ABSTRACT Methylation has been implied in a number of biological processes and has been shown to vary under environmental influences as well as in age. Most results on the correlation of methylation patterns with phenotypic characteristics of cells have been obtained by analysis of very few or even single genomic fragments for methylation. However, variation of methylation may more often than not be a phenomenon that affects multiple genomic loci. The role of methylation has been most conclusively demonstrated in complex disease, with cancer being the most prominent example. The influence of aging and environmental influences such as diet seems to be on global methylation patterns, in turn exerting local effects on groups of genes. Hence, methylation seems literally to be orchestrating complex genetic systems. It could, therefore, be considered an archetypal “genomics” parameter. In consequence, technologies used to analyze methylation patterns should be as industrialized as possible to capture the local events across the entire genome. Epigenomics’ research team is the first to have achieved the industrialized production of genome sequence–specific wide methylation data. Our microarray and mass-spectrometry–based detection platform currently allow the analysis of up to 50,000 methylation positions per day, for the first time making methylation data amenable to sophisticated information mining. The information content of methylation position has never been analyzed using the high-dimensional statistical methods that are recognized to be required for the analysis of, for example, mRNA expression profiles or proteomic data. As methylation patterns are nothing but a quasi-digital form of expression data, their information content must be evaluated using similar but adapted algorithms. This article presents a broad set of studies that demonstrate that methylation yields information that is comparable or even superior to the current state of the art, namely, mRNA profiling. We argue that the resulting robust, digital and—because of the highly stable nature of DNA as the analyte—more reproducible information could become the “gold standard” for clinical diagnostics and disease gene identification in age-related, environmentally influenced and epigenetic disease in general, substituting for mRNA expression. J. Nutr. 132: 2440S-2443S, 2002.

KEY WORDS: • DNA methylation • type 2 diabetes • microarray-based technology • methylation profiling

Type 2 diabetes mellitus comprises a heterogeneous group of mild forms of diabetes, compared to type 1 diabetes. Patients with type 2 diabetes can be grouped on the basis of body weight into obese and nonobese subtypes, but further subclassification based on clinical parameters such as circulating insulin levels and other laboratory parameters currently is not possible. Up to 85% of patients with type 2 diabetes are obese, with an abdominal distribution of fat and an abnormally high waist to hip ratio. One of the most prominent laboratory findings is an insensitivity of fat, muscle and liver cells to insulin.

Genetic predisposition is a major risk factor for developing type 2 diabetes. It has been suggested that polymorphisms in genes involved in insulin secretion and response might modify individual disease susceptibility; however, in large population-based studies only a few polymorphisms in such genes could be shown to influence the incidence of diabetes (1–3). Recent studies have shown that the development of obesity and type 2 diabetes is associated with changes in the expression levels of several genes (4–9). In a mouse model of type 2 diabetes and obesity, progression from a lean state to obesity and to overt hyperglycemia was found to be associated with changes in gene expression inverse to those seen in adipocyte differentiation (10). In a similar study, human gene expression in omental fat from lean and obese nondiabetic subjects and obese type 2 diabetic patients was analyzed (11). Over 200 cDNAs that showed potential differential expression in the
omental fat of lean versus obese nondiabetic subjects versus obese type 2 diabetic patients were identified. Furthermore, upregulation of several genes in response to insulin was completely abrogated in type 2 diabetic patients compared to control subjects, insulin-resistant nondiabetic obese patients and hyperglycemic type 1 diabetic subjects (12). Interestingly, several oral antidiabetic agents have been shown to correct altered expression found in animal models of type 2 diabetes (4,13).

DNA methylation during cellular differentiation

DNA methylation is a major regulator of transcriptional activity and is involved in the regulation of expression of a broad variety of genes. Changes in methylation patterns have been described for several genes during differentiation in a broad spectrum of cell types (14–18). For example, T helper cells differentiate into T helper 1 (Th1) and T helper 2 (Th2) T-cells following stimulation with antigen and the cytokines interleukin-12 (IL-12) and IL-4, respectively. Th1 and Th2 cells exhibit striking differences in their patterns of cytokine expression. The subtype-specific expression pattern is maintained in resting Th1 and Th2 cells in the absence of further stimulation. The persistence of these expression patterns is thought to be due to epigenetic changes in chromatin structure, locus accessibility and DNA methylation (19,20). Also, methylation patterns in the interferon (IFN)-γ promoter exhibit long-term faithful inheritance in T cells and their progeny, through >10 cell divisions and a clonal expansion of 5 orders of magnitude. Moreover, the demethylated IFN-γ promoter is faithfully inherited following the withdrawal of T cell stimulation and the loss of detectable IFN-γ mRNA (21). Therefore, DNA methylation represents a stable cellular memory for the maintenance of cytokine expression patterns and T cell memory and may contain more information about the differentiation state of a population than the actual expression patterns.

