR methods, and genome-wide DNA methylation profile from microarray or sequencing-based methods.

**Methods for Nucleotide Methylation Analysis**

There are multiple methods to study nucleotide methylation. We report a series of techniques that allow to quantify the 5-mC levels, mainly, and some other recent nucleotide modifications such as 5-hmC. Table 1 shows a scheme of the various techniques that can be used to quantify methylation according to the pretreatment that the DNA receives and the number of modified nucleotides and the resolution of the detection (global, specific-region or genome-wide methylation). All of these approaches are based on one of three main pretreatment techniques: sodium bisulfite conversion, enzyme restriction digestion or affinity enrichment. During bisulfite treatment, unmethylated cytosines in single-stranded DNA are deaminated to give uracil, while leaving methylated cytosine intact [16]. Bisulfite conversion is the most conventional approach for pretreatment and is considered the gold standard for determining DNA methylation status because it offers single-CpG resolution [17]. However, this treatment has some disadvantages such as typical calls for larger quantities of sample DNA, which can degrade following chemical treatment, it can be limited by incomplete conversion of all unmethylated cytosines to uracils, and bisulfite conversion cannot discriminate between methylcytosine and hydroxymethylcytosine. Moreover, enzyme restriction digestion treatment is based on the ability of some methylation-sensitive restriction endonucleases

**Table 1. DNA methylation analysis methods**

<table>
<thead>
<tr>
<th>Pretreatment</th>
<th>Resolution:</th>
<th>Global quantitative</th>
<th>Locus-specific PCR-based</th>
<th>Genome-wide array-based</th>
<th>NGS-based</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sodium bisulfite conversion</td>
<td>Alu/LINE-1-PCR-pyrosequencing</td>
<td>Infinium</td>
<td>RRBS</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Alu/LINE-1-HRM</td>
<td>Golden Gate</td>
<td>WGSBS</td>
<td>oxBS-Seq</td>
<td></td>
</tr>
<tr>
<td>Enzyme restriction digestion</td>
<td>HPLC</td>
<td>DMH</td>
<td>Methyl-Seq</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>LC-MS</td>
<td>HpaII-PCR</td>
<td>MCA-Seq</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>LUMA</td>
<td>MS-MLPA</td>
<td>MSC</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>MS-FLAG</td>
<td>HELP</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CHARM</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MMASS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affinity enrichment</td>
<td>5-mC ELISA</td>
<td>MeDIP-PCR</td>
<td>MeDIP-chip</td>
<td>MeDIP-Seq</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5-mHC ELISA</td>
<td>MIRA</td>
<td>MIRA-chip</td>
<td>MethylCap-Seq</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MIRA-Seq</td>
<td></td>
</tr>
</tbody>
</table>

HPLC = High-performance liquid chromatography; LC-MS = liquid chromatography-mass spectroscopy; LUMA = luminometric methylation assay; SMART-MSP = sensitive melting analysis after real-time methylation-specific PCR; MS-HRM = methylation-specific high-resolution melting analysis; CHARM = comprehensive high-throughput arrays for relative methylation; MMASS = microarray-based methylation assessment of single samples; -seq = followed by sequencing; WGSBS = whole-genome shotgun bisulfite sequencing; MSCC = methylation-sensitive cut counting; MethylCap = methyl-DNA binding domain capture technique.

In this review, we discuss the technologies that can be suitable, in a nutrigenomic context, for the quantitative global analysis of DNA methylation or the methylation levels at specific loci in DNA. These techniques include pyrosequencing™, primer extension or real-time PCR methods, and genome-wide DNA methylation profile from microarray or sequencing-based methods.
(HpaII, MspI and HhaI) and their corresponding isoenzymes (not sensitive to methylation) to distinguish methylated from unmethylated cytosines [18]. An important limitation is that all restriction enzyme-based techniques condition analysis methylation within recognition sites. Lastly, affinity enrichment treatment is the most recent and simplest way to enrich methylated DNA, which was first proved with the methyl-binding protein MECP2 [19]. Enrichment-based methods are based on immunoprecipitation of genomic DNA with specific antibodies for methylated cytosine. An important point regarding affinity-based techniques is that they measure the density of methylation in a specific region. Therefore, a methylated stretch of DNA where methylation CG target sites are sparse might be difficult to differentiate from an unmethylated region. Specifically, in mammalian genomes, CG density is generally low and CG-dense sequences are typically unmethylated [20].

The different combination of pretreatment methods and subsequent molecular biology techniques, such as DNA microarrays and high-throughput sequencing, generates a plethora of techniques for mapping DNA methylation in a feasible way on a genome-wide scale.

**Global Quantification of DNA Methylation**

In mammalian cells, DNA methylation is found at 70–80% of CpG dinucleotides. More recently, global DNA methylation has been examined in surrogate tissues such as leukocyte DNA. A lower level of leukocyte DNA methylation has been associated with an increased risk of several types of cancer [21, 22]. The processes underlying tumor behavior in long interspersed nucleotide element-1 (LINE-1) hypomethylation are currently unknown. Possible mechanisms include activation of retroviruses at transposons, which may cause genomic instability, in particular chromosomal instability [23]. This finding suggests that DNA methylation may serve as a surrogate biomarker for systemic genomic methylation and provide an independent risk factor for cancer development. In addition, there is widespread interest in finding correlations between the genomic 5-mC levels and diet, lifestyle and clinical disorders. In nutritional studies, it would be important to quantify global patterns of DNA methylation in order to evaluate the effect of nutrient intake, and a specific phenotype such as a different dietary pattern (i.e. diet characterized by high intake of vegetables and fruits) may protect against global DNA hypomethylation [24].

Several methods can be used to detect the total 5-mC content in the genome. DNA can be digested into single nucleotides and total genomic 5-mC can be quantitated by high-performance liquid chromatography, thin-layer chromatography, or liquid chromatography/mass spectroscopy, but these methods are very laborious techniques. To bypass these problems, several PCR-based methods [high-resolution melt (HRM) curve analysis technology or real-time PCR or advanced pyrosequencing-based technology] have been developed to estimate global DNA methylation by assessing repetitive DNA elements such as Alu elements and LINE-1 [25, 26]. Interestingly, it has been shown that a wide range of bioactive food components, such as folate, riboflavin, pyridoxine or methionine, regulate global DNA hypomethylation by modifying the levels of S-adenosyl-L-methionine, a methyl group donor, and S-adenosyl-L-homocysteine, which is an inhibitor of enzymes catalyzing the DNA methylation reaction [27]. In a recent report, Schernhammer et al. [28] showed that folate supplementation decreases the risk to develop LINE-1-hypomethylated colon cancers. In the present study, researchers analyzed 606 colon cancers reported in two huge cohorts (the Nurses’ Health Study with 121,701 women and the Health Professionals Follow-Up Study with 51,529 men) for LINE-1 methylation level, a good indicator of cellular 5-mC level using pyrosequencing technology. The average of the relative amounts of C in the four CpG sites was used as overall LINE-1 methylation level in a given sample.
In parallel, simple methods have been developed for rapid detection and quantitation of global 5-mC, such as a competitive enzyme immunoassay in urine, isolated cells or tissue DNA samples [29].

**Locus-Specific Quantification of Methylation**

As global methylation analyses provide no information on the genomic positions at which nucleotide methylation is altered, other high-resolution approaches are needed to relate such changes to functional outcomes. Site-specific DNA methylation levels normally tend to be lower than global methylation levels. It is currently estimated that 50–60% of genes have CpG islands [30], defined as a GC-rich region (GC content >50%) greater than 200 bp in length and having an observed CpG/expected CpG ratio of greater than or equal to 0.6 [31]. Furthermore, a majority of studies have focused on the methylation of CpG islands in gene promoters in order to correlate its levels with transcriptional expression.