DNA methylation in type 2 diabetes

There is evidence that this DNA methylation memory also is involved in maintaining gene expression patterns associated with insulin resistance in type 2 diabetes mellitus:

- Prenatal glucose and insulin levels influence the risk of developing type 2 diabetes later in life, independent of the maternal type of diabetes and therefore independent of genetic predisposition (22). This suggests the presence of a cellular memory in insulin target tissues such as adipose tissue, skeletal muscle and liver.
- Several genes involved in (glucose) metabolism have been shown to exhibit differential DNA methylation in their promoters, e.g., facilitative glucose transporter 4, the major glucose transporter in adipose and muscle tissues (15), and uncoupling protein 2 (23), a major candidate gene for the development of type 2 diabetes. Recent insights into the pathogenesis of transient neonatal diabetes, a rare subtype of diabetes that is characterized by transient hyperglycemia in the neonatal period and a predisposition for diabetes in adult life, provide a link between methylation, gene dosage effects and diabetes. Transient neonatal diabetes results from doubling of the gene dosage of genes on chromosome 6q24. Paternal uniparental isodisomy, duplication of the respective band on 6q24 and loss of methylation in this imprinted region all result in phenotypically indistinguishable transient neonatal diabetes (24).

In addition to targeted DNA methylation changes in response to external stimuli and during cellular differentiation, random DNA methylation changes have been shown to occur during aging of organisms in several tissue types (25–28). Accumulating age-related DNA methylation changes are involved in a number of different diseases, e.g., atherosclerosis and cancer. In the colon, for example, hypermethylation often starts in normal mucosa as a function of age and leads to field defects with an increased risk of developing colorectal cancer (acquired predisposition to colorectal neoplasia) (28). Also, methylation-associated inactivation of the estrogen receptor α gene in vascular tissue has been suggested to play a role in atherogenesis and aging of the cardiovascular system (25).

Interestingly, DNA methylation of the promoter region of the amyloid precursor protein gene, which is involved in the development of Alzheimer’s disease, is reduced with increasing age (27). Type 2 diabetes is strongly age-related: not only is its incidence increased in older populations, but also the metabolic situation of individual patients deteriorates over time. DNA methylation errors that accumulate with increasing age could provide an explanation of both phenomena.

A general defect in DNA methylation in diabetes is suggested by the recent observation that S-adenosylmethionine (SAM), the main physiologic donor of methyl groups, is decreased in erythrocytes of diabetic patients. In addition, decreased erythrocyte concentrations of SAM and other alterations were found to be associated with disease progression (29). Taken together, methylation plays an important role in regulating gene expression, most likely including the expression of those genes essential for the strict maintenance of normal blood glucose levels. Aberrant expression patterns that develop in response to diet (for review, see L. Poirier at this workshop) (3), increased body weight (10–12) and environmental factors (see F. J. Corrales review at this workshop) are likely to become “locked” by DNA methylation if they occur over a longer period of time. DNA methylation, therefore, is likely to be involved in the propagation of insulin resistance in insulin target tissues and, being a reversible modification, might also confer the adaptability of metabolism to loss of body weight. On the other hand, metabolism of methyl groups may be affected by diet (see reviews of S. J. James, L. Poirier and S.-W. Choi at this workshop), body weight and environmental factors (30), thus leading to untargeted, general hypomethylation of DNA in diabetic patients (29). Moreover, DNA methylation errors have been shown to accumulate over time, contributing to many age-related diseases. These errors could add to the development of type 2 diabetes by reducing gene responsiveness (i.e., gene expression) that needs to be adjusted to fast changing glucose levels.

DNA methylation profiling to identify new targets in type 2 diabetes

As outlined above, there is considerable evidence that aberrant DNA methylation patterns are laid down early in the development of insulin target tissues, which may lead to overt insulin resistance later in adult life. Also, DNA methylation patterns are subjected to age-related changes that may promote the development of metabolic diseases with age. As the underlying “expression memory” of insulin target tissues, DNA methylation information can be used to identify new and confirm known potential drug targets. In general, DNA methylation profiling has several advantages compared to other
Genomewide profiling approaches, e.g., mRNA expression analyses:

- DNA methylation picks up permanent expression changes rather than short-term alterations—fewer false positives will occur (e.g., temporary expression changes).
- DNA methylation detection picks up changes which cannot be picked up by any other method as it exploits a totally new information layer.
- DNA methylation is suited to being analyzed in very large populations, therefore subclasses can be detected—the pathogenesis of type 2 diabetes is likely to be heterogeneous.
- DNA methylation has technical advantages, including the lack of particular requirements for sample handling (e.g., transfer procedures from different hospitals), and tissue samples can be analyzed retrospectively (e.g., paraffin-embedded tissues).