Most gene-specific methylation analysis methods are primarily 'PCR-based'. Specifically primers are designed to amplify both methylated and unmethylated regions with equal efficiency, as the methylation level is analyzed by different approaches. In this section, we focused on key quantitative techniques that determine the levels of 5-mC in a particular region of the genome.

**Methylation-Specific PCR**

This method entails initial modification of DNA by sodium bisulfite treatment and subsequent amplification with specific primers for methylated versus unmethylated DNA. Methylation-specific PCR (MSP) requires only small quantities of DNA and it is sensitive to 0.1% methylated region of a given CpG island locus [32]. Regarding nutritional studies using this methodology, Cordero et al. [33] reported that leptin and TNF-alpha promoter methylation levels in adipose tissue biopsies from obese women could predict the response to a low-calorie diet.

One significant improvement of this assay is the quantitative real-time PCR method that is capable of quantitating DNA methylation by using DNA oligonucleotides that anneal differentially to bisulfite-converted DNA according to the methylation status. Moreover, the fluorescence-based technique called MethyLight, uses three oligonucleotide primers/probes (forward and reverse primers, and interpositioned probe) allowing for a high degree of specificity, sensitivity, and flexibility in methylation detection [34]. This method can be performed on DNA extracted from paraffin-embedded tissue (FFPE) [35].

Another modification of this assay is methylation-specific fluorescent amplicon generation (MS-FLAG) [36]. Furthermore, FLAG (fluorescent amplicon generation) is a homogeneous signal generation technology based on the exceptionally thermostable endonuclease PspGI. FLAG provides real-time signal generation during PCR by PspGI-mediated cleavage of quenched fluorophores at the 5′-end of double-stranded PCR products. The MS-FLAG method uses 5′-end methylation-sensitive primers added to a restriction endonuclease recognition sequence and, as more recognition sequences are generated, the stronger is the fluorescence intensity that is observed. If the primer target sequence is unmethylated, the primer will not be able to bind, which results in the absence of a detectable signal.

**Methylation-Sensitive HRM Analysis**

A further method to differentiate methylated from unmethylated bisulfite-treated DNA is methylation-sensitive HRM (MS-HRM), a quantitative real-time PCR-based technique initially designed to distinguish genetic polymorphisms [37]. The PCR products are analyzed
directly by dissociation curve analysis and the resulting release of an intercalating fluorescent dye during melting. The degree of methylation, as represented by the C-to-T content in the PCR product, determines the rapidity of melting and consequent release of the dye. This method allows direct quantitation in a single-well assay and is particularly useful for assessing methylation in short fragments of DNA (a range of 150 bp) derived from FFPE biopsies, which are often moderately to highly degraded and have very small amounts of tissue. MS-HRM also provides accurate distinction between mixtures of methylated and unmethylated DNAs, allowing discrimination between DNA from 0 to 100% methylated as long as the bisulfite conversion is completely homogeneous. However, heterogeneous methylation across the CpG sites could influence the melting profile of the PCR product and, therefore, the estimates of methylation obtained by this technique [38]. An interesting example of the use of MS-HRM and the genotype-epigenotype interaction was reported by Carroll and Parle-McDermott [39], who assessed whether DNA methylation in the proximal promoter of RUNX1, a transcription factor identified as a responder to changes in the folate/riboflavin status, correlated with the MTHFR 677C>T genotype. In that study, the MTHFR677-TT genotype group tended to exhibit 0% DNA methylation of their proximal RUNX1 promoter compared to CC or CT individuals, particularly in the context of nutritional status, which is in line with previous studies showing low levels of folate in TT individuals [39, 40].

**Pyrosequencing**

Pyrosequencing is a high-throughput quantitative method used for bisulfite sequencing. Similar to MSP and MS-HRM, pyrosequencing requires the use of bisulfite-converted DNA and has the same limitations derived from this treatment. Although MS-HRM is a simpler and relatively inexpensive approach for the quantification of methylation of specific CpG sites than pyrosequencing, this method is the gold standard technique for DNA methylation detection, providing also information on the methylation status of single-CpG sites. Using PCR, the DNA is amplified and tagged by a primer that is biotinylated. The pyrosequencing technology is based on the release of pyrophosphate when nucleotides incorporate into the sequencing primer only if it is complementary to the template DNA sequence. It is a method widely used in disease research and it has also been employed in different nutritional studies [41–44]. As an example, a recent study describes the consequences of maternal undernutrition during gestation and lactation on DNA methylation and expression of the leptin gene, an adipokine that is involved in the regulation of body weight and food intake [45]. The leptin promoter methylation differences in mice fed a low-protein diet compared with controls caused a lower body weight/adiposity and exhibit a higher food intake.

**Mass Spectrometry Methylation Assays (MassARRAY Assay)**

The MassARRAY assay provides a highly sensitive method of detection based on the difference in fragment size that has been cleaved according to the methylation status of multiple CpG sites. This strategy has the ability to sequence reads up to 600 bp, which is considerably longer than other methods [46], but it also requires the use of bisulfite-converted DNA, and primers designed in regions without CpG nucleotides. Selected DNA regions are amplified by PCR with tagged primers, in vitro transcribed into RNA, and base-specifically cleaved by an endoribonuclease. Mass spectra of cleavage products are obtained by MassARRAY analyses based on MALDI-TOF mass spectrometry. Fragments differ in mass depending on the sequence changes introduced by the initial bisulfite treatment. This assay is used for validations of whole epigenome screening experiments (e.g. arrays or NGS technology) and candidate gene promoter analyses, and is even suitable for analyses of DNA extracted from FFPE tissues. Studies have shown variable site specificity of the assay, demon-
strating ranging values between 75 to 95% concordance when validated with other sequencing methods [47, 48]. The MassARRAY assay has been used in nutritional studies because of its capacity to study large fragments of DNA [49–51]. Milagro et al. [49] describe how hypocaloric diet-induced weight loss in humans could alter the DNA methylation status of specific gene promoters (ATP10A, CD44 and WT1). Moreover, baseline DNA methylation patterns may be used as epigenetic markers that could help to predict weight loss.

Methylation-Specific Multiplex Ligation-Dependent Probe Amplification

Methylation-specific multiplex ligation-dependent probe amplification (MS-MLPA) is a semi-quantitative method for DNA methylation analysis. It is a variant of the MLPA technique, in which copy number detection is combined with the use of a methylation-sensitive restriction enzyme [52]. The MS-MLPA protocol is very similar to the standard MLPA method, except that each MS-MLPA reaction generates two samples: one undigested sample for copy number detection and one digested sample for methylation detection, and MS-MLPA probes for target sequence contain the restriction site of the methylation-sensitive endonuclease HhaI. This technique has been widely used for the detection of epigenetic alterations in imprinting diseases, such as the Prader-Willi, Angelman and Beckwith-Wiedemann, and RSS syndrome [53–55]. Furthermore, MS-MLPA is an excellent choice for quantitative analysis of methylation in archival FFPE samples. Both the quality and quantity of DNA may limit subsequent analyses, insomuch as it does not require a large amount of input, no bisulfite conversion is needed, and it has the ability to quantify multiple CpGs per sample at the same time. Pavicic et al. [56] determined the methylation level at 11 microRNAs (miRNA) loci associated with CpG islands by MS-MLPA assays in order to correlate with hypermethylation at classic tumor suppressor promoters in normal and tumor tissues from FFPE biopsies. MS-MLPA assay has not yet been performed in the nutrition field. Nonetheless, this technique would enable the development of nutritional custom panels of middle throughput quantifying the methylation level of different loci such as miRNA genes that could be epigenetically regulated through the associated CpG islands and some could additionally modulate the epigenome by putatively targeting the DNA methyltransferases, suggesting an epigenetic regulation of their expression.