Genomewide screening for new drug targets

In a first step, Epigenomics applies genomewide screening technologies to identify new potential drug candidates. Shortly, these technologies are based on restriction of genomic DNA with methylation-sensitive enzymes, which cut their recognition site only if it is unmethylated (31–34). All of these methods are completed unbiased; i.e., no assumptions regarding candidate genes have to be made. They are suitable to compare genomewide methylation patterns between two samples or two pools of samples, and result in discovery of several hundred differentially methylated sequences tags that can be mapped to promoter regions of genes in most cases.

Confirmation of potential targets in large populations: microarray-based methylation profiling

Epigenomics has developed a microarray-based technology to screen thousands of cytosine guanine dinucleotides (CpG) sites in large populations. The method is based on bisulfite treatment of genomic DNA from all samples, thereby converting all unmethylated cytosines to uracil whereas methylated cytosines are conserved (35). Regions of interest are then amplified by polymerase chain reaction (PCR) using fluorescently labeled primers conserving originally unmethylated CpG dinucleotides to thymine guanine (TG) and conserving originally methylated CpG sites. Primers are designed complementary to DNA segments containing no CpG dinucleotides. This allows unbiased amplification of both methylated and unmethylated alleles in one reaction. Multiplex PCR reactions are used to amplify several fragments in one reaction. All PCR products performed on an individual sample are mixed and hybridized to glass slides carrying a pair of immobilized oligonucleotides for each CpG position. Each of these detection oligonucleotides is designed to hybridize to the bisulfite-converted sequence around one CpG site that was either originally unmethylated (TG) or methylated cytosine guanine (CG). Hybridization conditions are selected to allow the detection of the single nucleotide differences between the TG and CG variants. Different bioinformatic approaches, such as learning algorithms and clustering analyses, are used to evaluate the highly complex data obtained by this approach (36,37).

This technology allows Epigenomics to assess differentially methylated sites identified in genomewide screening experiments in very large populations.

We have printed microarrays with methylation positions from a comprehensive list of candidate genes involved in metabolism. Along with methylated sequence tags derived from genomewide discovery studies, candidates currently are being tested on different tissues from large numbers of patients, thereby achieving both the possible confirmation of potential candidates, as well as the statistical validation of genes newly discovered from few samples through blind screening of the genome.

Further steps

For genes confirmed to be altered in a large percentage of diabetics, the functional relevance of these alterations for the disease needs to be determined. This comprises two steps: analysis of effects of methylation alteration on protein levels (overexpression or abrogation of expression) and experimental proof that differing protein levels are responsible for the observed phenotype (classical target validation)

Assessing known potential targets in large populations

Taking together all known genes involved in glucose and lipid metabolism, there already are a number of potential candidates for drug targets. The problem is that some might actually constitute excellent targets from a pharmacological point of view, but little is known about their contribution to type 2 diabetes in a larger population. DNA methylation can be used to assess their potential role in the general population and thereby help to reduce the number of genes to enter "classical" validation. In this scenario, the candidate genes will be printed on chips and analyzed in hundreds of patients and tissues.

SUMMARY

When taken as a whole, there is considerable indirect evidence that aberrant DNA methylation plays a role in the development of type 2 diabetes. Two mechanisms are likely to contribute to the pathogenesis: First, imprinting of feedback loops in insulin target tissues such as adipose, muscle tissue and liver. DNA methylation has been shown to function as a cellular memory, and it provides the major mechanism by which expression patterns and response to stimuli can become heritable over many cell generations. In many cell types, expression patterns and differentiation pathways are “locked” at a certain time point. Second, untargeted DNA methylation errors, e.g., hypomethylation changes associated with decreased levels of the methyl group donor SAM, have been shown to accumulate with increasing age—this could account for the increasing incidence of type 2 diabetes in older individuals.

Combining the array of technologies developed by Epigenomics, DNA methylation information can be used to screen for new gene targets. The target identification process involves several steps from genomewide screening, assessing the prevalence of alterations in larger populations to classical target validation and drug development steps. In addition, DNA methylation profiling will constitute an important tool to assess the prevalence of dysregulations of known potential candidates. Ultimately, this approach can lead to mechanism specific and therefore highly effective, oral antidiabetic drugs that are tailored for particular subgroups of patients.

LITERATURE CITED