Methylated DNA Affinity Enrichment Assay

In affinity-based approaches, such as the methylated DNA immunoprecipitation (MeDIP) [20] and the methylated CpG island recovery assay (MIRA) [57], fragmented DNA is immunoprecipitated using a monoclonal antibody to 5-mC or by methylated DNA-binding proteins (MBD2b/MBD3L1 complex). The resulting enrichment of methylated DNA in the immunoprecipitated fraction can be determined by PCR-based methods to assess the methylation state of specific regions. Alternatively, MeDIP can be combined with large-scale analyses using microarrays as a genome-wide experimental readout, as described below.

In addition, some technical improvements have allowed the quantification of methylated DNA from low amounts of cells or even to distinguish between 5-mC and 5-hmC by specific antibodies [58, 59].

An interesting example of the use of MeDIP followed by quantitative PCR was reported by Barres et al. [60]. With this method, they found that acute exercise induced a dose-dependent expression of PGC-1α, PDK4, and PPAR-δ in human soleus muscle, together with a marked hypomethylation in their promoters. Similarly, the same promoter had a markedly decreased methylation level in mouse soleus muscles 45 min after ex vivo contraction. In L6 myotubes, caffeine exposure enhanced an intramembranous charge movement in skeletal muscle by increasing cytoplasmic Ca2+ concentrations [61], and induced gene hypomethylation of PGC-1α, PDK4, and PPAR-δ in parallel with an increase in the respective mRNA
content, and this effect was blocked by dantrolene, an inhibitor of Ca2+ release. Collectively, their results provide evidence that acute exercise induces gene-specific DNA hypomethylation in human skeletal muscle mediated by caffeine [60].

**Genome-Wide DNA Methylation Analysis**

Similarly to the revolution that took place in transcriptomic studies with the development of gene expression microarrays, rapidly improving technologies are increasing the capacity to assess either locus-specific (epigenetic) or genome-wide DNA methylation (epigenomics).

The different principles used to analyze locus-specific DNA methylation patterns are nowadays combined with microarrays or NGS [62, 63]. However, microarrays are being gradually replaced by sequencing-based approaches, which hold the promise of providing biological insights and new avenues for translational research and clinical applications.

**Microarray-Based Methylation Analysis**

Three main classes of microarray-based methods have been developed to quantify 5-mC distribution in genomes: (1) those based on sodium bisulfite modification; (2) those utilizing methylation-sensitive restriction enzymes, and (3) those enriching for highly methylated regions using an antibody specific for 5-mC or methyl-binding proteins.

In bisulfite-based epigenotyping, the microarrays are designed using pairs of oligonucleotide hybridization probes targeting the CpG sites of interest. One probe is complementary to the unaltered methylated sequence, and the other is complementary to the unmethylated sequence. The Illumina Methylation Assay (Infinium or Golden Gate technologies) is one of the methods that apply the bisulfite sequencing technology on a microarray level to generate genome-wide methylation data, as it is able at once to identify more than 485,000 methylation sites per sample at single-nucleotide resolution.

An increasing number of reports are found using these technologies in the nutrition field. Recently, Jacobsen et al. [64] examined the methylation status of 27,578 CpG sites from 14,475 genes in skeletal muscle biopsies from 21 healthy young men after the ingestion of a short-term high-fat overfeeding (HFO) diet and a control diet. They used a bead array-based methylation assay, and described that a HFO diet introduced widespread DNA methylation changes affecting 6,508 genes (45%), with a maximum methylation change of 13%. The HFO-induced methylation changes were only partially and nonsignificantly reversed after 6–8 weeks. Milagro et al. [49] used a combination of a high-throughput microarray-based assay with a second more specific technique (MassARRAY assay), including more individuals than in the first technique, which allowed to find new putative biomarkers and to reliably validate them. With this dual approach, it has been proven that hypocaloric diets induce changes in the DNA methylation pattern in 25 overweight and obese men under a caloric restriction intervention. Moreover, a baseline DNA methylation profile may be used as an early indicator of response to the metabolic effects of the weight loss program [49, 65]. Alterations in the DNA methylation levels primarily affect genes involved in inflammation, the reproductive system and cancer.

One of the enzyme-based methods coupled to array-based analysis is methylated CpG island amplification with microarray hybridization (MCAM), which uses the differential methylation sensitivities and cutting behaviors of two endonucleases (Smal and Xmal) with a final array hybridization step [66]. An alternative approach is differential methylation hybridization (DMH), which involves digestion of one pool of DNA with a methylation-sensitive restriction enzyme and a mock digestion of another pool. This process produces parallel DNA pools that are amplified and labeled with different fluorescent dyes for two-
Color array hybridization [67]. The relative fluorescent signal intensities can be used to extract DNA methylation information at the corresponding loci on the array. An adaptation of DMH, known as MethylScope, utilizes the endonuclease McrBC to randomly cut sheared DNA [68] instead of DNA digested with MseI restriction enzymes whose recognition sequence (TTAA) is found frequently within bulk DNA, but is rarely found within CpG islands which remain intact after digestion [19]. The combination with the tiling array (a subtype of microarray that functions by hybridizing labeled DNA or RNA target molecules to probes fixed onto a solid surface) design and data processing is referred to as comprehensive high-throughput arrays for relative methylation [69]. A further variation of DMH, combined with array, is the method called microarray-based methylation assessment of single samples, which uses a cocktail of methylation-sensitive restriction enzymes to digest one pool of DNA and McrBC to digest the other pool [70]. Another method, known as HpaII tiny fragment enrichment by ligation-mediated PCR (HELP), uses ligation-mediated PCR for the amplification of HpaII or MspI genomic restriction fragments followed by array hybridization [71].

Finally, another affinity enrichment-based method involves the coupling of MeDIP with DNA microarrays to obtain relative methylation levels at the loci represented on the array. Defining a region as methylated from the MeDIP-chip data depends on several factors, including the sequence context of the loci (CpG-rich versus CpG-poor regions), the number of probes in a region, and their hybridization values. DNA methylation status is expressed relative to the input DNA or to another sample. This MeDIP-chip has been widely used in the analysis of cancer cell lines [20, 72] and stem cells [73, 74], among others. An interesting approach was used by Bell et al. [75] in the analysis of genotype-epigenotype interactions in the context of type 2 diabetes, finding an increased DNA methylation on the FTO obesity susceptibility haplotype by using methylated DNA immunoprecipitation on a targeted array (MeDIP-chip).

However, methylation changes identified by microarray experiments will continue to require validation by one or more of the locus-specific methods [76]. Moreover, the relative fineness of DNA methylation changes induced by nutritional exposures will motivate continued development of analytical methods that should be rapid, inexpensive, highly sensitive and quantitatively accurate. On the contrary, although methylation microarrays are a powerful tool for epigenetic studies, they have an important limitation in that they analyze only a small part of the CpG sites of the genome.

**NGS-Based Methylation Analysis**

NGS has revolutionized the genomic research field since its introduction in 2007 [77]. More recently, this technology has been adapted to epigenomic research including DNA methylation profiling at high resolution and low cost. Several sequencing strategies are available and the choice of platform and methodology is dependent on the scientific application, the cost, and the capacity for extensive bioinformatics analysis.

Sodium bisulfite conversion followed by massively parallel sequencing (Bisulfite-seq) has become an increasingly used method for performing epigenetic profiles in the human genome [78]. Bisulfite-seq is well suited to the investigation of epigenetic profile from clinical tissue samples [79, 80], and can be applied to very small quantities of DNA [81] including formalin-fixed samples [82]. Comprehensive mapping of DNA methylation in relevant clinical cohorts is likely to identify new disease genes and potential drug targets, help to establish the relevance of epigenetic alterations in disease and provide a rich source of potential biomarkers [83].

Whole-genome bisulfite sequencing provides coverage at a single-base pair resolution and is the most comprehensive technique, covering more than 90% of cytosines in the human genome [9]. This method requires the most extensive bioinformatics tools and is the most
expensive because more sequences are needed to cover the entire genome. An alternative to the whole-genome bisulfite sequencing is RRBS, which uses restriction fragment size selection to select a portion of the genome enriched for CpG islands and gene regulatory sequences [11, 84]. Specifically, this technology covers approximately 12% of all CpG dinucleotides and 84% of all CpG islands in the human genome [84].

As with microarray analysis, approaches focused on affinity-based sequencing, such as MeDIP-seq or MethylCap [methyl-DNA binding domain (MBD) capture technique], or enzymatic-based sequencing technologies, like Methyl-Seq and HELP-Seq have been described [85].

More recently, a new chemistry called oxidative bisulfite sequencing (oxBS-Seq) has allowed to sequence the recently discovered base 5-hmC in whole-genome studies [86]. Although there are other techniques to quantify this modification such as MeDIP (array or sequencing) or competitive enzyme immunoassay with specific 5-hmC antibody [29, 87], or oxBS-Seq development, 5-hmC can be detected at single-base resolution. This method is based on the selective chemical oxidation of 5-hmC to 5-fC, which enables bisulfite conversion of 5-fC to uracil. It has been applied to sequence 5-hmC in embryonic stem cell genomic DNA, and high levels of 5-hmC were found in CpG islands associated with transcriptional regulators and in long interspersed nuclear elements, suggesting that these regions might undergo epigenetic reprogramming. Furthermore, the global balance between 5-mC and 5-hmC in the genome has been described as a critical step for regulating gene expression to maintain cellular functions [88–90]. In this sense, the global quantification of 5-hmC is becoming crucial to understand dynamic epigenetic mechanisms and changes that occur in the cells. For example, it has been reported that the aging process is connected with changes in the epigenetic machinery in mouse hippocampus, and caloric restriction acts as an inducer of a complex interplay involving chromatin remodeling and the balance between methylation and hydroxymethylation [91].

Despite these advances, and probably because of its high cost, the study of methylation changes by using these new approaches is slowly finding its way into metabolic and nutritional studies. In spite of this, further studies in this area will begin to emerge in the coming years.

**Conclusion**

A large number of strategies have been used to examine DNA methylation, and these methods have progressed from small-scale candidate gene analysis to the ability to construct whole-genome methylation profiles. Clearly, there is no one method of DNA methylation analysis that is best suited for every application. Researchers should conduct a preliminary analysis focused on the type of validation and information provided by each technique in order to select the best method fitting for their research.

Furthermore, it has recently become apparent that, in addition to DNA methylation, there are other modifications of DNA, such as hydroxymethylation, that play a role in epigenetic regulation of gene expression. These findings suggest that DNA modifications are more dynamic than previously thought. In the next few years, given the speed with which NGS has been adopted in epigenetic research and turned into an effective tool for hypothesis testing and generation, important nutrient-gene interaction questions will likely find an answer. The main limitations will be the following: often in research the hypotheses must be clear and as simple as possible, and the experimental design adequate for the purpose, avoiding confounding factors associated with environmental and nutritional variables. This task is difficult in a field in which many dietary components (fat, methyl donors, polyphenols, amino
acids, minerals) and metabolic processes (obesity, inflammation, oxidative stress, hyperglycemia, stress, endocrine imbalances, etc.) have been reported to be involved in the epigenetic control of gene expression, especially during the in utero and early postnatal periods [84].

Acknowledgements

This work was supported by Línea Especial about Nutrition, Obesity and Health (University of Navarra LE/97). M.L.M. holds a Juan de la Cierva fellowship from the Spanish Ministry of Economy and Competitiveness. The CIBERObn scheme is also gratefully acknowledged.

Disclosure Statement

The authors declare that they have no competing interests.

References


